Discovery of Infectious Disease Biomarkers in Murine Anthrax Model Using Mass Spectrometry of the Low-Molecular-Mass Serum Proteome

Aarthi Narayanan¹, Weidong Zhou², Mark Ross³, Jane Tang¹, Lance Liotta³, Emanuel Petricoin², Fatah Kashanchi⁴, Charles Bailey¹, Serguei Popov⁴

¹National Center for Biodefense and Infectious Diseases  
²Center for Proteomics and Molecular Medicine, George Mason University, 10900 University Blvd Manassas, Virginia – 20110-2201  
³Noblis, 3150 Fairview Park Drive, Falls Church, Virginia – 22042  
⁴The George Washington University, 2121 Street, NW, Washington, DC 20052

Abstract

A novel mass spectrometry (MS)-based approach to the identification of host-derived biomarkers (BMs) in the circulating low-molecular-mass (LMM) fraction (<25 kDa) of blood proteome was tested in a murine model. DBA/2/J mice were challenged intraperitoneally with spores of either the toxigenic B. anthracis Sterne strain (pXO1+, pXO2-) that is virulent in DBA/2 mice or the non-toxigenic, non-virulent delta Sterne strain (pXO1-, pXO2-). Serum samples were obtained at multiple time points and separated by continuous flow denaturing gel electrophoresis followed by Coomassie staining to isolate the LMM archive for subsequent MS identification. Peptide fragments derived from more than 200 proteins displayed low-variance differential abundances between lethal and non-lethal challenges. Several proteins from the MS analysis were subjected to secondary verification by western blots. Serum abundances of 6 proteins (carbonic anhydrase 2, adenylate kinase 1, peroxiredoxin 2, UMP-CMP kinase, Ras-related C3 botulinum substrate 1, and dextrin) from a total of 10 tested proteins were strongly coincident with established anthrax disease and mortality thus making them potential candidates for host-derived anthrax disease associated BMs. These BMs were demonstrated to be “elastic” in that their abundance levels in sera of doxycycline-treated mice responded to the therapeutic intervention thus making them useful tools for monitoring efficacies of existing and novel treatment regimens.

Keywords: Anthrax; Biomarkers; Proteome; Mass Spectrometry

Abbreviations: MS: Mass spectrometry; BM: Biomarker

Introduction

Anthrax is a serious bacterial infection of humans and many animal species caused by Bacillus anthracis. Effective prophylaxis and treatment of inhalational and gastrointestinal anthrax remain a challenging task in spite of the modern advances in the antimicrobial therapies and vaccines. A quick progression of the disease takes place without specific clinical symptoms, and patients often seek medical help at the stage of disseminated infection, which historically has been highly lethal regardless of the treatment (Holty et al. 2006). The mortality rate in the U.S. inhalational anthrax attack in 2001 was reduced to 55%, which still remains unacceptably high, underscoring the need for new therapies and early diagnostics (CDC 2001; Jernigan et al. 2001). Identification of biomarkers (BMs), which can serve as diagnostic tools, and indicators of pathological process has become an active area of research in recent years. However, little information is currently available about potential anthrax BMs. The presence of the bacterial protective antigen in host serum has been suggested to be a useful BM in case of an exposure event (Hammamieh et al. 2008; Kobiler et al. 2006). Alterations in host gene profiles following an exposure to B. anthracis have also been pursued as a means of identification (Das et al. 2008). While all of the above methods offer the ability to successfully identify the pathogen in the case of an infection, no systematic effort has been reported to discover BMs relevant to the host response at various stages of anthrax progression.

In the current report we chose to test in a murine model a novel mass spectrometry (MS)-based approach for the identification of host-derived BMs in the circulating blood proteome that utilizes continuous flow denaturing electrophoresis to isolate and enrich the low molecular mass (LMM) analyte pool prior to MS analysis. This approach, while being successfully applied to cancer biomarker discovery (Camerini et al. 2007), has not been used for infectious disease biomarker discovery. While pathogen-derived BMs can offer a high degree of specificity, low abundance of such BMs at the early stages of infection may represent a considerable problem for a timely diagnosis. In contrast, host-derived BMs relevant to the innate immune response, or generated by host cell changes immediately following exposure to the infectious agent or agent-derived analytes, could undergo significant biological “amplification”


Copyright: © 2009 Narayanan A. et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

J Proteomics Bioinform  Volume 2(9) : 408-415 (2009) - 408
by host response cycles during the early stage of infection and therefore may have a distinct advantage over pathogen-derived BMs. An intrinsic broad-spectrum nature of the host-derived BMs may also be useful for a quick triaging of patients for further specific analyses.

Blood is often a primary choice for a BM discovery because it is expected to comprehensively reflect the health state of the organism associated with a systemic disease such as anthrax and can be easily procured. However, a global analysis of the blood proteome is a challenging task, especially in the case of high-molecular-mass or low-abundance proteins. Recently, the low-molecular-mass (LMM) proteome/peptidome of blood has received attention for BM discovery because of the potential that small proteins and fragments of larger proteins produced in the disease microenvironment can appear in the bloodstream by freely crossing the endothelial cell barrier (Anderson, Anderson 2002; Faca et al. 2008; Liotta et al. 2003; Petricoin et al. 2006) or produced within the circulation as part of the innate infectious process. MS is a powerful and a sensitive approach that has been employed in discovery of BMs relevant to prostatic and ovarian cancers and other non-neoplastic diseases such as Alzheimer’s disease (Diamandis EP 2004; Geho et al. 2006; Ornstein, et al. 2004; Petricoin, et al. 2002; Petricoin, Liotta 2003; Sardana et al. 2008). We suggested that MS, which is most sensitive in the range of analytes with molecular mass < 25 kDa, would be especially suitable for the discovery of LMM host-derived BMs of infectious diseases.

Initially, we identified a large number of preliminary BM candidates correlating with the progression of murine anthrax and virulence in comparative challenge experiments with virulent and non-virulent B. anthracis strains. Several proteins from the MS analysis were verified using western blots. Finally, we demonstrated that these proteins are remarkably ’elastic’ in that their abundance levels in serum respond to antibiotic therapeutic intervention, thus opening up additional avenues for use of these BMs in monitoring efficacies of treatment regimens.

Materials and Methods

Animal Experiments

Female DBA/2j mice were purchased from Jackson Laboratory (Bar Harbor, ME) and maintained at Biocin Inc. (Rockville, MD). Animals were allowed to acclimate to their surroundings for one week before experiments were performed. The George Mason University Institutional Animal Care and Use Committee and Biocon Animal Care and Use Committee/Institutional Review Board have approved all protocols that were used in the animal experiments. Mice were challenged intraperitoneally (i.p) with a water suspension of 5 x 10° spores of B. anthracis non-encapsulated Sterne strain 34F2 (pXO1, pXO2) or the non-toxigenic delta Sterne strain (pXO1, pXO2). Eight mice were used per Sterne strain challenge group, and three mice per group in all other experiments. The Sterne and the delta Sterne strains were obtained from Colorado Serum Company (Boulder, CO) and National Center for Biodefense and Infectious Diseases (Manassas, VA), respectively (Bradburne, et al. 2008). For MS studies, mice were sacrificed at 24, 48 and 72 hours post exposure, and 200 to 500 µl of blood were drawn from each animal via retro-orbital sinus. Blood samples were allowed to clot overnight at 4°C and sera were separated by two centrifugations at 400g for 10 min each. Samples from each group corresponding to a given time point were pooled for MS analysis. For studies of the individual animal responses, blood was obtained from the tail vein. At the end of the experimental period the surviving animals were terminally bled as described above. Doxycycline (Sigma) was administered i.p. on days 1 to 4 post infection at the dose of 10 mg/kg of body weight in sterile water. Control groups were injected with sterile water only.

Sample Preparation and Nanoflow Reversed-phase Liquid Chromatography Tandem MS

LMM proteins were fractionated using a Prep Cell gel electrophoresis device from BioRad and the fractions were then used for electrophoresis by SDS-PAGE (4-20% Tris-glycine, Invitrogen). After Coomassie staining using standard procedures, each lane of the gel was sliced into 4 parts in order to reduce the complexity of further analysis. For each piece, in-gel digestion was performed using trypsin, and the resultant peptides were eluted from the gel (Camerini et al. 2007; Sardana et al. 2008; Shevchenko, et al. 1996) for analysis by a reversed-phase liquid chromatography nanospray tandem MS using high-resolution LTQ-Orbitrap spectrometer (ThermoFisher). The reverse-phase column was slurry-packed in-house with 5 µm, 200-Å pore size C18 resin (Michrom BioResources, CA) in 100 µm x 10 cm fused silica capillary (Polymeric Technologies, Phoenix, AZ) with a laser-pulled tip. After sample injection the column was washed for 5 min at 200 nl/min with 0.1% formic acid, peptides were eluted using a 50-min linear gradient from 0 to 40% acetonitrile and an additional step of 80% acetonitrile (all in 0.1% formic acid) for 5 min. The LTQ-Orbitrap MS was operated in a data-dependent mode in which each full MS scan was followed by five MS-MS scans where the five most abundant molecular ions were dynamically selected and fragmented by collision-induced dissociation (CID) using normalized collision energy of 35%. Tandem mass spectra were matched against the National Center for Biotechnology Information mouse database through the Sequest Bioworks software (ThermoFisher) using full tryptic cleavage constraints and static cysteine alkylation by iodoacetamide. For a peptide to be considered legitimately identified, it had to be the top number one matched and had to achieve cross correlation scores of 1.9 for [M+H]+, 2.2 for [M+2H]2+, 3.5 for [M+3H]3+, ΔCn > 0.1, and a maximum probability of randomized identification of 0.01. The MS data were filtered to improve the quality of the data set prior to BM selection. The initial set of proteins was limited to those that could be confidently identified, and further screened to remove proteins with few non-zero peptide hits. The peptide hit numbers were normalized on the scale of 0 to 1, and proteins with high inter-sample abundance variability were excluded from consideration. For further analysis, MS data were classified into 10 groups based on common patterns of expression using K-means cluster analysis performed using the data mining tool Weka (Witten, Frank 2005).

Western Blot Analysis

Total protein concentration of the serum samples was determined by using colorimetric Bradford reagent (BioRad) and a Microquant plate reader (Bio-Tek Instruments, Inc.). Bovine
Results

Identification of Differentially Expressed LMM Protein BM Candidates

Mice were infected with spores of either the toxigenic Sterne strain or the non-toxigenic delta Sterne strain of Bacillus anthracis in order to identify LMM blood-borne BMs that are specific to the toxigenic exposure. The Sterne strain is lethal in DBA/2 mice, while the delta Sterne strain causes a self-limiting infectious process due to the absence of the toxigenic plasmid. We therefore expected that a comparison of serum protein profiles between mice challenged with these strains would help reveal the BMs specific for the virulent infection and therefore distinct from the general innate responses common to different infectious agents. In the conditions of our experiments, the Sterne strain infectious process is characterized by an onset of mortality beginning day 3-post challenge. Therefore, blood samples taken at this time are likely to closely reflect pre-mortal conditions of challenged animals. At later time points, a quick progression of disease makes it difficult to collect a sufficient number of samples and may even result in skewed data due to analysis of survivors mice representing a cohort of animals with higher resistance. Serum samples from the infected mice taken during the course of infection at 0, 24, 48 and 72 h post challenge were pooled and the LMM protein fraction with a molecular mass below 25kDa was separated by a preparative cell gel electrophoresis with the aid of a molecular marker run alongside. The preparative cell gel electrophoresis was used as the method of choice as it offers higher accuracy of separation and prevents loss of sample during the fractionation procedure when compared with cut-off spin filters (Camerini et al. 2007). The gel was cut into four pieces, and proteins from each of the gel regions were subjected to an in-gel digestion with trypsin followed by high-mass-accuracy LC-MS-MS as described in Materials and Methods. It is important to note that the peptides identified in the tryptic hydrolysates have originated either from the intact protein molecules present in the sample, their fragments generated in the course of the infectious process, or degradation of the sample during its preparation (Yi et al. 2007). The resultant spectra were analyzed using Sequest Bioworks software. Spectral counting has been used as an approach to estimate relative abundances of differentially expressed candidate proteins (Faca et al. 2008). To reduce the impact of noise on BM selection, proteins with low spectral counts were screened out of the 1200 total identified proteins. 213 proteins with 10 or more total peptide hits were selected for further evaluation.

Owing to the elaborate nature of the experiment, the indicated MS data set that was obtained from one MS run was analyzed for inter-sample variability to estimate reliability. As an estimate of the data variability, for each protein the standard deviations of hits for control mock-injected mice at 0 time point were found and normalized by dividing by the hit counts averaged across all time point readings. 208 proteins showing relatively low variance between 0 and 1.5 were subjected to further analysis. Supplementary Table shows the identity of the first the identity of the first 100 proteins sorted by the highest difference between the total numbers of Sterne and delta Sterne hits relative to Sterne ones in order to reveal the proteins with the property to respond to virulent infection by Sterne strain but not to abortive, self-limiting infection by delta Sterne strain. K-means cluster analysis performed using the data for 208 proteins with low variance identified proteins that displayed common trends in their responses to the infection. This also helped to further narrow the number of BM candidates. MS data were first normalized to a scale from 0 to 1 by dividing all peptide hit numbers for a protein by the highest value among all time points. Following normalization, delta Sterne hit values for each protein were subtracted from Sterne values and sorted into 10 groups based on the maximum similarity within members of each group and maximum distance between the centers of the groups.

<table>
<thead>
<tr>
<th>BM</th>
<th>% occurrence</th>
<th>Average % occurrence (AM ± SD)</th>
<th>Expected range of occurrence (AM ± 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRX2</td>
<td>n/a</td>
<td>100% ± 7.5</td>
<td>95.6% ± 8.48</td>
</tr>
<tr>
<td>UCK</td>
<td>100</td>
<td>93.3% ± 11.5</td>
<td>93.3% ± 13.05</td>
</tr>
<tr>
<td>CA2</td>
<td>80</td>
<td>78.3% ± 20.2</td>
<td>78.3% ± 22.85</td>
</tr>
<tr>
<td>RAC1</td>
<td>60</td>
<td>74% ± 13.5</td>
<td>74% ± 15.27</td>
</tr>
<tr>
<td>DST</td>
<td>60</td>
<td>72.3% ± 9.2</td>
<td>72.3% ± 10.51</td>
</tr>
</tbody>
</table>

* For each BM candidate, the number of times an increasing trend was observed in each experiments is shown; n, the number of mice analyzed in each experiment; AM, arithmetic mean; SD, standard deviation of mean; CI, confidence interval. n/a, not analyzed.

Table. BM occurrence in 3 challenge experiments*.
Figure 1: A. Protein abundance profiles of candidate BMs in Sterne exposed mice. The six western blot panels represent profiles of the candidate BM proteins in sera of randomly selected mice exposed to either the Sterne strain of *B. anthracis*, or the non-lethal delta Sterne (d-Sterne) strain. One microliter of serum was suspended in SDS running buffer for evaluation of protein levels for each of the proteins indicated. All samples were normalized by taking into consideration total protein levels of each of the serum samples. The control panel for each of the proteins tested refers to the protein level in sera of uninfected mice. The * in the Sterne lanes indicate unavailability of serum samples post day 3 as Sterne-infected mice die by that time. The delta Sterne-infected mice and control uninfected mice were followed up to 10 days post infection. For all proteins, gels were cut approximately at the 30 kDa mark as determined using Magic mark XP western protein standard (Invitrogen) and used for western blots.

B. Quantitative representations of fold differences in the abundances of the candidate proteins upon lethal and non-lethal exposure conditions. Quantitations were carried out by measuring band intensities in western blots using Quantity One software. All data points are represented with arithmetic means and corresponding standard deviations (n=4 for Sterne-exposed mice; n=3 for delta Sterne-exposed mice).
Figure 2: A. Abundance profiles of candidate BMs in Sterne-exposed mice treated with doxycycline. Sera obtained at time points indicated were probed in western blots with antibodies to each BM candidate. Western blot bands corresponding to all indicated proteins except CP were obtained by excising the gel at approximately below the 30 kDa mark as mentioned earlier. For CP, the bands were obtained by using the portion of the gel above the 80 kDa mark of the Magic Mark XP western blot protein standard.

B. Abundances of the candidate proteins in Sterne-exposed mice treated with doxycycline relative to control unchallenged, antibiotic-treated mice. Error bars show standard deviations of arithmetic means (n=3).
Ten proteins from our MS results, namely carboxic anhydrase 2 (CA2), adenylate kinase 1 (AK1), peroxyredoxin 2 (PRX2), UMP-CMP kinase (UCK), destrin (DST), ceruloplasmin (CP), Ras-related C3 botulinum substrate 1 (RAC1), Necdin, ADP ribosylation factor (ARF), Serum amyloid A (SAA) and ubiquitin conjugating enzyme E2N were chosen for confirmatory studies with western blots based on the (i) strong trends in our cluster analyses, (ii) low variance between multiple samples, (iii) commercial availability of highly-specific antibodies. In order to account for the variability of responses between animals, all further analyses were carried out with blood samples obtained from individual mice. From the list of 10 candidate BMs, 6 proteins displayed an expected increase in serum level as a result of the lethal Sterne exposure, but did not increase in mice exposed to the non-lethal anthrax strain (Fig. 1A). Quantitative reflections of Fig. 1A for all tested samples normalized by the total protein concentration of the serum are presented in Fig.1B. The MS spectra corresponding to these proteins were also subjected to manual verification to confirm their identification (Supplemental Fig. S1). While levels of CA2 and AK1 increased considerably upon exposure to toxigenic strain, PRX2 and UCK were moderate responders, RAC1 showed a moderate to a weak response and DST was a weak responder. The trend reflected in Fig. 2 correlated strongly with physical symptoms of established anthrax disease prior to death (data not shown). The Table shows the results of 3 independent challenge experiments in which blood samples drawn from each individual mouse were tested with a panel of antibodies against the selected BM candidates. PRX2, UCK, and CA2 are the most reliable BM candidates. Although more rigorous analysis would be necessary for the potential BM to be accepted for clinical practice, the results of Table 2 indicate that simultaneous analysis of these three proteins could be a reliable indicator of infection with a false negative rate below 1%.

Effect of Antibiotic Intervention on Serum Protein BM Profiles

BM s for monitoring the progress of treatment and host recovery process are highly desirable. We tested if the BMs we identified were suitable for these purposes. Sterne-infected mice were treated with the antibiotic doxycycline starting one day post exposure for up to 5 days. Doxycycline is the antibiotic of choice for treatment of human anthrax and is highly effective in protection of mice up to 48 h post lethal exposure to B. anthracis (S. Popov, unpublished observation). A group of uninfected control mice was also treated with antibiotic to control for its direct effect on the levels of BMs. Blood samples from tail vein were obtained for each mouse. All Sterne-exposed mice that were not treated with the antibiotic succumbed to the infection by 5 days post the initial exposure, while all treated Sterne-exposed mice survived. As shown in Fig. 2, all tested BMs in the Sterne-exposed antibiotic-treated mice displayed serum levels similar to those in control unexposed mice receiving the antibiotic only. Ceruloplasmin (CP) was used as a negative control attesting to the specificity of the observations with the positive BM candidates.

Discussion

In the present study we employed a novel LMM enrichment and isolation strategy using continuous flow electrophoresis coupled to MS to explore a utility of the LMM proteome for the discovery of host-derived infectious disease BMs. We focused our discovery efforts on the identification of candidate BMs specific to a murine anthrax model whereby toxigenic and non-toxigenic strain responses were compared. The fact that a LMM fraction of serum proteome displayed an altered composition in response to infection is an important proof of principle that this source contains significant information reflecting the health state of organism and can be tapped for the BM discovery in blood samples of human patients.

Our results revealed peptides corresponding to a large number of proteins with a wide range of abundances. We performed extensive verification of six proteins to confirm the alterations in abundances in response to the lethal Sterne strain. Although our evaluation was focused on the proteins with highest number of spectra, which have been generally shown to be well-correlating with protein levels (Faca et al. 2008) independent MS experiments (not shown) confirmed a reproducible identification of analytes with low peptide hits, thus indicating a possibility to further explore a large number of low-abundant analytes. K-means cluster analysis of the MS data proved useful to categorize proteins into distinct response groups during the selection of candidate BMs such as those with a sustained increase in the lethally-infected mice relative to the survivors. While the selected BMs were validated with reasonable success rate using western blots as early as 24 h post infection, they also displayed an elastic nature in their responsiveness to the antibiotic treatment. As the sensitivity of the immunoassays could be further increased, the human analogs with similar properties could be quite useful in a clinical setting. Significantly, a panel of the three most reliable BMs, namely CA2 (about 10 fold increase), PRX (about 5 fold increase), and UCK (about 5 fold increase), is expected to have a low false-negative rate.

The appearance of these BMs in the serum proteome following exposure to Bacillus anthracis provides some insights regarding the mechanism of anthrax pathogenesis. CA2 is a physiologically important enzyme that among other functions catalyzes the reversible hydration of carbon dioxide to bicarbonate. There are several possible explanations for the increased abundance of CA2 in anthrax. Acidosis is observed in the case of anthrax pathogenesis in humans (Mina et al. 2002), and CA2 might be an important contributor in the regulation of pH homeostasis. The proinflammatory cytokine IL-1 upregulated during anthrax infection in mice is known to increase expression of CA2 in murine cell lines (Popov et al. 2004). Alternatively, increased bicarbonate levels in the host are required for the expression of B. anthracis virulence factors (Klichko et al. 2003; Wilson, et al. 2008). In support of the CA2 role in anthrax pathogenesis, inhibitors that prevent the interconversion of carbon dioxide and bicarbonate significantly affect virulence of Streptococcus, Clostridium botulinum and Vibrio cholerae (Wilson et al. 2008).

PRX2, RAC1, and DST are either directly or indirectly involved in oxidative stress responses. Oxidative stress responses are likely to be important regulators of Bacillus anthracis virulence and infectivity (Shatalin et al. 2008). RAC1 is a key component in the regulation of reactive oxygen species (ROS) by the NADPH oxidase complex in phagocytic and nonphagocytic cells with a wide range of functions in the immune and hypoxic.
responses, oxidative modifications of proteins, etc (Moldovan et al. 2006; Tobar et al. 2008). ROS are emerging as a key regulator of actin-based cellular network remodeling seen in many bacterial infections (Eswarappa et al. 2008; McGhie et al. 2004). Relevant to RAC1, DST is a molecule that is intimately involved in actin-mediated networks and therefore is likely to be influenced by oxidative stress (McGhie et al. 2004). PRX2 is an integral component of the redox system designed to combat the destructive effects of reactive oxygen species (Bindoli et al. 2008; Low et al. 2008). It seems likely that increased expression of PRX2 reflects a development of oxidative stress in anthrax. Along these lines, expression of adenylate kinases essential for meeting the energy demands of tissues in the conditions of stress such an infectious disease is also expected to increase. Although the upregulation of UCK, a cellular enzyme that converts uridine and cytidine monophosphates to diphosphates (Hutter et al. 2000), reflects strong pathophysiological alterations in the host, its link to the mechanism of disease is currently unknown.

It is obvious that not all of the BMs identified in our study are likely to be unique to anthrax. While research in this direction is ongoing, it is important to consider that although several of these proteins are likely to appear in multiple infectious scenarios, specific combinations of them could be unique to a given agent. On the other hand, the advantage of having BMs common to multiple infectious processes is likely to be seen in complex scenarios such as hospital-acquired infections, situations of bioterrorism, where multiple antigens are possible and are likely to be of uncertain origin. Common targets can be monitored to evaluate exposure to infectious bacterial agents. As an added advantage, generic antibiotic treatments can be employed and their effectiveness can be evaluated using the same BMs.

In summary, our LMM-focused blood-borne BM discovery using a novel LMM enrichment and isolation procedure based on continuous flow denaturing electrophoresis coupled with MS as a global proteome tool successfully generated several protein BM candidates displaying altered distribution in the mouse serum proteome during the course of lethal infection. Subsequent verification of selected candidate analytes using Western blotting confirmed the Sterne-specific differential abundance.

Efforts are underway to evaluate relevance of the identified murine BMs to B. anthracis infection in other animal models such as rabbits and non-human primates using virulent challenge strains in blinded validation studies. Follow up comparative studies are also planned to determine the BM appearance in the inhalation anthrax. It would also be very important to validate our BM candidates in the cases of lethal human exposures to anthrax.

Acknowledgements

This work has been supported by the USA Department of Energy grant DE-FCS2-04NA25455 and Noblis Sponsor Research. The authors thank Li Dong for technical assistance with animal experiments and Ann Kennedy for assistance with statistical studies.

References


