Published in final edited form as:

J Sep Sci. 2012 September ; 35(17): 2153–2161. doi:10.1002/jssc.201200216.

Multiple liquid chromatography separations and nanoESI-ion trap detection of plasma proteins in search of stroke biomarkers: A pilot study

Phanichand Kodali¹, Agnese Jurkevica¹, Julio Landero¹, Christopher Kuhlmann¹, Joseph Caruso¹, and Opeolu Adeoye²

¹Department of Chemistry, University of Cincinnati, Cincinnati, OH, USA

²Department of Emergency Medicine, University of Cincinnati, Cincinnati, OH, USA

Abstract

Stroke is the most common cause of morbidity and death in the Western world, following ischemic heart disease and cancer. Stroke can be of two types, ischemic or hemorrhagic, with ischemic stroke accounting for approximately 85% of the total number of strokes. Well-recognized environmental risk factors for stroke include hypertension, smoking, diabetes mellitus, atrial fibrillation, and atherosclerosis. Computed tomography (CT) scanning is used to diagnose hemorrhagic stroke but is relatively ineffective and may remain normal in patients with mild ischemic strokes. Magnetic Resonance Imaging (MRI) is more sensitive in detecting ischemia than CT, especially in the diagnosis of mild stroke but it is still not 100% sensitive or precise. A simple and low-cost, rapid blood test to confirm a clinical and imaging diagnosis of ischemic stroke would be extremely useful. Based on this, the central idea of this paper is to develop a method that would be applicable to a statistically viable sample set to provide candidate biomarkers for distinguishing stroke types. In search of these candidate biomarkers, different analytical separation techniques have been used to screen for major differences in the proteomes of patients plasma samples with proteomics for identification.

Keywords

Biomarkers; HPLC; Plasma; Stroke

1 Introduction

A major challenge in current biological research is the development and application of analytical methods that generate information about biological systems at the molecular level. Advances in gene expression studies have identified cellular alterations and biomarkers associated with disease diagnosis, etiology, progression, outcome, and response to therapy. However, a genomic approach requires validation of the initial gene expression studies, followed by confirmation of any relevant changes in protein expression as a consequence of

Correspondence: Dr. Joseph A. Caruso, Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221, USA, joseph.caruso@uc.edu, Fax: +513-556-9239.

The authors have declared no conflict of interest.

the gene expression changes [1,2]. Any proteins found would further require confirmation and assay development to allow clinical utilization. While proteomic analyses target proteins, challenges remain regarding the preferred use of a “proximal” fluid/tissue (i.e. brain fluid/tissue) to guide the search for candidate biomarkers [3]. Thus, other techniques that readily identify biologically relevant candidate biomarkers should be pursued.

Stroke is the 3rd leading cause of death and the leading cause of disability in the United States, accounting for about \$74 billion in costs in 2010 [4]. It is a heterogeneous disease and an overwhelming condition encircling a wide range of pathophysiological entities that include thrombosis, hemorrhage, and embolism. Stroke can be further classified into different subtypes including atherothrombotic, small vessel disease (sporadic), and cardiac emboli falling under ischemic stroke; vascular malformation, bleeding diathesis, and cerebral amyloid angiopathy under hemorrhagic stroke; and finally spinal cord stroke [5]. A variety of imaging techniques are available and the primary goal for the neurological imaging studies is to differentiate between hemorrhagic and ischemic stroke or exclude controls. Magnetic resonance imaging (MRI) and computed tomography (CT) are normally used in the initial acute estimation of stroke. CT has advantages such as widespread availability and the timely manner in which the scan can be completed. Although MRI-based techniques have improved resolution, some of the obstacles to routine implementation include limited availability and high cost of the scanners. There are inherent limitations to CT and MRI, despite advances in the field of neurological imaging. There is no proven therapy for hemorrhagic stroke, and recombinant tissue plasminogen activator (rt-PA) is the only approved therapy for ischemic stroke and is given to only ~3% of ~700 000 ischemic stroke patients annually [6, 7]. A contributor to these low rates of treatment is the inability to rapidly and reliably distinguish ischemic from hemorrhagic stroke without advanced imaging.

Development of blood biomarkers that distinguish ischemic from hemorrhagic stroke was identified as a key unresolved question by the National Institutes of Health (NIH) (2006) [3]. Several studies are being carried out in the field of blood biomarkers for stroke. Whiteley et al. published systematic reviews on diagnosis of ischemic stroke using blood biomarkers where they list certain number of proteins that are predicted to be potential biomarkers for stroke based on their biological actions [8–10]. Research is also being done on investigating the use of plasma biomarkers to guide stroke etiology diagnosis and classification [11, 12]. Clinical studies were also done using selected plasma biomarkers in the diagnosis of acute ischemic stroke [13,14] and it also explained that how concept of proteomics is helpful in search of biomarkers for stroke and cardiovascular disease [15].

In several fields of medicine, the use of plasma biomarkers is getting increasingly popular. Human blood plasma is a complex mixture of virtually all proteins found in humans, which originate from blood cells and tissue through either active secretion or leakage; hence, plasma should contain early markers for all possible diseases. The use of multiple chromatographic separation dimensions is a well-known approach to reduce the complexity of the plasma samples [16–19]. NanoLC-ESI-MSⁿ represents the state-of-the-art instrumentation when it comes to proteomics and metabolomics [20–23] (nano high-performance liquid chromatography with electrospray ionization mass spectrometry – the

MSⁿ indicates additional mass spectra are taken from specific peptide peaks from the first MS). The enhanced separation capabilities, and superior signal-to-noise ratio, applied to small amounts of samples make microfluidic devices a very useful tool as the last step in a shotgun proteomic study [24–26]. Earlier studies also suggested that identification of novel protein biomarkers for disease can be achieved using multiple chromatographies for separation and mass spectrometry for plasma/serum samples [27–29]. Using a comprehensive analytical approach, we conducted a case-control double-blind feasibility study using advanced mass spectrometry techniques to detect and identify proteins in human plasma as candidate novel biomarkers for stroke diagnosis. This study is therefore, directed to establish method protocols to indicate putative biomarkers rather than as probably or highly likely. We demonstrate that in this paper with the broader statistical study on the horizon.

2 Experimental

2.1 Overview

Stroke patients who presented to The University of Cincinnati Hospital (TUH) emergency department (ED) were enrolled in a plasma banking project. Blood draws were performed in the ED, and demographic and clinical information (time of onset, severity, etc.) were recorded.

Informed consent and Institutional Review Board (IRB) approval were obtained. For this preliminary methods development report, we analyzed six plasma samples collected within 12 h of symptom onset. Four samples involved two cases each with hypertension and ischemic or hemorrhagic stroke. Based on the limited available samples, each case of ischemic stroke was matched with a case of hemorrhagic stroke by age, race, and gender. The remaining two samples were controls that presented to the same ED. Demographics on the selected cases and stroke controls presented (can be seen in Supporting Information Table S1).

Twenty-two proteins such as human serum albumin (HSA), immunoglobulins (IgG), haptoglobin, transferrin, and lipoproteins account for 99% of the protein mass of plasma. A multiple liquid chromatography separations of the plasma proteins have been developed. Affinity chromatography allows removing the two more abundant proteins HSA and IgG, which represent $\approx 70\%$ in mass of the total protein concentration in plasma, while gel filtration can eliminate the low molecular weight compounds and give an extended resolution of the proteins present in the samples. One of the most powerful techniques in liquid chromatography is by a reverse-phase separation based on lipophilic interaction. It can resolve very complex mixtures and the conditions of the separation can be improved by changing the flow rate, temperature, and/or pore size of the stationary phase, and changing the organic solvent gradient. By using lipophilic interaction as the last step of protein separation, significant information about the major changes in protein levels and types between different samples is obtainable. Shotgun proteomic studies (bottom-up proteomics) were performed as a well-established method for biomarker studies via proteomics, [24–26] with nanoliquid chromatography electrospray ionization, ion trap mass spectrometry (nanoLC-CHIP/ITMS). The power and reliability of nanoLC-CHIP/ITMS was used to

identify the tryptic peptides from the single-peak RP-HPLC-DAD-detected proteins with different abundances between the samples.

2.2 Apparatus

An Agilent 1100 series HPLC system equipped with a binary pump, vacuum membrane degasser, thermostated auto sampler, column oven, and diode array detector with a semimicro flow UV-Vis cell was used for all chromatographic analysis. The entire system was controlled using Chemstation software (Agilent Technologies, Santa Clara, CA, USA). The human Multiple Affinity Removal System column 4.6×50 mm (MARS) was purchased from Agilent Technologies to deplete HSA and IgG. For the gel filtration separation, a TSK Gel 3000SW 7.5×300 mm (Tosoh Bioscience, Germany) was used. Reverse-phase separations were carried out on Phenomenex Jupiter C4, 4.6×300 mm, 300-Å pore size analytical column with a C4 precolumn cartridge. For the sample concentration, a freeze-dryer (Millrock Technology, NY, USA) was used.

Protein identifications were carried out on a 6300 series MSD Ion Trap XCT Ultra system, with a capillary binary pump for loading the sample into a microfluidic HPLC-Chip column and with a nano flow binary pump to provide the analytical flow for the RP separation on a Zorbax SB 300A C18 column (Agilent Technologies). The ionization system utilized was the microfluidic chip that is automatically loaded and positioned into the MS nanospray chamber. Two modes of operation were used, full scan mode to identify tryptic digested peptides, and multiple reaction monitoring (MRM) to compare the levels of the peptides between samples.

2.3 Reagents and solutions

Double deionized (DDI) water was prepared by passing distilled water through a NanoPure (18 MΩ) treatment system (Barnstead, Boston, MA, USA) and was used to prepare all solutions used in the experiments. Acetonitrile (ACN), methanol, ammonium acetate, urea, isopropyl alcohol, and formic acid were analytical grade purchased from Sigma (St. Louis, MO, USA). Ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide, and acetic acid were purchased from Fluka (St. Louis, MO, USA). The protease used for the peptide mapping of the proteins was the modified sequence grade trypsin from Promega (Madison, WI, USA). Buffer A and buffer B mobile phases for the MARS column and for the sample concentration after the MARS system, spin concentrators for proteins (5-kDa Molecular weight cut-off filters [MWCO]) from Agilent Technologies (Santa Clara, CA, USA) were used. For calibrating the size-exclusion chromatography (SEC) column, the SEC standard (Bio-Rad Laboratories, Hercules, CA, USA) mixture of molecular weight markers ranging from 1300 to 670 000 Da was used.

2.4 Multiple chromatography separations

In order to reduce the sample complexity, the two most abundant proteins were removed. In a recent comparative study, the MARS system was shown to have comparatively high performance compared with other traditional depletion methods available in the market, both in terms of proteome coverage and reproducibility [30,31]. Thus, we used the MARS system for separation. The MARS column was used as directed by the specific instructions

from vendor. In brief, the HPLC system was flushed with isopropyl alcohol for 10 min followed by flushing with water for 2 h, then two blanks with and without column were injected into the system. Plasma samples of 200 μ L each were diluted four times with buffer A (which is mobile phase A) and 400 μ L was injected into the system. The first fraction between 2 and 8 min was collected for further separation and the HSA and IgG retained in the column were eluted using buffer B. This process was carried out twice per sample in order to have enough protein concentration in the depleted fraction. After the fraction collection from the MARS system, the volume was reduced using a 5-kDa MWCO spin filter and the 3 mL collected were reduced to 400 μ L by spinning for 30 min at $5000 \times g$ at 4°C.

The subsequent HPLC separation was by SEC that allows removal of low molecular weight compounds present in the plasma, while preserving a size-based fraction of the proteins in the samples. It is also important to mention that the mobile phase used in the affinity chromatography to remove the HSA and IgG is a proprietary solution with a high concentration of nonvolatile salts, so using a volatile buffer in the SEC process is a good way to prepare the fractions collected for preconcentration by freeze drying. The concentrated samples collected after spinning were then introduced into the HPLC system, using a TSK 3000SW column with a precolumn 0.45- μ m inline filter. The mobile phase used was 50 mM ammonium acetate pH 6.5 with 5% of methanol at a flow rate of 0.7 mL min⁻¹. The addition of methanol decreased the unspecific interaction of the sample with the stationary phase [32]. The sample was loaded using a 200- μ L loop and injected twice per sample. The column was calibrated with a UV detector (wavelength, 280 nm) by using a gel filtration standard mixture (molecular weight of thyroglobulin, 670 kDa; molecular weight of g-globulin, 158 kDa; molecular weight of ovalbumin, 44 kDa; molecular weight of myoglobin, 17 kDa; molecular weight of vitamin B12, 1.3 kDa [mixture from Bio-Rad Laboratories]), with $r = 0.998$.

It was observed that column cleaning was essential after several injections because of resolution loss. For the analysis, the column was cleaned after two injections (between every sample). To clean the column, it was flushed with 10 mM β -mercaptoethanol for 25 min as reported, [33] followed by 30 min of equilibration with the analytical mobile phase.

The final HPLC separation was via reverse-phased chromatography. To maximize the interactions between the analytes and the stationary phase, the proteins were denatured with urea and at 50°C. The analysis was carried out with UV detection at 280 nm with diode array detection. The four fractions collected from the SEC separation were freeze-dried for 12 h until dryness with a collector temperature of -50°C and at <200 mbar.

Protein denaturing was done as follows: the pellet was resuspended in 200 μ L of 6 M urea and 10 μ L of acetic acid. The mixture was heated at 80°C for 3 min and cooled in ice right after 90 μ L of water was added to the sample and particulates were removed by filtration through a 0.22- μ m pore size spin filter before introduction to the LC system.

The stationary phase selected for the reversed phase protein separation was a C4 300 Å pore size. Standard protein separation conditions were mobile phase A: 0.1% FA (formic acid) in

water and mobile phase B: 0.08% FA in ACN. One hundred microliters of sample were injected each time; the separation was carried out at 1 mL min⁻¹ and 50°C to improve the resolution. The gradient used was as follows: 0 min 5% of B, 3 min 5% of B, 45 min 40% of B, 50 min 75% of B, 55 min 75% of B, 60 min 5% of B. A post runtime of 15 min at 1 mL min⁻¹ allowed the regeneration of the initial column conditions.

2.5 Tryptic digestion and protein identification

The single peak signals from the RP-HPLC were collected and concentrated by speed-vac and freeze drying. A tryptic digestion was carried out as follows: the pellets were resuspended in 25 µL of 50 mM ammonium bicarbonate, 2 µL of 100 mM DTT were added as reducing buffer and the mixture was heated at 95°C for 5 min; this step unfolds the proteins and reduces the disulfide bonds. After cooling the sample, an alkylation was carried out to protect the thiol groups of the cysteine residues by adding 3 µL of 100 mM iodoacetamide. The mixture was incubated in the dark for 20 min at room temperature. This avoids the formation of new disulfide bonds that complicate the peptide assignment from the MS/MS spectra. After the alkylation, 1 µL of modified sequence grade trypsin solution was added and incubated at 37°C for 2 h. Then, 2 µL of additional trypsin was added to complete the reaction, followed by incubation at 35°C for 8 h. Two microliters of formic acid were added to stop the reaction, and the solution was ultrafiltrated through a 5-kDa filter to eliminate the undigested proteins and the unreacted trypsin.

2.6 nanoHPLC-ESI-IT-MS² analysis for protein identification

An Agilent 6300 Series HPLC-CHIP-ESI-Ion Trap XCTsystem (Agilent Technologies) coupled to an Agilent model 1200 LC, equipped with both capillary and nano pump was used for mass identification. Three microliters of sample was loaded via the capillary pump onto the on-chip enrichment column. The chip used consists of a Zorbax 300SB C18 enrichment column (4 mm × 75 µm, 5 µm) and a Zorbax 300SB C18 analytical column (150 mm × 75 µm, 5 µm). Samples were loaded onto the enrichment column at a flow rate of 3 µL min⁻¹ with a 97:3 ratio of solvent A (0.1% FA [v/v] in water) and B (90% ACN, 0.1% FA [v/v] in water). After the enrichment column was loaded, the on-chip microfluidics switched to the analytical column at a flow rate of 0.3 µL min⁻¹. The following gradient conditions were used in the analysis: 0–5 min, 10% B; 5–80 min, 40% B; 80–95 min, 70% B; 95–100 min, 70% B; 100–110 min, 3% B; 110–115 min, 3% B. Full scan mass spectra were acquired over the *m/z* range 150–2200 in the positive ion mode. For MS/MS experiments, experimental conditions consisted of: *m/z* range: 150–2200; isolation width: 2 *m/z* units, fragmentation energy: 30–200%, fragmentation time: 40 ms. For the MRM analysis, the parent ions and its transitions were selected based on the full scan results for each protein of interest. The basic MRM criteria [34, 35] were taken into account, in brief, only non-miss cleaved peptides were used, the Glu and Cys terminal residues were avoided, the highest peptide score fragments were selected as transitions to monitor. Also, the elution in a central part of the chromatogram was preferred avoiding the highly hydrophilic or hydrophobic peptides.

The MS/MS data obtained from the experiments were exported to the online MASCOT database search engine, and submitted with the following parameters: taxonomy (mammals),

enzyme (trypsin), missed cleavages (two), fixed modifications (carbamidomethyl), peptide tolerance (2 Da), MS/MS tolerance (0.8 Da), peptide charge (+1, +2, +3), and instrument (ESI-TRAP). Mascot then searched against Uniprot database and the reported hits were validated doing blast analyses of the reported peptides.

3 Results and discussion

The goal of this project was to test the feasibility of using a multidimensional liquid chromatography separation and a bottom-up proteomics strategy to detect and identify differentially expressed proteins in the plasma of stroke and control patients. Clearly, without such a test to verify the suitability of the techniques, it would not be appropriate to apply to fully statistically viable sample numbers. We present test data from our methods development, so more extensive studies may be undertaken by those who may be interested and capable of using statistically viable numbers of samples.

The immunodepletion of HSA and IgG applied to the plasma samples by using the MARS system effectively removed the two more abundant proteins present in the plasma; a typical chromatogram of the separation is shown in Fig. 1A. As can be seen, the fraction collected has a high protein concentration and the HSA and IgG are eluted from the column in a single peak. The presence of HSA and IgG in the discarded fraction was confirmed by analyzing that fraction by SEC. Two signals were observed, collected, and analyzed as described in the procedures section above to obtain the protein identification. Only these two proteins were reported as valid hits from the MASCOT search engine against the Uniprot database (results not shown).

SEC was selected as the second separation due to its high capacity, allowing loading of the depleted serum fractions in two injections with good results. Further, it is one of the most commonly used separation techniques on biomolecules such as proteins [36, 37]. It is a valuable tool used as a prepurification step in a multiple chromatography protein separation scheme for proteomic analysis. This second LC separation shows four regions on the SEC chromatogram; a typical chromatogram is shown in Fig. 1B. At this point, all the chromatograms based on the UV detection at 280 nm of the analyzed samples looked the same. This was an expected result because of the low chromatographic resolution associated with SEC and the high degree of similarity expected for the protein levels in the samples. Although the resolution did not allow a single protein separation, it decreased the number of total proteins by separating them into four molecular weight fractions. It was also possible to discard the low molecular weight compounds eluted at the end of the SEC chromatogram. Also by using a volatile buffer such as ammonium acetate, it was possible to use freeze drying as an inexpensive preconcentration method in comparison to the MWCO.

The freeze-dried fractions were analyzed on the reversed-phase column after denaturing (as described in the previous section). A set of chromatograms obtained after running the four SEC fractions, which were separated in the C4 column, is presented in Fig. 2. The chromatograms are consistent in the number and intensities for a large number of signals. The UV spectra show a protein pattern (280 nm), and the overlaps of the chromatograms from the six samples analyzed on each of the SEC fractions show very similar

chromatograms. The fact that the majority of the proteins present in the samples are the same was expected and the regions of the chromatograms that showed clear differences were selected as the targets for candidate protein biomarker identification.

The protein identification analysis was carried out on the second and third SEC fractions once the reverse-phase separations were completed. These fractions were selected to perform the last step of the proposed method in order to decrease the number of samples to be analyzed by the nanoLC-ITMS.

To validate the selection of the chromatographic regions for further use, a statistical analysis was performed. A table with the UV absorbance expressed as mAU of the six samples was taken for the signals with different intensity in the reverse-phase chromatograms and an ANOVA analysis was carried out to choose the signals with differences in intensity at $P = 0.1$. The regions selected in the chromatograms are shown in Table 1. These signals were collected from the HPLC flow as single peak signals. The reproducibility of the chromatograms between the injections of the same sample ($3 \times 100 \mu\text{L}$ each) can be seen in Supporting Information Fig. S1. It is seen that these are acceptable.

The possibility of having coeluted proteins in a single chromatographic signal was concerned; in order to exclude this possibility, two approaches were performed. The use of a diode array detector instead of a single wavelength UV-Vis detector allows evaluating the peak purity factor. This is based on the measurement of similarity of the UV-Vis spectra in different regions of the chromatographic peak. A minimum of 12 spectra were collected per peak, and the purity factor calculated exceeded the nonparametric value of 980.

The fact that coeluted proteins in a reverse-phase column can have similar UV-Vis absorbance spectra was contemplated. To overcome this possible situation, the absorbance at 280 nm of the proteins in study was compared to the ion intensity obtained from MRM measurements of the parent ion associated to the peptides reported from the MASCOT search engine. In Table 2, the reported peptides are selected to follow by MRM mode in the nanoHPLC-ESI-ITMS-MS instrument, the parent ion and the two transitions followed for each case. The ion intensity of the parent ions shows a consistent pattern when compared with the UV intensity of the analyzed chromatographic peaks (data not shown).

Using six plasma samples in small amounts, a double-blind study (both the sample provider and analysts were not notified about the types of sample sets) was carried out as a method development step in both stroke and control patients. To summarize our findings, multidimensional LC, followed by nanoLC-CHIP/ITMS experiments identified possible disease-specific protein differences in the plasma concentrations of 12 proteins.

Various studies were done on some of these proteins shown in Table 1. Studies were done to be familiar with the role of low-density receptor-related protein (LRP1) in neuronal synapse formation and function in Alzheimer's disease (AD) [38]. Investigations were carried out to know the effect of the interaction between tissue-type plasminogen activator (tPA) and microglial LRP1 on matrix metalloproteinase-9 (MMP-9) activity after middle cerebral artery occlusion (MCAO) [39]. Interaction between tPA and LRP plays a role in the regulation of the permeability of the neurovascular unit (NVU) during cerebral ischemia

[40]. LRP1B is a novel candidate tumor suppressor gene that is inactivated by genetic and transcript alterations in nearly 50% of nonsmall cell lung cancer cell lines [41]. Endothelin-1 is produced by endothelial cells, and endothelin-3 is derived from neuronal tissue. In theory, endothelin-mediated vasoconstriction may enhance ischemic neuronal damage. Ischemic stroke is associated with acute and marked increases in plasma levels of endothelin-1 [42]. Studies are also done on endothelin-converting enzyme (ECE) inhibitors for the treatment of subarachnoid hemorrhage-induced vasospasm [43]. The analysis of the levels of expression and distribution of ECE-1 in the brain under normal and pathologic conditions could be important in neurodegeneration and pathogenesis of Alzheimer disease [44].

Research was done on glucocorticoids (GC) to examine the role of stroke-induced changes in GC and cytokines in mediating poststroke depression-like behavior in mice. The pathophysiological processes that occur after stroke contribute to the etiology of poststroke depression (PSD) [45]. Various studies were done to see the expression in the brain of two heat shock proteins, the 70-kDa Hsp (Hsp70) and the 27-kDa Hsp (Hsp27), is notable because both proteins are highly inducible in glial cells and neurons following a wide range of noxious stimuli including ischemia, epileptic seizure, and hyperthermia [46]. Studies were done on Serine/threonine-protein kinase LATS2 for structure, expression, and chromosome mapping of *LATS2*, a mammalian homolog of the *Drosophila* tumor suppressor gene *lats/wart* [47].

Three proteins with biological and probably biomarker relevance were (1) endothelin converting enzyme-like 1, which is highly expressed in the brain, processes endothelin 1 and regulates vasoconstriction, (2) E3 ubiquitin-protein ligase is also highly expressed in the brain and participates in the removal/detoxification of abnormal proteins, and (3) Na/K/Ca exchanger 3 precursor is a transport protein that is abundant in the brain. In the present study, these proteins were found with different levels along the samples, for most of them, there was no significant correlation of the levels with the disease studied, perhaps because of the small number of samples provided for this study. But in the case of the Endothelin-converting enzyme-like 1, E3 ubiquitin-protein ligase, and Na/K/Ca exchanger 3 precursor, a positive correlation between the protein levels and the ischemic stroke was found.

These three easily exceed the MASCOT relevance score of 51 and suggest further confidence that this current study provides the analytical methodology necessary to carry out a more extensive study. Further, these were successfully subjected to blast analysis, which tests for the uniqueness of a peptide in a particular protein, allowing further confidence in the results. An example of the tryptic peptide mapping performed is shown in Fig. 3, in brief, from the total ion chromatogram, it was possible to extract the ion chromatogram, based in the m/z of the precursor ion corresponding to the reported MASCOT peptide. The chromatographic behavior of the extracted ion as a single peak signal and its fragmentation mass spectra was evaluated and the assignment of b and y ions is shown. This procedure, along with the peptide blast report, was used as quality control of the protein identification process.

As shown in Table 1, the differing areas were quite different between control, ischemic, and hemorrhagic patients of various sample sets. The three proteins mentioned above can be

predicted as potential biomarkers based on their biological significance, differences found in areas of various sample sets and as they are more abundantly located in brain, but more studies are needed to be done further to consider these proteins as biomarkers. The central idea of this paper is to come with an analytical method that can differentially express proteins in the plasma of stroke and control patients, which we were able to achieve by using different chromatographic techniques.

Although the number of samples analyzed was small, the proof of concept is clear, but the full significance of those proteins found for possible biomarkers must wait until a comprehensive study is complete. It is demonstrated that this method, used appropriately, can find disease-related proteins at different levels between the samples; thereby indicating the methodology is suitable for a larger scale study.

3.1 Validation studies

In order to validate the proposed method, a nonhuman protein (Thaumatococcus daniellii) at three different concentrations (0.05, 0.2, and 1 mg/mL) was spiked in three pooled plasma samples. After spiking the samples, the same procedure was employed as mentioned above in the chromatography methods. In brief, the first step was immunodepletion of the plasma samples to remove two most abundant proteins (HSA and IgG) followed by SEC based on the size of the proteins. After these two separations, the next step was reverse-phase chromatography to see the differences in proteins. In parallel, a standard of the thaumatin protein was analyzed by reverse phase to match the retention time. Later, the identification based on retention time of the standard was confirmed by collecting the chromatographic fraction from the spiked plasma sample and performing the protein identification protocol used for the proteins of interest. Extracted ion chromatogram and MS/MS fragmentation of the collected fraction can be seen in Supporting Information Fig. S2. The recovery of the nonhuman protein was calculated by comparing the intensity of the signals obtained from standard at the same concentrations.

The intensity of the UV absorbance signals in the reverse-phase chromatogram of the spiked protein in the human plasma increase in a semilinear way, going from 0.05 to 1 mg/mL. After the enzymatic digestion and identification of the thaumatin in the spiked samples, an MRM measurement of 813.4 and 1045.3 peptides was performed, using the 1085.6 → 813.4 and 858.4 → 1045.3 transitions respectively, and the intensity of the peptide ion signal was correspond to the UV intensity of the reverse-phase chromatograms (data not shown).

4 Concluding remarks

Considering the limited amount of sample used in this approach (200 µL), the possibility of using this high-throughput multiple-LC-separation scheme to show major differences in the plasma proteome from stroke-related patients is the major advantage of the method developed. The use of a microfluidic LC-MS/MS system allowed us to have positive protein identifications in the last chromatographic fraction with these low amounts of original sample.

This level of LC separation was useful to find relative differences in the protein concentration between plasma samples. Even when not all the proteins were resolved at a single peak signal, it was possible to select regions with differences in the protein levels. The major advantage of the methodology investigated is the feasibility of direct comparison of a large number of proteins present in the samples from simple chromatograms. There are reports of intense separation steps when looking for a large number of proteins in plasma [18, 19, 48], but most of these are based on separation of the tryptic peptides generated from the bulk of proteins. The number of proteins identified is large but the relative abundance of each protein is unknown, because the enzymatic digestion of very complex samples as plasma or serum is not quantitative and therefore the MS intensity of the peptides formed is not always a direct quantitative response. In contrast, this methodology can focus the identification efforts applied generally to the different protein concentrations and determine if those are biologically relevant and can be used as differential diagnostic tools. The simplification of the matrix by the chromatographic separation scheme proposed allows using a label-free semiquantitative MRM experiments to confirm the feasibility of UV absorbance as a comparative tool.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

CT	computed tomography
ED	emergency department
IRB	institutional review board
MARS	multiple affinity removal system
MRI	magnetic resonance imaging
MWCO	molecular weight cut-off filters
NIH	National Institute of Health
TUH	University of Cincinnati Hospital

References

1. Stamova B, Xu H, Jickling G, Bushnell C, Tian Y, Ander BP, Zhan X, Liu D, Turner R, Adamczyk P, Khoury JC, Pancioli A, Jauch E, Broderick JP, Sharp FR. Stroke. 2010; 41:2171–2177. [PubMed: 20798371]
2. Zhan X, Jickling GC, Tian Y, Stamova B, Xu H, Ander BP, Turner RJ, Mesias M, Verro P, Bushnell C, Johnston SC, Sharp FR. Neurology. 2011; 77:1718–1724. [PubMed: 21998319]
3. Rifai N, Gillette MA, Carr SA. Nat Biotechnol. 2006; 24:971–983. [PubMed: 16900146]
4. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, Ferguson TB, Ford E, Furie K, Gillespie C, Go A, Greenlund K, Haase N, Hailpern S, Ho PM, Howard V, Kissela B, Kittner S, Lackland D, Lisabeth L, Marelli A, McDermott MM, Meigs J, Mozaffarian D, Mussolino M, Nichol G, Roger V, Rosamond W, Sacco R, Sorlie P, Stafford R, Thom T, Wasserthiel-Smoller S, Wong ND, Wylie-Rosett J. Circulation. 2010; 121:948–954. [PubMed: 20177011]

5. Amarenco P, Bogousslavsky J, Caplan LR, Donnan GA, Hennerici MG. Cerebrovascular Dis. 2009; 27:493–501.
6. Kleindorfer D, Lindsell CJ, Brass L, Koroshetz W, Broderick JP. Stroke. 2008; 39:924–928. [PubMed: 18239184]
7. Adeoye O, Hornung R, Khatir P, Kleindorfer D. Stroke. 2011; 42:1952–1955. [PubMed: 21636813]
8. Whiteley W, Tseng MC, Sandercock P. Stroke. 2008; 39:2902–2909. [PubMed: 18658039]
9. Whiteley W, Chong WL, Sengupta A, Sandercock P. Stroke. 2009; 40:e380–e389. [PubMed: 19286602]
10. Saenger AK, Christenson RH. Clin Chem. 2010; 56:21–33. [PubMed: 19926776]
11. Garcia-Berrocso T, Fernandez-Cadenas I, Delgado P, Rosell A, Montaner J. Curr Cardiol Rev. 2010; 6:194–201. [PubMed: 21804778]
12. Montaner J, Perea-Gainza M, Delgado P, Ribo M, Chacon P, Rosell A, Quintana M, Palacios ME, Molina CA, Alvarez-Sabin J. Stroke. 2008; 39:2280–2287. [PubMed: 18535284]
13. Kim MH, Kang SY, Kim MC, Lee WI. Ann Clin Lab Sci. 2010; 40:336–341. [PubMed: 20947807]
14. Reynolds MA, Kirchick HJ, Dahlen JR, Anderberg JM, McPherson PH, Nakamura KK, Laskowitz DT, Valkirs GE, Buechler KF. Clin Chem. 2003; 49:1733–1739. [PubMed: 14500614]
15. Anderson L. J Physiol. 2005; 563:23–60. [PubMed: 15611012]
16. Lee HJ, Lee EY, Kwon MS, Paik YK. Curr Opin Chem Biol. 2006; 10:42–49. [PubMed: 16418010]
17. Horvatovich P, Hoekman B, Govorukhina N, Bischoff R. J Sep Sci. 2010; 33:1421–1437. [PubMed: 20486207]
18. Garcia A, Prabhakar S, Brock CJ, Pearce AC, Dwek RA, Watson SP, Hebestreit HF, Zitzmann N. Proteomics. 2004; 4:656–668. [PubMed: 14997489]
19. Issaq HJ. Electrophoresis. 2001; 22:3629–3638. [PubMed: 11699900]
20. Gokce E, Andrews GL, Dean RA, Muddiman DC. J Chromatogr B. 2011; 879:610–614.
21. Levin Y, Jaros JA, Schwarz E, Bahn S. J Proteomics. 2010; 73:689–695. [PubMed: 19896566]
22. Boutin M, Berthelette C, Gervais FG, Scholand MB, Hoidal J, Leppert MF, Bateman KP, Thibault P. Anal Chem. 2009; 81:1881–1887. [PubMed: 19178285]
23. Young JB, Van Soest V, Neyer DW, Hunter CL, Hebert N. J Biomol Tech. 2011; 22(Supplement):S58.
24. McDonald WH, Yates JR 3rd. Dis Markers. 2002; 18:99–105. [PubMed: 12364816]
25. Wu CC, MacCoss MJ. Curr Opin Mol Ther. 2002; 4:242–250. [PubMed: 12139310]
26. McDonald WH, Yates JR 3rd. Curr Opin Mol Ther. 2003; 5:302–309. [PubMed: 12870441]
27. Adkins JN, Varnum SM, Auberry KJ, Moore RJ, Angell NH, Smith RD, Springer DL, Pounds JG. Mol Cell Proteomics. 2002; 1:947–955. [PubMed: 12543931]
28. Tirumalai RS, Chan KC, Prieto DA, Issaq HJ, Conrads TP, Veenstra TD. Mol Cell Proteomics. 2003; 2:1096–1103. [PubMed: 12917320]
29. Pieper R, Su Q, Gatlin CL, Huang ST, Anderson NL, Steiner S. Proteomics. 2003; 3:422–432. [PubMed: 12687610]
30. Dayarathna MK, Hancock WS, Hincapie M. J Sep Sci. 2008; 31:1156–1166. [PubMed: 18306431]
31. Whiteaker JR, Zhang H, Eng JK, Fang R, Piening BD, Feng LC, Lorentzen TD, Schoenherr RM, Keane JF, Holzman T, Fitzgibbon M, Lin C, Cooke K, Liu T, Camp DG 2nd, Anderson L, Watts J, Smith RD, McIntosh MW, Paulovich AG. J Proteome Res. 2007; 6:828–836. [PubMed: 17269739]
32. Anspach B, Gierlich HU, Unger KK. J Chromatogr. 1988; 443:45–54. [PubMed: 2459149]
33. Doker S, Mounicou S, Dogan M, Lobinski R. Metallomics. 2010; 2:549–555. [PubMed: 21072339]
34. Lin S, Shaler TA, Becker CH. Anal Chem. 2006; 78:5762–5767. [PubMed: 16906721]
35. Han B, Higgs RE. Brief Funct Genomic Proteomic. 2008; 7:340–354. [PubMed: 18579614]
36. Lecchi P, Gupte AR, Perez RE, Stockert LV, Abramson FP. J Biochem Biophys Methods. 2003; 56:141–152. [PubMed: 12834973]

37. Alvarez-Manilla G, Atwood J 3rd, Guo Y, Warren NL, Orlando R, Pierce M. *J Proteome Res.* 2006; 5:701–708. [PubMed: 16512686]
38. Harris-White ME, Frautschy SA. *Curr Drug Targets CNS Neurol Disord.* 2005; 4:469–480. [PubMed: 16266282]
39. Zhang C, An J, Haile WB, Echeverry R, Strickland DK, Yepes M. *J Cerebral Blood Flow Metab.* 2009; 29:1946–1954.
40. Polavarapu R, Gongora MC, Yi H, Ranganathan S, Lawrence DA, Strickland D, Yepes M. *Blood.* 2007; 109:3270–3278. [PubMed: 17170123]
41. Liu CX, Musco S, Lisitsina NM, Yaklichkin SY, Lisitsyn NA. *Genomics.* 2000; 69:271–274. [PubMed: 11031110]
42. Ziv I, Fleminger G, Djaldetti R, Achiron A, Melamed E, Sokolovsky M. *Stroke.* 1992; 23:1014–1016. [PubMed: 1615534]
43. Lin CL, Winardi W, Jeng AY, Kwan AL. *Neurol Res.* 2006; 28:721–729. [PubMed: 17164035]
44. Fisk L, Nalivaeva NN, Turner AJ. *Exp Biol Med.* 2006; 231:1048–1053.
45. Craft TK, DeVries AC. *Biol Psychiatry.* 2006; 60:812–818. [PubMed: 16730336]
46. Franklin TB, Krueger-Naug AM, Clarke DB, Arrigo AP, Currie RW. *Int J Hyperthermia.* 2005; 21:379–392. [PubMed: 16048836]
47. Yabuta N, Fujii T, Copeland NG, Gilbert DJ, Jenkins NA, Nishiguchi H, Endo Y, Toji S, Tanaka H, Nishimune Y, Nojima H. *Genomics.* 2000; 63:263–270. [PubMed: 10673337]
48. Chen R, Pan S, Brentnall TA, Aebersold R. *Mol Cell Proteomics.* 2005; 4:523–533. [PubMed: 15684406]

Figure 1.

(A) Chromatogram obtained from the immunodepletion of HSA and IgG from the plasma samples, the fraction collected for further separations was between 2.5 and 8.5 min for each sample. The fraction at 15 min is the HSA and IgG that were separated from plasma samples after immunodepletion (affinity column). (B) Typical SEC chromatogram, as a second LC separation. The molecular mass markers are on top, and the four fractions collected are shown. The four fractions collected were further separated on C4 reverse-phase column as a third LC separation.

Figure 2.

Typical set of six chromatograms obtained from SEC F2 fraction using the reverse-phase C4 column. Overlaid are six plasma samples obtained from C4 reverse-phase column. Rectangles (in black) show a typical peak selection taken for ESI-MS.

Figure 3. Peptide mapping approach used to assign the protein IDs. (A) Total ion chromatogram obtained by nanoLC-ESI-MS/MS, for one of the peaks in the reverse-phase separation, after tryptic digestion and identified by MASCOT as endothelin converting enzyme-like 1. (B) Extracted ion chromatogram at the m/z 882.7 ± 0.3 , it shows the signal of the peptide LTYGTIAAIGEQN^{EE}R as [M-H]²⁺ ion with good chromatographic behavior. (C) MS/MS fragmentation of the ion m/z 882.7, and the assigned b, and y ions for the sequence of the matched peptide.

Table 1

Proteins found at different levels between samples. The first column is the SEC fraction, the second is the retention time from the reverse-phase separation as a single peak signal; the next three columns are the signal areas associated with the proteins found in each group of samples; the intracerebral hemorrhage (ICH) and acute ischemic stroke (AIS). The last two columns are the protein scores and peptide matches. The areas are the average of the two samples per group and calculated as mAu^* minute at 280 nm. The three proteins shown in bold black are those discussed in the text that have putative implications for AIS or ICH.

Protein name	SEC fraction	RP(t_R)	Controls	ICH	AIS	Protein score	Peptides matched
Serine/threonine-protein kinase LATS2- <i>Homo sapiens</i> (Human)	2	15.2	110	1270	672	47	2
Low-density lipoprotein receptor-related protein 1B precursor – <i>Homo sapiens</i> (Human)	2	31.2	67	146	98	58	2
Kinesin heavy chain isoform 5A – <i>Homo sapiens</i> (Human)	2	13.7	79	311	145	55	2
Low-density lipoprotein receptor-related protein 2 precursor (Megalin) – <i>Homo sapiens</i> (Human)	2	14.9	2200	1190	2030	46	2
Serine/threonine-protein kinase LATS1 – <i>Homo sapiens</i> (Human)	2	17.8	246	194	534	51	2
Endothelin-converting enzyme-like 1 – <i>Homo sapiens</i> (Human)	3	8.1	15	16	21	78	3
Glucocorticoid-induced transcript 1 protein – <i>Homo sapiens</i> (Human)	3	14.8	583	1360	2670	52	2
Heat shock 70-kDa protein 12B – <i>Homo sapiens</i> (Human)	3	17.9	226	217	256	94	4
Sodium/potassium/calcium exchanger 3 precursor (Na(+)/K(+)/Ca(2+) – <i>Homo sapiens</i> (Human)	3	27.2	59	153	108	82	2
T-complex protein 1 subunit alpha B – <i>Homo sapiens</i> (Human)	3	16.2	3550	2372	3860	54	2
E3 ubiquitin-protein ligase UHRF1 – <i>Homo sapiens</i> (Human)	3	19.8	1420	699	1480	57	3
Cell division cycle-associated protein 7 (Protein JPO1) – <i>Homo sapiens</i> (Human)	3	14.0	nd*	333	nd*	49	2

* no detection

Table 2

Summary of the design for the MRM experiments used to confirm the semiquantitative comparison of the selected proteins levels

Protein name	Charge	Peptide sequence	Parent ion (m/z)	Product ion A (m/z)	Product ion B (m/z)
Endothelin converting enzyme-like 1	2	VLSLLGADAVEQK	672.5	760.4	817.4
	2	LTYGTIAAIGEQNEER	883.4	675.3	804.3
	2	LYDNFTVYNQR	717.5	927.5	1041.5
E3 ubiquitin-protein ligase UHRF1	2	NDASEVVLAGER	630.8	743.4	872.5
	2	DDDEPGPWTK	580.7	685.4	814.4
Sodium/potassium/calcium exchanger 3 precursor	2	QTHTVDSLRS	572.8	676.4	777.4
	2	HTVENGTGPSSAPDR	763.4	887.4	944.4
	2	AYTNGESEVAIK	641.9	775.4	832.4