*Clinical Proteomics*

ISSN 1542-6416/06/02:67–90/$30.00 (Online)

# Original Article

## Comparison of Methods to Examine the Endogenous Peptides of Fetal Calf Serum

*Declan Williams,[1] Peihong Zhu,[1] Peter Bowden,[1] Catherine Stacey,[2] Mike McDonell,[2] Paul Kowalski,[2] Jane Marie Kowalski,[2] Ken Evans,[3] Eleftherios P. Diamandis,[3,4] K. W. Michael Siu,[3] and John Marshall[1,3,*]*

[1]Department of Chemistry and Biology, Faculty of Engineering and Applied Science, Ryerson University, Toronto, ON, Canada; [2]Bruker Daltonics, Inc., Bellerica, MA; [3]Ontario Cancer Biomarker Network, MaRS, Toronto Ontario, ON, Canada; and [4]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital and Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada

### Abstract

There is a great desire to relate the patterns of endogenous peptides in blood to human disease and drug response. The best practices for the preparation of blood fluids for analysis are not clear and also relatively few of the peptides in blood have been identified by tandem mass spectrometry. We compared a number of sample preparation methods to extract endogenous peptides including C18 reversed phase, precipitation, and ultrafiltration. We examined the results of these sample preparation methods by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) and liquid chromatography-tandem mass spectrometry (MS/MS) using MALDI-TOF/TOF and electrospray ionization-ion trap. Peptides from solid-phase extraction with C18 in the range of hundreds of femtomoles per spot were detected from the equivalent of 1 µL of serum by MALDI-TOF. We observed endogenous serum peptides from fibrinogen α- and β-chain, complement C3, α-2-HS-glycoprotein, albumin, serine (or cysteine) proteinase inhibitor, factor VIII, plasminogen, immunoglobulin, and other abundant blood proteins. However, we also recorded significant MS/MS spectra from tumor necrosis factor-α-, major histocompatibility complex-, and angiotensin-related peptides, as well

*Author to whom all correspondence and reprint requests should be addressed:
John Marshall, Department of Chemistry and Biology, Faculty of Engineering and Applied Science, Ryerson University, 350 Victoria Street, Toronto, ON, Canada M5B 2K3; and Ontario Cancer Biomarker Network, MaRS, 101 College Street, Toronto Ontario, ON, Canada.
E-mail: 4marshal@ryerson.ca

as peptides from collagens and other low-abundance proteins. Amino acid substitutions were detected and a phosphorylated peptide was also observed. This is the first time the endogenous peptides of fetal serum have been examined by MS and where peptides from low-abundance proteins, phosphopeptides, and amino acid substitutions were detected.

**Key Words:** Fetal calf serum; liquid chromatography-electrospray ionization mass spectrometry; matrix-assisted laser desorption/ionization time-of-flight; MALDI; ESI tandem mass spectrometry.

## Introduction

A mass spectrometer (MS) can be used to identify the peptides in serum, plasma, or other biological fluids *(1,2)* as an aid in diagnosis of disease or to identify physiologically relevant molecules *(3,4)*. The patterns of low-molecular mass peptides in samples from normal and diseased populations have been compared using time-of-flight (TOF) MS of unfractionated serum bound to a chromatographic surface on the matrix-assisted laser desorption/ionization (MALDI) target *(5)*. The peptide/protein patterns in unfractionated serum have been shown to vary considerably during sample handling making their clinical use as biomarkers difficult *(6)*. Serum fractionated by C18 partition chromatography produced peptide patterns by MALDI-TOF in agreement with classical biochemical and immunological investigations *(7)*. Some of the endogenous peptides from serum have been identified by tandem MS (MS/MS) *(7–10)*. Fetal serum contains growth factors that can stimulate cell lines to grow in culture. Some of these factors might be low-molecular mass polypeptides. Proteolytic fragments of larger molecules that circulate in the blood may have profound regulatory effects. For example, the proteolytic processing of collagen and related molecules produces fragments with roles in angiogenesis, cell growth, and complex regulatory processes that remain to be better defined *(11)*. Proteolytic processing of fibrinogens releases fragments that have been associated with acute events, such as coronary thrombosis *(12)* or the effects of cancer *(13–15)*.

To date, there has been no comparison of the methods used to extract polypeptides from blood fluids for MS/MS analysis. Richter et al. used hemofiltrate depleted of larger proteins to construct a database of circulating human peptides *(16)*. Reversed-phase chromatography has been used to efficiently collect peptides for serum profiling *(3,7,17)*. Polypeptides have also been isolated by precipitation with organic solvents *(10)* or from the carrier protein albumin *(18,19)*. Recently, membrane filter plates have been used to solid-phase extract (SPE) the endogenous peptides from serum prior to identification using TOF/TOF *(9)*. A key step in developing such methodology will be to determine what pool of peptides in serum are bound to carrier proteins, such as albumin *(18–20)*. We compared sample preparation methods for detecting peptides from serum by MS. Polypeptides from fetal serum were extracted by a variety of methods and compared by MALDI-TOF MS prior to liquid chromatography (LC)-MS/MS with electrospray ionization (ESI)-ion trap and MALDI-TOF/TOF identification.

## Materials

1. Tissue culture grade fetal bovine calf serum was obtained from GIBCO (Burlington, Ontario, Canada) and is produced by rapid clotting in bags, separation, and cooling to near freezing point followed by utrafiltration. We did not subsequently heat inactivate the serum.
2. Magnetic beads with C18, SAX, WAX, SCX, HIC, and other bonded phases were obtained from Bruker Daltonics (Billerica, NJ).

3. C18 resins were also obtained from Millipore (Bedford, MA) or Vydac (Hesperia, CA).
4. Amicon ultrafiltration membranes were obtained from Millipore.
5. Capillary C18 reversed-phase columns were obtained from Vydac.
6. P10 gel, DEAE-Blue resin, and sodium dodecyl sulfate-polyacrylamide gel electrphoresis standards were obtained from Bio-Rad Laboratories (Hercules, CA).
7. Solvents were obtained from Caledon Laboratories (Georgetown, Ontario, Canada).
8. Cibacron blue resin and all other salts and reagents were obtained from Sigma-Aldrich-Fluka (St. Louis, MO) except where indicated.

## Methods

### Chromatography

Capacities were calculated and loading and elution conditions were in accordance with manufacturer's instructions except where indicated. Partition chromatography and SPE with C18 reversed-phase resin was run by gravity drip except where indicated.

### Preparative C18 Chromatography

Batch chromatographic separations with C18 magnetic beads were performed according the manufacturer's instructions and buffers (Bruker), as previously described ***(17)***. Briefly, beads were equilibrated with the appropriate binding buffer before incubation in a volume of serum (equal to the packed volume) diluted in 10 vol of binding buffer. The serum sample was mixed with binding buffer and bound to the beads before washing in the binding buffer. Beads were collected with a magnetic 96-well plate holder before eluting with the appropriate buffer and collecting the eluant from the magnetic plate into 1-mL centrifuge tubes. Brief centrifugation at 12,000*g* ensured complete removal of particles. To accurately estimate the peptides in the equivalent of 1-µL of serum and for loading gels, 50 µL of serum were diluted in binding buffer for immobilization onto 100 µL of C18 resin-packed volume and then eluted in 100 µL. Two microliters of the eluant (equivalent to 1 µL of serum) were spotted on MALDI plates. Alternatively, the eluate was dried in a speed vac.

### Organic Solvent Precipitation

Serum samples were precipitated with 5 vol of acetonitrile on ice for 15 min prior to spinning at 12,000*g* for 15 min ***(10)***. The supernatant and pellet were separated and the supernatant was dried under vacuum. Some samples were treated with 2 m*M* dithiothreitol (DTT) prior to organic solvent precipitation or SPE with C18.

### Ammonium Sulfate Precipitation

Serum was precipitated using 4 vol of saturated ammonium sulfate on ice for 15 min followed by centrifugation at 15,000*g* for 15 min. The supernatant and pellet were collected into separate tubes.

### Dye Affinity Chromatography

DEAE-Blue and Cibacron Blue resins were loaded with sample in hypotonic buffer (20 m*M* Tris, pH 8.0, with 40 m*M* NaCl) or physiological buffer (1X phosphate-buffered saline[PBS], pH 7.3) before washing in loading buffer and elution with 1 *M* NaCl, as previously described ***(21)***.

### Size Exclusion Chromatography

P10- and G25-size exclusion media were used in the batch mode. Serum or serum diluted fivefold in water or PBS was equilibrated with an equal volume of exclusion media on ice. The resin was collected by brief centrifugation, rapidly resuspended in water, repelleted, and the superanatant discarded. The resin was then resuspended in an equal volume of water and allowed to equilibrate on a Ferris wheel for 15 min before collecting the

supernatant. The process was repeated twice to yield three fractions. The peptides were then loaded on gels or collected over C18 for MALDI or LC-ESI-MS/MS analysis.

### Ultrafiltration

Amicon ultrafiltration membranes were used as previously described (20). Briefly, sera were diluted to a final concentration of 50 m*M* ammonium bicarbonate, pH 8.3, with 20% v/v acetonitrile. The solution was ultrafiltered through YM 30-kDa nominal molecular-weight-limit filters under centrifugation at 12,000*g* and the filtrate was collected by C18 (20).

### Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

Endogenous peptide samples collected from all prefractionation schemes were prepared for MALDI analysis by collection over C18 reversed phase resin and spotted onto a stainless steel target for an Ultraflex MALDI-TOF/TOF (Bruker). The eluted peptides were dried onto stainless steel targets and, after allowing all the samples to dry evenly, the energy-absorbing matrix cyano-4-hydroxy cinnamic acid (CHCA) was applied for collection of MS spectra. An automatic acquisition mode was used to collect data from the various treatments under identical parameters.

### Liquid Chromatography-Electrospray Ionization Tandem Mass Spectrometry

Endogenous peptide samples collected from all prefractionation schemes were prepared for ESI analysis by collection over C18 reversed-phase resin and analyzed over a 90-min gradient from 5 to 65% acetonitrile at a flow rate of 2 mL/min with an Agilent 1100 series capillary pump (Agilent, Palo Alto, CA) through an Agilent 150 × 0.3 mm C18 column. Peptides separated by capillary high-performance LC were ionized via a metal needle electrospray head in a Bruker HCT 100 ion trap.

### Liquid Chromatography-Matrix-Assisted Laser Desorption/Ionization-Tandem Mass Spectrometry

Endogenous peptide samples collected from prefractionation schemes were analyzed over a 30-min gradient from 5 to 65% acetonitrile at a flow rate of 2 mL/min with an Agilent 1100 series capillary pump through an Agilent 150 × 0.3 mm capillary C18 column. Peptides were separated onto a stainless steel MALDI plate for Ultraflex MALDI-TOF/ TOF analysis, as described here previously. MS/MS spectra were collected using CHCA as a matrix.

### Correlation Analysis

The resulting ESI-MS/MS spectra were analyzed by X-Tandem (22) with –1 Da and +3 Da from parent peptide as tolerance of the parent mass or likely isotopes in the search parameters and 0.4 Da error in the fragments. The MALDI-MS/MS data were searched by MASCOT (23) with the default settings. The default-expected value for significance by X-Tandem is E-1, but in this study we only accepted peptides with a score of E-2 or lower. The database was the bovine RefSeq database downloaded from the National Center for Biotechnology Information. Because no exogenous enzymes were used to digest these samples, we performed the search with no enzyme. Searches of nonphysiological databases yielded few significant hits at these expectation values.

## Results

Serum was prepared by SPE with C18 reversed phase or by precipitation with acetonitrile or ammonium sulfate. We compared the profile of the polypeptides in the extracts. We also examined serum dealbuminized with DEAE-Blue and Cibacron Blue dye-binding resins. Size exclusion chromatography resins, such as P10 and G25, and ultrafiltration membranes were also compared. The extracts from these methods were further treated with C18

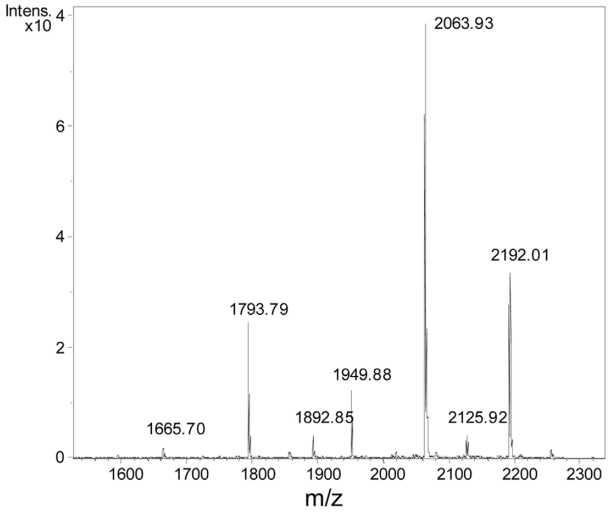


Fig. 1. MALDI-TOF spectrum of peptides collected by C18 recorded with a TOF-TOF.

reversed phase prior to LC-MS/MS. The LC-MS/MS was performed by both LC-ion trap and LC-TOF-TOF.

## Solid Phase Extraction With C18

We examined the collection of serum peptides with C18 for extraction from the solid phase before resolution on MALDI-TOF. We observed that DTT treatment prior to SPE altered a few aspects of the MALDI-TOF spectrum (not shown). We found that reversed phase extraction of nonreduced samples could be used to produce samples that permitted the direct detection of endogenous peptides when the sample was deposited on a single MALDI target spot **(Fig. 1)**.

## Comparison of Methods

### Acetonitrile Precipitation

We precipitated serum with acetonitrile and compared the polypeptides in the supernatant vs the pellet, by MALDI-TOF ***(10)***. About 1 µg of protein and peptides was collected by reversed-phase C18 resin and loaded on metal plates with CHCA matrix. We observed the presence of peptides in the organic supernatant but the signal-to-noise ratios were much higher in the pellet by MALDI-TOF **(Fig. 2A,B)**.

### Ammonium Sulfate

Proteins from ammonium sulfate precipitates typically retain enzyme activity and, therefore, are not strongly denatured. A wide range of peptides were observed in the supernatant and showed some differences with the pool of peptides in the pellet, but the major peptides were shared and there were much higher signal-to-noise ratios in the pellet **(Fig. 2C,D)**.

### Dye-Binding Chromatography

We determined if endogenous peptides were selectively bound to albumin by dye affinity chromatography in hypotonic (20 m*M* Tris, pH 8.0, with 40 m*M* NaCl) and physiological buffers (1X PBS, pH 7.4) using both DEAE Blue and Cibacron Blue resin. We observed that many peptides eluted in the bovine serum albumin (BSA)-depleted fraction of serum that flowed through the dye-binding column. MALDI-TOF MS of the flow-through and retained fractions showed that the dye resin did not selectively retain many peptides **(Fig. 2E,F)** (Cibacron Blue result shown).

### Size Exclusion

The components of fetal bovine serum were separated on the basis of their capacity to penetrate the size exclusion resins P10 and G25. A similar range of peptides was observed in the fraction that penetrated the beads compared to the fraction that did not penetrate the beads (**Fig. 2G,H**) (P10 result shown).

### Ultrafiltration

To determine if peptides could be separated from high-mass components, we employed ultrafiltation with and without the presence of 20% acetonitrile. The presence of acetonitrile did not seem to affect strongly the distribution or intensity of small peptides **(Fig. 2I,J)**.

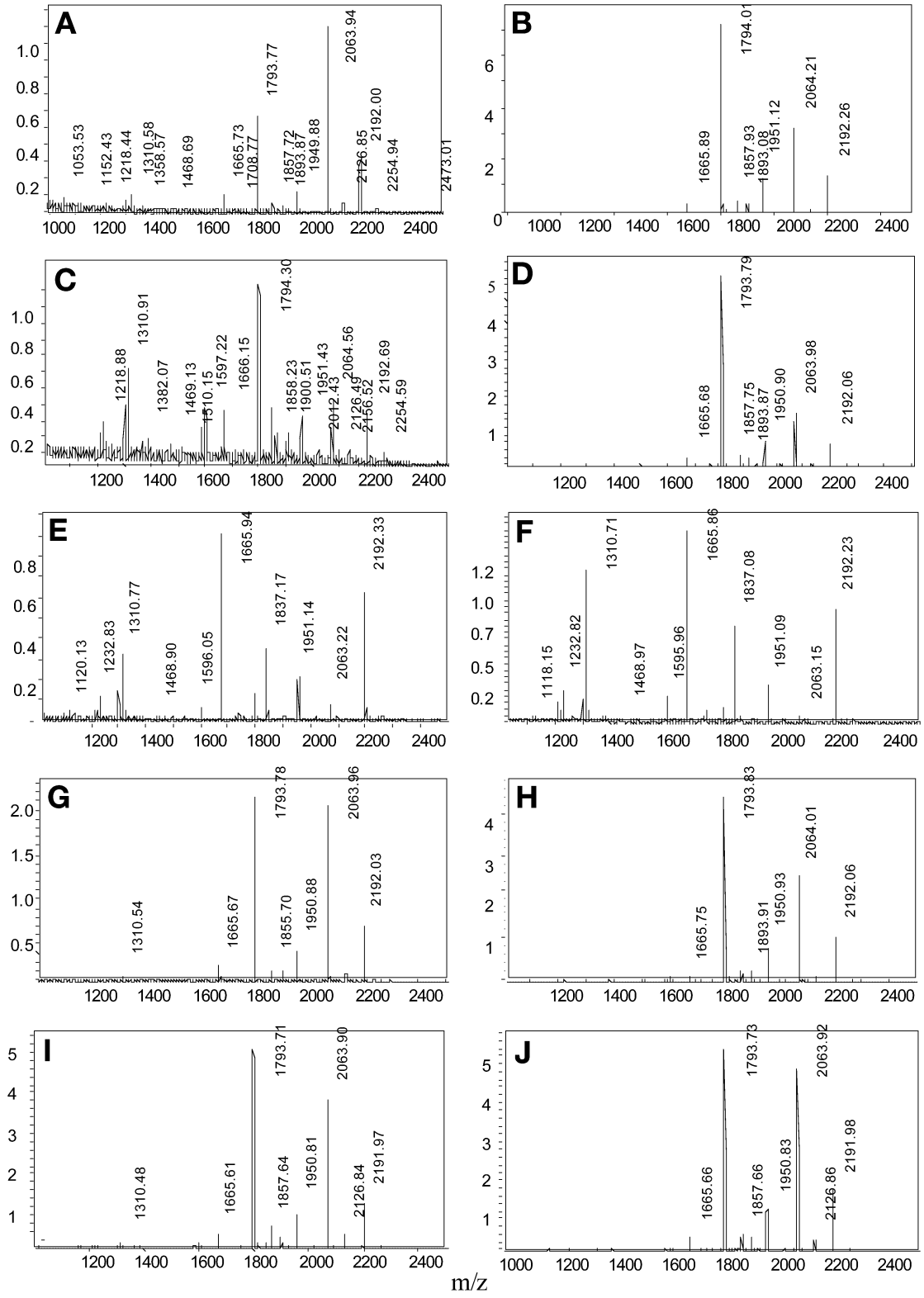
A
1053.53
1152.43
1218.44
1310.58
1358.57
1468.69
1665.73
1708.77
1793.77
1857.72
1893.87
1949.88
2063.94
2126.85
2192.00
2254.94
2473.01
B
1665.89
1794.01
1857.93
1893.08
1951.12
2064.21
2192.26
C
1218.88
1310.91
1382.07
1469.13
1510.15
1597.22
1666.15
1794.30
1858.23
1900.51
1951.43
2012.43
2064.56
2126.49
2156.52
2192.69
2254.59
D
1665.68
1793.79
1857.75
1893.87
1950.90
2063.98
2192.06
E
1120.13
1232.83
1310.77
1468.90
1596.05
1665.94
1837.17
1951.14
2063.22
2192.33
F
1118.15
1232.82
1310.71
1468.97
1595.96
1665.86
1837.08
1951.09
2063.15
2192.23
G
1310.54
1665.67
1793.78
1855.70
1950.88
2063.96
2192.03
H
1665.75
1793.83
1893.91
1950.93
2064.01
2192.06
I
1310.48
1665.61
1793.71
1857.64
1950.81
2063.90
2126.84
2191.97
J
1665.66
1793.73
1857.66
1950.83
2063.92
2126.86
2191.98
m/z

Table 1
The Protein Concentration of the Sample Preparation Schemes Described in the Material and Methods Prior to MALDI and ESI Analyses

| Preparation (mg/mL) | Protein concentration (mg/mL) |
|---|---|
| AS Supernatant | 0.26 |
| AS Pellet | 8.83 |
| P10 Excluded | 9.95 |
| P10 Penetrating | 1.08 |
| G25 Excluded | 9.72 |
| G25 Penetrating | 2.93 |
| ACN Supernatant | 0.25 |
| ACN Pellet | 9.2 |
| YM30 Cake | 3.19 |
| YM30 Filtrate | 0.26 |
| YM30 ACN Cake | 3.23 |
| YM30 ACN Filtrate | 0.27 |
| CIBA Flowthrough | 0.38 |
| C18 Eluant | 0.57 |
| C18 Flowthrough | 2.85 |

AS, anmonium sulfate; ACN, acetonitrile.

The protein concentrations of the various fractions are shown in **Table 1**.

### *Sample Detection Limit*

In general, the C18 method was the simplest and fastest way to collect peptides from serum. We found that the majority of peptides in serum were observed at less than m/z 2300 **(Fig. 3A)**. We examined the sensitivity of detecting serum peptides and found that peptides extracted from the equivalent of 1 µL of serum produced strong signal-to-noise ratios and showed a complex peptide mass distribution with many analytes well resolved by MALDI-TOF **(Fig. 3B)**. We compared the spectra from 1 µL of serum to that of the matrix alone and observed that the matrix ions were not prominently displayed in the MALDI spectra of serum. The use of an equivalent to 1 µL of serum resulted in robust detection; no significant advantage was observed with amounts larger than 5 µL.

### *Analyte Detection Limit*

We used a human fibrinogen β peptide as both an external and internal standard to estimate the levels of endogenous bovine fibrinogen β peptides in 2 µL of serum on a MALDI target spot. We observed that as little as 25 fmol of the human Glu-fibrinogen peptide (EGVNDNEEGFFSAR) could be detected in the serum peptide mixture, along with other peptides, including those of a bovine fibrinogen β-chain peptide (QFPTDYDEGQDDRP). We observed that 300–400 fmol of the Glu-fibrinogen peptide provided roughly a similar response compared to the endogenous fibrinogen β-chain peptide from the equivalent of 2 µL of serum **(Fig. 4A)**. When the instrument was externally standardized with purified human Glu-fibrinogen, we observed roughly the same response from 100 fmol of the human Glu-fibrinogen as the endogenous levels of bovine β fibrinogen peptide. In a separate experiment, we confirmed the presence of β fibrinogen peptides by LC-ESI. In **Fig. 4B**,**C**, we show one of the most intense ions that mapped to fibrinogen B peptide and the fragmentation spectra showing +1 ions that matched well with the predicted series from the peptide TDYDEGQDDRP.

Fig. 2. *(Opposite page)* The selective enrichment of low-mass polypetides by SPE after precipitation, chromatography, and ultrafiltration. **(A)** MALDI-TOF spectra of the acetonitrile supernatant and **(B)** pellet; **(C)** MALDI-TOF spectra of the ammonium sulfate supernatant and **(D)** pellet; **(E)** MALDI-TOF spectra of the depleted flow through volume and **(F)** retained albumin-containing fractions from Cibacron Blue chromatography in hypotonic salts at pH 8.0. **(G)** MALDI-TOF spectra of the excluded volume and **(H)** the penetrating volume of P10 chromatography; **(I)** MALDI-TOF spectra of the ammonium bicarbonate filtrate vs **(J)** the filtrate with ammonium bicarbonate and 20% acetonitrile.

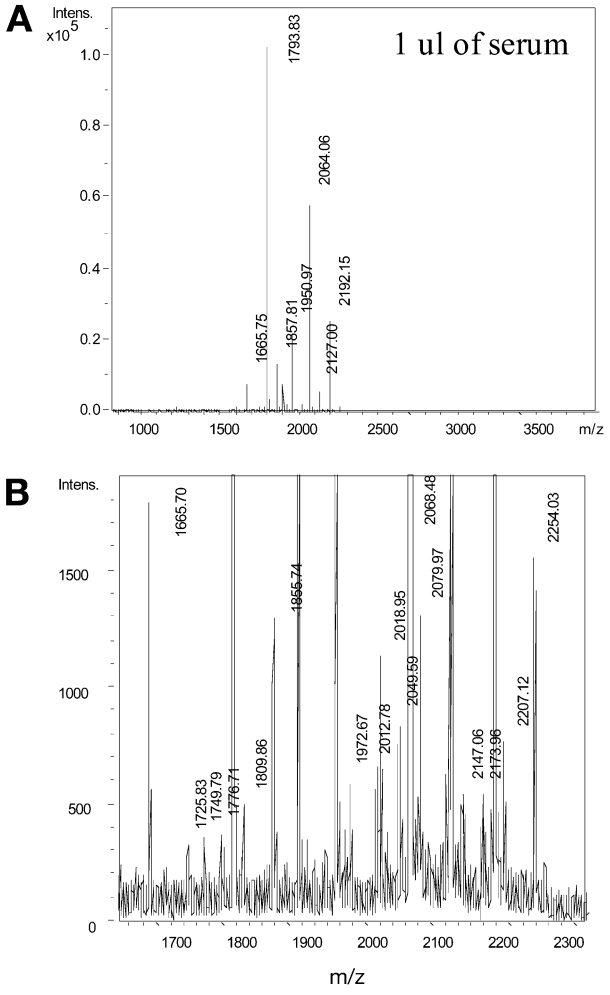


Fig. 3. The distribution of peptides from fetal calf serum after collection by SPE with C18 reversed phase from 1 μL of serum. One hundred microliters of serum were prepared by SPE with C18 and 1/100 of the product loaded onto a single MALDI target spot. **(A)** SPE with C18 yields peptides with a mass distribution of about 1000 to 2300 Da with the major peptides in the range 1600 to 2200 m/z. **(B)** Closer look at the lower-abundant peptides in 1 μL of serum. Closer examination of the spectra reveals that there are some lower-abundance peptides in the spectrum. Internal standardization experiments showed that detection of exogenous peptides was lost into the grass between 10 to 25 fmol on a single spot (not shown).

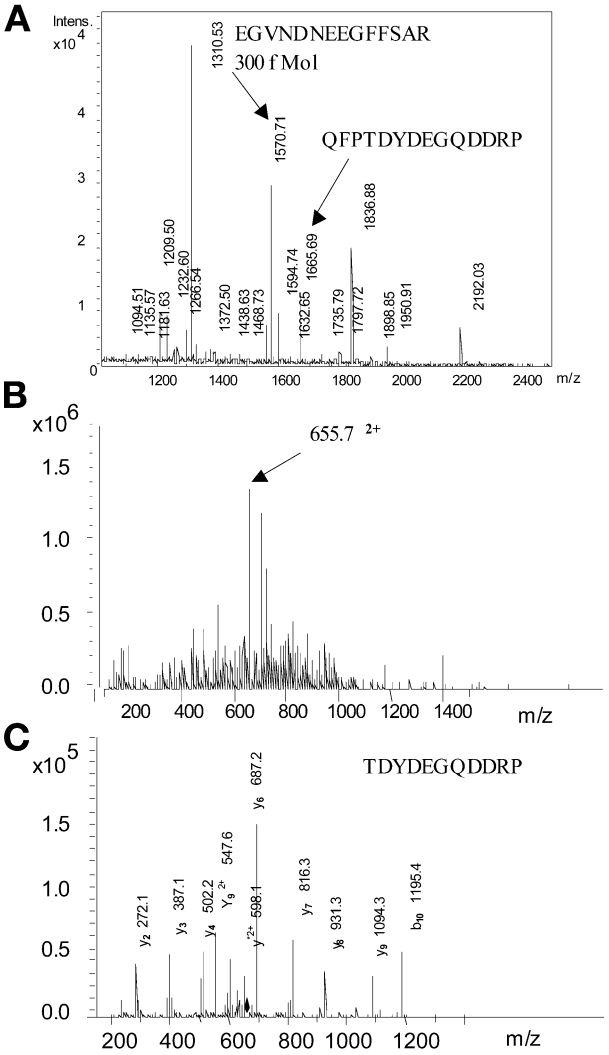


Fig. 4. MS and MS/MS spectra of a fibrinogen beta peptide collected by SPE with C18 reversed-phase chromatography. **(A)** The internal standardization of the MALDI spectrum from C18 SPE showing endogenous fibrinogen β peptide and exogenously added human glu-fibrinogen peptide. **(B)** The resolution of fibrinogen β peptide by LC-ESI ion trap. The peptide eluted at approx 44.5 min into the 90-min gradient. The MS spectra showing the parent 2+ ion at 655.7 m/z. **(C)** The collision-induced dissociation fragmentation spectra showing the remaining parent ion and product ions that correlated with fibrinogen β peptide.

### *MALDI Reproducibility*

We recorded the MALDI-TOF spectra of peptides collected from serum by SPE with C18 chromatography (**Fig. 5**). The peptides that eluted from reversed-phase resin with 25% acetonitrile prominently included fibrinogen peptide A (FPA), a previously demonstrated

marker of myocardial infarction in serum (7) (Fig. 5, inset). The peptides corresponding to FPA are shown in the box and are enlarged in the inset. We were interested in estimating the reproducibility of MALDI peptide patterns. We show the results from six consecutive MALDI spots from six previously frozen aliquots of the same serum. We observed that extracted peptides spotted onto MALDI targets produced a reproducible pattern of ions, including larger peptides, such as the β fibrinogen fragments and FPA (Fig. 5). From this experiment, we found that the relative variation in the intensity of α fibrinogen peptides is not large but the absolute variation is significant.

### *Identification of C18 Peptides by Offline LC-MALDI-TOF/TOF*

To better detect and identify lower-abundance components of MALDI-TOF spectra, we prepared serum by SPE with C18 and separated the components by offline LC onto metal plates prior to MALDI-TOF/TOF. The peptides were separated over a 30-min LC gradient from 5 to 65% acetonitrile at 2 μL/min onto metal plates and subsequently mixed with CHCA and examined by MALDI-TOF/TOF. Separation of the peptides by LC resulted in MALDI spots containing individual peptides and groups of chromatographically similar peptides (Fig. 6). Offline separation permitted identification of many peptides that were not prominent in the spectra of the batch C18 sample. The peptides LVKSQVNDNEEGFFSARGH at 2191 D, TDYDEGQDDRPKVGLGA at 1835 D, and TDYDEGQDDRP at 1310 Da were identified from the parent ion mass and fragmentation patterns produced by the TOF/TOF (not shown). A 1466-Da FPA was essentially purified onto a single MALDI target spot and subsequent MS/MS fragmentation resulted in a clear set of fragments corresponding to the FPA-like peptide TKESDFLAEGGGVR (Fig. 7A). Similarly,

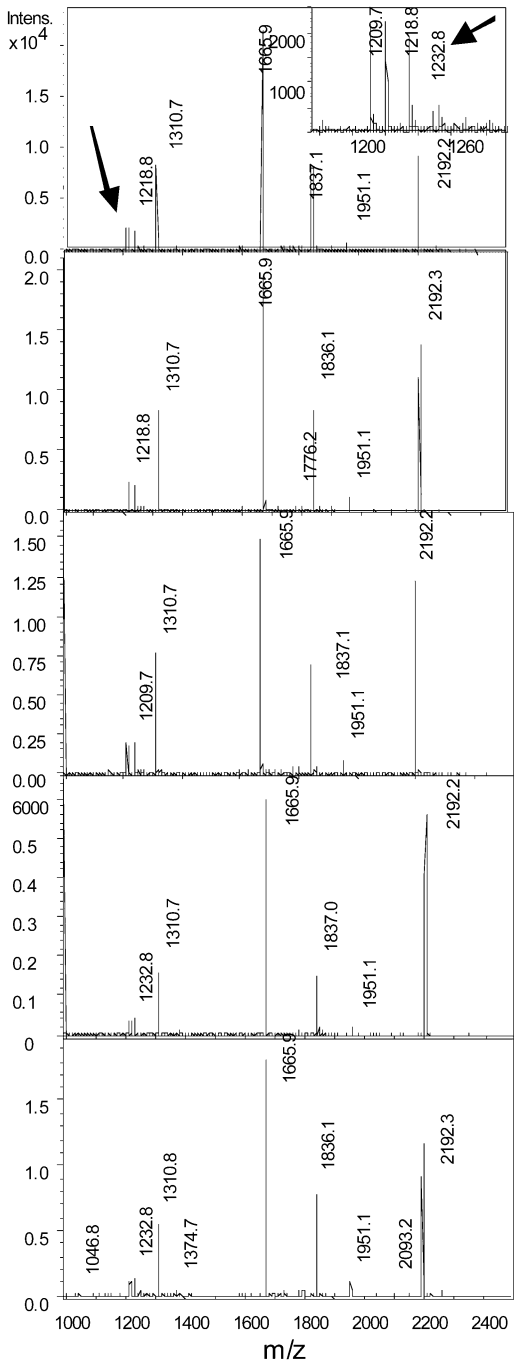


Fig. 5. The reproducibility of the MALDI spectra from serum prepared by SPE with C18 reversed-phase chromatography recorded with a Bruker TOF. A sample was broken into 100-μL aliquots and then frozen. Frozen aliquots were thawed, SPE with C18, and prepared for MALDI-TOF analysis. The position of an FPA cluster centered on 1218 m/z is shown by the black arrow and detailed view shown in the inset of the top panel.

a peptide at 1518 m/z showed MS/MS spectra correlating with the peptide SDPAGGEFLAEGGGVR **(Fig. 7B)**; a peptide at 1207 m/z correlated to EGDFLAEGGGVR and another homolog of FPA DPPSGDFLTEGGGV at 1232 Da was also observed (not shown). We detected a phosphopeptide $D_P$SGEGDFLAE GGGVR at 1546 m/z with a MOWSE score of 94 **(Fig. 7C)**. The parent molecule, fibrinogen, is an abundant protein in serum and its presence has previously been confirmed by MS of serum digested with trypsin and separated by two-dimensional-LC *(21,24,25)*. In addition, the presence of angiotensin-like peptide II DRVYIHPF at 1046 Da, which differs from the common MS standard DRVYVHPFHL, was apparently very weakly detected with a MOWSE score of 43 **(Fig. 7D)**. The different isoforms, posttranslational modifications of FPA we have detected here with automatic MS/MS identification by TANDEM or MASCOT or previously, are summarized in **Table 2**.

### *Liquid Chromatography-Electrospray Tandem Mass Spectrometry*

We tested extractions with C18 bound to paramagnetic beads in the 96-well format *(17)* or with more labor-intensive methods using commercial preparative microcolumns (ZipTips) and by batch collection over-magnetic and non-magnetic C18 resin and found the results to be similar. The TIC profile for the C18 beads is shown in **Fig. 8A**. C18 resin seemed to efficiently capture FPA **(Fig. 8B,C)** with confident correlation scores by MASCOT or X-TANDEM using no enzyme searching parameters, consistent with previous results *(3,7,17)*. We found that, for the identification of the fibrinogens and other serum proteins, C18 was a good choice and the combination of C18 with batch columns (SCX, WAX, SAX, HIC, and others), ultrafiltration, or organic solvent precipitation resulted in the identification of common serum proteins including fibrinogen α- and β-chain. In general,

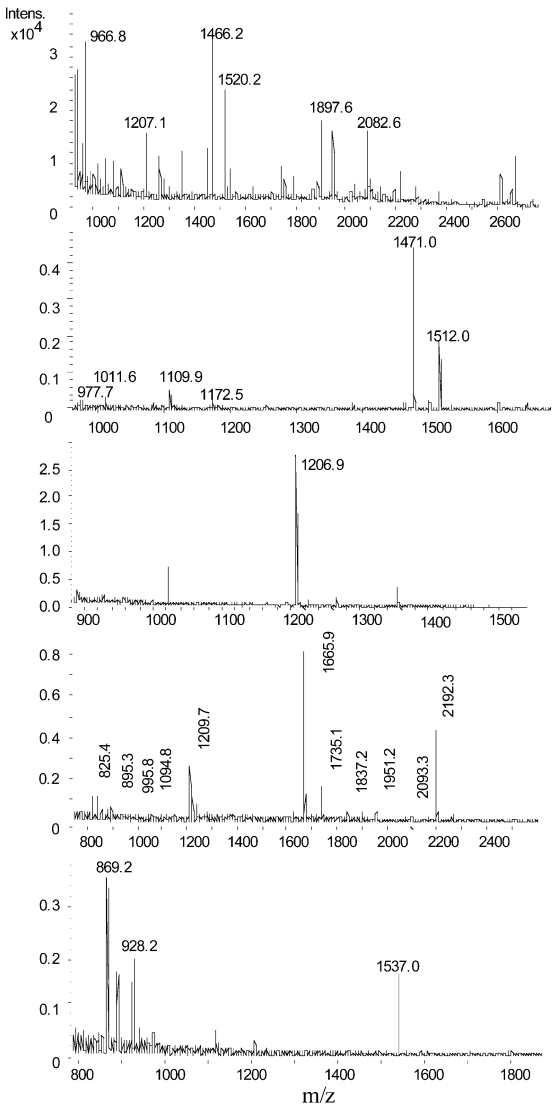


Fig. 6. Offline LC separation and detection of lower-abundance endogenous peptides from fetal bovine serum. Panels show selected offline LC-MALDI target MS spectrum illustrating the detection of lower-abundance polypeptides after SPE with C18 followed by capillary C18 separation of 1 µg of extracted peptides in more than 30 MALDI target spots over the course of 30 min from 5 to 65% acetonitrile.

poorer results were obtained with samples prepared with dye-binding or size-exclusion media **(Table 3)** and from ammonium sulfate supernatants. In agreement with others, we found that polypeptides may cofractionate with serum

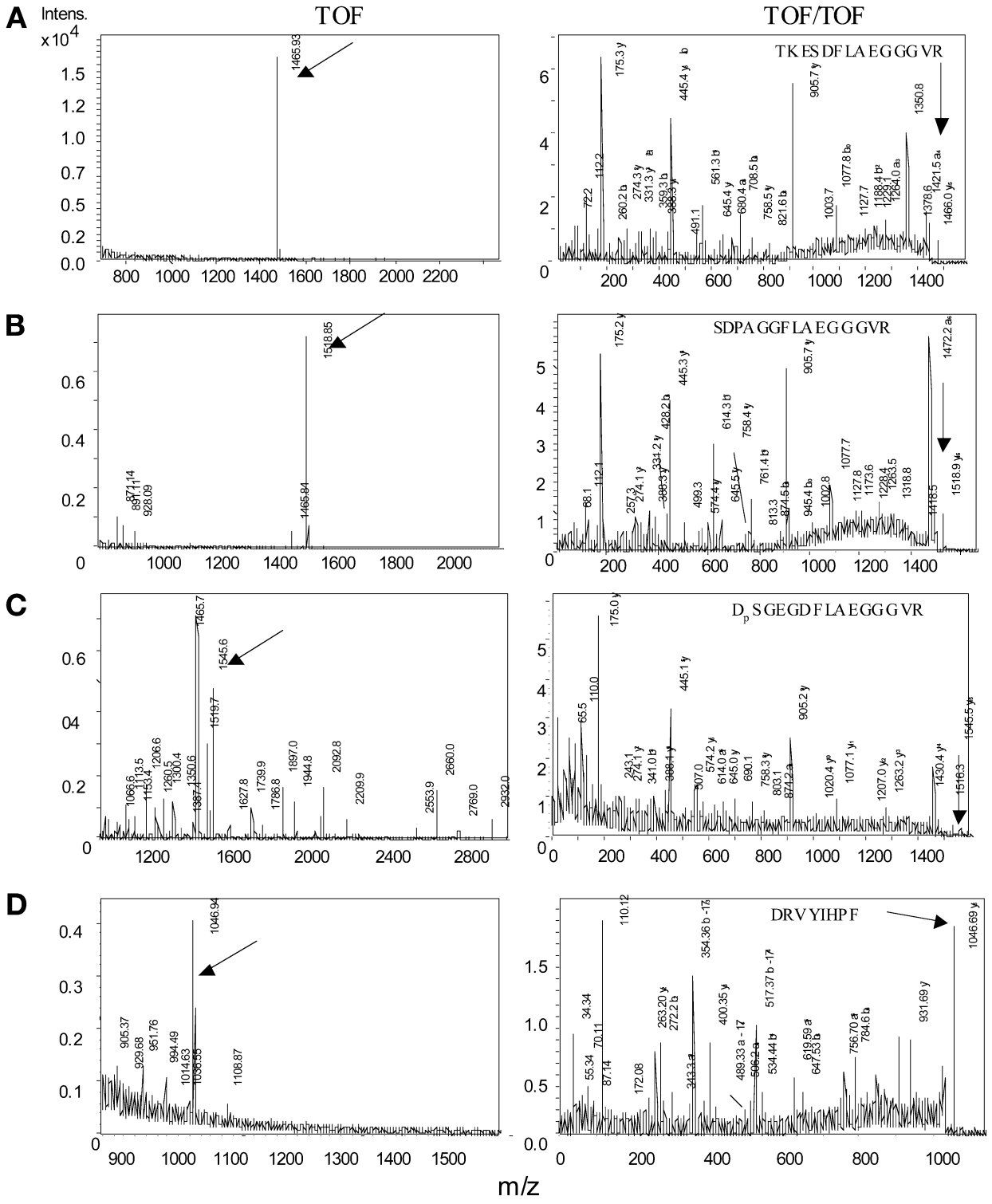


Fig. 7. Peptides separated over the course of offline capillary C18 LC over metal target plates analyzed by MALDI-TOF with fragmentation and identification by TOF/TOF. Panels: **(A)** the 1466 m/z peptide isolated on a MALDI target spot by offline analytical C18 LC separation with MS/MS spectra indicating the peptide TKESDFLAEGGGVR; **(B)** peptide at 1518 m/z with MS/MS spectra indicating the peptide SDPAGGE-FLAEGGGVR; **(C)** detection of a phsophorylated form of FPA $D_P$SGEGDFLAEGGGVR; **(D)** putative identification of angiotensin-like peptide II DRVYIHPF at 1046 m/z.

Table 2
Sequence Variations and Amino Acid Substitutions in FPA as Detected by MADLI-TOF/TOF and ESI-Ion Trap Analysis of Fetal Calf Serum and MALDI Analysis of Human Serum

| | | |
|---|---|---|
| Fetal calf serum | | |
| **MALDI** | | Da |
| | **E**GDFL**A**EGGGVR | 1206.6 |
| | S**DPAGGE**FL**A**EGGGVR | 1518.7 |
| | **TKESD**FL**A**EGGGVR | 1465.7 |
| | $DS_p$GEGDFL**A**EGGGVR | 1545.6 |
| **ESI** | | |
| | LKTEDGSDPPSGDFL**T**EGGGV | 1836.8 |
| Human serum | | |
| **MALDI** | | |
| | ADSGEGDFLAEGGGVR | 1536.7 |
| | DSGEGDFLAEGGGVR | 1465.7 |
| | GEGDFLAEGGGVR | 1263.6 |
| | **SES**AFLAEGGGVR | 1279.6 |
| | EDGFLAEGGGVR | 1206.6 |
| | GDFLAEGGGVR | 1077.5 |
| | DFLAEGGGVR | 1020.5 |
| **ESI** | | |
| | ADSGEGDFLAEGGGVR | 1536.7 |

Note that MALDI analysis of fetal calf serum identified the FL**A**EGGGVR sequence whereas ESI analysis seemed to have identified the FL**T**EGGGVR. In contrast, both MALDI and ESI analysis of human serum revealed the FL**A**EGGGVR sequence (*see* Marshall et al. 2003 ***[7]***).

albumin ***(18,26)***. MALDI ionization seemed to produce the best spectra from FPA isoforms containing the sequence FL**A**EGGGVR, similar to MALDI experiments with human serum. In LC-ESI ion-trap experiments the best MS/MS spectra were usually obtained from peptides containing the sequence FL**T**EGGGVR, in contrast to previously published human LC-ESI-MS/MS data ***(7)***.

### *Correlation Analysis of Peptides*

We used X-TANDEM ***(22)*** to correlate endogenous serum peptides to canonical blood proteins and proteins that are apparently not in high abundance in blood fluids ***(27)***. This algorithm is designed to detect peptides and modified peptides, but does not contemplate the detection of peptides joined by cystine bridges. We present these correlations with log expectation values of –2 or lower in **Table 3**, but we do not accept these results as unequivocal for the reasons outlined in the Discussion and previously ***(8)***. We found some peptides from known blood proteins, such as complement C3 and α-2-HS-glycoprotein, albumin, serine (or cysteine) proteinase inhibitor, globulins, and other well-known blood proteins such as major histocompatibility complex molecules, clotting factor VIII, and plasminogen **(Table 3)**.

We also detected single peptides from many other blood proteins including immunoglobulins, C1Q, α-2-macroglobulin, apolipoproteins, soluble receptors, disintegrins, fibronectins, thyroglobulin, ankyrins, proteoglycan, cadherin,

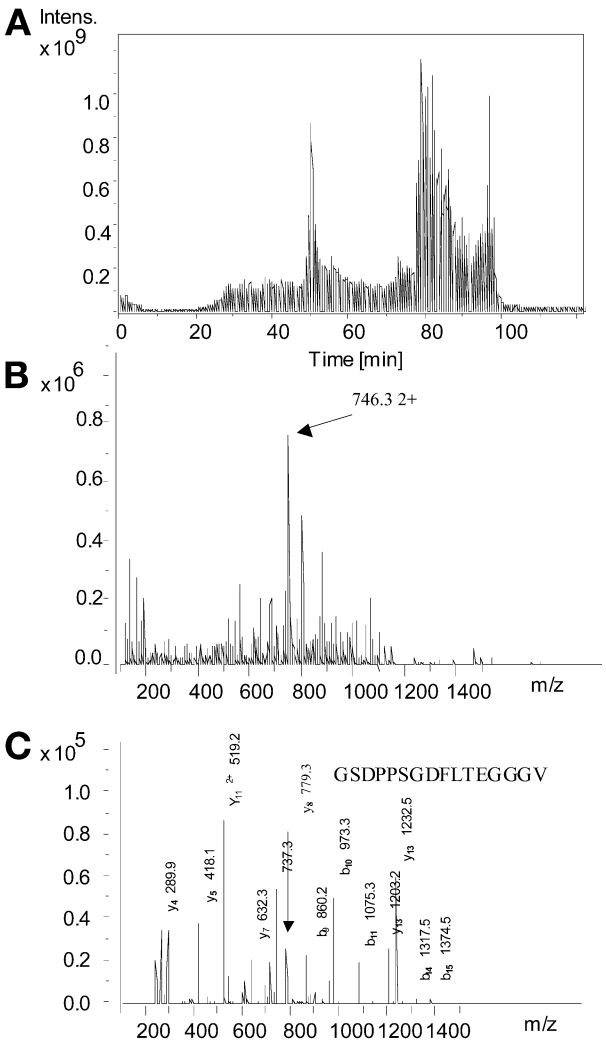


Fig. 8. Capillary LC-ESI ion trap analysis of endogenous peptides by SPE with C18. Panels: **(A)** the LC-ESI total ion current of peptides obtained from fetal calf serum using C18 SPE; **(B)** the resolution of FPA peptide by LC-ESI-ion trap. The peptide eluted at approx44.5 min into the 90-min gradient. The MS spectra showing the parent 2+ ion at 746 m/z; **(C)** the collision-induced dissociation fragmentation spectra showing the remaining parent ion and the product ions that correlate with FPA.

cellular adhesion, cell signaling, and many other proteins (not shown) including peptides mapping to known cancer biomarkers or associated with mechanisms or metabolic response to cancer. To illustrate this we show one peptide from five different isoforms from the collagen superfamily (**Table 4**) and we obtained one apparently significant peptide from the functionally and structurally related protein scavenger receptor B (**Table 2**). Few of these peptides were from the common high-abundance structural collagens I, II, and III, which would be expected if these peptides merely represented contamination from tissues. Rather, they were from the lower abundance collagens IV, V, VII, and XXVII that are associated with diverse regulatory functions on the cell surface or in the extracellular space (**Table 3**) *(11)*. In addition, we observed peptides that could be weakly correlated to known antibody variable chains by MASCOT and with MOWSE scores more than 40 to peptides from many proteins that are not definitively known to exist in blood (not shown).

## Discussion

We identified endogenous peptides from sera and have used MALDI MS/MS and LC-MALDI MS/MS to identify families of fragments from albumin, serum amyloid A, apolipoprotein E and CIII, clusterin, complement C3 and 4a, and fibrinogen α (*see* Marshall et al. 2004 *[8]* and US Patents especially 20020161180, 20020161187, 20060024671, and many others originally filed in April 2001) and these results have recently been confirmed by two other groups *(9,28)*. Similarly, we first showed that the mass spectra of serum peptides that differ between physiological states change with time and were the result of the ex vivo action of endo- and exopeptidases and that fresh plasma has fewer peptides than fresh serum *(7)* and these observations and hypothesis have been confirmed by other groups *(9,28)*. However, our position is that this type of pattern analysis work is only for discovery experiments. Biomarkers discovered by MS and LC-MS should show differences by using standard analysis of variance of each peptide mass and followed by leave-one-out

Table 3
A Comparison of the Peptides Identified From Fetal Calf Serum Using X-Tandem Algorithm After SPE C18 Alone or C18 Collection of Peptides After Various Extraction Procedures

| | AA | | | | | Log protein expectation value | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Description | Start | Stop | AA | M+H | Peptide | C18 | FILT | EXCL | DYE | AS | ORG SUP | ORG PELL | CO |
| Albumin | 596 | 607 | 122.7 | PKLVVSTQTALA | | –2.0 | –18.1 | | | | | –3.0 | |
| | 421 | 433 | 1479.7 | LGEYGFQNALIVR | | | | | | | | | |
| | 347 | 359 | 1567.7 | DAFLGSFLYEYSR | | | | | | | | | |
| Alpha-2-Hs-Glyco | 19 | 31 | 1369.7 | IPLDPVAGYKEPA | | –4.5 | –5.4 | | | | | –2.8 | |
| | 334 | 348 | 1474.8 | TPIVGQPSIPGGPVR | | | | | | | | | |
| | 33 | 50 | 1948.9 | DDPDTEQAALAAVDYINK | | | | | | | | | |
| Casein Alpha–Si | 189 | 201 | 1373.6 | TDAPSFSDIPNPI | | –3.8 | | | | | | | |
| Hemoglobin, Gamma | 31 | 40 | 1308.7 | LVVYPWTQRF | | –2.5 | –6.3 | | | | | | |
| | 40 | 58 | 2014.9 | FFESFGDLSSADAILGNPK | | | | | | | | | |
| Keratin 1 | 149 | 160 | 1302.7 | SLDLDSIIAEVK | | | –5.3 | | | | | | |
| Keratin 3 | 149 | 160 | 1302.7 | SLDLDSIIAEVK | | | –5.3 | | | | | | |
| Keratin 5 | 333 | 344 | 1302.7 | SLDLDSIIAEVK | | | –5.3 | | | | | | |
| Keratin 6 | 325 | 336 | 1329.7 | NLDLDSIIAEVK | | | –3.5 | | | | | | |
| Keratin 7 | 146 | 157 | 1302.7 | SLDLDSIIAEVK | | | –5.3 | | | | | | |
| Plasminogen | 27 | 45 | 2080 | DLLDDYVNTQGASLLSLSR | | | –4.1 | | | | | | |
| 116 Kda U5 Snrnp | 636 | 657 | 2253.1 | AMQPGVAEDGMGAHQSLKPVLT | | | –2.4 | | | | –2.4 | | |
| Avian V-Rel | 303 | 317 | 1223.6 | VAGNGGAGAAVGEPP | | –2.1 | | | | | | | |
| Brain Rptpmam4 | 11 | 23 | 1262.5 | CGYSVALGTSGFT | | | | | | | | | |
| C7orf12, Partial | 374 | 386 | 1588.8 | GWMQLVILIYHIS | | | | | | | | | |
| Cancer Candidate3 | 31 | 48 | 1438.6 | RGGGSCSGSVGGGGSGSL | | –2.0 | –2.0 | –2.0 | | | | | |
| Cochlear Otoferlin | 16 | 27 | 1307.5 | MDPVVCVEVGDE | | –2.0 | | | | | | | |
| Collagen Alpha 2(V) | 87 | 110 | 2289.1 | AQMAGLDEKSGLGSQVGLMPGSVG | | | | | –2.1 | –2.5 | | | |
| Complement C3 | 523 | 531 | 1031.5 | ILWESASLL | | –2.9 | –9.3 | –2.9 | | | | | |
| Epidermal Langerh | 55 | 68 | 1412.5 | DLADPSSSQDGSFS | | –2.1 | | | | | | | |
| Epsilon I Globin | 33 | 42 | 1308.7 | LVVYPWTQRF | | –2.5 | | | | | | | |
| Factor Xiii | 329 | 340 | 1305.7 | PPTVELQGLVPR | | –3.9 | | | | | | | |
| Fanconi Anemia E | 111 | 121 | 1198.5 | SLPSADGDHHY | | | | | | | | | –2 |
| Fibrinogen Alpha | 177 | 187 | 1038.4 | SGDFLTEGGGV | | –1341 | –224.1 | –49.8 | –84.7 | 24.7 | –92.1 | –82.2 | –40 |
| | 171 | 181 | 1106.4 | DGSDPPSGDFL | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 724 | 734 | 1224.5 | DEAESLEDLGF | | | | | | | |
| | 700 | 710 | 1225.6 | QFVSSSTTVNR | | | | | | | |
| | 175 | 187 | 1232.5 | PPSGDFLTEGGGV | | | | | | | |
| | 170 | 181 | 1235.5 | EDGSDPPSGDFL | | | | | | | |
| | 171 | 183 | 1336.5 | DGSDPPSGDFLTE | | | | | | | |
| | 174 | 187 | 1347.6 | DPPSGDFLTEGGGV | | | | | | | |
| | 723 | 734 | 1353.5 | EDEAESLEDLGF | | | | | | | |
| | 171 | 184 | 1393.5 | DGSDPPSGDFLTEG | | | | | | | |
| | 173 | 187 | 1434.6 | SDPPSGDFLTEGGGV | | | | | | | |
| | 171 | 185 | 1450.5 | DGSDPPSGDFLTEGG | | | | | | | |
| | 170 | 183 | 1465.5 | EDGSDPPSGDFLTE | | | | | | | |
| | 172 | 187 | 1491.6 | GSDPPSGDFLTEGGGV | | | | | | | |
| | 171 | 187 | 1606.6 | DGSDPPSGDFLTEGGGV | | | | | | | |
| | 722 | 735 | 1612.7 | MEDEAESLEDLGFK | | | | | | | |
| | 675 | 692 | 1667.7 | TGLAPEFAALGESGSSSS | | | | | | | |
| | 170 | 187 | 1735.7 | EDGSDPPSGDFLTEGGGV | | | | | | | |
| | 700 | 717 | 1809.8 | QFVSSSTTVNRGGSAIES | | | | | | | |
| | 169 | 187 | 1836.7 | TEDGSDPPSGDFLTEGGGV | | | | | | | |
| | 167 | 187 | 2077.9 | LKTEDGSDPPSGDFLTEGGGV | | | | | | | |
| | 179 | 187 | 894.4 | DFLTEGGGV | | | | | | | |
| | 173 | 181 | 934.4 | SDPPSGDFL | | | | | | | |
| | 171 | 180 | 993.3 | DGSDPPSGDF | | | | | | | |
| Fibrinogen Beta | 33 | 41 | 1094.4 | YDEGQDDRP | | –58.0 | –48.9 | –25.7 | –2.4 | –6.0 | –6.4 |
| | 33 | 41 | 1094.4 | YDEGQDDRP | | | | | | | |
| | 32 | 41 | 1209.4 | DYDEGQDDRP | | | | | | | |
| | 31 | 41 | 1310.5 | TDYDEGQDDRP | | | | | | | |
| | 28 | 41 | 1665.6 | QFPTDYDEGQDDRP | | | | | | | |
| | 28 | 42 | 1793.7 | QFPTDYDEGQDDRPK | | | | | | | |
| | 28 | 44 | 1949.8 | QFPTDYDEGQDDRPKVG | | | | | | | |
| | 28 | 47 | 2190.9 | QFPTDYDEGQDDRPKVGLGA | | | | | | | |
| Gamma Globin | 31 | 40 | 1308.7 | LVVYPWTQRF | | –2.5 | –6.3 | | | | |
| Hemoglobin Alpha | 192 | 203 | 1279.7 | FLANVSTVLTSK | | –3.5 | | | | | |
| | 177 | 186 | 1083.5 | LPSDFTPAVH | | | | | | | |
| Hemoglobin Beta | 56 | 65 | 1308.7 | LVVYPWTQRF | | –2.5 | | | | | |
| Hemoglobin, Gamma | 31 | 40 | 1308.7 | LVVYPWTQRF | | –2.5 | –6.3 | | | | |
| Hsmp-1 | 172 | 189 | 1386.6 | GPGHGPGPGGGSGQGPGG | –2.5 | –2.5 | | –2.6 | | –2.5 | |

*(Continued)*

Table 3 *(Continued)*

| Description | AA Start | AA Stop | AA | M+H | Peptide | Log protein expectation value C18 | FILT | EXCL | DYE | AS | ORG SUP | ORG PELL | CO |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mgc 15396 | 148 | 165 | 1695.8 | | PPASPPADGSPPLLPPGE | | | | | –2.0 | | | |
| Hypothetical | 51 | 65 | 1735.9 | | KFVEEQNTKKSEVAV | –2.4 | –2.0 | | | | | | |
| Kiaa 0303 | 520 | 531 | 1273.6 | | PVLPSNLSRGMS | | –2.2 | | | | | –2.7 | |
| Kiaa 0809 | 520 | 531 | 1273.6 | | PVLPSNLSRGMS | | –2.2 | | | | | | |
| Kiaa 1185 | 73 | 90 | 1666.7 | | PSGDNPATGCGSFGTIVS | –4.3 | | | | | | | |
| Luc7 Homolog | 61 | 73 | 1382.5 | | HSGMMCGCGKRAE | –2.4 | | | | | | | |
| Glutamater 6 | 55 | 70 | 1606.9 | | MAGVPALLPLLLALAQ | –2.0 | | | | | | | |
| Mhc Class I–Like | 251 | 268 | 1928.9 | | LPSSYDDKVAALSHVSCH | –2.1 | | | | | | | |
| Myc Induced Antigen | 70 | 80 | 1100.5 | | FYPGGGGGLEF | –2.1 | | | | | | | |
| Nedd4 – Binding | 66 | 82 | 1675.7 | | PGPNTGLMTGPDGKGMN | –2.0 | | | | | | | |
| Piccolo; Aczonin | 436 | 454 | 1833.8 | | QPMGMARAAAGPLPPISAD | –2.0 | | | | | | | |
| Tubulin, Beta 2 | 5 | 20 | 1623.6 | | EQGGSCGPVCGSLPFW | | –2.0 | | | | | | |
| IV Alpha 6 Collagen | 100 | 115 | 1427.7 | | GEPGLPGPPGLLGLPG | –3.7 | –2.1 | –2.6 | | –2.2 | –2.5 | –2.0 | –2.6 |
| Scavenger Receptor | 373 | 386 | 1498.7 | | DIHPVTGIPMNCSV | –2.2 | | | | | | | |
| S Proteinase Inhib.A | 403 | 416 | 1456.8 | | SPLFVGKVVNPTQA | | –2.5 | | | | | | |
| | 323 | 337 | 1593.8 | | SVLGDVGITEVFSDR | | | | | | | | |
| | 398 | 409 | 1250.6 | | DTDTGALLFIGK | | | | | | | | |
| S Proteinase Inhib.F | 398 | 409 | 1250.6 | | DTDTGALLFIGK | | | –2.9 | | | | –3.2 | |
| Small Cond K Chan | 83 | 96 | 891.3 | | GGGGGSTGGGGGGG | | | | | | –3.2 | | |
| Ring Finger 6 | 350 | 365 | 1450.7 | | LRSDAGPPGGATGPPT | | | | | | –2.9 | –2.9 | |
| Ring Finger 1 | 243 | 254 | 1284.6 | | GSVYSIAVTNHH | | | | –2.4 | | | | |
| C7orf12, Partial | 374 | 386 | 1588.8 | | GWMQLVILIYHIS | | | | | | –2.2 | | |

FILT, YM30; DYE, DEAE Blue; AS, ammonium sulfate; ORG SUP, organic supernatant; ORG PELL, organic pellet; CO, other collumns.

Table 4
Endogenous Fragments of the Collagen Superfamily Detected by LC-ESI MS/MS of Fetal Calf Serum

| RefSeq | Description | AA start | AA stop | M + H | Protein expect | Peptide |
|---|---|---|---|---|---|---|
| XP_595835.1 | Collagen α 2(V) | 87 | 110 | 2289.1 | –2.1 | AQMAGLDEKSGLGSQVGLMPGSVG |
| XP_604487.1 | Type IV α 6 collagen | 100 | 115 | 1427.8 | –3.7 | GEPGLPGPPGLLGLPG |
| XP_583692.1 | Type VII collagen | 2343 | 2358 | 1386.7 | –1.3 | PGPPGSAGVPGVPGLK |
| XP_591171.1 | Type IV collagen α 5 | 20 | 39 | 1839.9 | –1.2 | GFQGPPGPPGIPGPIGQPGL |
| XP_602219.1 | Collagen, type XXVII, α 1 | 294 | 305 | 1180.6 | –1.1 | SPVSPASSPRPV |

The endogenous peptides were extracted from fetal calf serum as described in the materials and methods and separated at 2 μL/min through a metal needle into a Bruker HCT ion trap. Expect, expectation value log probability associated with peptide.

cross validation (*7*). The peptides identified by MS/MS should be confirmed by a second method, such as immunology or the use of synthetic standards. In order to be useful in standard MS work flows, the peptides would have to be quantified by LC-ESI triple-stage quadrupole or with internally or externally standardized MS. All these important points remain to be given consideration, but, in a recent analysis of this field, it was suggested that the use of antibodies presenting on a surface-enhanced laser desorption/ionization may provide greater specificity (*29*) as previously contemplated (*30*). The Human Proteome Organization concluded that plasma was the preferred blood sample (*31*) but has recently indicated there was no difference (*32*). Villanueva et al. (*28*) conclude that serum is preferable for peptide profiling. We have used serum for peptide profiling and protein discovery (*7*,*8*). However, the best sample for a specific purpose will likely be determined on a case-by-case basis, and both serum and plasma may be useful. In the present article, we turn our attention to the use of ESI that might be more amenable to quantitative analysis.

It may be possible to use endogenous low-abundance peptides of serum and plasma as biomarkers. At present, there is considerable debate as to whether the so-called peptidome even exists (*33*) and in what form (*18*). Peptides specifically cleaved from parent molecules that are designed to act as paracrine or endocrine signals are likely to be present briefly in small amounts. MS have detection limits in the attomol range for pure substances; but, in complex mixtures, only a small number of analytes attain useful ionization in any one instance (*8*). Fibrinogen peptides have a concentration in plasma in the nanomoles per liter or the femtomoles per microliter. Levels in serum should be greater. The robust limit for automatic MS/MS identification using our capillary LC-ESI-ion trap was about 20 fmol for a pure known peptide. A 100 pmol mixture of three digested standard proteins were very strongly detected and identified. We easily detected the fibrinogens at apparent concentrations of about 300 fmol per 1 μL of serum (~300 n*M*). We used 100 μL of starting serum for our extractions and identifications and therefore had sufficient material to

identify fibrinogens with expectation values as low as E-50 and other known blood proteins with values as high as E-2. It has been suggested that physiologically relevant molecules in blood may only exist at levels less than the detection limits of MS (*34*,*35*). However, because the binding constants ($K_d$) of physiologically relevant proteins can be in the micromolar-to-nanomolar range, it would appear that some biologically relevant molecules in the nano- to picomolar range might be discovered from small amounts of serum.

Simple methods that have high yields and good resolving power are required to discover low abundance endogenous peptides. C18 chromatography worked well alone, or for collecting peptides previously fractionated by precipitation, filtration, or column chromatography. Thus, the optimization of preparative chromatographic materials, reagents and protocols for the extraction of peptides from biological blood fluids may prove to be a key step in advancing the identification of endogenous peptides and warrants much greater scrutiny than it has received to date (*36*). The large yield of peptides achieved by SPE permitted coupling of this method with other separation methods in serial combination, to explore more deeply the endogenous peptides of serum. C18 could be used to collect peptides from all other methods examined. We found that SPE with C18 provided the high yield and relative reproducibility required for discovery experiments. SPE with C18 resin worked in a variety of formats to extract and identify FPA. C18 microcolumns or 96-well plates with C18 magnetic beads could be used to rapidly prepare serum for identification of the potential biomarker FPA. C18 extraction occurs at low pH and this process will not only extract peptides that are freely dissolved in serum, but also those that may be noncovalently bound to high-mass proteins.

Albumin has many cysteine residues, but we did not observe stronger signals or different peptides from serum treated with DTT. It has been previously suggested that low-mass peptides may noncovalently bind to higher-molecular mass proteins, specifically BSA, and that these interactions could be disrupted by organic solvents used in ultrafiltration (*20*). We examined this concept using organic solvent and ammonium sulfate precipitation, dye affinity chromatography, size exclusion chromatography, and ultrafiltration. We used precipitation with acetonitrile and observed that a pool of polypeptides did not coprecipitate with the large-molecular mass proteins. The inclusion of albumin isolation prior to C18 extraction resulted in sample loss and reduced the number of peptides identified by LC-MS/MS. In agreement with previous results, we found that ultrafiltration could be used to detect peptides from blood (*20*). We concluded that not all peptides remain strongly bound to albumin at all times. However, peptides may potentially be bound to albumin or other lipoprotein complexes (*14*,*36*,*37*).

The intense MALDI spectral elements from unfractionated serum or plasma represent peptides from abundant circulating proteins of the blood (*7*,*9*). Biomarker peptide patterns have been discerned from unfractionated serum or plasma for many diseases or physiological states using TOF-MS, but these techniques can also discern the differences in sample collection procedures used by sites or individuals (*6*). We have previously proposed that the differences observed in MALDI spectra between physiological and disease states could reflect the activity of disease-specific proteases or enzymes released or activated in the blood that act on blood proteins (*7*). Although it seems possible that blood peptides might one day be used for robust diagnosis via MS, it is highly unlikely that this will be achieved without prefractionation, followed by targeted assays to quantify the peptides (*6*,*33*). It is possible that some of the peptides in plasma or serum may have been generated in vivo. However, previous time-course and inhibitor studies show that MALDI patterns of abundant serum

peptides change rapidly after blood collection to reach ex vivo steady states that represent the balance between the generation of peptides by protease activity and the degradation of peptides by amino and carboxy peptidases (*7*). Recently, an independent group repeated our original experiments (*7*) in which blood fluids were incubated ex vivo and they came to the same conclusion that blood peptide levels reflect the balance of endopeptidase and exopeptidase activities and that this principle might be used to detect disease specific peptidase activities (*28*). However, we argue that standard analytical identification, quantification, and statistical analysis should be used along with synthetic chromatography standards and immunological confirmation (*7*).

The use of SPE together with other preparation schemes, followed by LC-MS/MS, revealed multiple isoforms of FPA bearing different amino acid substitutions detectable by LC-MALDI-MS/MS. For myocardial infarction, the disease mechanism has been linked to the generation of common serum peptides, such as FPA and C3f, and confirmed by immunological methods (*7*). In the present study, we were able to identify numerous amino acid substitutions, and even posttranslational modification of the FPA peptide. This may indicate that, when combined with LC-separation, these techniques will permit the acquisition of information about specific peptide isoforms and modifications from biological samples. Developing sample preparation methods for assigning identity to low-molecular mass peptides is of central importance. These methods will help to transform mass spectra-based diagnosis from a relative technique of comparing unknown peak patterns (*5*) to a more robust method of quantifying known peptides.

The concept of a "peptidome" (*38*) brings up several additional considerations for the standard of proof associated with protein identifications (*39*). Serum is known to contain a variety of proteolytic enzymes including endo- and exopeptidases. Serum contains tryptic, chymotryptic, metalloproteinase, and other proteolytic activities. Because of the effect of constitutive amino- and carboxyl-exopeptidase activity (*7*), and perhaps the apparent presence of multiple proteases in sera (*24*), searches need to be performed with no enzyme specified. Previous studies have used MS analysis of serum and, apparently, fit the parent mass values to the predicted cleavage products of the major blood proteins in which a minority of peptides were positively identified by MS/MS. In practical experience, it is actually quite challenging to get unbiased and automatic identification of endogenous blood peptides by MS/MS compared to tryptic digests. In this article, we only present unbiased automatic MS/MS identifications with no inference of additional peptides by parent mass. Peptides such as angiotensin are single discrete polypeptide messengers that circulate in the blood and act as hormones or regulatory agents. In contrast to intentional tryptic digests, where there is a high probability of finding multiple peptides from the same digested protein, there is no guarantee of finding multiple peptides from the same protein in the endogenous peptides of serum. Confident identification of single peptides of low abundance is well known to be problematic and this difficulty is further compounded by the weak and complex signals associated with the endogenous peptides of serum. The confidence in single peptides would likely be somewhat increased by replicate identification in time (for example, Cancer Susceptibility was detected in multiple samples *see* Candidate 3, was detected in multiple samples, *see* **Table 2**) but this is likely not sufficient evidence because the same misidentification may be reproducible. Common immunological methods for proteins such as enzyme-linked immunosorbent assay or Western blot may be difficult to apply to small peptides. Single antibody methods, such as dot blots or affinity purification prior to MA, may provide the necessary confirmation (*2*).

Because there are only about 22,000 proteins in RefSeq, but there may be approx 1 billion variable chains associated with immunoglobulins that, as a class, are a highly abundant component of serum, we cannot entirely rule out the possibility that peptides from variable chains could confound mass spectral identifications in the absence of highly accurate mass measurements.

Detection and identification of blood peptides may be confounded by biochemical modifications. It remains possible that some peptides were difficult to analyse or shifted in parent mass by posttranslational modifications in addition to C- and N-terminal exopeptidase action that may remove basic charges and add acidic charges, such as phosphate of sulfate, that may effect ionization. It is known that removing phosphorylations may increase the ionization and detection of peptides ***(40)***. The fibrinopeptides, for example, are well known to be modified by both phosphorylation and sulfation ***(41–45)*** and this may have partially inhibited their detection. It has been suggested that sulfonation may be useful under some circumstances ***(46)***.

Questionable identifications of low-abundance peptides could be confirmed by similar chromatographic characteristics and fragmentation of synthetic peptides, but this approach is not useful for high-throughput discovery. A better solution could be the use of high-resolution MS, such as Fourier transform ion cyclotron resonance. For peptidomics, highly accurate parent masses combined with MS/MS analysis may also be required to provide highly confident predictions ***(47,48)***. If the endogenous peptide is indeed a powerful regulatory molecule, then it may be possible to detect its effects by classical bioassays. A practical means of building confidence in putative identifications may be comparison of endogenous peptides identified with lists of proteins identified from tryptic digestion of serum. For example, the rab/rac binding protein piccolo does not seem to be a blood protein, but similar proteins have been observed in blood ***(26,49)***. Other proteins that also seem to make little intuitive sense, such as the Nedd 4 binding protein, actually show significant homology to tumor necrosis factor-α, which is known to circulate remain unconfirmed.

A library of known endogenous peptide analytes might be created using LC-MS/MS and related to the changes in the intensity of a particular parent ion in spectra from normal and disease samples ***(16)***. In this article, we show that with optimal sample preparation, the peptides of biological fluids could be identified by LC-MS/MS using ESI and MALDI ionization. Bovine fetal serum is a rich source of growth factors. From this fluid, we correlated several peptides to isoforms of the collagen superfamily, a group known to be processed proteolytically to release powerful regulatory molecules ***(11)***. The ongoing gains in sensitivity of MS indicate that the number of endogenous peptides identified by MS/MS from biological fluids will likely increase in the future. To be useful, the peptides must be unambiguously identified by mass-accurate MS and MS/MS and quantified by triple-quadrupole MS. Their practical applicability in diagnostics remains an open question.

## Acknowledgments

We kindly thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the Heart and Stroke Foundation of Canada, and the Ontario Biomarker Network for operating grants to JGM in support of this work.

## References

1 Oleschuk, R. D., McComb, M. E., Chow, A., et al. (2000) Characterization of plasma proteins adsorbed onto biomaterials. By MALDI-TOFMS. *Biomaterials* **21**, 1701–1710.

2 Weinberger, S. R., Morris, T. S., and Pawlak, M. (2000) Recent trends in protein biochip technology. *Pharmacogenomics* **1,** 395–416.

3 Tammen, H., Hess, R., Uckert, S., et al. (2002) Detection of low-molecular-mass plasma peptides in the cavernous and systemic blood of healthy men during penile flaccidity and rigidity—an experimental approach using the novel differential peptide display technology. *Urology* **59,** 784–789.

4 Guo, J., Yang, E. C., Desouza, L., et al. (2005) A strategy for high-resolution protein identification in surface-enhanced laser desorption/ionization mass spectrometry: Calgranulin A and chaperonin 10 as protein markers for endometrial carcinoma. *Proteomics* **5,** 1953–1966.

5 Ardekani, A. M., Liotta, L. A., and Petricoin, E. F., 3rd. (2002) Clinical potential of proteomics in the diagnosis of ovarian cancer. *Expert Rev. Mol. Diagn.* **2,** 312–320.

6 Karsan, A., Eigl, B. J., Flibotte, S., et al. (2005) Analytical and preanalytical biases in serum proteomic pattern analysis for breast cancer diagnosis. *Clin. Chem.* **51,** 1525–1528.

7 Marshall, J., Kupchak, P., Zhu, W., et al. (2003) Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J. Proteome Res.* **2,** 361–372.

8 Marshall, J., Jankowski, A., Furesz, S., et al. (2004) Human serum proteins preseparated by electrophoresis or chromatography followed by tandem mass spectrometry. *J. Proteome Res.* **3,** 364–382.

9 Koomen, J. M., Li, D., Xiao, L. C., et al. (2005) Direct tandem mass spectrometry reveals limitations in protein profiling experiments for plasma biomarker discovery. *J. Proteome Res.* **4,** 972–981.

10 Chertov, O., Biragyn, A., Kwak, L. W., et al. (2004) Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry. *Proteomics* **4,** 1195–1203.

11 Ricard-Blum, S. and Ruggiero, F. (2005) The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathol. Biol.* **53,** 430–442.

12 Eisenberg, P. R., Sherman, L. A., Schectman, K., Perez, J., Sobel, B. E., and Jaffe, A. S. (1985) Fibrinopeptide A: a marker of acute coronary thrombosis. *Circulation* **71,** 912–918.

13 Michael, I. P., Sotiropoulou, G., Pampalakis, G., et al. (2005) Biochemical and enzymatic characterization of human kallikrein 5 (hK5), a novel serine protease potentially involved in cancer progression. *J. Biol. Chem.* **280,** 14,628–14,635.

14 Orvisky, E., Drake, S. K., Martin, B. M., et al. (2006) Enrichment of low molecular weight fraction of serum for MS analysis of peptides associated with hepatocellular carcinoma. *Proteomics* **6,** 2895–2902.

15 Theodorescu, D., Wittke, S., Ross, M. M., et al. (2006) Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. *Lancet Oncol.* **7,** 230–240.

16 Richter, R., Schulz-Knappe, P., Schrader, M., et al. (1999) Composition of the peptide fraction in human blood plasma: database of circulating human peptides. *J. Chromatogr. B Biomed. Sci. Appl.* **726,** 25–35.

17 Villanueva, J., Philip, J., Entenberg, D., et al. (2004) Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal. Chem.* **76,** 1560–1570.

18 Lowenthal, M. S., Mehta, A. I., Frogale, K., et al. (2005)Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clin. Chem.* **51,** 1933–1945.

19 Lopez, M. F., Mikulskis, A., Kuzdzal, S., et al. (2005) High-resolution serum proteomic profiling of Alzheimer disease samples reveals disease-specific, carrier-protein-bound mass signatures. *Clin. Chem.* **51,** 1946–1954.

20 Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., and Veenstra, T. D. (2003) Characterization of the low molecular weight human serum proteome. *Mol. Cell Proteomics* **2,** 1096–1103.

21 Marshall, J., Jankowski, A., Furesz, S., et al. (2004) Human serum proteins preseparated by electrophoresis or chromatography followed by tandem mass spectrometry. *J. Proteome Res.* **3,** 364–382.

22 Craig, R. and Beavis, R. C. (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20,** 1466–1467.

23 Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–3567.

24 Adkins, J. N., Varnum, S. M., Auberry, K. J., et al. Toward a human blood serum proteome: analysis by multidimensional separation

coupled with mass spectrometry. *Mol. Cell Proteomics* **1,** 947–955.
25 Chan, K., Lucas, D. A., Hise, D., et al. (2004) Analysis of the human serum proteome. *Clin. Proteomics* **1,** 101–225.
26 Shen, Y., Kim, J., Strittmatter, E. F., et al. (2005) Characterization of the human blood plasma proteome. *Proteomics* **5,** 4034–4045.
27 Kapp, E. A., Schutz, F., Connolly, L. M., et al. (2005) An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. *Proteomics* **5,** 3475–3490.
28 Villanueva, J., Shaffer, D. R., Philip, J., et al. (2006) Differential exoprotease activities confer tumor–specific serum peptidome patterns. *J. Clin. Invest.* **116,** 271–284.
29 Liotta, L. A. and Petricoin, E. F. Serum peptidome for cancer detection: spinning biologic trash into diagnostic gold. *J. Clin. Invest.* **116,** 26–30.
30 Weinberger, S. R., Boschetti, E., Santambien, P., and Brenac, V. (2002) Surface-enhanced laser desorption-ionization retentate chromatography mass spectrometry (SELDI-RC-MS): a new method for rapid development of process chromatography conditions. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **782,** 307–316.
31 Omenn, G. S., States, D. J., Adamski, M., et al. (2005) Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* **5,** 3226–3245.
32 States, D. J., Omenn, G. S., Blackwell, T. W., et al. Challenges in deriving high–confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat. Biotechnol.* **24,** 333–338.
33 Diamandis, E. P. (2003) Point: proteomic patterns in biological fluids: do they represent the future of cancer diagnostics? *Clin. Chem.* **49,** 1272–1275.
34 Pieper, R., Gatlin, C. L., Makusky, A. J., et al. (2003) The human serum proteome: display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. *Proteomics* **3,** 1345–1364.
35 Anderson, N. L. and Anderson, N. G. (2003) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell Proteomics* **2,** 50.
36 Sviridov, D., Meilinger, B., Drake, S. K., Hoehn, G. T., and Hortin, G. L. (2006) Coelution of other proteins with albumin during size-exclusion HPLC: implications for analysis of urinary albumin. *Clin. Chem.* **52,** 389–397.
37 Hortin, G. L., Shen, R. F., Martin, B. M., and Remaley, A. T. (2006) Diverse range of small peptides associated with high-density lipoprotein. *Biochem. Biophys. Res. Commun.* **340,** 909–915.
38 Verhaert, P., Uttenweiler-Joseph, S., de Vries, M., Loboda, A., Ens, W., and Standing, K. G. (2001) Matrix-assisted laser desorption/ionization quadrupole time-of-light mass spectrometry: an elegant tool for peptidomics. *Proteomics* **1,** 118–131.
39 Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K., and Nesvizhskii, A. (2004) The need for guidelines in publication of peptide and protein identification data: working group on publication guidelines for peptide and protein identification data. *Mol. Cell Proteomics* **3,** 531–533.
40 Corthals, G. L., Aebersold, R., and Goodlett, D. R. (2005) Identification of phosphorylation sites using microimmobilized metal affinity chromatography. *Methods Enzymol.* **405,** 66–81.
41 Barling, P. M., Palmer, D. J., and Christie, D. L. (1986) Preparation of desulphated bovine fibrinopeptide B and demonstration of its sulphation in vitro by an enzyme system from neuroblastoma–glioma hybrid cells. *Int. J. Biochem.* **18,** 137–141.
42 Krajewski, T. and Blomback, B. (1968) The location of tyrosine-O-sulphate in fibrinopeptides. *Acta. Chem. Scand.* **22,** 1339–1346.
43 Lucas, J. and Henschen, A. (1986) Identification and assay of phosphoserine and tyrosine-O-sulphate in fibrinopeptides by reversed-phase high-performance liquid chromatography. *J. Chromatogr.* **369,** 357–364.
44 Maurer, M. C., Peng, J. L., An, S. S., Trosset, J. Y., Henschen–Edman, A., and Scheraga, H. A. (1998) Structural examination of the influence of phosphorylation on the binding of fibrinopeptide A to bovine thrombin. *Biochemistry* **37,** 5888–5902.
45 Seydewitz, H. H., Matthias, F. R., Schondorf, T. H., and Witt, I. (1987) Increase in the degree

of phosphorylation of circulating fibrinogen under thrombolytic therapy with urokinase. *Thromb. Res.* **46,** 437–445.

46 Lee, Y. H., Kim, M. S., Choie, W. S., Min, H. K., and Lee, S. W. (2004) Highly informative proteome analysis by combining improved N-terminal sulfonation for de novo peptide sequencing and online capillary reverse-phase liquid chromatography/tandem mass spectrometry. *Proteomics* **4,** 1684–1694.

47 Hunt, D. F., Shabanowitz, J., Yates, J. R., 3rd, Zhu, N. Z., Russell, D. H., and Castro, M. E. (1987) Tandem quadrupole Fourier-transform mass spectrometry of oligopeptides and small proteins. *Proc. Natl. Acad. Sci. USA* **84,** 620–623.

48 Bergquist, J., Palmblad, M., Wetterhall, M., Hakansson, P., and Markides, K. E. (2002) Peptide mapping of proteins in human body fluids using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mass Spectrom. Rev.* **21,** 2–15.

49 Ping, P., Vondriska, T. M., Creighton, C. J., et al. (2005) A functional annotation of subproteomes in human plasma. *Proteomics* **5,** 3506–3519.