



The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title: Use of Novel Chemistry & Microwave Instrumentation to Improve Body Fluid Assay Sensitivity & Speed while Reducing Costs

Author(s): Donald Siegel, Ph.D.

Document Number: 251915

Date Received: July 2018

Award Number: 2012-DN-BX-K044

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

USE OF NOVEL CHEMISTRY & MICROWAVES TO IMPROVE BODY FLUID ASSAY SENSITIVITY & SPEED WHILE REDUCING COSTS

Final Technical Report NIJ Grant 2012-DN-BX-K044

Principal Investigator:

Donald Siegel, Ph.D.

Principal Scientist

Office of Chief Medical Examiner

421 East 26th Street

New York, NY 10016

Tel: 212-323-1434

Fax: 212-323-1560

Email: dsiegel@ocme.nyc.gov

Web: www.nyc.gov/ocme

TABLE OF CONTENTS

ABBREVIATIONS	4
ABSTRACT	6
EXECUTIVE SUMMARY	6
SYNOPSIS OF THE PROBLEM	6
PURPOSE	7
RESEARCH DESIGN.....	7
FINDINGS	8
CONCLUSIONS INCLUDING IMPLICATIONS FOR POLICY AND PRACTICE.....	28
INTRODUCTION	31
STATEMENT OF THE PROBLEM	31
LITERATURE REVIEW AND CITATIONS	32
RATIONALE FOR RESEARCH	35
RESULT SECTIONS	
SECTION 1 – COMBINATORIAL PEPTIDE LIGAND CHROMATOGRAPHY–DYNAMIC RANGE REDUCTION	36
COMBINATORIAL PEPTIDE LIGAND CHROMATOGRAPHY	36
<i>What is CPLC</i>	36
<i>Experimental Design</i>	36
<i>Sample Preparation Methods</i>	37
<i>Hemoglobin, Albumin & Immunoglobulin Depletion</i>	37
<i>Polyacrylamide Gel Electrophoresis</i>	38
<i>Mass Spectrometry Methods</i>	38
<i>Results</i>	38
<i>Comparison of Extraction and Elution Buffers</i>	38
<i>Comparison of Hb/Alb/IgG Depletion and IEF to CPLC</i>	42
DETERMINATION OF CPLC COLUMN BINDING CAPACITY	44
<i>Methods</i>	44
<i>Results</i>	45
BODY FLUID MIXTURES WITH CPLC.....	48
<i>Methods</i>	49
<i>Results</i>	49
SECTION 2 – REDUCING ASSAY TIME: ORGANIC SOLVENTS, TEMPERATURE & MICROWAVE-ASSISTED PROTEIN DIGESTION	53
EXPERIMENT 1 – ESTABLISHING THE ASSAY SYSTEM.....	54
EXPERIMENT 2 – EVALUATION OF TRYPSIN DIGESTION TEMPERATURE OVER TIME.....	58
EXPERIMENT 3 – EVALUATING THE EFFECTS OF ORGANIC SOLVENTS ON TRYPSIN DIGESTION	59
EXPERIMENT 4 – EVALUATING MICROWAVE-ASSISTED ENZYMATIC DIGESTION.....	60
EXPERIMENT 5 – EVALUATING MICROWAVE POWER AND NEED FOR WATER BATH.....	63
EXPERIMENT 6 – COMBINED DIGESTION CONDITIONS AT STANDARD TRYPSIN CONCENTRATION	64
SECTION 3 – MENSTRUAL BLOOD AND VAGINAL FLUID MARKER STABILITY	65
MATERIALS & METHODS	65
RESULTS.....	67
<i>Vaginal Fluid</i>	67
<i>Menstrual Blood</i>	70

EFFECTS OF REAGENTS BLUESTAR AND LUMINOL ON VAGINAL FLUID MARKERS.....	72
<i>Materials & Methods</i>	72
<i>Results</i>	73
THE VAGINAL FLUID MICROBIOME	74
SECTION 4 – SIMULTANEOUS PROTEIN & DNA EXTRACTION & PURIFICATION	75
METHODS.....	75
<i>AllPrep DNA/RNA/Protein Mini Kit Method</i>	75
<i>FASP Method</i>	76
<i>NYC OCME Method</i>	76
RESULTS.....	77
<i>Quantitative and Qualitative Protein Analysis</i>	77
<i>Quantitative and Qualitative DNA Analysis</i>	81
DISCUSSION	83
CONCLUSION.....	84
CONCLUSIONS, IMPLICATIONS FOR POLICY, PRACTICE& FUTURE RESEARCH.....	84
CONCLUSIONS	84
IMPLICATIONS FOR POLICY AND PRACTICE.....	86
IMPLICATIONS FOR FURTHER RESEARCH.....	87
REFERENCES.....	88
DISSEMINATION OF RESEARCH FINDINGS	94
PUBLICATIONS.....	94
PRESENTATIONS	94

ABBREVIATIONS

A2ML1 = alpha-2-macroglobulin-like 1 precursor
ABC = ammonium bicarbonate
ACN = acetonitrile
Alb = albumin
AP = acid phosphatase
BCA = bicinchoninic acid
BSA = bovine serum albumin
CPLC = combinatorial ligand peptide chromatography
CST2 = cystatin SA
DTT = dithiothreitol
ENO = enolase
FASP = filter-aided sample preparation
GAPDH = glyceraldehydes-3-phosphate dehydrogenase
GnHCl = guanidine hydrochloride
Hb = hemoglobin
HCCA = α -cyano 4-hydroxycinnamic acid
HPLC = high performance liquid chromatography
IAA = iodoacetamide
IgG = immunoglobulin G
K2C8 = keratin, type II cytoskeletal 8
LC-MALDI = liquid chromatography matrix-assisted laser desorption ionization
LEG1 = LEG1 homolog precursor
MiGE = microfluidic gel electrophoresis
MS = mass spectrometry
MW = molecular weight

MWCO = molecular weight cut-off

NP-40 = nonyl phenoxypolyethoxyethanol

PAGE = polyacrylamide gel electrophoresis

PAEP = progesterone-associated endometrial protein (glycodelin)

PMSF = phenylmethylsulfonyl fluoride

RLT = Qiagen lysing buffer for cells and tissue

SDS = sodium dodecyl sulfate

SLC4A1 = solute carrier family 4 member 1; 2 anion transport protein

SPRR3 = small proline-rich protein 3

SUP = supernatants

TCEP = TCEP tris(2-carboxyethyl)phosphine

TFA = trifluoroacetic acid

TMPRSS11B = Transmembrane protease, serine 11B (TM11B)

TMPRSS11D = Transmembrane protease, serine 11D (TM11D)

UCA= 8M Urea, 2% CHAPS and 5% Acetic acid extraction buffer

WH = whole homogenates

ABSTRACT

Body fluid identification plays a vital role in forensic investigations contributing to both, case evidence as well as directing the course of further investigations - particularly with respect to sexual assaults. Yet the methods used for body fluid detection have, for the most part, failed to keep pace with scientific and technological advances. As a consequence, current testing is a diverse mixture of techniques that vary dramatically in sensitivity, reproducibility, speed and cost. Most are not confirmatory, and there are many body fluids for which no routine tests are currently available, e.g. menstrual blood and vaginal fluid.

With the aid of previous NIJ support, we developed a proteomic mass spectrometry (**MS**) assay for blood, saliva and semen. This single test is confirmatory and can automatically identify any one of these body fluids in an unknown sample without a priori knowledge of what they might be. Species identification is also automatically determined. Our current system can process ~ 750 samples per week. Many more samples could be processed, but trypsin digestion of proteins, a prerequisite for nearly all MS assays, requires overnight incubation. We have also developed confirmatory assays for menstrual blood, vaginal fluid as well as body fluid mixtures. These tests, however, require protein purification prior to MS and consequently require significantly more time.

Recently, advances in protein chromatography and microwave technology have made significant gains not only in dramatically reducing sample preparation and digestion times (in some cases from days to minutes), but also improving the quality of analytes and consequently, test results. Current body fluid testing in common casework requires multiple assay systems: Kastle-Meyer as a presumptive test for blood, amylase agar plates as a presumptive test for saliva tests, and AP and p30 ELISA presumptive tests for sperm. This type of consecutive testing is time and sample consuming, and expensive. We are beginning validation testing of our MS assay. The overall goal of this application is to test and evaluate new technologies and integrate them into our current assays in order to make them faster, more accurate, more sensitive and less expensive. In addition, we also evaluated the simultaneous extraction of protein and DNA from samples in order to conserve sample use.

EXECUTIVE SUMMARY – This summary is divided into five sections: 1) SYNOPSIS OF THE PROBLEM, 2) PURPOSE, 3) RESEARCH DESIGN, 4) FINDINGS and 5) CONCLUSIONS INCLUDING IMPLICATIONS FOR POLICY AND PRACTICE.

1. SYNOPSIS OF THE PROBLEM: Body fluid identification plays a vital role in forensic investigations contributing to case evidence as well as in directing the course of further investigations - particularly with respect to sexual assaults. Yet the methods used for body fluid detection have, for the most part, failed to keep pace with scientific and technological advances. As a consequence, current testing is a diverse mixture of techniques that vary dramatically in sensitivity, reproducibility, speed, cost and above all specificity – i.e. their ability to confirm that test results are real. Additionally, there are many body fluids for which no routine tests are currently available, e.g. menstrual blood and vaginal fluid.

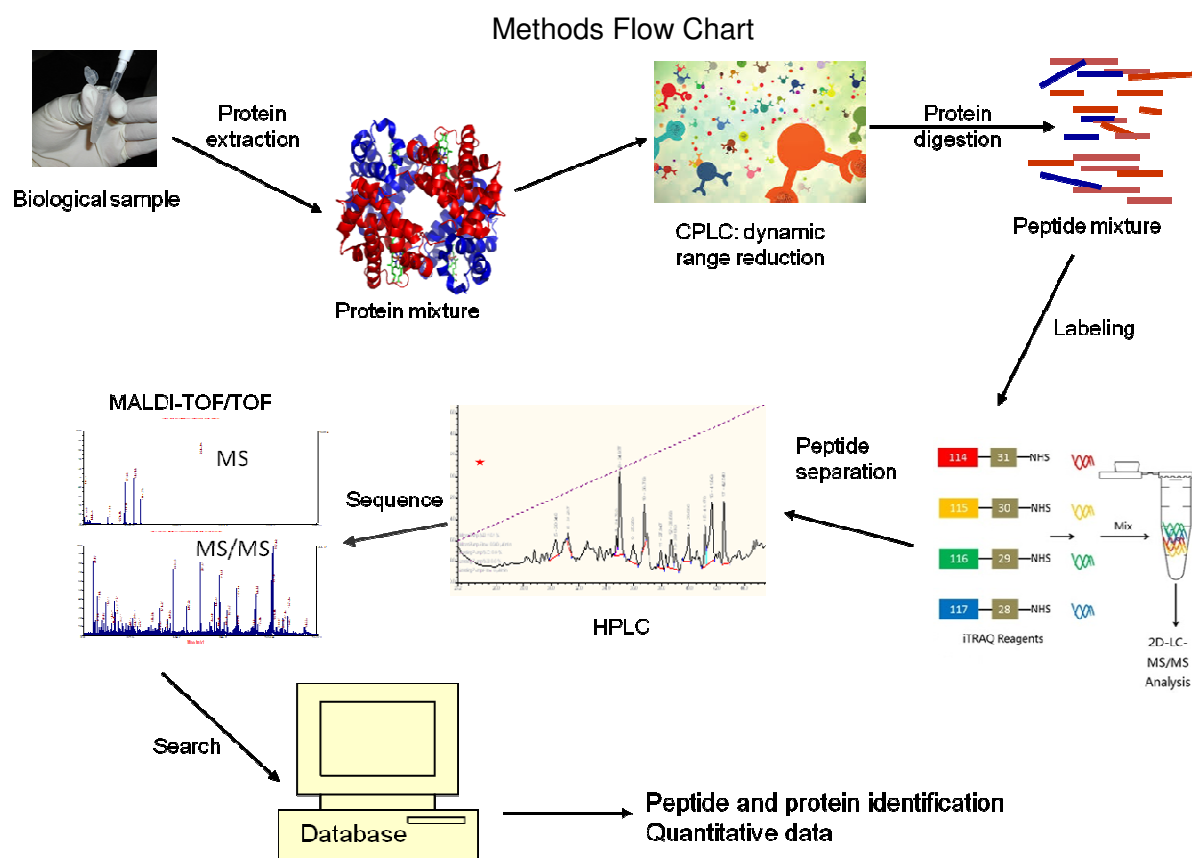
Adoption of a mass spectrometry (**MS**) proteomics assay for body fluid detection would result in a single, uniform assay, using the same extraction, purification and detections methods

for all body fluids, eliminating multiple techniques, saving time, money and sample, while giving the same confirmatory results for all body fluids. Above all, it would allow for the establishment of common, uniform standards throughout the forensic community in a manner similar to what is now in place for DNA testing.

2. PURPOSE: Recently, advances in protein chromatography and microwave technology purport to reduce sample preparation and digestion times. The goal of this application was to test and evaluate these new technologies and integrate them, when applicable, into our current assays in order to make them faster, more accurate, more sensitive and less expensive. In addition, we plan to i) evaluate the simultaneous extraction of protein and DNA from samples in order to conserve sample use, ii) evaluate the stability of menstrual blood and vaginal fluid markers over time and iii) assess the potential use of recently published vaginal fluid flora [1,2] as additional markers for our vaginal fluid assay.

3. RESEARCH DESIGN: Our goal was to improve the speed, sensitivity, and scope of MS proteomic assays to identify body fluid. Toward that end this research covers several techniques and sample types. As such the work is divided into four sections: **i)** evaluation of combinatorial peptide-ligand chromatography as a means of rapidly preparing complex biological fluids (menstrual blood and body fluid mixtures) for MS analysis, **ii)** reduce trypsin digestion time by evaluating the use of organic solvents and of microwave technology, **iii)** determine the stability of menstrual blood and vaginal fluid markers over time, evaluate newly identified microbial vaginal fluid markers, and determine the effects of forensic reagents Bluestar and luminol on these markers, and **iv)** determine the efficacy of simultaneous protein and DNA extraction for use in body fluid and STR testing to reduce sample consumption.

• **OVERALL METHODS:** This work is a series of smaller investigations, each with the goal of refining the process of forensic body fluid identification by proteomic mass spectrometry. The figure below shows the overall workflow common to nearly all proteomic experiments described herein. Design and methods specific to each research goal, along with findings for each, are detailed in the following sections.



4. FINDINGS: Findings are separated in the four experimental sections described above: **i)** evaluation of combinatorial peptide-ligand chromatography for protein dynamic range reduction, **ii)** evaluation of organic solvents, temperature and microwave-assisted protein digestion to reduce sample preparation time, **iii)** determination of menstrual blood and vaginal fluid marker stability over time as well as evaluation of microbial vaginal fluid markers in comparison to intrinsic vaginal fluid markers, and **iv)** evaluation of the efficacy of simultaneous protein and DNA extraction for use in body fluid and STR testing to reduce sample consumption.

i) COMBINATORIAL PEPTIDE LIGAND CHROMATOGRAPHY (CPLC) – REDUCTION OF DYNAMIC RANGE
The evaluation of combinatorial peptide ligand chromatography was a specific aim of this application. Our demonstration of its usefulness in improving protein marker identification led to the technique's incorporation in our search for menstrual blood markers. Consequently, part of the general discussion of how CPLC works was reported earlier (NIJ Grant 2010-DN-BX-K192).

Combinatorial peptide-ligand chromatography is a method to reduce the dynamic range of proteins in a sample. The dynamic range is the difference between the most abundant and least abundant proteins present. When this range is too great, proteins at the lowest end may not be detected. This particularly impacts samples where marker proteins may be a minor component of total proteins, e.g. menstrual blood markers or markers in extreme mixtures of body fluids.

CPLC non-specifically enriches all proteins equally, effectively capturing similar amounts of high and low abundant proteins in a single step. We evaluated combinatorial peptide ligand chromatography as a means of rapidly reducing the dynamic range of complex biological fluids (menstrual blood and body fluid mixtures) in order to improve body fluid marker detection in MS analysis.

- **COMPARISON OF EXTRACTION AND ELUTION BUFFERS:** We chose to use menstrual blood for our initial evaluation of CPLC since proteins that differentiate menstrual blood from venous blood represent only a small fraction of total blood protein. Different sample extraction and CPLC matrix elution buffers were evaluated to determine which combination would allow the greatest number of proteins to be recovered and identified following CPLC enrichment. Three extraction buffers were evaluated: i) PBS, ii) PBS plus 0.1% SDS and iii) PBS plus 0.5% NP-40, as well as two elution buffers: i) 8 M Urea, 2% CHAPS and 5% acetic acid (**UCA**), and ii) 2% SDS in PBS.

Qualitative analysis of different buffer systems by SDS PAGE showed that the combination of PBS with 0.5% NP-40 as extraction buffer and 2% SDS as elution buffer gave the stronger protein bands on the gels than the other buffer systems. HPLC-MALDI TOF/TOF analysis confirms these results, with the 0.5% NP-40 / 2% SDS combination resulting in the most proteins identified.

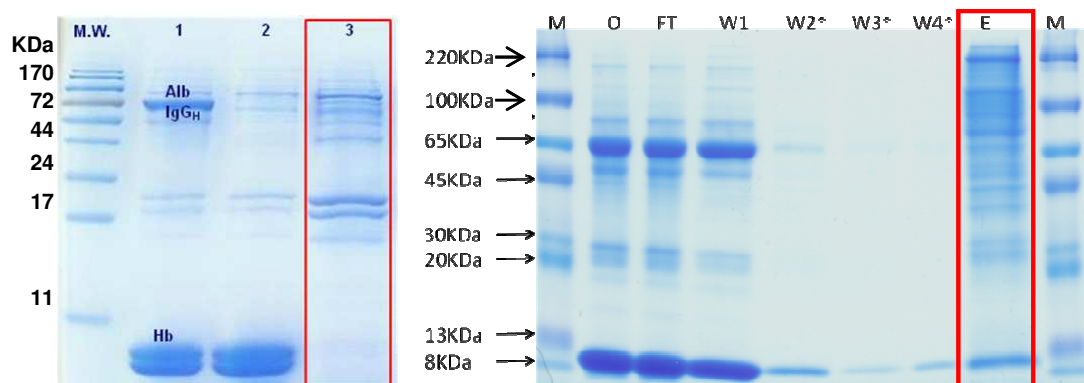
Table 1 Number of proteins identified MALDI TOF/TOF using different extraction/elution solution combinations

Confidence	PBS / UCA	PBS / 2% SDS	0.5%NP40 / 2% SDS	0.1%SDS / 2% SDS
>2.0 (99%)	157	216	237	96
>1.3 (95%)	215	262	312	124

- **COMPARISON OF HB/ALB/IGG DEPLETION METHODS TO CPLC:** An alternative to CPLC for reducing protein dynamic range in blood samples is specific depletion of hemoglobin (Hb) albumin (Alb) and immunoglobulin G (IgG) – the three most abundant proteins in blood. To evaluate the effectiveness of CPLC in dynamic range reduction a comparison was made to Hb, Alb and IgG depletion by both methods.

A qualitative comparison of samples processed by depletion and CPLC was made using SDS PAGE gels. Lane 1 on the left side of **Figure 1** shows the protein banding pattern of menstrual blood without any treatment. As may be seen, Hb, Alb and IgG constitute the bulk of protein bands. Lane 2 shows the pattern after depletion of Alb and IgG, and Lane 3 the enrichment of other proteins after depletion of Hb. The right side of **Figure 1** shows dynamic range reduction by CPLC. Here the original starting menstrual blood sample is shown in Lane O. CPLC column flow-through and four column washes in Lanes FT and W1-W4 respectively. Lane E shows the increase in the number of previously unseen bands in the CPLC eluate.

FIGURE 1 SDS PAGE of menstrual blood sample processed by depletion (left) and CPLC (right).



Left (depletion): Lane 1 = starting material, 2 = Alb and IgG depletion, 3 = Hb depletion, M.W. = molecular weight standards. Right (CPLC): O = original starting material, FT = flow-through, W1-W4 = column washes, E = eluate, M = molecular weight standards.

MALDI TOF/TOF mass spectrometry was performed on depleted and CPLC fractions to identify proteins. While 470 distinct proteins were found in the depleted fraction, and 428 in the CPLC fraction, only 59% (275) of depletion proteins were distinct from those found in venous blood, while 73% (312) of CPLC proteins were different. This difference is likely due to the enrichment process of the CPLC method. Finally, in addition to its ability to reduce protein dynamic range, CPLC is appreciably faster than traditional the depletion method, since only one chromatographic step is necessary in CPLC a sample processing.

- **DETERMINATION OF CPLC COLUMN BINDING CAPACITY:** Unlike traditional protein chromatographies where typically only 10% of a column's binding capacity is utilized in order to maximize purification, CPLC requires column saturation so that even the smallest amounts of sample components have an opportunity to completely bind to the matrix. With the preloaded matrix spin columns (ProteoMiner, Bio-Rad Hercules, CA) the manufacturer suggests a starting point of 10 mg protein for column saturation. Because of limited amounts of menstrual blood available, we empirically measured CPLC column binding capacities for blood and menstrual blood. Saturation was determined by comparing the amounts of eluted protein against the amounts applied – i.e. once saturation occurs column binding should be maximized and eluted protein should plateau. Data analysis revealed a relatively linear relationship between applied and eluted samples up to approximately 7,500 μg applied protein (linear regression). Samples with more than 7,500 μg applied protein cluster between 150 and 300 total μg protein recovered, suggesting that saturation occurs at about 7,500 μg applied protein. On average, approximately 2% of applied protein was recovered in CPLC eluates from samples where more than 7,500 μg of protein were applied to the columns.

- **BODY FLUID MIXTURES WITH CPLC:** Identifying mixtures containing more than one type of body fluid can be challenging, especially when mixture ratios are wide. We evaluated reduction of dynamic range using CPLC as a method for improving detection of specific protein markers for

the minor component of a body fluid mixture. We focused on blood, semen and saliva, three body fluids commonly encountered in forensic casework and for which we have previously established specific protein markers [3]. Semen and saliva were mixed with blood in different volume ratios (saliva/blood, 5:1, 1:1, 1:2, 1:5, 1:10; semen/blood, 1:5, 1:15, 1:25, 1:40, 1:80). However, while volume ratios give an easily understood comparison of body fluid mixtures, they can be deceptive in terms of marker ratios which are more accurately described by the amount of protein in each body fluid, which can be remarkably different, (Tables 2a & 2b) and by the relative proportion of markers in total protein.

Table 2a Typical amounts of protein (μg) extracted per μl of body fluid using PBS with 0.5% NP-40 and 1.4 mm ceramic beads in a reciprocating homogenizer.

Blood $\mu\text{g}/\mu\text{l}$ (Ratio Blood:Blood)	Semen $\mu\text{g}/\mu\text{l}$ (Ratio Semen:Blood)	Saliva $\mu\text{g}/\mu\text{l}$ (Ratio Saliva:Blood)
134 (1)	54 (0.40)	4 (0.03)

For example, a 1:10 volume ratio of semen to blood has 1:25 protein ratio, and a 1:10 volume ratio of saliva to blood has approximately a 1:333 protein ratio (Tables 2b).

Table 2b Body Fluid Mixture Ratios by Volume & μg Protein

Mixture	Ratios Volume (μg Protein)				
saliva/blood	5:1 (1:7)	1:1 (1:33)	1:2 (1:67)	1:5 (1:167)	1:10 (1:333)
semen/blood	1:5 (1:13)	1:15 (1:38)	1:25 (1:63)	1:40 (1:100)	1:80 (1:200)

Following mixture extractions, sample proteins were processed by CPLC. The amount of protein eluted from the column ranged from 155 – 259 μg , with mean of 190 μg and standard deviation of 35 μg , which are consistent with the results obtained from the menstrual blood and venous blood samples. Results following HPLC-MALDI analysis of body fluid mixtures (Table 3) clearly show that reduction of protein dynamic range with CPLC allows more proteins to be detected in complex samples. Four to five times more proteins were identified in the CPLC eluate than in the original protein solution.

Table 3 Total proteins identified before and after CPLC treatment in body fluid mixtures.

Mixture Ratio Volume (μg Protein)		Total Proteins Identified	
		Before CPLC	After CPLC
1:10 saliva: blood	(1:333)	25	166
1:5 saliva: blood	(1:167)	39	221
1:2 saliva: blood	(1:67)	42	226
1:1 saliva: blood	(1:33)	43	238
5:1 saliva: blood	(1:7)	38	109
1:5 semen: blood	(1:13)	56	143
1:15 semen: blood	(1:38)	26	175
1:25 semen: blood	(1:63)	47	123

1:40 semen:blood	(1:100)	41	134
1:80 semen:blood	(1:200)	36	182

However, for both saliva/blood and semen/blood mixtures no specific marker protein enrichment was observed after CPLC. In contrast, mass spectrometry revealed marker loss during the process. The number of peptides identified for α -amylase 1, the most abundant saliva-specific marker protein, was reduced after CPLC. Detection of semenogelin 1 and 2 (the most abundant semen-specific markers) in a 1:5 semen:blood mixture was improved with CPLC treatment. However with smaller relative amounts of semen in blood, semenogelin 1 and 2 were detected at low levels without CPLC treatment and detection was not improved after CPLC treatment.

- **CONCLUSIONS:** These results indicate that while CPLC treatment is effective in reducing the dynamic range of complex biological samples and increasing the number of proteins identified (e.g. menstrual blood), CPLC treatment of these particular mixtures does not improve detection of marker target proteins. For saliva in a mixture with blood, it appears that amylase is out-competed for binding on the CPLC column as it is found in the flow through fraction. Similarly, semenogelin does not appear in the CPLC eluates. Why these marker proteins cannot find high affinity CPLC aptamers to bind to may be due to similarities in their structure/function – e.g. both form large complexes with other proteins in their respective body fluids. Amylase, along with proline-rich proteins, statherin and histatin, has been shown to form complexes with highly glycosylated mucin 5 > 40×10^6 Da [4,5]. Semenogelins 1 and 2 (glycosylated) are cross-linked by cystine bond and transglutaminases [6,7]. These carbohydrate-coated multimer complexes may sequester amylase and semenogelins, preventing them from finding sufficient numbers of high affinity aptamers to bind to and thus effectively reduce their enrichment. Additionally, the large macromolecular structures (> 40×10^6 Da) formed by markers amylase (saliva) and semenogelin (semen) with other glycoproteins would unlikely be able to resist passage through CPLC columns – as is seen with amylase in the flow-through, and suggest three approaches to improve body fluid mixture sensitivity: i) use of more powerful detergents in sample solubilization to help solubilize the macro complexes, ii) reduce cystine bonds with a reducing agent such as DTT – again to destabilize the large protein complexes, and iii) remove the large carbohydrate trees that help link protein complexes together by deglycosylation using commercially available kits (e.g. EDEGLY and NDEGLY from Sigma-Aldrich). Clearly, additional research will be required to identify mechanisms.

ii) EVALUATING ORGANIC SOLVENTS, TEMPERATURE AND MICROWAVE-ASSISTED PROTEIN DIGESTION TO REDUCE SAMPLE PREPARATION TIME

A limiting aspect for a high throughput mass spectrometry assay is the current standard method of overnight trypsin digestion. The ability to substantially reduce trypsin incubation time could allow for single day sample processing hundreds of samples as mass spectrometry instrument time per sample occurs in micro-seconds. Here three parameters that can effect trypsin digestion were examined: i) inclusion of organic solvents, ii) incubation temperature, and iii) microwave radiation. All assays were performed on blood with digestion of hemoglobin (Hb) and albumin (Alb), the two most abundant proteins in blood, monitored for digestion.

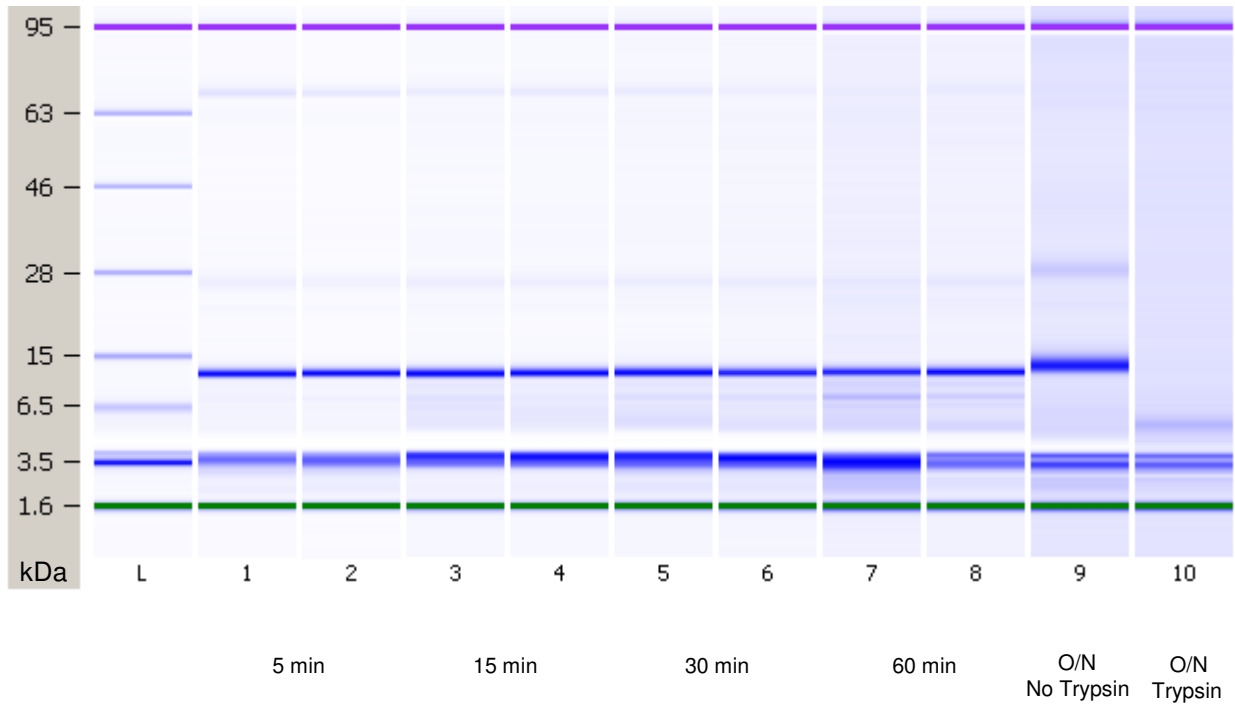
EXPERIMENT 1 – ESTABLISHING THE ASSAY SYSTEM

- **METHODS:** Blood was homogenized in a 1:10 (v/v) ratio with RIPA buffer (25mM Tris-HCl pH 7.6, 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, Pierce Rockford, IL), centrifuged at 18,000 g for 30 min at 4°C and supernatant protein precipitated with a 1:7 (v/v) sample to acetone ratio. Pellets were resuspended in 20 µl of 500 mM ammonium bicarbonate (**ABC**). Concentration of samples was measured by bicinchoninic acid (**BCA**) and Bradford assays (Sigma). Reduction was performed using one part dithiothreitol (**DTT**) to ten parts of sample and incubated at 37°C for one hr. Samples were alkylated with a final concentration of 15 mM iodoacetamide at room temperature for 30 min. A 2,000:1 (µg:µg) ratio of sample protein to trypsin was used. Positive controls were treated at the same sample:trypsin µg ratio but digested overnight in a 37°C incubator (current standard) unless otherwise indicated. Negative controls were treated the same as positive controls except an equal volume of 50 mM ABC was added in place of trypsin. Three µl of 100 mM PMSF was added to all reactions after microwave or incubation and samples immediately placed at -20°C. Samples were thawed and kept on ice until assayed. The extent of trypsin digestion was measured by MiGE electropherograms quantified by time corrected area under the curve (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA, using Agilent Protein 80 Kits).

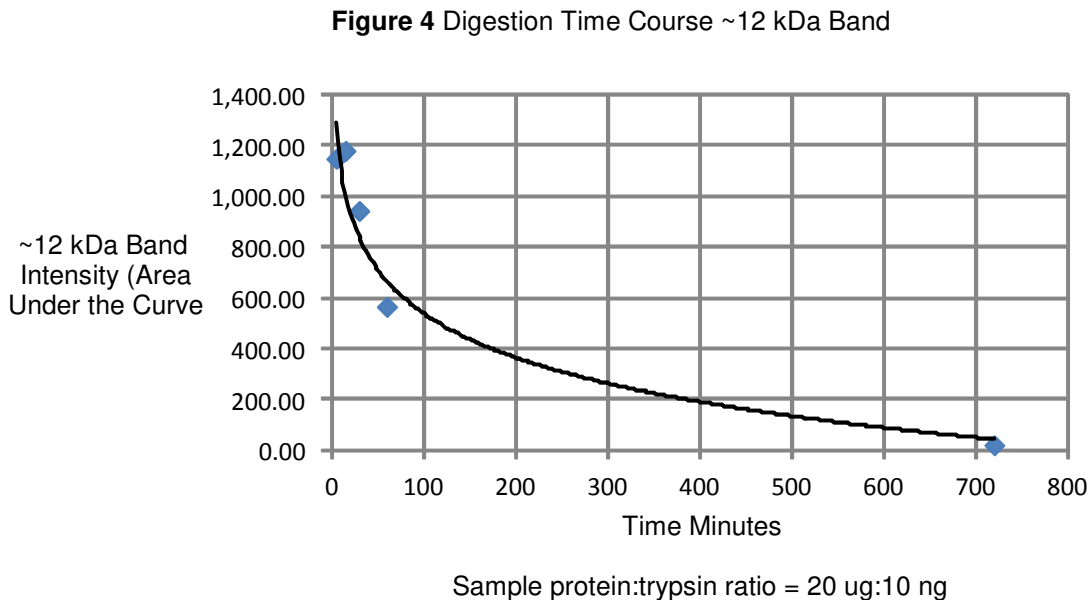
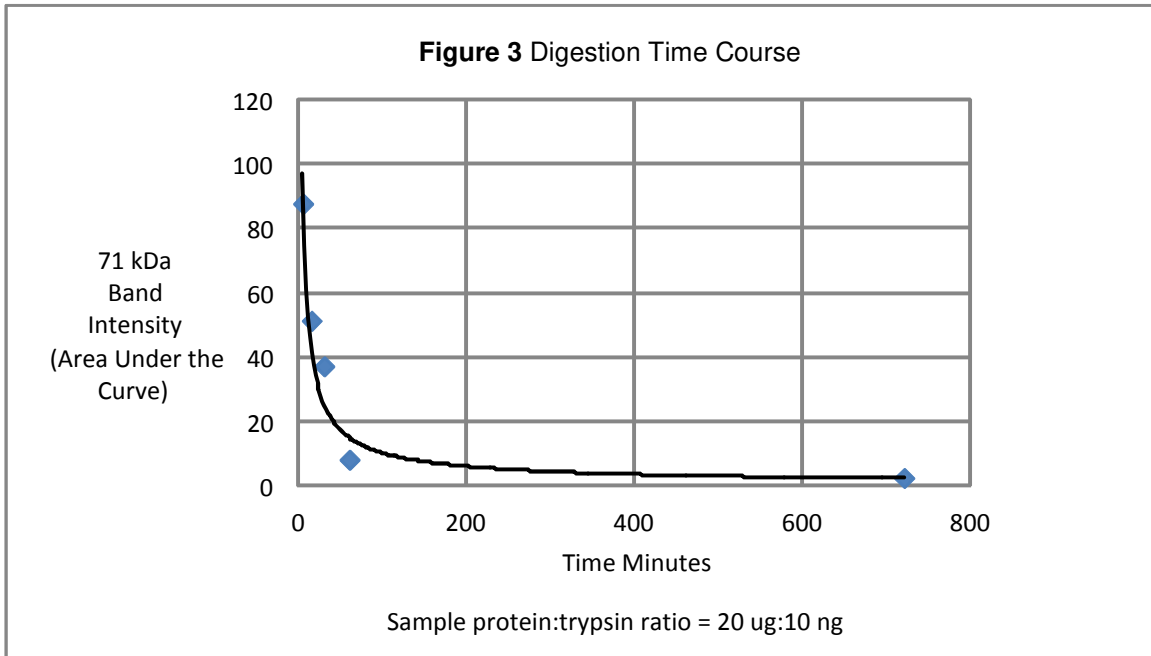
- **EXPERIMENTAL DESIGN:** In order to evaluate the effects of organic solvents, microwaves, time and temperature on trypsin digestion efficiency, it was first necessary to establish an assay system able to following substrate marker (Hb and Alb) digestion over time.

- **RESULTS:** **Figure 2** shows a microfluidic gel electrophoresis (**MiGE**) image (2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA) of increasing digestion of blood proteins at 5, 15, 30 and 60 minutes at a sample protein:trypsin ratio of 2,000:1 (µg/µg). The protein band at ~71 kDa (**Alb**) shows diminishing intensity over time as does the band at ~12.5 kDa (**Hb**) until overnight digestion when both are nearly gone (Lane 10). Quantitation of the ~71 and ~12.5 kDa bands (time corrected area under the curve; average of duplicate samples) shows this more clearly in **Figures 3 and 4**. These data demonstrate that at a 2,000:1 (µg:µg) ratio of sample protein to trypsin, changes in protein digestion can be observed.

Figure 2 100 X diluted trypsin digestion of blood proteins over time.



Equal amounts of blood supernatant protein incubated with and without trypsin at a 2,000:1 ratio ($\mu\text{g}/\mu\text{g}$) respectively at 37°C for varying times and overnight. Protein bands at ~71 kDa (likely albumin) and 12 kDa (likely hemoglobin) shows diminishing intensity over digestion time. L = ladder, O/N = overnight.

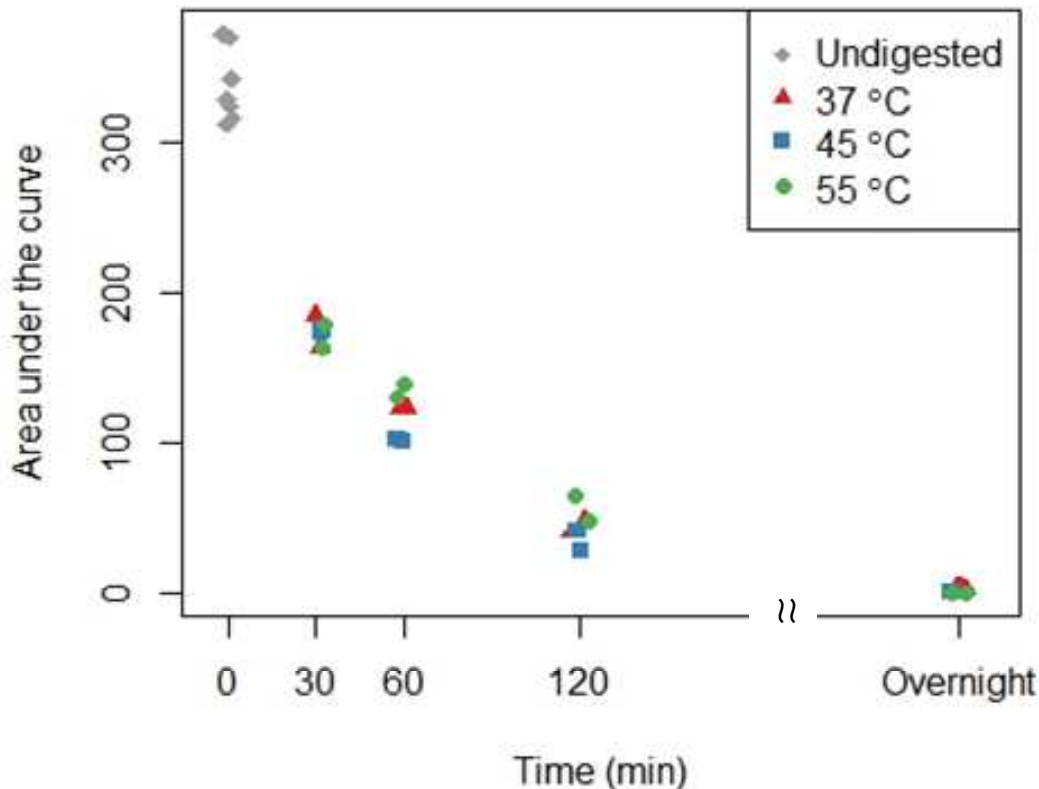


EXPERIMENT 2 – EVALUATION OF TRYPSIN DIGESTION TEMPERATURE OVER TIME

To identify optimal digestion temperature, blood samples were incubated at the 2,000:1 ($\mu\text{g}/\mu\text{g}$) sample protein to trypsin ratio at 37°C, 45°C and 55°C and followed over time. Samples were run on MiGE and the 12.5 kDa Hb band quantified by time corrected area under the curve from resulting electropherograms. **Figure 5** shows the continuous digestion the Hb band at 30, 60, 120 min and overnight. Undigested control (gray diamonds) is a sample with trypsin at zero

time quenched with PMSF on ice. Results show little difference of the effects of incubation temperature on digestion.

Figure 5 Comparison of Digestion of the 12.5 kDa Hb Band at 37°C, 45°C, and 55°C Over Time



EXPERIMENT 3 – EVALUATING THE EFFECTS OF ORGANIC SOLVENTS ON TRYPSIN DIGESTION

- Five different solvents were tested at varying concentrations to identify the optimal ratio of solvent to ABC for maximum digestion. Digestions were performed with a 2,000:1 (µg/µg) ratio of sample protein to trypsin in a 37°C incubator for 20 min. Table 4 shows the range of solvents tested and the percentages at which maximum digestion occurred (green shading). For acetonitrile 5% and 10% gave similar results. Once optimal organic solvent concentrations were determined, direct comparison experiments between microwave and incubator were performed for each solvent (see below).

Table 4 Organic Solvents and Percent Range Tested

Organic	Percent in ABC						
	5	10	15	20	25	30	35
Acetone	5	10	15	20	25	30	35
Acetonitrile	2	5	10	15	20	25	-
Dimethyl Sulfoxide	2	5	10	15	20	25	-
Isopropanol	2	5	10	15	-	-	-

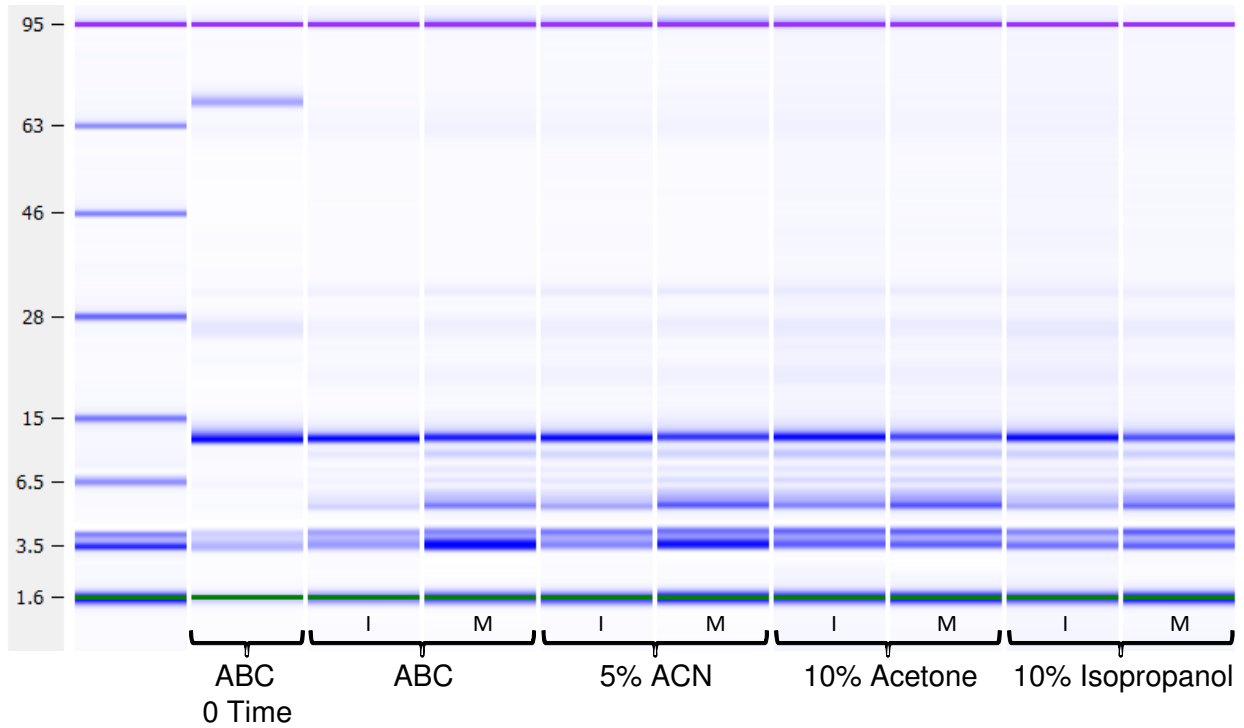
Methanol	2	5	10	20	30	40	50
----------	---	---	----	----	----	----	----

Digestion in a 37°C incubator for 20 min with a 2,000:1 ($\mu\text{g}/\mu\text{g}$) ratio of sample protein to trypsin. Green shadings are percentages that gave the best digestion results.

EXPERIMENT 4 – EVALUATING MICROWAVE-ASSISTED ENZYMATIC DIGESTION

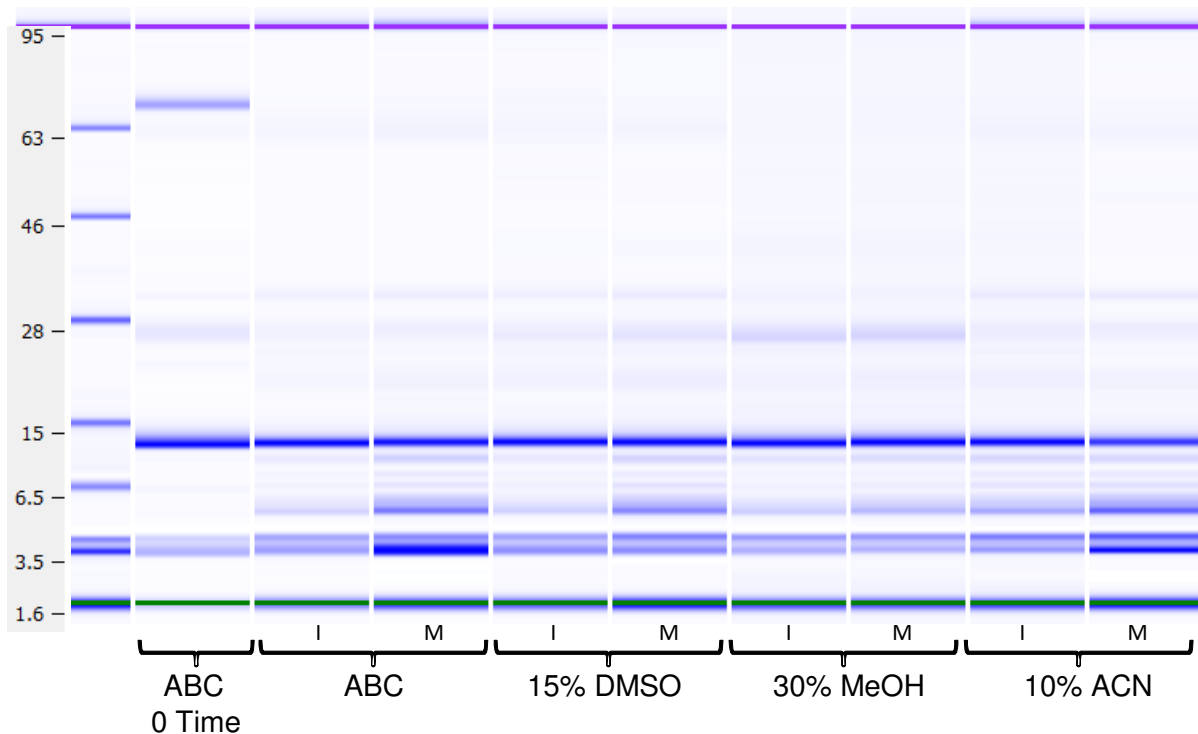
- *CEM MARS6 MICROWAVE*: Enzymatic digestions in the CEM MARS6 microwave are performed in 1.5 ml microfuge tubes. The tubes are held in a small closed container with 25 ml of water below the tubes that does not touch the tubes. This is to ensure that sample tubes are directly exposed microwaves. The 25 ml water bath below the tubes acts as a heat sink capturing microwave energy until the desired temperature is reached above the water bath in the closed container. Temperature is continuously monitored by a probe inserted into a small hole in the cap of a microfuge tube filled with the same volume of buffer as in the digestion tubes. The temperature monitoring tube is also above the water. As temperature and power are dependent variables, microwave input (watts) automatically turns on and off to maintain desired set temperature. Consequently, sample exposure to microwave radiation varies as temperature is maintained. Setting a constant temperature with different power inputs exposes the sample to greater microwave energy but for shorter amounts of time.
- *EXPERIMENTAL DESIGN*: A direct comparison of incubation vs microwave-assisted digestion was performed at the optimal solvent concentrations for the five organic described above (*Experiment 2*). Incubation was at 37°C; microwave temperature was held constant at 37°C at 50 watts. Digestion time was 20 minutes.
- *RESULTS*: As may be seen in **Figures 6 and 7**, microwave-assisted trypsin digested samples (M) appear to show less intense 12.5 kDa (Hb) bands (particularly in ACN, acetone and isopropanol) than non-microwave treated samples, suggesting increased digestion due to microwaves. (Increases in the number and intensity of intermediate digestion products between 3.5 kDa and 10 kDa in the microwave treated samples also suggest improved digestion.) This was confirmed by analyzing electropherogram time corrected area under the curve for Hb digestion in each sample.

Figure 6 MiGE Comparison of Incubator and Microwave-Assisted Trypsin Digestion of Blood in Different Organic Solvents at 37°C (50 Watts) for 20 Minutes



I= incubator, M = microwave, ABC = ammonium bicarbonate, ACN = acetonitrile, Iso = isopropanol.
Sample to trypsin ratio 2,000:1

Figure 7 MiGE Comparison of Incubator and Microwave-Assisted Trypsin Digestion of Blood in Different Organic Solvents at 37°C (50 Watts) for 20 Minutes

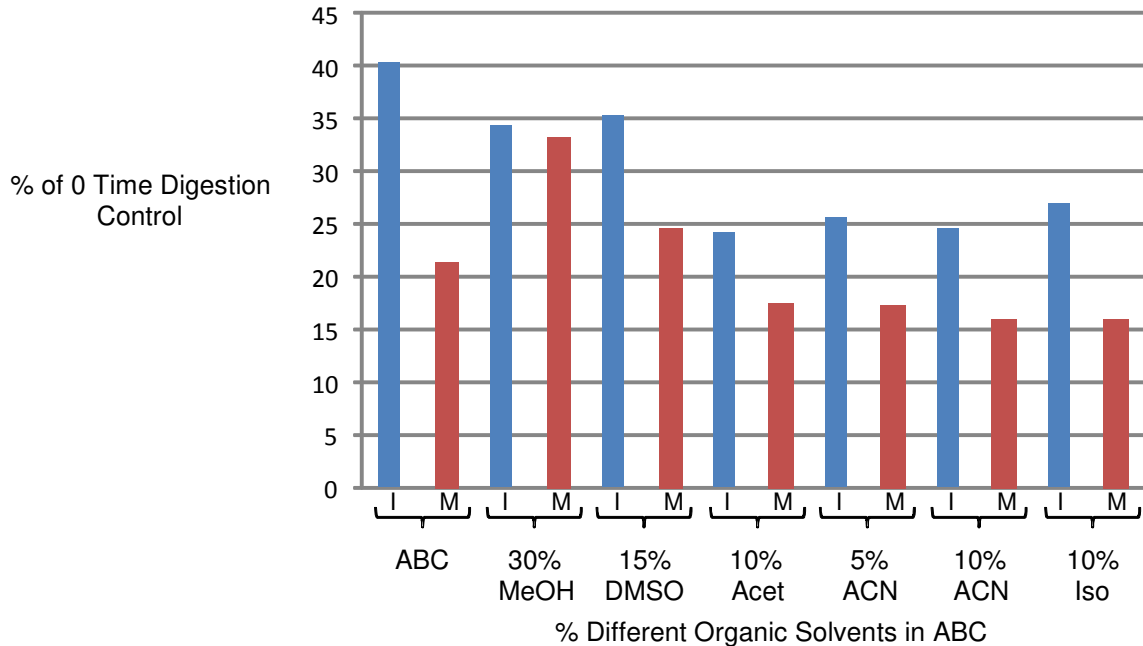


I= incubator, M = microwave, ABC = ammonium bicarbonate, DMSO = dimethyl sulfoxide, MeOH = Methanol, ACN = acetonitrile. Sample to trypsin ratio 2,000:1

Figure 8 shows the percent of the ~12.5 kDa band (Hb) remaining after 20 min digestion at 37°C compared to the zero time negative control for all samples in each organic with and without microwave treatment. First, with respect to the effects of organic solvents on digestion of incubator samples alone (blue bars), all organics demonstrated some improvement with the data segregating into two main groups. The first, 30% MeOH and 15% DMSO reduced the Hb band to about 35% of the starting material (negative control), compared to ABC alone which reduced the Hb band to ~40% of starting material - about a 12.5% improvement. The second group of organics, 10% acetone, 5% ACN, 10% ACN and 10% isopropanol reduced the Hb band to about 25% of the ABC control - approximately a 37.5% improvement.

With respect to the effects of microwaves on trypsin digestion (red bars), with the exception of methanol, microwaves improved digestion efficiency in all cases. Most notably in the ABC no organic control which showed an improvement of nearly 50% digestion. However, in the 10% acetone, 5% ACN, 10% ACN and 10% isopropanol samples, microwaves demonstrated an approximate 34% improvement compared to their non-microwave counterparts.

Figure 8 Comparison of Incubator and Microwave-Assisted Trypsin Digestion of ~12.5 kDa Band in Different Organic Solvents at 37°C for 20 Minutes

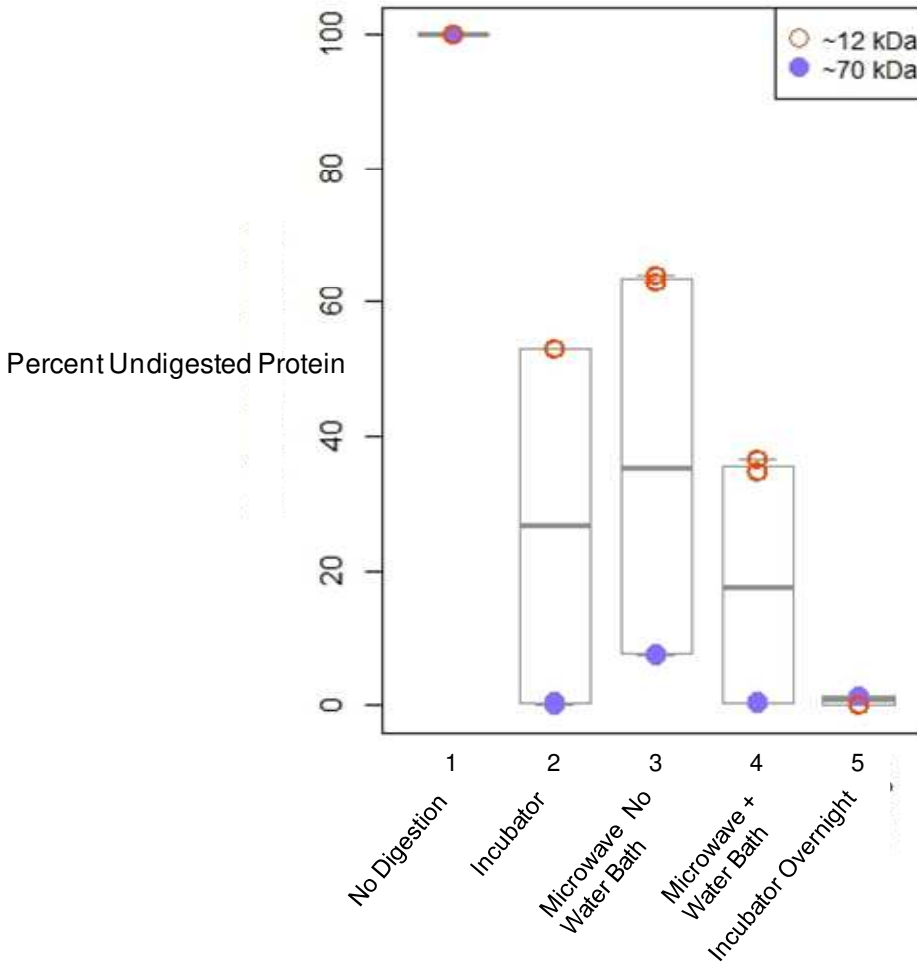


I= incubator, M = microwave, ABC = ammonium bicarbonate, MeOH = Methanol, DMSO = dimethyl sulfoxide, Acet= acetone, ACN = acetonitrile, Iso= isopropanol

EXPERIMENT 5 – EVALUATING MICROWAVE POWER

To determine if the level of microwave power has an effect on digestion rate, blood samples were again assayed at 37°C but using microwaves at 850 watts. Lanes 2 (incubator) and 4 (microwave) of **Figure 9** again demonstrate that the use of microwaves improves protein digestion in a manner similar to what was observed using 50 watts at 37°C (**Figure 8**).

Figure 9 Microwave – Assisted Trypsin Digestion at 37°C and 850 Watts for 15 Minutes in the Presence and Absence of a Water Bath



EXPERIMENT 6 – COMBINED DIGESTION CONDITIONS AT STANDARD TRYPsin CONCENTRATION

The ultimate goal of these experiments was to identify trypsin digestion conditions that can substantially reduce digestion time. Consequently, we returned to the standard ratio of sample protein to trypsin (i.e. 20 µg sample protein to 1 µg trypsin) and repeated the digestion. As may be seen in Table 5, nearly complete Hb and Alb digestion can be obtained at 37°C in 10% isopropanol with in 30 min. Increasing temperature to 45°C or 55°C had little effect (as shown earlier) and digestion at 60 and 120 min made only modest difference.

Table 5 % Protein Remaining at 20:1 ($\mu\text{g}/\mu\text{g}$)
 Sample Trypsin Ratio
 Temperature

Time	37°C	45°C	55°C
30 min	0.58	0.42	0.32
1 hr	0.96	0.69	0.44
2 hrs	0.32	0.62	0.39
Overnight	0.09		

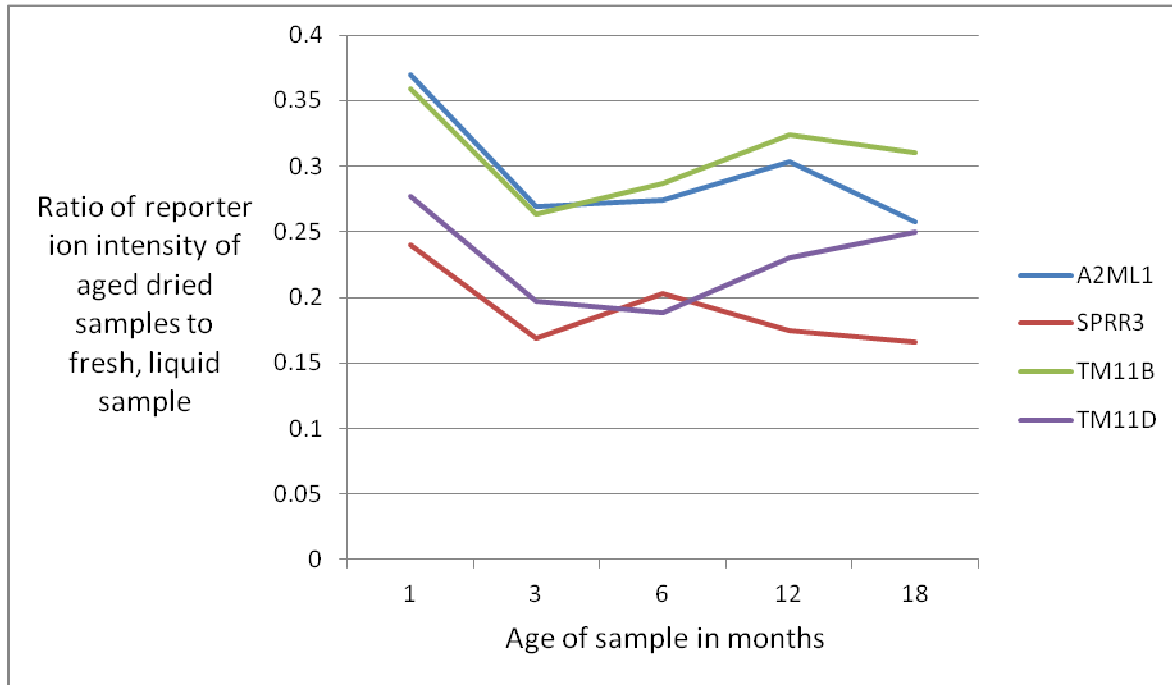
In conclusion, the combination of organic solvents and use of microwaves can extensively reduce the time for Hb and albumin digestion. Near overnight digestion levels can be achieved in 10% isopropanol at 37°C in approximately two hours using the 2,000:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio. Using a 20:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio under these same conditions, these same levels can be achieved in approximately 30 minutes.

iii) MENSTRUAL BLOOD & VAGINAL FLUID MARKER STABILITY: One of the criteria for selecting a body fluid assay (DNA, RNA or protein) is the stability of its markers over time and their resistance to environmental insults. We have previously shown little or no loss of protein marker signals with samples aged up to two year, or when in contact with the common forensic reagents luminol and Bluestar up to 30 days. Here we evaluate menstrual blood and vaginal fluid marker proteins in a similar way.

- **METHODS:** We tested the stability of vaginal fluid and menstrual blood specific protein markers over time by creating a series of dried samples of each fluid and extracting proteins from these samples after an aging period of 1 week and 1, 3, 6, 12 and 18 months at room temperature and protected from sunlight. Criteria for assessment of stability included qualitative banding pattern in SDS PAGE gels, specific markers identified by HPLC-MALDI TOF/TOF analysis, and reporter ion intensity in samples labeled with isobaric tags. Menstrual blood samples were run on CPLC columns prior to HPLC-MALDI TOF/TOF analysis (due to the small volumes eluted from CPLC columns, samples quantities were not sufficient to run on SDS PAGE gels).

- **VAGINAL FLUID:** Vaginal fluid samples labeled with isobaric tags were analyzed by HPLC-MALDI. **Figure 10** shows the ratio of reporter ion intensity for 4 vaginal fluid marker proteins: alpha-2 macroglobulin-like protein (A2ML1), SPRR3, small proline-rich protein 3 (SPRR3); transmembrane protease serine 11B (TM11B), transmembrane protease serine 11D (TM11D), as well as A2ML1, from dried, aged samples relative to the same fresh, liquid sample from the same individual. The ion intensities of these markers in dried, aged samples range from approximately a sixth to a third of the signal from the fresh sample; however signals from dried samples do not noticeably decrease over the course of 18 months.

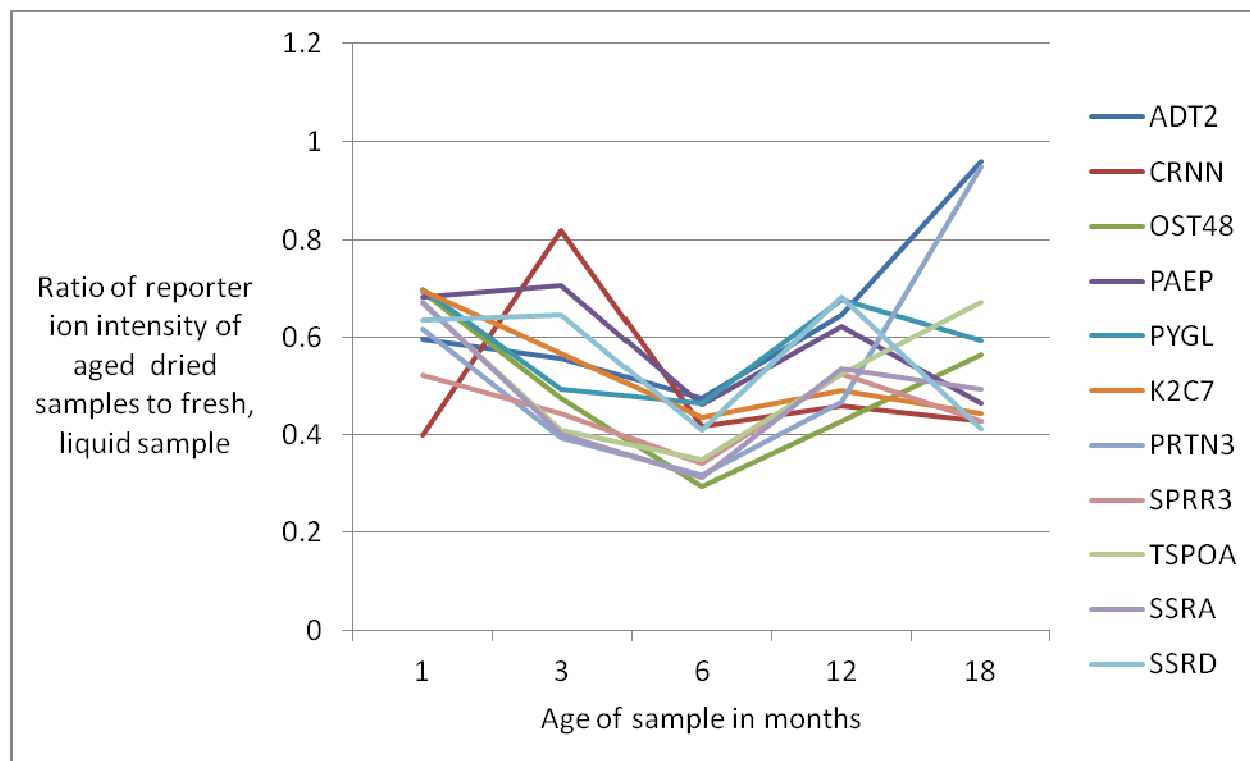
Figure 10 Ratio of reporter ion intensity of dried, aged samples to fresh, liquid sample from the same individual (533D19) for 4 protein markers that are specific to, or enriched in, vaginal fluid.



A2ML1: alpha-2 macroglobulin like protein; SPRR3, small proline-rich protein 3; TM11B, transmembrane protease serine 11B; TM11D, transmembrane protease serine 11D

• **MENSTRUAL BLOOD:** **Figure 11** shows the ratio of reporter ion intensity for nine protein markers that are specific to or enriched in menstrual blood from dried, aged samples relative to the same fresh, liquid sample from the same individual. The ion intensities of these markers in dried, aged samples range from approximately a third to nearly 100% of the signal from the fresh sample, and signals from dried samples do not decrease over the course of 18 months.

Figure 11 Ratio of reporter ion intensity of dried, aged samples to fresh, liquid sample from the same individual (118D2) for 11 protein markers that are specific to, or enriched in, menstrual blood.



ADT2, ADP/ATP translocase; CRNN, cornulin; OST48, dolichyl-diphosphooligosaccharide protein glycosyltransferase 48 kDa subunit; PAEP, glycodelin; PYGL, glycogen phosphorylase; K2C7, Keratin type II cytoskeletal 7; PRTN3, myeloblastin; SPRR3, small proline rich protein 3; TSPOA, translocator protein; SSRA, translocon-associated protein subunit alpha; SSRD, translocon-associated protein subunit delta

• **EFFECTS OF BLUESTAR AND LUMINOL ON VAGINAL FLUID MARKERS:** Bluestar and luminol are common crime scene reagents used to identify bloodstains. In our previous study we demonstrated little effect of Bluestar and luminol on the stability of blood, saliva and semen marker proteins at 1, 9 and 29 days (NIJ Grant 2010-DN-BX-K192). We tested the effects of Bluestar and luminol on aged vaginal fluid samples applied to cotton swabs by adding 10 μ l of Bluestar or luminol to the stains, and extracted proteins after a period of 1, 9, and 29 days after application of Bluestar or luminol.

Due to the age of the samples, fewer proteins were identified than would be expected from fresher sample. Despite this, at least one, and in most cases more than one, vaginal fluid specific marker was identified in all samples, both those treated with Bluestar and luminol and untreated positive controls. Treatment of dried vaginal fluid stains with Bluestar or luminol did not affect the overall number of proteins detected or the number of specific markers detected.

- **VAGINAL FLUID MICROBIOME:** In order to identify microbial proteins, 96 samples of vaginal fluid taken on two separate occasions from 48 women were extracted, digested with trypsin and processed by HPLC-MALDI TOF/TOF.

Proteins from *Lactobacillus*, which have been shown to be present in the human vaginal tract [1,2], were found in all samples. The two most abundant *Lactobacillus* proteins identified were glyceraldehyde-3-phosphate dehydrogenase (**GAPDH**) and enolase (**ENO**). *Lactobacillus* glucose-6-phosphate isomerase, triosephosphate isomerase, D-lactate dehydrogenase and phosphate acyltransferase were identified as well.

iv) SIMULTANEOUS PROTEIN & DNA EXTRACTION AND PURIFICATION - When limiting amounts of biological samples are available, there may be insufficient material for both DNA and protein extractions and analyses. Here we evaluate the simultaneous extraction of DNA and proteins from a single sample to determine the feasibility of concurrent STR and body fluid identifications. This section is divided into five parts: **iv.i)** EXPERIMENTAL DESIGN – OVERVIEW, **iv.ii)** METHODS, **iv.iii)** FINDINGS - PROTEIN ANALYSIS, **iv.iv)** FINDINGS - DNA ANALYSIS, **iv.v)** DISCUSSION

iv.i) EXPERIMENTAL DESIGN - OVERVIEW: The AllPrep DNA/RNA/Protein purification kit from Qiagen (Valencia, CA) was used for simultaneous protein and DNA extractions from neat and diluted samples of blood and saliva. To compare protein and DNA recoveries from the AllPrep kit to protein and DNA recoveries from single method extractions, results from the AllPrep kit were compared to methods currently used in the lab: for protein the routine OCME method and that of Mann et al. [8], and for DNA, DNA IQ method.

iv.ii) METHODS: Comparison of methods was done using neat, 10x and 100x dilutions of the blood and saliva samples. Following extraction, protein fractions were quantified using the BCA assay with BSA as a standard. The extracted DNA samples were quantified using Plexor HY System (Promega, Madison, WI) on the Rotor-Gene Q (Qiagen, Valencia, CA).

iv.iii) FINDINGS - PROTEIN ANALYSIS: In order to compare the efficiency of protein recovery by AllPrep and individual methods, concentrations were calculated as micrograms per microliter of the original sample, whether neat or diluted, since the specific volumes of sample used for each method varied. Both protein-only extraction and purification methods recovered appreciably more protein than the AllPrep kit for both blood and saliva samples (**Figure 12**). Protein concentrations of saliva sample processed with AllPrep were too low to be accurately measured by BCA assay, and therefore no data was available for comparison.

Undiluted and 100x diluted blood samples and undiluted and 10x diluted saliva samples extracted with AllPrep were analyzed with HPLC-MALDI TOF/TOF. Tables 6 and 7 list the body fluid specific protein markers found in each of these samples and the number of peptides identified with >95% confidence for each.

iv.iv) FINDINGS - DNA ANALYSIS: DNA was analyzed and STR profiles generated using GeneMapper Software (Applied Biosystems, Foster City, CA). A DNA profile with true allelic peaks for at least 7 out of 13 STR loci is considered a usable profile. Table 8 indicates which AllPrep processed samples produced a useable DNA profile in comparison with the DNA IQ processed samples. Although separate protein and DNA extraction and purification methods yielded best results, dual protein/DNA extraction with the AllPrep kit was successful. MS

proteomic analysis with AllPrep gave better results than the DNA as body fluid specific protein markers were identified for neat and diluted blood samples and neat saliva samples. Useable DNA profiles were obtained for the neat blood and saliva samples as well as for the 10x diluted saliva samples. Our results show that AllPrep dual extraction could be an effective tool for obtaining both proteomic and genetic information from a single sample. However, for limiting samples, AllPrep is not as sensitive as separate extraction methods.

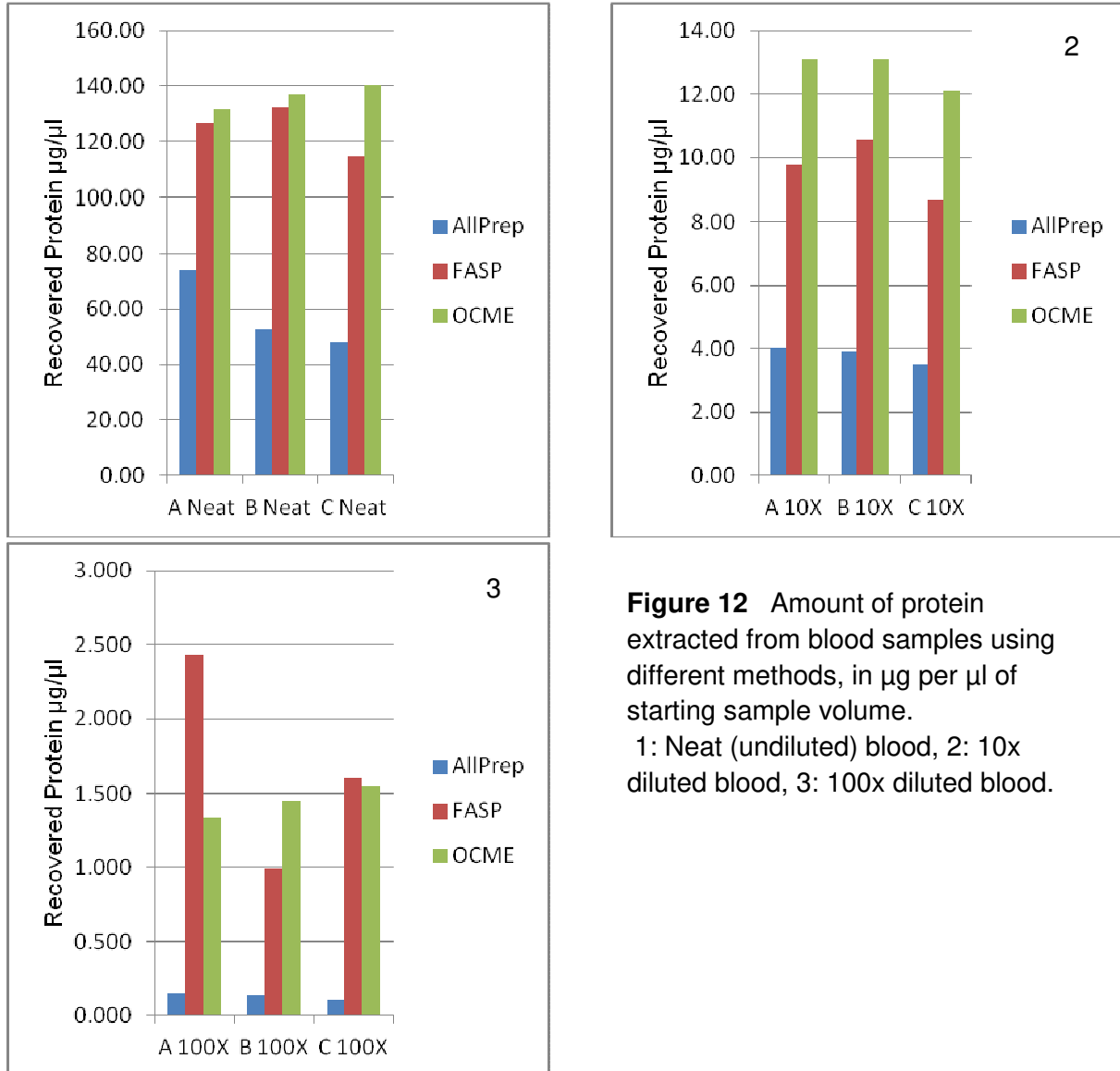


Figure 12 Amount of protein extracted from blood samples using different methods, in µg per µl of starting sample volume.
 1: Neat (undiluted) blood, 2: 10x diluted blood, 3: 100x diluted blood.

Table 6 Blood-specific protein markers found in blood samples extracted with AllPrep

	No. peptides at >95% confidence					
	Neat A	Neat B	Neat C	100x A	100x B	100x C
Hemoglobin subunit beta (HBB)	91	125	121	56	108	105
Hemoglobin subunit alpha (HBA1/HBA2)	72	73	85	36	58	59
Band 3 anion transport protein (SLC4A1)	3	4	5	-	-	-
Spectrin alpha chain, erythrocytic 1 (SPTA1)	1	-	-	-	-	-

- No peptides identified

Table 7 Saliva-specific protein markers found in saliva samples extracted with AllPrep

	No. peptides at >95% confidence		
	Neat D	Neat E	Neat F
Alpha-amylase 1 (AMY1A)	48	44	52
Cystatin-SA (CST2)	24	15	17
Protein LEG1 homolog (LEG1/C6orf58)	1	1	1

Table 8 Determination of whether or not a usable DNA profile was obtained from each of the AllPrep processed samples in comparison to the DNA IQ processed samples.

	Neat Blood A	Neat Blood B	Neat Blood C	10x Blood A	10x Blood B	10x Blood C	100x Blood A	100x Blood B	100x Blood C	Neat Saliva D	Neat Saliva E	Neat Saliva F	10x Saliva D	10x Saliva E	10x Saliva F	100x Saliva D	100x Saliva E	100x Saliva F
All Prep	Usable	Usable	Usable	Not Usable	Not Usable	Not Usable	Not Usable	Not Usable	Not Usable	Usable	Usable	Usable	Usable	Usable	Usable	Not Usable	Not Usable	Not Usable
DNA IQ	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Usable	Failed	Usable

Usable Profile
 Not Usable Profile
 Failed to amplify due to an instrumentation error

iv.v) DISCUSSION: Both protein-only extraction and purification methods recovered appreciably more protein than the AllPrep kit for both blood and saliva samples. The FASP method and NYC OCME method both recovered twice as much protein as the AllPrep kit for neat blood. For 10x diluted blood, FASP recovered twice as much, while the NYC OCME method recovered almost three times as much protein. The 100x diluted protein results were variable and both FASP and NYC OCME methods recovered 6-25 times more protein. AllPrep saliva protein concentrations were too low to be measured by the BCA assay, meaning there was minimal recovery of protein. The difference in recovery by these methods is likely due to the different chemistries and the number of steps involved in the processing methods. The AllPrep kit required multiple reagents and multiple steps in order to get the final protein fraction for digestion suggesting that there was more opportunity for protein loss than in the other protocols.

HPLC-MALDI TOF/TOF analysis identified blood-specific protein markers hemoglobin subunit beta (HBB) and hemoglobin subunit alpha (HBA1/HBA2) in both undiluted and 100x diluted blood samples extracted with AllPrep. Additional blood-specific markers band 3 anion transport protein (SLC4A1) and spectrin alpha chain, erythrocytic 1 (SPTA1) were found in some undiluted samples. Saliva-specific protein markers α -amylase 1 (AMY1A), cystatin-SA (CST2), and protein LEG1 homolog (LEG1/C6orf58) were found in all undiluted saliva samples extracted with AllPrep, however no saliva-specific markers were found in the 10x diluted saliva samples.

Almost all the DNA IQ extracted samples produced full, usable DNA profiles. Neat blood, neat saliva, and 10x diluted saliva were the only AllPrep samples that produced useable profiles. One possible cause is the AllPrep reagent present in the purified DNA sample interfering with the DNA analysis kit.

Although separate protein and DNA extraction and purification methods yielded the best results, the results of dual protein/DNA extraction with the AllPrep kit were good enough to be used in forensic casework when testing options are limited. The MS proteomic analysis was more successful than the DNA analysis in that body fluid specific protein markers were identified for neat and diluted blood samples and neat saliva samples. Useable DNA profiles were obtained for the neat blood and saliva samples as well as for the 10x saliva samples. Different simultaneous protein/DNA extraction kits, or altering instrument parameters, may ameliorate the AllPrep results and produce useable profiles from more diluted samples. Our results show that AllPrep dual extraction could be an effective tool for obtaining both proteomic and genetic information from a single sample, especially when dealing with small evidence samples that may only allow for one extraction to be conducted. However, there is opportunity for improvement.

5. OVERALL CONCLUSIONS AND IMPLICATIONS FOR POLICY AND PRACTICE

CONCLUSIONS - The overall goal of this work was to assess methods that could aid in reducing body fluid sample preparation time prior to mass spectrometry analysis (which is rapid) in order to improve throughput. This goal was achieved through the evaluation of CPLC for dynamic range reduction, and the use of organic solvents and microwaves to reduce trypsin digestion time which has traditionally been done overnight. A second goal was to determine the effects of long term aging and the chemicals luminol and Bluestar (commonly used forensic reagents that can be caustic) on the stability of menstrual blood and vaginal fluid protein markers. This aim also included an evaluation of the mass spectrometry method to simultaneously identify vaginal fluid microbiome markers along with our identified vaginal fluid human markers. The final goal was to evaluate the ability to simultaneously extract proteins and DNA from a sample in order to reduce sample consumption for body fluid and STR identify when sample is limiting. Success of each of these objectives is discussed sequentially below.

1. REDUCTION OF SAMPLE PROCESSING TIME – CPLC DYNAMIC RANGE REDUCTION AND USE OF ORGANIC SOLVENTS AND MICROWAVES TO REDUCE TRYPSIN DIGESTION TIME

Results demonstrated that CPLC treatment is effective in reducing the dynamic range of complex biological samples and increasing the number of proteins identified. This was particularly true for menstrual blood. However, while CPLC treatment of mixed body fluids

(blood, saliva and semen) also showed a protein dynamic range reduction, in these mixtures marker proteins detection did not improve. For saliva in a mixture with blood, it appears that amylase is out-competed for binding on the CPLC column as it is found in the flow through fraction. Similarly, semenogelin does not appear in the CPLC eluates. Why these marker proteins cannot find high affinity CPLC aptamers to bind to may be due to similarities in their structure/function – e.g. both form large complexes with other proteins in their respective body fluids. Amylase, along with specific markers proline-rich proteins, statherin and histatin, has been shown to complex with highly glycosylated mucin 5 to form complexes $> 40 \times 10^6$ Da [4,5]. Similarly semenogelins 1 and 2 (glycosylated) are cross-linked by cystine bond and transglutaminases [6,7], forming large complexes. These carbohydrate-coated multimer complexes may sequester amylase and semenogelins, preventing them from finding sufficient numbers high affinity aptamers to bind to and thus effectively reducing their enrichment. Clearly additional research will be required to identify mechanisms. However, for the forensic identification of menstrual blood CPLC notably reduced assay time.

Results from experiments aimed at reducing trypsin digestion time demonstrated that the combination of organic solvents and use of microwaves can dramatically reduce the time for Hb and albumin digestion. Near overnight digestion levels were achieved in 10% isopropanol at 37°C in approximately two hours using the 2,000:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio. Using a 20:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio under these same conditions achieved similar results in approximately 30 minutes, making these conditions ideal for high throughput body fluid testing.

2. STABILITY OF MENSTRUAL BLOOD AND VAGINAL FLUID MARKER PROTEINS & IDENTIFICATION OF VAGINAL FLUID MICROBIOME MARKERS

Menstrual blood and vaginal fluid protein markers were shown to be stable for at least 18 months (the end of the trial period) under conditions tested. Bluestar and luminol showed no noteworthy effect on protein stability for the 30 days over which these samples were exposed to these reagents. Additionally, we confirmed that many *Lactobacillus* proteins that have been used to identify the vaginal fluid from the vaginal fluid microbiome can also be detected by mass spectrometry of vaginal fluid samples. However, we believe that the identification of human specific vaginal fluid marker proteins is a more definitive test for the presence of vaginal fluid.

3. SIMULTANEOUS EXTRACTION OF PROTEINS AND DNA FROM A SINGLE SAMPLE

Results demonstrated that protein and DNA can be simultaneous extracted from a single sample and analyzed for body fluid and STR identity markers. Although separate protein and DNA extraction and purification methods yielded the best results, the results of dual protein/DNA extraction with the Qiagen AllPrep kit were good enough to be used in forensic casework when testing options are limited. The MS proteomic analysis was more successful than the DNA analysis in that body fluid specific protein markers were identified for neat and diluted blood samples and neat saliva samples. Useable DNA profiles were obtained for the neat blood and saliva samples as well as for the 10x saliva samples. Different simultaneous protein/DNA extraction kits, or altering instrument parameters, may ameliorate the AllPrep results and produce useable profiles from more diluted samples. Our results show that AllPrep dual extraction could be an effective tool for obtaining both proteomic and genetic information from a single sample, especially when dealing with small evidence samples that may only allow

for one extraction to be conducted. However, there is opportunity for improvement to this method with additional research needed.

IMPLICATIONS FOR POLICY AND PRACTICE - Current methods used for body fluid identification encompass a wide range of technologies including chemical, biochemical, immunochemical, enzymatic, spectrophotometric and others. Some rely on technologies that are a decades old, some are sensitive, other not, some rapid, other laborious hardly any are confirmatory. This assortment of techniques presents practical laboratory as well as policy problems as courts and juries come to expect test results that meet the high standards set by DNA testing.

The practical laboratory problems are many; some tests require biological activity, meaning that sample degradation in the field or in storage limit their usefulness. The need to perform multiple tests on an unknown sample to find out which body fluid is present is both time and sample consuming and requires multiple, and often expensive, instruments, as well as trained personnel to operate them. There are currently no regularly performed test for menstrual blood and vaginal fluid, and species identification is rarely performed because of expense and time.

Recent progress in the field of proteomics has demonstrated that a single methodology, mass spectrometry, can rapidly, and cost effectively, identify all body fluids in a sample simultaneously and confirmatively. Use of this technology will have implications for both the policy and practice of forensic science. Results from the work performed in this application can unify all body fluid identifications into a single sensitive method and will establish tests for menstrual blood, vaginal fluid and body fluids mixtures. The 2009 National Research Council's report *Strengthening Forensic Science in the United States: A Path Forward* has identified the need to bring all fields of forensics to the level of quality routinely provided by DNA testing, and policies to establish common standards and practices for body fluid testing are inevitable. Because mass spectrometry identifies the amino acid sequences of body fluid specific peptides, it offers a high level of certainty for results. Also, as the use of MS becomes more prevalent, its methods more streamlined and user friendly, sample extraction and preparations "kits" may become available and in our opinion mass spectrometry may become, in practice, the gold standard of body fluid testing.

We believe adoption of MS for body fluid detection would result in a single, uniform assay, using the same extraction, purification and detections methods for all body fluids, eliminating multiple techniques, saving time, money and sample, and giving the same confirmatory results for all body fluids. Above all, it would allow for the establishment of common, uniform standards throughout the forensic community in a manner similar to what is now in place for DNA testing. Its adoption would be beneficial not only to forensic science, but the criminal justice system, as the source of DNA (testing of which is approaching the level of a few cells) will be able to be identified – meaning that a jury will be more confident that a person's DNA was not the results of shed epithelial cells or transfer from a touched object, but from a specific body fluid – e.g. blood, saliva, semen, menstrual blood or vaginal fluid. This knowledge can only improve the quality of justice.

We have identified multiple unique and enriched protein markers in blood, saliva and semen. The strength of these protein markers rests in the fact that they perform specific biological tasks necessary for these body fluids to perform their functions (e.g. hemoglobin in blood, semenogelin in semen and amylase in saliva) and consequently these markers are well

established in their respective body fluids in the fields of biology, medicine and forensics. Further, we have identified multiple proteins in each body fluid and multiple peptides from each protein. The presence of multiple proteins and peptides imparts greater confidence to the accuracy of the assay, especially when trying to deconvolute body fluid mixtures. In addition, we have also identified multiple functional markers in menstrual blood and vaginal fluid which will allow confident identification of these body fluids. Indeed, because of our ability to identify menstrual blood specific markers, we can distinguish a mixture of venous blood and vaginal fluid from menstrual blood. Such tests may prove decisive in criminal cases, especially sexual assaults.

This work has focused on making these markers useful to the forensic practitioner in a way that is confirmatory, fast and inexpensive. Microwave digesting significantly reduced protein digestion time from overnight to less than one hour, and CPLC reduces menstrual blood processing time from day to hours. We believe that further work, described in the conclusion of the body fluid mixtures section, will improve mixture deconvolution detection levels even more than our current assay's ability.

The identification of multiple *functional* proteins unique to or enriched in these five body fluids will directly contribute to both the policy and practice of forensic investigations. With regard to policy, we believe that the demonstration of a panel of *functional* proteins markers will bring a solid scientific approach to serology analysis, making it a reliable confirmatory test. With respect to practice, the NYC OCME, the largest public forensic laboratory in the nation, is moving forward with confirmatory proteomic serology test because it is accurate, reliable, fast, amenable to high throughput and less expensive than the other molecular alternatives currently available.

INTRODUCTION – This section is divided into three parts: 1) STATEMENT OF THE PROBLEM, 2) LITERATURE REVIEW AND CITATIONS, AND 3) RATIONALE FOR RESEARCH

1. STATEMENT OF PROBLEM – Body fluid identification plays a vital role in forensic investigations contributing to case evidence as well as in directing the course of further investigations - particularly with respect to sexual assaults. Yet the methods used for body fluid detection have, for the most part, failed to keep pace with scientific and technological advances. As a consequence, current testing is a diverse mixture of techniques that vary dramatically in sensitivity, reproducibility, speed, cost and above all specificity – i.e. their ability to confirm that test results are real. For example, neither the hemoglobin (Kastle-Meyer) nor amylase tests are confirmatory; both give false positives for a variety of naturally occurring compounds, both rely on the subjective evaluation of qualitative results, both use technology that is nearly 100 years old (Patzelt, 2004) and both are in common use today. Even the more quantitative immunoassays like the enzyme-linked immunosorbent assay (ELISA) suffer from the inherent limitations of antibody specificity, affinity, narrow working ranges (i.e. antigen-antibody concentration ratios – the Hook effect), and manufacturing variability, as well as relatively high costs and long assay times – nor are they confirmatory. Additionally, there are many body fluids for which no routine tests are currently available, e.g. menstrual blood and vaginal fluid.

Unlike DNA testing, which relies on a single, dominant technology for testing all samples regardless of their source, there is no single uniform methodology to simultaneously evaluate an unknown forensic sample for all possible body fluids. Consequently, each body fluid test must be performed separately, consuming both time *and* evidence, as well as requiring multiple instruments and laboratory personnel trained in the nuances of each method – e.g. the subjective evaluation of the clearing of an opaque field around a suspected amylase sample in an agarose gel, or the relative appearance of a band on an immuno stick. In short, a single reproducible, confirmatory test that can identify all body fluids simultaneously is sorely lacking.

What would constitute an ideal body fluid assay? A single method able to simultaneously detect all body fluids accurately, reproducibly, rapidly, with high sensitivity, at low cost and above all - it must be confirmatory. Movement in this direction is the subject of much current research, particularly in the areas of mRNA, micro-RNA and DNA methylation (which are reviewed below). However, we believe a mass spectrometry (**MS**) proteomics assay offers several advantages over these methods: i) proteins are less susceptible to environmental degradation than RNA and DNA, ii) partial degradation of proteins does not preclude marker identification (nucleic acid assays require unbroken sequences from primer to primer), iii) multiple function-specific protein markers are present in each body fluid and simultaneously detected – increasing confidence that test results are correct, iv) MS assay sensitivity is greater than that of mRNA assays [3], v) mass spectrometry is unbiased with regard to sample type, i.e. no a priori knowledge of a sample is necessary prior to assay (e.g. no specific body fluid or species primers), MS simply identifies those proteins that are present, vi) species information is inherent in the sample and reported with results and vii) the assay is amenable to high throughput, and costs are low.

2. LITERATURE REVIEW & CITATIONS - This review focuses on the three most commonly studied confirmatory molecular methods currently being investigated for body fluid identification. All involve nucleic acids.

- **DNA METHYLATION PROFILING: EPIGENETIC ANALYSIS** – The advent of epigenetics (changes in DNA that can be heritable and/or effect gene expression but do not affect DNA sequence) has opened another area of research into molecular targets for body fluid identification [39]. This is particularly the case for DNA methylation. The DNA methylation mechanism most often studied is that of cytosine. The biology behind cytosine methylation is regulation of gene expression. Methylation of cytosine in areas of the chromosome rich in CpGs (typically upstream of genes) can down regulate gene expression [80]. As different cells, tissues and body fluids express different proteins for their particular function, they have, by necessity, different patterns of DNA methylation. Several techniques are commonly employed to detect cytosine methylation including: sequencing, single-nucleotide primer extension, restriction endonuclease pattern analysis, and methylcytosine immuno-precipitation among others [82]. All depend on the chemical conversion of unmethylated cytosines by bisulfite to uracils, which are replaced during PCR to thymines. Numerous studies have identified specific CpG sites that can be used to distinguish blood, saliva, semen, vaginal fluid and menstrual blood. [39,76,78,79,81,83,85,86,87]. As with other body fluid molecular identification methods, multiple markers are employed. For example Silva et al. [78] have identified three markers differentially expressed in blood, saliva and semen and which can recognize mixtures of two and three of these body fluids, although until additional markers identified, they are not yet able to deconvolute mixtures. Their results show good replication between laboratories.

An advantage of this method is the greater stability of DNA compared to RNA, and while the method has matured (samples in the ng to sub ng level can be detected depending on the locus

[78]), chemical and biological obstacles remain. Bisulfite conversion is an aggressive chemical reaction that can lead to DNA degradation, consequently bisulfite treatment is a balance between conversion and degradation. Leontiou et al. [77] reported DNA loss between 33-55%. This could be important when DNA evidence is in low amounts. Identification of menstrual and venous bloods also remains difficult for this method. Genome-wide methylation profiling of a variety of body fluids, including menstrual blood and vaginal fluid showed fewer significant CpGs when compared to other body fluids [76], as well as expression of vaginal fluid markers in blood and semen. A bioinformatic approach to estimate body fluid probabilities might prove helpful. Perhaps more importantly, however, are the changes in methylation patterns known to occur with age [41], gender [42], ethnicity [44], health and disease [43,45], as well as environmental conditions [46], and diet [48]. Lee et al. [76] reported that one of their vaginal fluid makers was elevated in blood, saliva and semen more commonly in males over the age of 50. Indeed, it is these very unique methylation patterns that are also leading research into DNA methylation as a forensic tool to determine age, health and even social behaviors such as smoking and alcohol consumption [84]. Methylation patterns can also differ between the same tissues in different individuals, as well as within the same tissue of a single individual [44]. Consequently significant numbers of samples will need to be tested to ensure a consistent core group of markers impervious to gender, aging, health and environmental conditions.

• **MICRORNA PROFILING** – Numerous studies have focused on the use of micro-RNAs (**miRNA**) for body fluid identification [18,28,29,64,74,75,88]. miRNAs are a class of small, non-coding RNAs (18-24 nucleotides in length) that primarily function in the regulation of protein expression by binding to the 3' untranslated region of mRNAs and blocking translation [30]. Mature miRNAs associate with Argonaute proteins (a family of endoribonucleases) that, when recruited to an mRNA can induce degradation [63]. It is estimated that ~30% of human mRNAs [31] are regulated by miRNAs and that a single miRNA can bind ~200 different mRNA transcripts [30]. Conversely, a single transcript can bind more than one miRNA [32, 63]. (The complexity of this regulatory system can make unambiguous interpretation of data difficult, see below.) Nearly 2,000 high confidence human miRNAs are listed in the miRBase v.21 database (mirbase.com).

Because miRNA expression patterns have been shown to vary between tissues they have become a focus of both clinical and forensics research, the former to distinguish diseased from non-diseased tissues (a relatively simpler task), the latter to distinguish body fluids. However, because it is the pattern of multiple commonly expressed miRNAs, not the presence of a single uniquely expressed miRNA that distinguishes tissues and body fluids, data interpretation requires the evaluation of a complex fingerprint. The techniques typically used to study miRNAs are reverse transcription followed by quantitative PCR or microarrays; both methods require care and attention to detail.

To date several groups have examined the five most commonly encountered forensic body fluids (blood, saliva, semen, menstrual blood and vaginal fluid) and have identified multiple, but *different* miRNA markers [18,73,74,75] reexamined three of the published markers using more rigorous analytical tools, but could confirm only one, miR-16 in blood.

It is not clear if these dissimilar findings are the result of technical differences (RT-qPCR vs microarray) or inherent biological obstacles, i.e. the actual sources of miRNAs. For example, Pigati et al. [32,34] have demonstrated the release of different miRNAs from normal and diseased mammary epithelial cells into the blood, milk and ductal fluids of women. One of the miRNAs released from malignant cells was miR-451, which was previously identified as a blood and menstrual blood miRNA marker [18]. Release of different miRNAs into the circulatory system occurs in both lactating and non-lactating women, and has also been shown to occur in lung and prostate cancers, and gliomas [34]. Further, Wang et al. have demonstrated altered miRNA expression profiles in seminal plasma from patients with unobstructed azoospermia, asthenozoospermia and oligospermia when compared to normal controls. Finally, evidence

from the literature demonstrates that epigenetic effects (in response to chemotherapeutic drugs [35,37] and trauma [38]) can substantially influence miRNA expression profiles in cell culture.

These data demonstrate that both normal and diseased cells routinely release miRNAs into blood and other body fluids, and pose at least a theoretical problem of body fluid identification in the presence of disease, trauma and perhaps even changes that can occur with age, puberty, menstrual cycle, therapeutic treatments, as well as with gender and ethnicity. At the moment, most of these are hypothetical possibilities and may not preclude use of miRNAs as body markers. However, because miRNA markers are not unique to the function of specific body fluids (as are hemoglobin to blood, and semenogelin to semen), and because they are not always the most abundant miRNAs found in particular body fluids [18], they may prove difficult to validate. For these same reasons, the use of miRNAs may also pose significant obstacles for identifying body fluid mixtures.

Conflicting data have also been published regarding the stability of miRNAs. One study, titled "Robust microRNA stability in degraded RNA preparations from human tissue and cell samples" [20] appears to contradict a second "RNA degradation compromises reliability of microRNA expression profiling" [21]. These different outcomes may be due to different sample types and methods of preparation as miRNAs shows greater stability in vivo when associated with their Argonaute protein, than ex vivo [63]. However, it does appear that miRNAs are inherently more stable than mRNA, likely due to their small size which makes them less of a target for RNases.

Importantly for this discussion, Hanson et al. [64], identified three miRNAs from ten menstrual blood samples that they used in a logistic regression model to demonstrate their ability to distinguish menstrual from peripheral blood. (Recently Li et al. [65] confirmed one marker, miR144-3P.) Only one false negative was reported, although the menstrual blood sample size was limited. Menstrual blood data had a wider spread than the three other body fluids tested (blood, saliva and semen). This may have resulted from different amounts of material collected during swabbing with cotton swabs, different day of menses, or different individuals. However, samples all tested positive. Of the three other body fluids tested, none gave a false positive result alone or in mixtures of two. However, a blood and vaginal fluid mixture was not tested.

• **MESSENGER RNA EXPRESSION PROFILING** - Body fluid identification by mRNA expression profiling, has been studied by numerous laboratories [10,11,15,17,19,66,67,69,70]. The goal of this technique is to identify body fluid specific mRNA markers that can be used to unambiguously identify each body fluid. The technology employed is typically reverse transcription (**RT**) followed by quantitative (real time) PCR (**RT-qPCR**) or capillary electrophoresis (**CE**). Ballantyne and others have identified specific mRNA and developed multiplex RT-qPCR/CE assays to detect body fluid-specific mRNAs for blood, saliva, semen, vaginal secretions and menstrual blood from single or mixed stains [10, 15, 16, 17, 66, 67, 68, 69, 72]. The European DNA Profiling Group has shown RNA/DNA co-analysis from menstrual blood and vaginal fluid [66].

There are, however, several potential drawbacks to this method that may limit its usefulness. Primary among these is the stability of mRNA and the ubiquitous nature of RNases which digest it [18,20,21,22]. RNases are generally small, stable molecules that function at neutral pH, and can be difficult to inactivate [23,71]. RNases are not only present within cells, where they function in cellular metabolism, but are part of the body's defense mechanism and are secreted into such body fluids as tears, blood, saliva and perspiration [24]. They are also produced by microorganisms. Studies have shown that RNA stability varies between tissues [17,25] and is dependent on storage conditions [27]. Consequently, digestion of mRNA by RNases from within cells and body fluids, as well as possible RNase contamination of evidence from collection through analysis by perspiration (fingertips) and/or microorganisms, has the potential to limit the usefulness of this method. Even partial degradation of a sample can be detrimental

to the assay, as the entire region of amplification from the 5' primer to the 3' primer must be intact for the assay to be effective. In addition, polymorphisms at primer binding sites can reduce or eliminate amplification. Use of smaller amplicons might benefit this approach, however, unless designed to cross exons, precautions against contaminations by genomic DNA will be required.

Another potential problem is that not all body fluid "specific" markers are unique to a body fluid, although they may be predominantly expressed there [17,26,67]. This is particularly a problem for PCR assays which can amplify extremely small amounts of nucleic acids, and therefore potentially detect signals from contaminating tissues/body fluids. Consequently, quantitative RT-PCR assays using several markers for each body fluid will need to be established in order to construct body fluid "mRNA fingerprints."

Currently, mRNA body fluid assays sensitivity appears to be nearly one order of magnitude less than that attained by MALDI mass spectrometry for blood, saliva and semen [3,16,69].

3. RATIONAL FOR RESEARCH –

The revolution that has swept proteomics over the past decade rivals what occurred in genomics in the previous decade. Much of this progress is due to advances in mass spectrometry instrumentation which is much faster, simpler, more accurate and more sensitive than previous generations of instruments [49]. Similar advances have occurred in nano-liquid chromatography and equally as important in the power of bioinformatics to rapidly search and recover annotated protein information from ever increasing databases. While these advances have propelled proteomics to the forefront of basic research and pharmaceutical discovery [50,51], they are also being eyed for routine rapid testing in the clinic [52,53], particularly as a diagnostic tool to identify diseases and monitor therapeutic treatments [54] from easily obtainable body fluids including: blood [55], saliva [56], semen [57], and vaginal fluid [58].

These same MS tools are directly applicable to the forensic identification of body fluids, and we have used them, with the support of NIJ, to establish a rapid, sensitive and confirmatory test for three simple body fluids, blood, semen and saliva [3] which is beginning validation testing at the NYC OCME. Confirmatory MS assays for more complex samples, e.g. menstrual blood, vaginal fluid and body fluid mixtures, are under development, but because their marker proteins represent only a tiny fraction of total sample protein, they require additional purification steps prior to mass spectrometry. Additionally, all current MS assays require overnight (12 hour) digestion with trypsin prior to analysis. It is these pre-MS preparation steps that are hampering the implementation of routine rapid testing of complex body fluids (i.e. samples in which markers are in low abundance) both forensically well as in the clinic.

Recently, advances in protein chromatography and microwave technology have made significant gains not only in dramatically reducing sample preparation and digestion times (in some cases from days to minutes), but also improving the quality of analytes and consequently, test results. The goals of this application were to test and evaluate these new technologies and integrate them into our current assays in order to improve throughput. Additionally, we evaluated the simultaneous extraction of protein and DNA from single samples with the aim of reducing sample consumption for body fluid and STR identification. Finally, we evaluated the stability of menstrual blood and vaginal fluid markers over time, as well as assess the potential use of recently published vaginal fluid flora [1,2] as additional markers for our vaginal fluid assay.

RESULT SECTIONS - Results presented in this report are divided into four sections: 1) COMBINATORIAL PEPTIDE LIGAND CHROMATOGRAPHY – DYNAMIC RANGE REDUCTION, 2) MICROWAVE-ASSISTED ENZYME DIGESTION, 3) MENSTRUAL BLOOD AND VAGINAL FLUID MARKER STABILITY AND 4) SIMULTANEOUS PROTEIN & DNA EXTRACTION & PURIFICATION

SECTION 1 COMBINATORIAL PEPTIDE LIGAND CHROMATOGRAPHY – DYNAMIC RANGE REDUCTION

The evaluation of combinatorial peptide ligand chromatography (CPLC) was a specific aim of this application. Our demonstration of its usefulness in improving protein marker identification led to the technique's incorporation in our search for menstrual blood markers. Consequently, part of the general discussion of how CPLC works was reported earlier (NIJ Grant 2010-DN-BX-K192).

- **WHAT IS COMBINATORIAL PEPTIDE-LIGAND CHROMATOGRAPHY?** CPLC is a method to reduce the dynamic range of proteins in a sample. The dynamic range is the difference between the most abundant and least abundant proteins present. When this range is too great, proteins at the lowest end may not be detected. In many body fluids the protein dynamic range can be many orders of magnitude. In blood, for example, it is estimated that abundant proteins (e.g. hemoglobin, albumin, immunoglobulins, etc.) are present at 9-10 orders of magnitude more than rarer proteins [59]. Reducing the protein dynamic range of a sample can significantly improve detection of rare but important marker proteins. This is particularly true of menstrual blood which is composed predominately of circulating blood proteins, but can prove equally important for body fluid mixtures which can contain extreme ratios of two or more body fluids.

CPLC is a relatively new method to reduce protein dynamic range [52,60]. The concept is not dissimilar to typical non-specific protein separation techniques – but with a catch. It non-specifically enriches all proteins equally, effectively capturing similar amounts of high abundant and low abundant proteins in a single step. CPLC employs random six-mer amino acid aptamers (synthesized with all 20 amino acids in all possible combinations) linked to a bead matrix packed in micro-spin columns. Each bead has thousands of the same aptamer bound to it and all beads are present in equal amounts. As different proteins have different affinities for different bead aptamers, each protein will bind most strongly to its cognate bead(s). This means that proteins of low abundance will accumulate as they pass through the column, while proteins of high abundance (after saturating their associated bead(s)) will pass through the column effectively reducing the overall protein dynamic range by concentrating low abundance proteins while maintaining representatives of nearly all proteins contained within the original sample. In order to ensure reproducibility and maximum dynamic range compression column binding capacity needs to be saturated.

The goal of this work is to evaluate combinatorial peptide ligand chromatography as a means of rapidly reducing the dynamic range complex biological fluids (menstrual blood and body fluid mixtures) in order to improve body fluid marker detection in MS analysis.

- **EXPERIMENTAL DESIGN:** We chose to use menstrual blood for our initial evaluation of CPLC since i) menstrual blood markers represent only a fraction of total menstrual blood proteins and ii) we have already identified some menstrual blood candidate markers using other methods. If successful, this method would significantly reduce sample processing time compared to

traditional hemoglobin (**Hb**) albumin (**Alb**) and immunoglobulin G (**IgG**) depletion followed by isoelectric focusing (**IEF**) and in-gel digestion, which take several days.

We evaluated the use of different detergents used during the CPLC process in order to optimize protein extraction and recovery of proteins after CPLC treatment. Because detergents can help solubilize cell membranes, they may release more and different types of proteins during initial extraction than non-detergent containing solutions. In addition, we wanted to test detergent containing elution buffers to find out which would release the most proteins from the CPLC matrix.

- **SAMPLE PREPARATION METHODS:** Menstrual blood samples were collected from volunteers using menstrual blood collection cups. Menstrual blood samples were vigorously vortexed to help break apart membranous tissue fragments. CPLC beads (ProteoMiner) were purchased from Bio-Rad, Hercules, CA. Because the chemical properties of sample extraction buffers function in both solubilizing sample proteins as well as influencing their ability to bind to the CPLC beads, we evaluated three different extraction buffers: i) phosphate buffered saline (PBS), ii) PBS plus 0.1% SDS and iii) PBS plus 0.5% NP-40. We also evaluated two column elution buffers: i) 8 M Urea, 2% CHAPS in 5% acetic acid (**UCA**), and ii) 2% SDS in PBS. A cocktail of protease inhibitors was added to all extraction buffers (Pierce™ Protease Inhibitor Tablets, EDTA-free).

Menstrual blood samples were extracted in 15 volumes of extraction buffer using Lysis Matrix D (1.4 mm ceramic beads) in a reciprocating homogenizer at 6 m/sec for 30 sec (FastPrep24, MP Biomedical, Santa Ana, CA). Samples were centrifuged at 18,000 g x 30 min at 4°C. The supernatant was transferred to a clean tube and protein concentration measured by bicinchoninic acid (BCA) assay with bovine serum albumin (BSA) as a standard. As per manufacturer instructions, CPLC columns were washed twice with PBS, loaded with 10 mg of sample protein and incubated for 2 hrs at room temperature. Columns were then washed three times with phosphate-buffered saline (PBS). For elution with UCA, 20 µl of elution buffer was added to the beads and incubated at room temperature with occasional light vortexing over a period of 15 min. Eluates were collected by centrifuge at 1,000 g x 60 sec. For elution using 2% SDS in PBS, 60 µl of elution buffer was added to the beads and incubated at 90°C x 15 min. This elution solution also was collected by centrifugation at 1,000 g x 60 sec. The BCA assay with BSA as a standard was used to measure protein concentrations of the flow through, all washes, and the eluate. SDS PAGE analysis was performed on all fractions.

- **HEMOGLOBIN, ALBUMIN AND IMMUNOGLOBULIN DEPLETION:** Hemoglobin (Hb) was depleted from samples using NTA agarose beads according to the method of Ringrose et al [61]. Approximately 800 µl Ni-NTA agarose beads were added to a volume of menstrual blood containing 20 mg of protein in lysis buffer. The suspension was gently mixed end-over-end for 1 hour at room temperature and beads separated from the depleted sample by centrifugation. The beads were washed twice with lysis buffer, centrifuged, and the washes combined with the original supernatant. The combined HB-depleted fraction typically contained about 10% (2 mg) of the original sample's protein. Hb was eluted from the Ni-NTA column using lysis buffer containing 250 mM imidazole. Following Hb depletion, the sample was depleted of albumin and IgG using the ProteoPrep Immunoaffinity Albumin and IgG Depletion kit (Sigma-Aldrich, St.

Louis, MO) following the manufacturer's instructions. Bound proteins were eluted using manufacturer's elution buffer (40 mM Trizma Base, 7.0 M thiourea, and 1% C7BzO detergent, pH 10.4).

- **POLYACRYLAMIDE GEL ELECTROPHORESIS:** Precast 4-20% SDS PAGE gels (Bio-Rad, Hercules, CA) were used for qualitative analysis of protein extracts and CPLC fractions. Samples were quantified by BCA as described above. Twenty μg protein (unless stated otherwise) were denatured in 100 mM beta-mercaptoethanol and 2% SDS at 95°C for 10 min. before being loaded on gels. Gels were run at 110 V constant voltage for 60 min. then fixed, stained with Coomassie and destained following manufacturer's instructions (Coomassie Brilliant Blue R-250, Bio-Rad).

- **MASS SPECTROMETRY METHODS:** All samples were prepared for mass spectrometry analysis as follows: 20 μg of protein were reduced and alkylated using 5 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP) and 15 mM iodoacetamide. Eluates containing 2% SDS were cleaned with detergent removal spin columns (0.5 ml, Pierce) as directed by manufacturer. Samples were digested with 1 μg of trypsin overnight at 37°C. Two micrograms of digested peptides were loaded for LC-MALDI. Peptides were first desalted with an inline C18 trap column (300 μm x 5 mm, 5 μm , 100 Å, Dionex, Sunnyvale, CA) and separated on an analytical a C18 column with an 85 min acetonitrile (ACN) elution gradient of 2-40% ACN containing 0.1% TFA at a flow rate of 300 nl/min using an Ultimate 3000 system (Dionex, Sunnyvale, CA). Eluted peptides were mixed in-line with α -cyano-4-hydroxycinnamic acid (HCCA) matrix (5 mg/ml in 60% ACN, 0.1% TFA) and 192 aliquots were spotted on a MALDI plate. MS and MS/MS data were collected on an Sciex 4800 Matrix-assisted laser desorption ionization time of flight/time of flight (MALDI TOF/TOF) Analyzer (Framingham, MA). MS data was acquired at a laser repetition rate of 200 Hz with 600 laser shots/spectrum (40 laser shots/sub-spectrum). MS/MS data was acquired at 200 Hz in 1kV MS/MS mode with 50 laser shots/sub-spectrum under the following TOF/TOF Series Explorer Stop Conditions: maximum shots per spectrum 2500; minimum shots per spectrum 1000; number of MS/MS fragments 8; signal to noise ratio for each fragment was set at 75. The top 30 strongest peaks were selected for MS/MS analysis.

- **RESULTS**

- I) COMPARISON OF EXTRACTION AND ELUTION BUFFERS:** Three different extraction buffers were evaluated: i) PBS, ii) PBS plus 0.1% SDS and iii) PBS plus 0.5% NP-40, and two elution buffers: i) 8M Urea, 2% CHAPS in 5% acetic acid (**UCA**), and ii) 2% SDS in PBS to determine which combination would allow the greatest number of proteins to be recovered and identified after CPLC treatment.

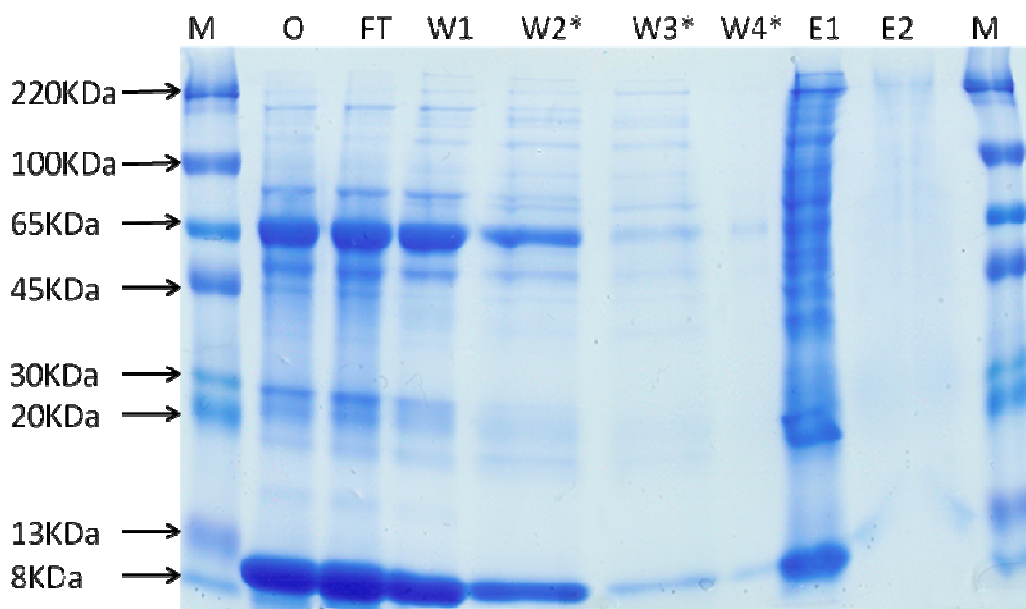
Qualitative analysis of different buffer systems by SDS PAGE (**Figures 13 – 16**) reveals that regardless of the buffer conditions, CPLC is effective in reducing protein dynamic range. This can be seen by comparing the banding patterns of the original starting material of each sample "O" with the final elution "E". As may be seen in each gel, the abundant Hb, Alb and IgG bands predominant in the starting material are all reduced in the eluates. However, this does not mean that the buffer systems are equivalent.

We evaluated the efficacy of SDS as a CPLC matrix elution reagent by comparing it to UCA (the manufacturer's suggested elution buffer). The SDS PAGE gel of a sample extracted with PBS and eluted from the CPLC columns with 2% SDS (**Figure 13**) shows more visible protein bands in the eluate fraction (E1) than the gel of sample extracted with PBS and eluted with UCA (**Figure 14**). This result was confirmed by MALDI TOF/TOF analysis (Table 9) showing more proteins eluted by 2% SDS.

A comparison of the eluates (E) of a sample extracted only with PBS (**Figure 13**) and a sample extracted with PBS plus 0.5% NP-40 (**Figure 15**) shows what appear to be similar elution patterns. However, MALDI TOF/TOF analysis of these eluates shows that NP-40 extracts more total protein (Table 9). The addition of 0.1% SDS as a solubilizing reagent is not as effective as NP-40, with fewer proteins identified and a less complex elution pattern (**Figure 16** and Table 9). This is likely because even though SDS, an ionic detergent, may solubilize more proteins, they are unable to bind to the CPLC matrix.

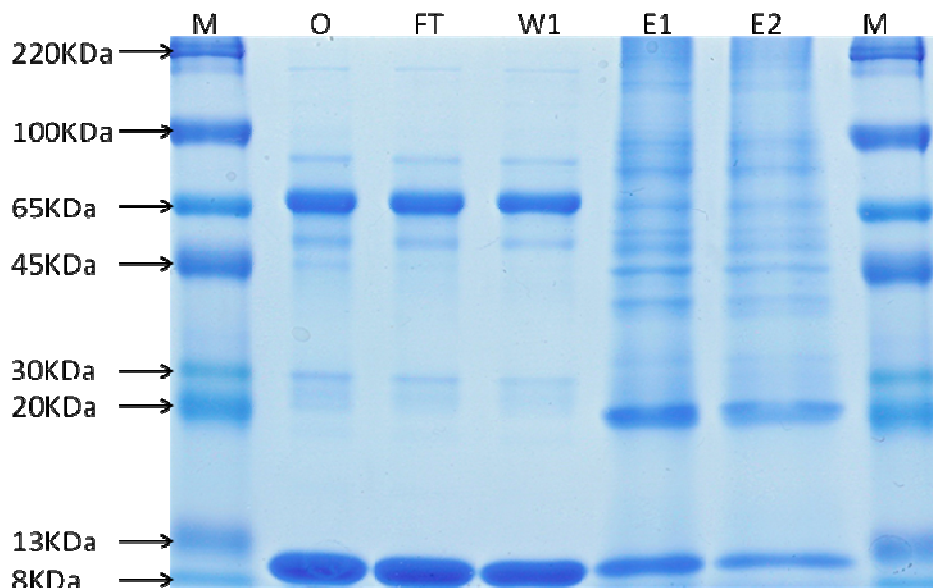
Based on these results we have chosen to use 0.5% NP-40 as an extraction buffer and 2% SDS as the CPLC matrix elution buffer.

Figure 13 SDS PAGE analysis of CPLC menstrual blood fractions



Extraction/Elution Buffers = PBS/2% SDS (1st elution) followed by UCA (2nd elution). M: Marker; O: Original starting material in PBS; FT: Flow-Through W1-W3: PBS washes 1-3; W4: Water Wash; E1: Protein eluted by 2% SDS in PBS; E2: Protein eluted by UCA. * Unknown protein amount was loaded because protein concentration in W2, W3 and W4 was too low to measure.

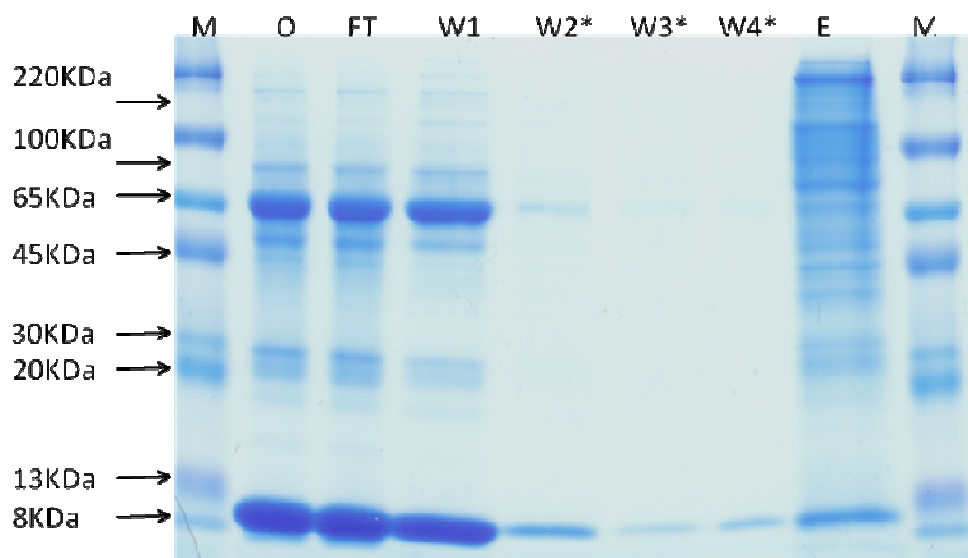
Figure 14 SDS PAGE analysis of CPLC menstrual blood fraction.



Extraction/Elution Buffers = PBS/UCA (1st elution & 2nd elution).

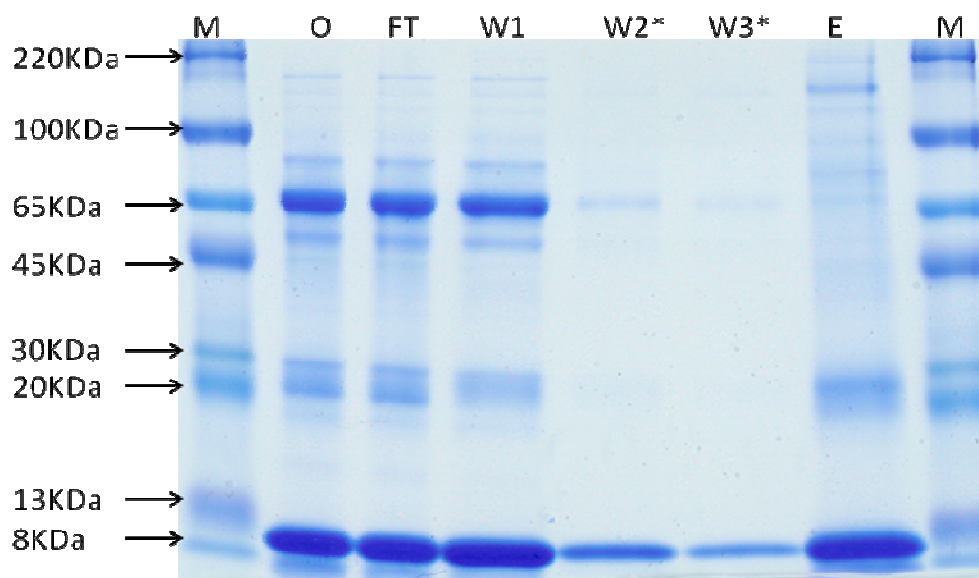
M: Marker; O: Original starting material in PBS; FT: Flow-Through; W1 PBS Wash; E1: Protein eluted by UCA; E2: Protein eluted by UCA. Washes W2-W4 not shown as little protein is seen in these washes (see **Figure 13**).

Figure 15 SDS PAGE analysis of CPLC menstrual blood fraction



Extraction/Elution Buffers = 0.5% NP40 in PBS/2% SDS. M: Marker; O: Original starting material in PBS; FT: Flow-Through W1-W3: PBS washes 1-3; W4: Water Wash; E: Protein eluted by 2% SDS. * Unknown protein amount was loaded because protein concentration in W2, W3 and W4 was too low to measure.

Figure 16 SDS PAGE analysis of CPLC menstrual blood fraction.



Extraction/Elution Buffers = 0.1% SDS in PBS/2% SDS in PBS. M: Marker; O: Original starting material in PBS; FT: Flow-Through W1-W3: PBS washes 1-3; W4: Water Wash; E: Protein eluted by 2% SDS. * Unknown protein amount was loaded because protein concentration in W2, W3 and W4 was too low to measure.

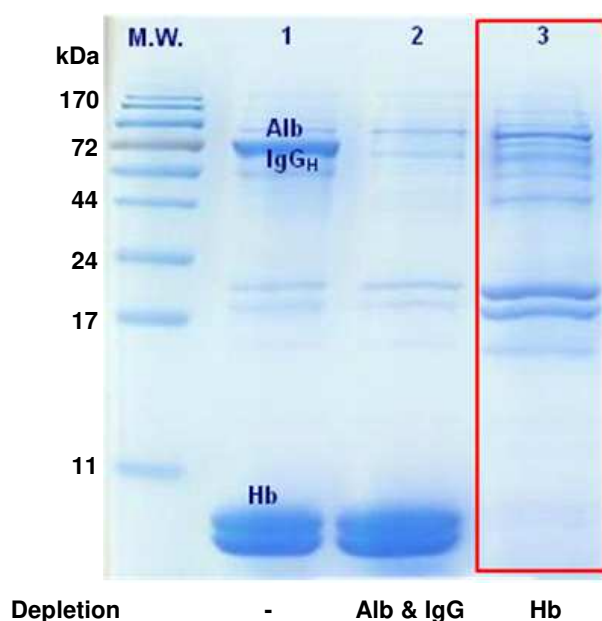
Table 9 Number of proteins identified MALDI TOF/TOF from 2 µg of different eluates

Confidence	PBS / UCA	PBS / 2% SDS	0.5%NP40 / 2% SDS	0.1%SDS / 2% SDS
------------	-----------	--------------	-------------------	------------------

>2.0 (99%)	157	216	237	96
>1.3 (95%)	215	262	312	124

II) COMPARISON OF HB/ALB/IGG DEPLETION AND IEF TO CPLC – **Figures 17 and 18** demonstrate the qualitative differences between Hb/Alb/IgG depletion of menstrual blood and CPLC treatment of menstrual blood. **Figure 17**, Lane 1 shows the banding pattern of 20 µg of the untreated menstrual blood starting material. Lane 3 shows 20 µg of the Hb/Alb/IgG depleted sample. As may be seen there is nearly complete removal of Hb, and a considerable reduction of Alb and IgG bands. Consequently, the depletion method allows proteins present in lesser amounts to be visualized. However, while the dynamic range is reduced, there is no protein enrichment, i.e. proteins present in extremely small amounts may remain outside the level of detection.

FIGURE 17 SDS PAGE of menstrual blood sample processed by traditional depletion methods for removal of Hb/Alb and IgG (20 µg proteins loaded/lane)



M.W. = molecular weight standards, 1 = menstrual blood starting material, 2 = Alb/IgG depleted sample, 3 = Hb depleted sample.

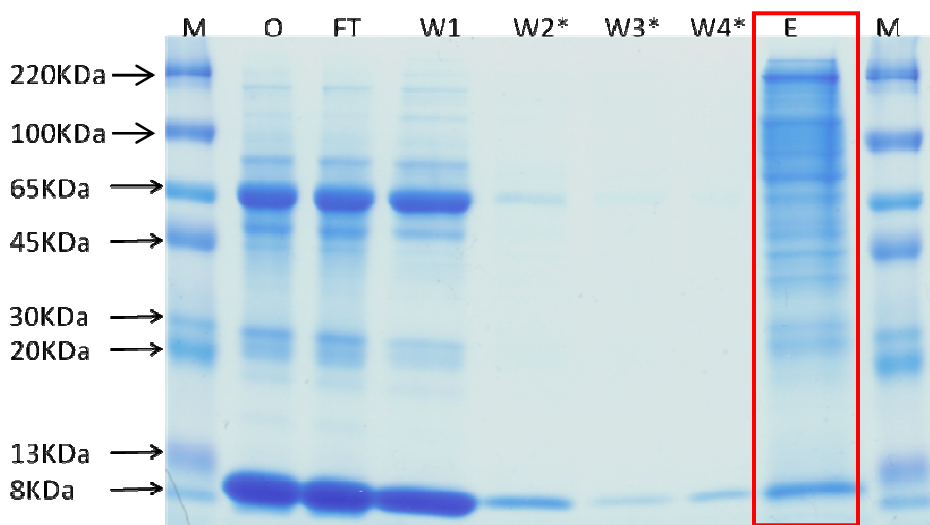
Figure 18 shows the results of a menstrual blood sample processed by CPLC. The banding pattern of the original starting material may be seen in Lane O where again, the largest bands represent Hb, Alb and IgG. The CPLC matrix is intentionally overloaded with 10 mg of menstrual blood starting material to promote maximum binding of all protein species regardless of their relative amounts. This is obvious from the non-bound flow-through fraction Lane FT which has a nearly identical pattern as Lane O. The first wash, W1 shows a similar pattern, but subsequent washes Lane W2-W4 show that most non-specifically bound protein are removed. The proteins eluted from the column, Lane E, show a near continuous banding pattern of

proteins not seen in the starting material, Lane O, and more complex pattern than seen by Hb/Alb/IgG depletion (**Figure 18** Lane 3). While the Hb band is still present, there has been an obvious reduction in protein dynamic range.

Table 10 shows that 470 distinct proteins were identified with MALDI TOF/TOF analysis for the depletion-IEF method and the CPLC method identified 428 proteins. Only 59% (275) of depletion-IEF proteins were distinct from those found in venous blood, while 73% (312) of CPLC proteins were different. This may be due to the enrichment process of the CPLC method.

It should also be noted that these two methods complement each other as the depletion-IEF method identified 298 proteins not found by the CPLC method, and CPLC identified 256 proteins not found by depletion-IEF. One hundred and seventy-two proteins were found by both methods. It is well established that multiple MS analyses of the same sample identify a small percentage of different proteins [62]. These data are graphically represented in the Venn diagram in **Figure 19**. Because fewer novel proteins were found in the depleted-IEF sample some of the 298 unique proteins are likely to be found in venous blood.

FIGURE 18 SDS PAGE of menstrual blood sample processed by CPLC. 20 µg proteins loaded/lane except W2-W4 where maximum volumes were loaded.



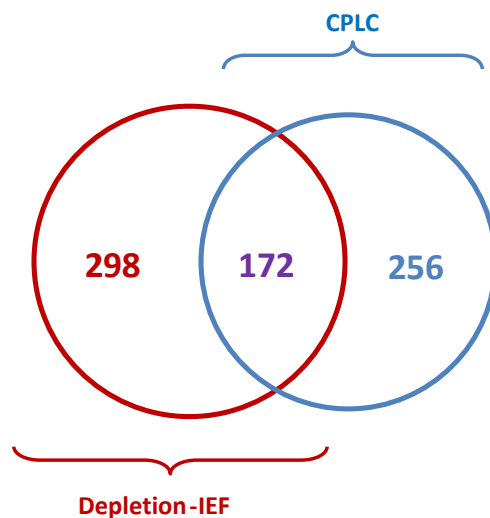
SDS PAGE analysis of menstrual blood after CPLC. M: molecular weight standards; O: Original menstrual blood starting material extracted with 0.5% NP40 in PBS buffer; FT: Flow Through; W1-W3: Washes; W4: Water Wash; E: Eluate with 2% SDS in PBS. *: unknown protein amount was loaded because protein concentration in was too low to be measured.

Table 10 Comparison of proteins identified using standard depletions methods followed by IEF compared with CPLC without IEF.

Method	Proteins Identified	Proteins Not Found in Venous Blood	IEF Proteins not found in CPLC	CPLC Proteins not found in IEF	Shared by IEF & CPLC
Depletion-IEF	470	275 (~59% novel)	298	-	172

CPLC	428	312 (~73% novel)	-	256	172
------	-----	------------------	---	-----	-----

Figure 19 Venn diagram showing the number of distinct and common proteins identified by the depletion-IEF and CPLC method.



Finally, in addition to its ability to reduce protein dynamic range, CPLC is significantly faster than the traditional depletion-IEF method. Since only one chromatographic step is necessary in CPLC a sample can be processed in several hours, while depletion-IEF typically requires several days to process.

DETERMINATION OF CPLC COLUMN BINDING CAPACITY: While CPLC saturation varies with the type and complexity of the protein sample used, the manufacturer suggests a starting point of 10 mg protein for the pre-loaded CPLC spin columns (ProteoMiner, Bio-Rad Hercules, CA). We measured the amount of protein eluted from CPLC column loaded with different amounts of menstrual blood or venous blood protein in order to determine the column binding capacity when used with menstrual blood or venous blood.

- **METHODS:** Samples were homogenized and extracted according to the methods described in the previous section, in PBS containing 0.5% NP-40. Extract concentrations were measured by the BCA assay using BSA as the standard. Due to the differences in menstrual blood samples (day of the menstruation, volumes collected, blood/mucus content, etc.) protein content varied widely. However, all samples with greater than 2 mg protein were used to determine CPLC column binding capacity. Of the 193 menstrual blood samples collected, 29 (15%) had less than 2 mg protein. The remaining 164 (85%) samples were processed through CPLC (Table 9). Applied protein ranged from ~2-12 mg (average 7.8 mg; standard deviation 2.7 mg). According manufacturer instructions (ProteoMiner, Bio-Rad), CPLC columns were washed twice with PBS, loaded with sample protein and incubated for 2 hrs at room temperature. Columns were then washed three times with phosphate-buffered saline (PBS). Protein was eluted with 2% sodium

dodecyl sulfate (SDS) in PBS at 90°C for 15 minutes. Concentration of eluted protein was measured by the BCA assay using bovine serum albumin BSA as the standard.

Table 11 Menstrual & Venous Blood Samples

Sample	Collected	Extracted	Evaluated by SDS MiGE (Agilent)	Run on CPLC		
				Insufficient Protein	164 (85%)	
					<7,500 µg	≥7,500 µg
Menstrual Blood	193	193	193	29 (15%)	52* (32%)	112* (68%)
Venous Blood	45	45	45	-	13 (29%)	32 (71%)

* Four ≤7,500 and one >7,500 columns failed; there was insufficient protein to repeat CPLC.

• **RESULTS:** Saturation was determined by comparing the amounts of eluted protein against the amounts applied – i.e. once saturation occurs column binding should be maximized and eluate protein should plateau. **Figure 20** shows a scatter plot applied vs. eluate of all menstrual blood samples.

A breakdown of the data (**Figure 21**) show a relatively linear relationship between applied and eluted samples up to approximately 7,500 µg applied protein (linear regression). **Figure 22** shows most samples with more than 7,500 µg applied protein appear to cluster between 150 and 300 total µg protein recovered, suggesting that saturation occurs at about 7,500 µg applied protein.

Figure 20 Determining CPLC Column Saturation for Menstrual Blood

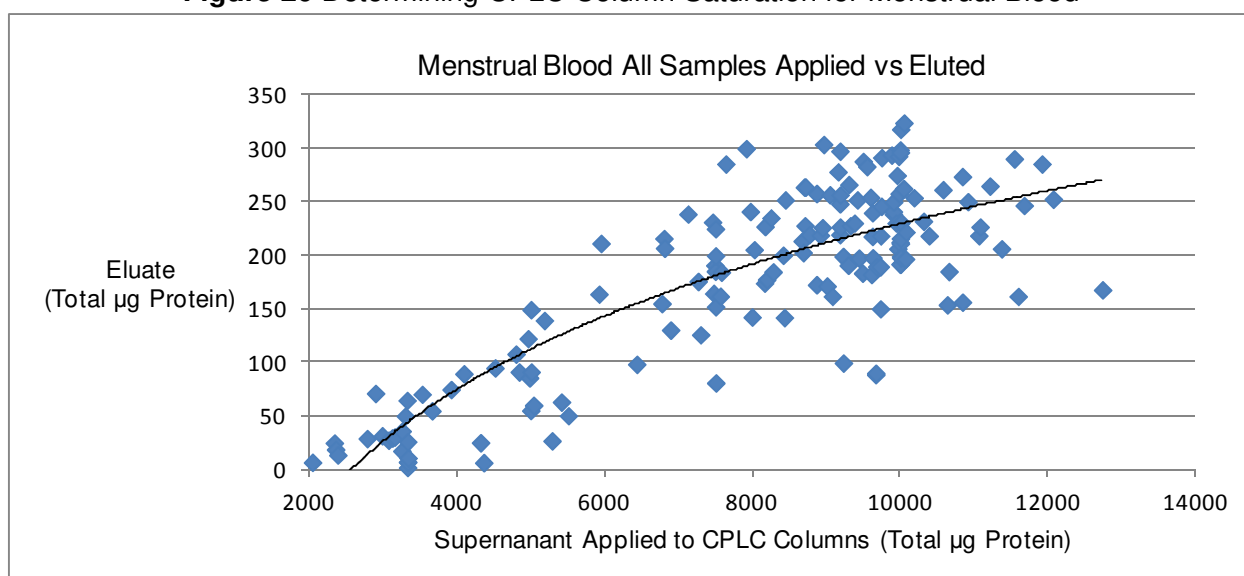


Figure 21 CPLC Linear Binding Range for Menstrual Blood

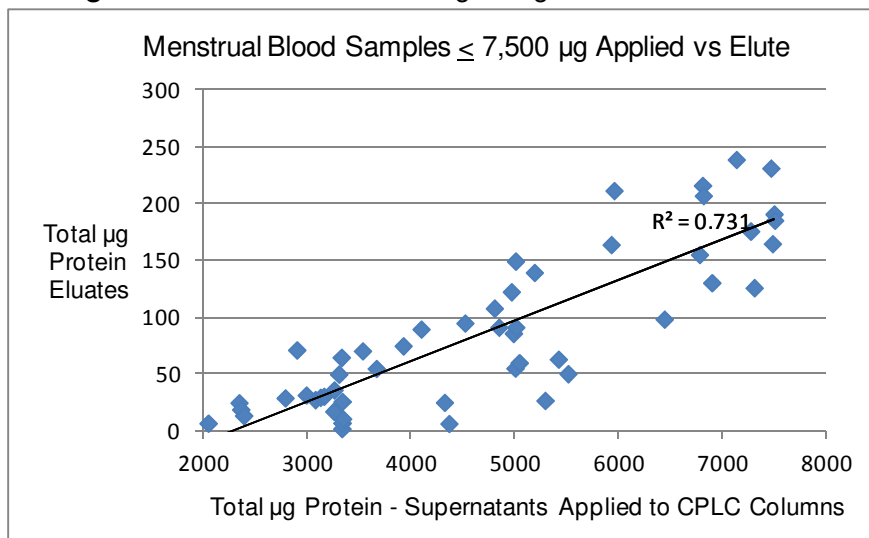
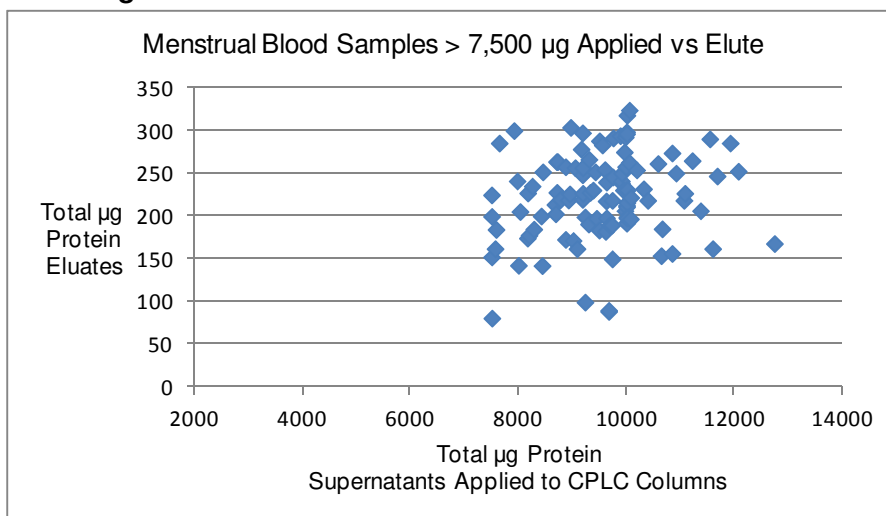


Figure 22 CPLC Protein Saturation for Menstrual Blood



Venous blood showed similar results (**Figures 23-25**) with a linear relationship between applied protein to eluted protein up to approximately 7,500 μg and CPLC column saturation at protein amounts greater than 7,500 μg .

Figure 23 Determining CPLC Column Saturation for Venous Blood

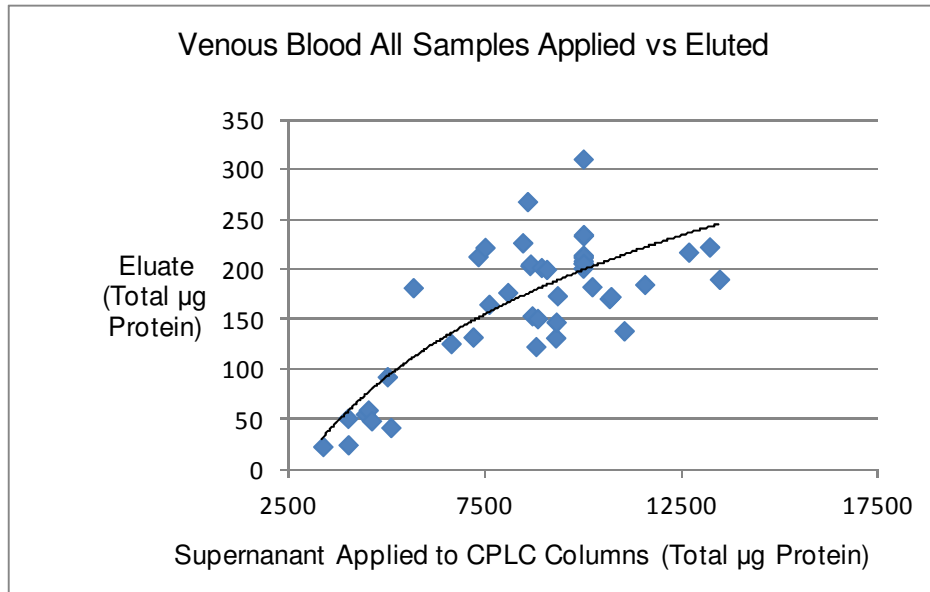


Figure 24 CPLC Linear Binding Range for Venous Blood

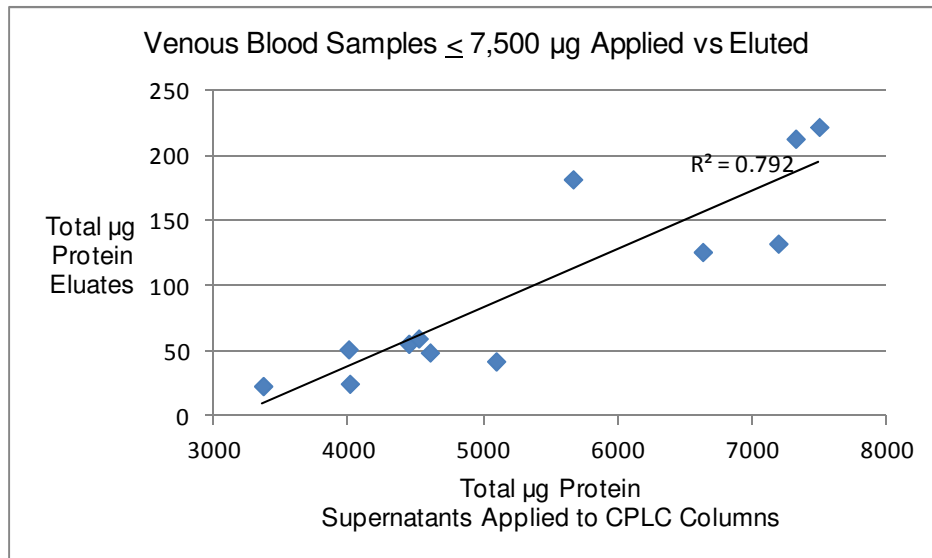
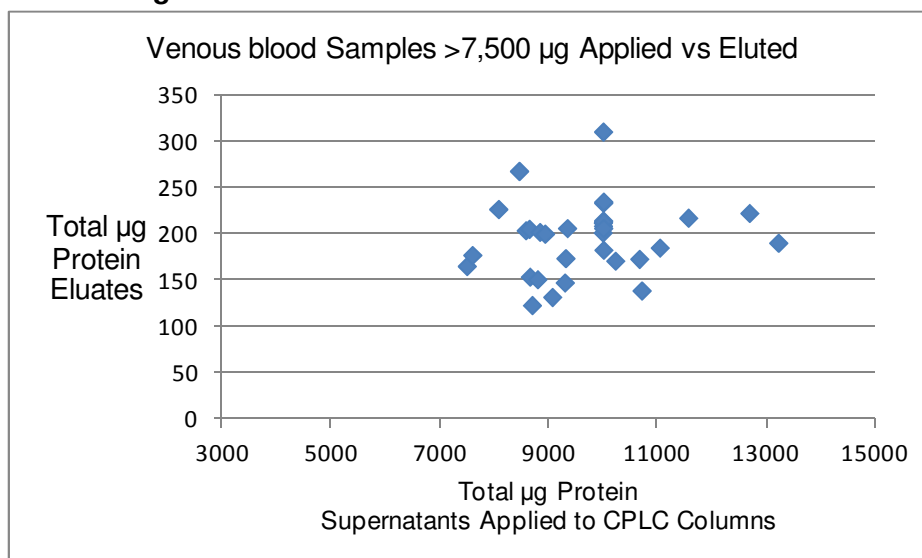


Figure 25 CPLC Protein Saturation for Venous Blood



Thus, assuming CPLC column saturation for menstrual blood and venous blood requires ~7,500 µg protein, 107 (55%) of 193 menstrual blood samples and 32 (71%) of 45 blood samples met saturation criteria (see Table 11). On average, approximately 2.3 percent of applied protein was recovered in CPLC elutes from samples where more than 7,500 µg protein were applied to the columns. Approximately 2.0 percent of applied protein was recovered from venous blood samples where $\geq 7,500$ µg protein were applied (Table 12).

Table 12 Statistical Analyses of Menstrual and Venous Blood Samples Meeting CPLC Saturation Criteria (>7,500 µg Applied Protein)

Body Fluid/ Fraction	Number of Samples >7,500 µg Protein	Range µg Protein	Mean µg Protein	Std Deviation	Median µg Protein
Menstrual Blood Applied	107	7,502 - 12,744	9,500	1,100	9,600
Menstrual Blood Eluted	107	81 - 234	220	50	230
Venous Blood Applied	32	7,595 - 13,470	9,900	1,400	10,000
Venous Blood Eluted	32	123 – 311	200	40	200

BODY FLUID MIXTURES WITH CPLC: Identifying mixtures containing more than one type of body fluid is a common challenge in forensic casework. We evaluated reduction of dynamic range using CPLC as a method for improving detection of specific protein markers for the minor

component of a body fluid mixture. We focused on blood, semen and saliva, three body fluids encountered commonly in forensic casework and for which we have previously established specific protein markers [3]. Because blood has a much higher concentration of protein than either semen or saliva (Table 13, [3]), detection of semen or saliva as the minor component of a mixture with blood poses a particular challenge.

Table 13 Amount of protein (μg) extracted per μl of body fluid using PBS with 0.5% NP-40 and 1.4 mm ceramic beads in a reciprocating homogenizer

Blood $\mu\text{g}/\mu\text{l}$ (Ratio:Blood)	Semen $\mu\text{g}/\mu\text{l}$ (Ratio:Blood)	Saliva $\mu\text{g}/\mu\text{l}$ (Ratio:Blood)
134 (1)	54 (0.40)	4 (0.03)

- **METHODS:** Semen and saliva were mixed with blood in different volume ratios (Table 14). Mixtures were homogenized and extracted in PBS containing 0.5% NP-40 as described in previous sections. Supernatant concentrations were measured by the BCA assay using BSA as the standard. As these samples were not in limited quantity, we were able to apply the manufacturer recommended 10 mg of protein from each body fluid mixture sample to CPLC columns. Application and elution of sample from CPLC columns was performed according to manufacturer's instruction, as described in previous sections using 2% SDS as the elution buffer. Concentration of eluted protein was measured by the bicinchoninic acid (BCA) assay using bovine serum albumin (BSA) as the standard. Qualitative analysis of all mixtures was performed by SDS microfluidic gel electrophoresis (SDS MiGE). Peptide separation by HPLC and MALDI TOF/TOF analysis were carried out on all mixtures as described in previous sections.

Table 14 Body Fluid Mixture Ratios by Volume & μg Protein

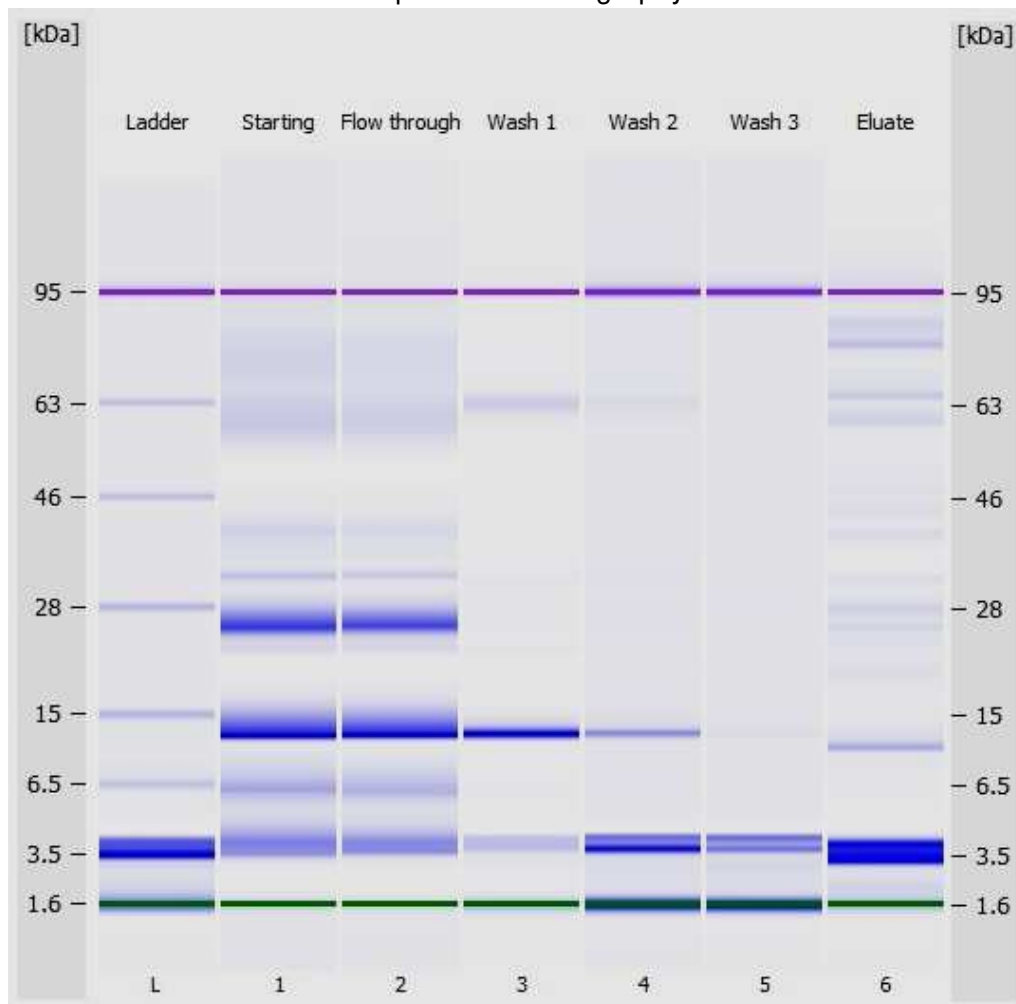
Mixture	Ratios Volume (μg Protein)				
saliva/blood	5:1 (1:7)	1:1 (1:33)	1:2 (1:67)	1:5 (1:167)	1:10 (1:333)
semen/blood	1:5 (1:13)	1:15 (1:38)	1:25 (1:63)	1:40 (1:100)	1:80 (1:200)

- **RESULTS:** The amount of protein eluted from the columns ranged from 155 – 259 μg , with a mean of 190 μg and standard deviation of 35 μg , which are consistent with the results obtained from menstrual and venous blood samples described above.

Figures 26 and 27 show a typical example of protein dynamic range reduction following CPLC in a saliva/blood mixture and a semen/blood mixture, as analyzed by SDS MiGE. The pseudo gel images show equal volumes of starting material (supernatant), pass-through, washes and eluate loaded onto a microfluidic gel chip with resolution between 5-80 kDa. The chip shows that the bulk of applied supernatant proteins do not bind to the CPLC column as evidenced by the similar band intensities of the loading material, pass-through and even the first wash. This confirms that the columns were overloaded and likely saturated as is necessary for protein dynamic range reduction. The most prominent band in these samples (Lanes 1-3) at ~15 kDa is hemoglobin (Hb). The eluate shows a dramatic reduction in dynamic range with the

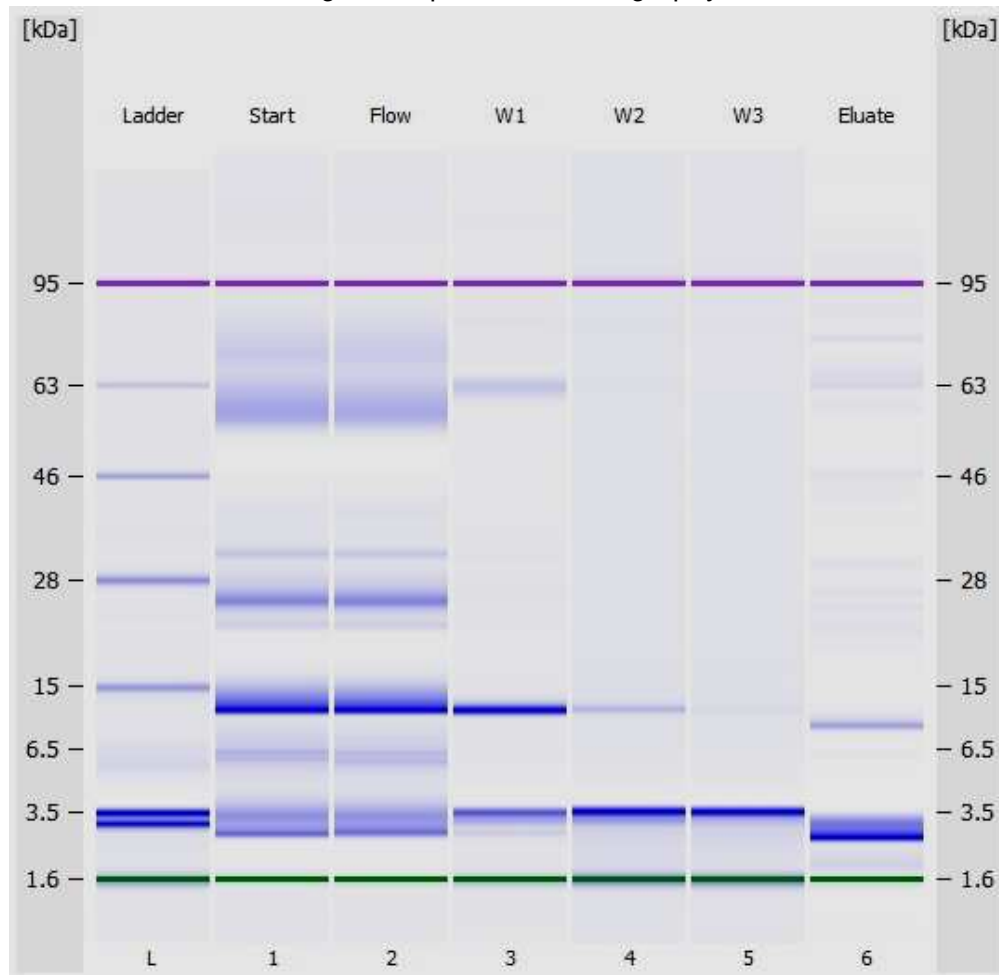
Hb dramatically reduced. Importantly, many protein bands that were not detectable in the starting material are now easily seen.

Figure 26 Protein Dynamic Range Reduction of a Saliva/Blood Mixture by Combinatorial Ligand Peptide Chromatography



Pseudo SDS-PAGE image of a 1:1 (volume) saliva: blood mixture evaluated by SDS microfluidic gel electrophoresis before and after protein dynamic range reduction. 10.3 mg of supernatant protein (Lane 1) were applied to the CPLC column; 184 μ g protein (~1.8%) were eluted (Lane 6).

Figure 27 Protein Dynamic Range Reduction of a Semen/Blood Mixture by Combinatorial Ligand Peptide Chromatography



Pseudo SDS-PAGE image of a 1:5 semen: blood mixture evaluated by SDS microfluidic gel electrophoresis before and after protein dynamic range reduction. 10.0 mg of supernatant protein (Lane 1) were applied to the CPLC column; 179 μ g protein (~1.8%) were eluted (Lane 6).

Results from HPLC-MALDI analysis of body fluid mixtures (Table 15) confirmed the qualitative assessment from MiGE gel images that reduction of protein dynamic range with CPLC allows more proteins to be detected in complex samples. Four to five times more proteins were identified in the CPLC eluate than in the original protein solution. CPLC is an efficient way to remove high abundance proteins and enrich low abundance proteins in complex biological samples.

Table 15 Total proteins identified before and after CPLC treatment in body fluid mixtures.

Mixture	Total Proteins Identified
---------	---------------------------

Volume (μg Protein)		Before CPLC	After CPLC
1:10 saliva:blood	(1:333)	25	166
1:5 saliva:blood	(1:167)	39	221
1:2 saliva:blood	(1:67)	42	226
1:1 saliva:blood	(1:33)	43	238
5:1 saliva:blood	(1:7)	38	109
1:5 semen:blood	(1:13)	56	143
1:15 semen:blood	(1:38)	26	175
1:25 semen:blood	(1:63)	47	123
1:40 semen:blood	(1:100)	41	134
1:80 semen:blood	(1:200)	36	182

However, for both saliva/blood and semen/blood mixtures no specific marker protein enrichment was observed after CPLC. In contrast, mass spectrometry revealed marker loss during the process. As shown in Table 16, the number of peptides identified for α -amylase 1, the most abundant saliva-specific marker protein, was reduced after CPLC. One possible reason is that α -amylase does not bind well to the CPLC column. This hypothesis was confirmed by HPLC-MALDI analysis which demonstrated α -amylase 1 in the flow-through (data not shown). No enrichment of specific marker proteins was seen in mixtures of semen and blood as well (Table 17). Detection of semenogelin 1 and 2 in a 1:5 semen:blood mixture was improved with CPLC treatment. However with smaller relative amounts of semen in blood, semenogelin 1 and 2 were detected at low levels without CPLC treatment and detection was not improved after CPLC treatment.

Table 16 Detection of α -amylase 1 in saliva:blood mixtures before and after CPLC

Saliva:Blood Ratio	Before CPLC			After CPLC		
	amylase detected	# of peptides	% peptide coverage	amylase detected	# of peptides	% peptide coverage
1:10	no	-	-	no	-	-
1:5	no	-	-	no	-	-
1:2	yes	2	36.6	no	-	-
1:1	yes*	1	16.8	yes*	0	4.3
5:1	yes	12	40.3	yes*	1	6.1

of peptides = Number of peptide detected with 95% confidence.

% peptide coverage = Percent of protein covered by detected peptides.

* Detected peptides are insufficient to differentiate α -amylase 1 (saliva specific) from α -amylase 2B (not saliva specific)

Table 17 Detection of semen-specific protein markers in semen:blood mixtures before and after CPLC

Semen/Blood Ratio	1:5	1:5	1:15	1:15	1:25	1:25	1:40	1:40	1:80	1:80
-------------------	-----	-----	------	------	------	------	------	------	------	------

CPLC (Before/After)	B	A	B	A	B	A	B	A	B	A
# proteins at 95% conf.	56	143	26	175	47	123	41	134	36	182
Semenogelin 1	31	73	16	6	6	4	8	3	2	1
Semenogelin 2	32	68	16	7	5	0	7	2	3	0
Prostatic acid phosphatase	1	10	0	1	0	0	0	0	0	0
Prostate specific antigen	2	5	0	0	0	0	0	0	0	0

of peptides = Number of peptide detected with 95% confidence.

CONCLUSIONS: These results indicate that while CPLC treatment is effective in reducing the dynamic range of complex biological samples and increasing the number of proteins identified, CPLC treatment of these particular mixtures does not improve detection of marker target proteins. For saliva in a mixture with blood, it appears that amylase is out-competed for binding on the CPLC column as it is found in the flow through fraction. Similarly, semenogelin does not appear in the CPLC eluates. Why these marker proteins cannot find high affinity CPLC aptamers to bind to may be due to similarities in their structure/function – e.g. both form large complexes with other proteins in their respective body fluids. Amylase, along with proline-rich proteins, statherin and histatin, has been shown to complex with highly glycosylated mucin 5 to form complexes $> 40 \times 10^6$ Da [4,4,5]. Semenogelins 1 and 2 (glycosylated) are cross-linked by cystine bond and transglutaminases [6,7]. These carbohydrate-coated multimer complexes may sequester amylase and semenogelins, preventing them from finding sufficient numbers high affinity aptamers to bind to and thus effectively reducing their enrichment. Additionally, the large macromolecular structures ($>40 \times 10^6$ Da) formed by markers amylase (saliva) and semenogelin (semen) with other glycoproteins would unlikely be able to resist passage through CPLC columns – as is seen with amylase in the flow-through, and suggest three approaches to improve body fluid mixture sensitivity: i) use of more powerful detergents in sample solubilization to help solubilize the macro complexes, ii) reduce cystine bonds with a reducing agent such as DTT – again to destabilize the large protein complexes, and iii) remove the large carbohydrate trees that help link protein complexes together by deglycosylation using of commercially available kits (e.g. EDEGLY and NDEGLY from Sigma-Aldrich). Clearly additional research will be required to identify mechanisms.

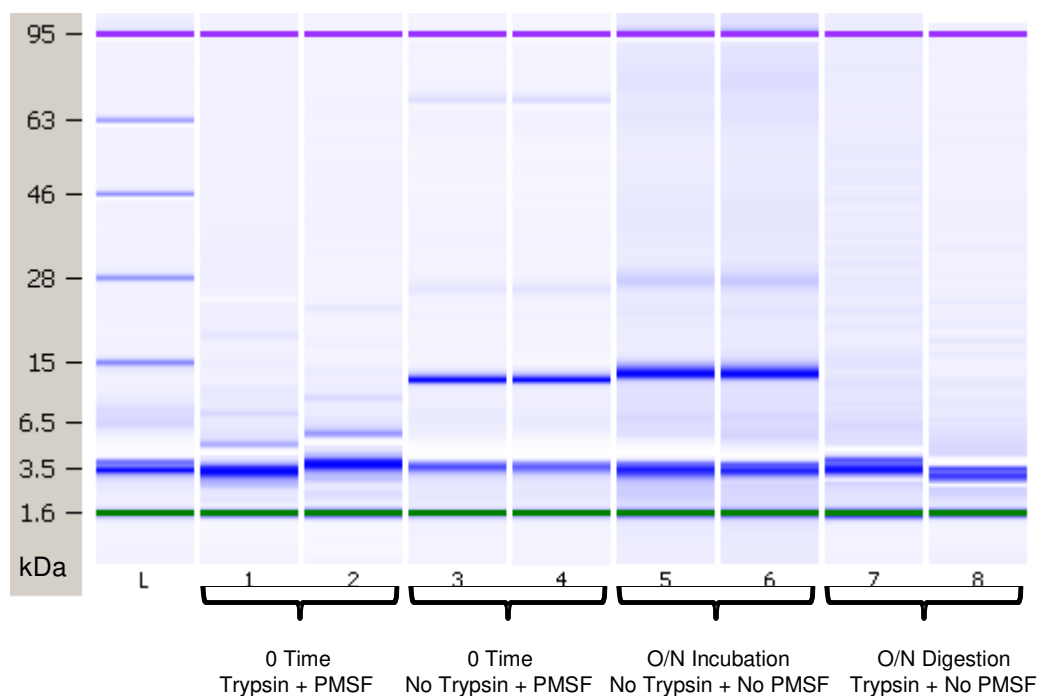
SECTION 2 REDUCING ASSAY TIME: ORGANIC SOLVENTS, TEMPERATURE & MICROWAVE-ASSISTED PROTEIN DIGESTION

As previously discussed one of the limiting aspects for a high throughput mass spectrometry assay is the current standard method of overnight trypsin digestion. The ability to substantially reduce trypsin incubation time could allow for single day sample processing and analysis, as mass spectrometry instrument time per sample occurs in micro-seconds. Here three parameters that can effect trypsin digestion were examined: i) inclusion of organic solvents, ii) incubation temperature, and iii) microwave radiation. All assays were performed on blood with digestion of hemoglobin and albumin, the two most abundant proteins in blood, used as markers for digestion.

EXPERIMENT 1 – ESTABLISHING THE ASSAY SYSTEM

• **EXPERIMENTAL DESIGN:** In order to evaluate the effects of organic solvents, microwaves, time and temperature on trypsin digestion efficiency, it is first necessary to identify a concentration of trypsin that would show progressive substrate digestion over time – i.e. no protein digestion at time zero and increasing digestion until the reaction was complete. While this may seem obvious, common mass spectrometry proteomics practice employs excessive trypsin (a 20:1 $\mu\text{g}/\mu\text{g}$ sample protein to trypsin ratio) to ensure maximum digestion – even if it requires overnight incubation. However, as shown in the microfluidic gel electrophoresis image (**MiGE**, 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA) **Figure 28**, at the 20:1 ratio, even in the presence of trypsin inhibitor phenylmethanesulfonyl fluoride (**PMSF**) and at 4°C, samples show partial digestion at time zero (lanes 1 & 2). Overnight digestion at 37°C, lanes 7 & 8, shows nearly complete protein digestion. (Note: The diminution of the ~71 kDa band in the sample incubated overnight without trypsin is likely due to endogenous proteases in the blood sample.)

Figure 28 Trypsin digestion of blood proteins at time zero and overnight with negative controls.

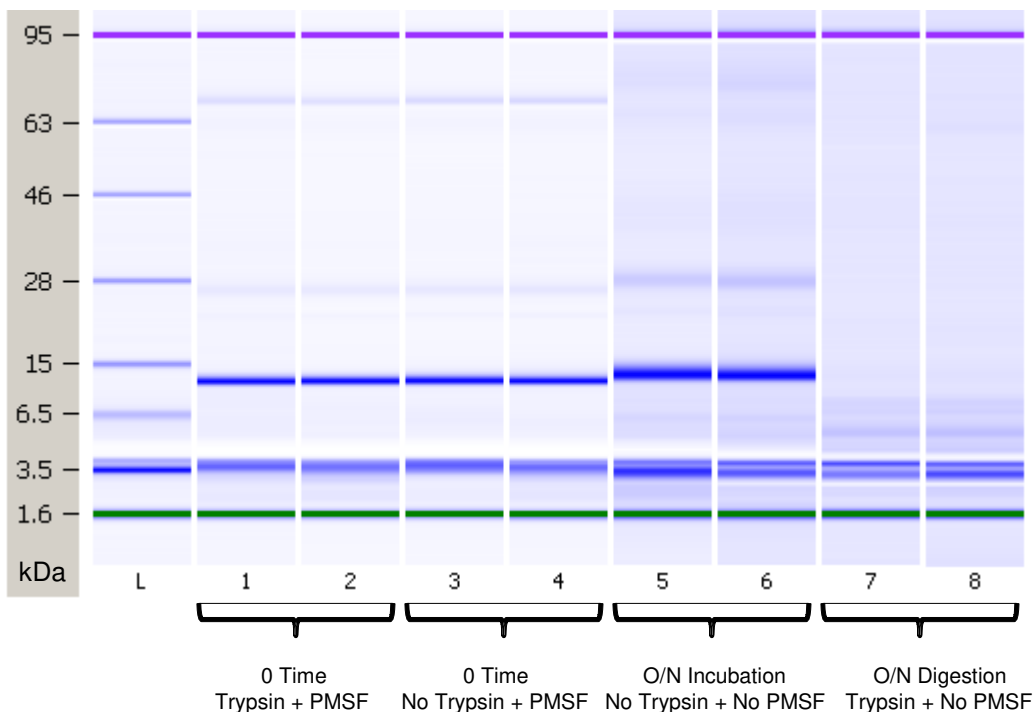


Equal amounts of blood supernatant protein incubated with and without trypsin at a 20:1 ratio ($\mu\text{g}/\mu\text{g}$) respectively at 37°C at 0 time and overnight. Undigested band at ~71 kDa is likely albumin, at ~12 kDa likely hemoglobin. L = ladder, O/N = overnight, PMSF is a trypsin inhibitor.

In order to evaluate the effects of microwaves and organics on trypsin digestion, serial dilutions of trypsin were made to identify a concentration of trypsin that showed no protein digestion at time zero with increasing digestion over time. **Figure 29** shows that when the amount of trypsin is reduced 100 fold (20 μg sample:10 ng trypsin), no digestion takes place at time zero (compare lanes 1 & 2 to lanes 3 & 4). Again, overnight digestion at 37°C, lanes 7 & 8, shows nearly complete protein digestion. (Again, the diminution of the ~71 kDa band in the

sample incubated overnight without trypsin is likely due to endogenous proteases in the sample.)

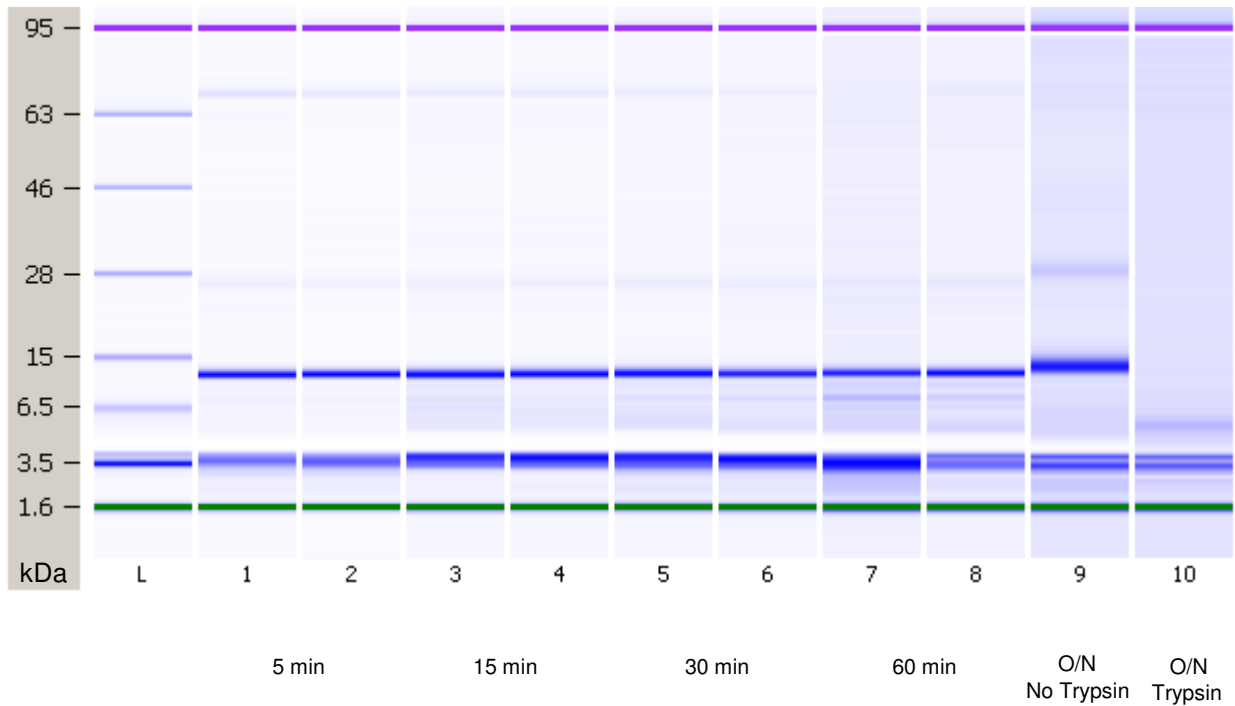
Figure 29 Trypsin digestion of blood proteins at time zero and overnight with negative controls at 100 X less trypsin



Equal amounts of blood supernatant protein incubated with (and without trypsin) at a 2,000:1 ratio ($\mu\text{g}/\mu\text{g}$) respectively at 37°C at 0 time and overnight. Undigested band at ~71 kDa is likely albumin, at ~12 kDa likely hemoglobin. L = ladder, O/N = overnight, PMSF is a trypsin inhibitor.

Figure 30 shows a MiGE image of increasing digestion of blood proteins at 5, 15, 30 and 60 minutes with 100 X diluted trypsin. The protein band at ~71 kDa (likely albumin (**Alb**)) shows diminishing intensity over time as does the band at ~12.5 kDa (hemoglobin (**Hb**)) until overnight digestion when both are nearly gone (Lane 10). Quantitation of the ~71 and ~12.5 kDa bands (time corrected area under the curve; average of duplicate samples) shows this more clearly in **Figures 31 and 32**. These data demonstrate that at a 2,000:1 ($\mu\text{g}:\mu\text{g}$) ratio of sample protein to trypsin, changes in protein digestion can be observed.

Figure 30 100 X diluted trypsin digestion of blood proteins over time.



Equal amounts of blood supernatant protein incubated with and without trypsin at a 2,000:1 ratio ($\mu\text{g}/\mu\text{g}$) respectively at 37°C for varying times and overnight. Protein bands at ~71 kDa (likely albumin) and 12 kDa (likely hemoglobin) shows diminishing intensity over digestion time. L = ladder, O/N = overnight.

Figure 31 Digestion Time Course ~71 kDa Band

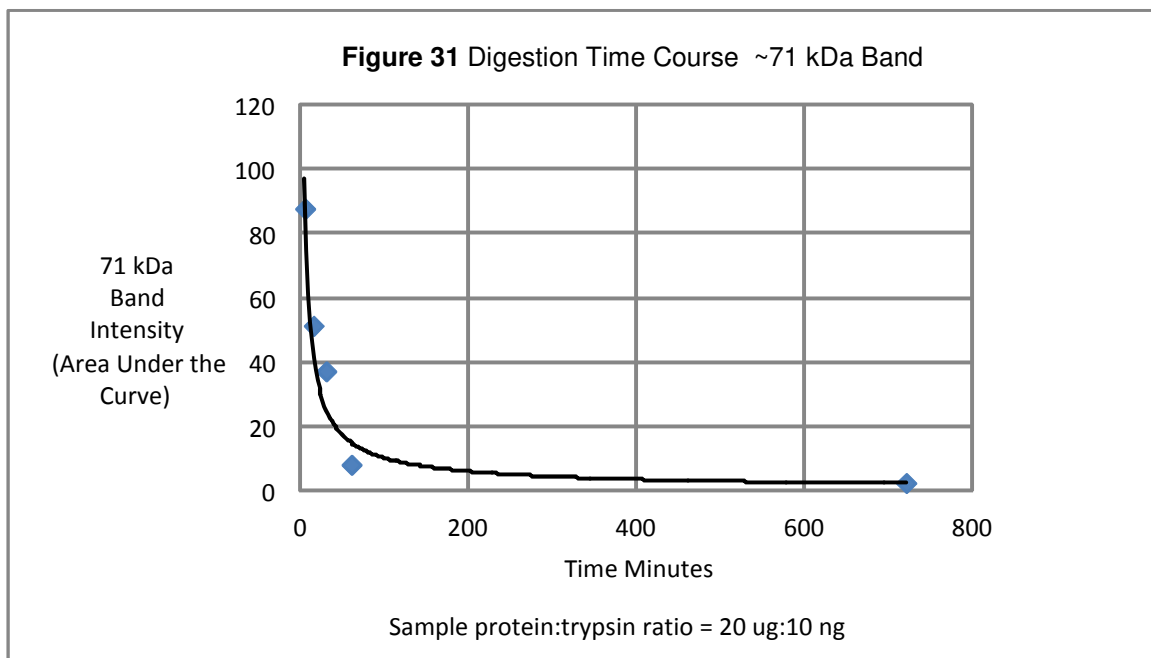
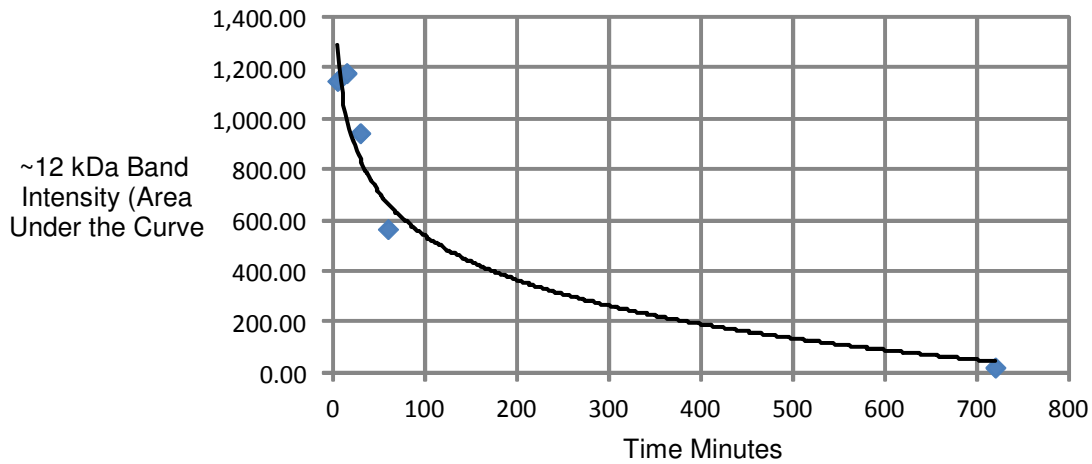


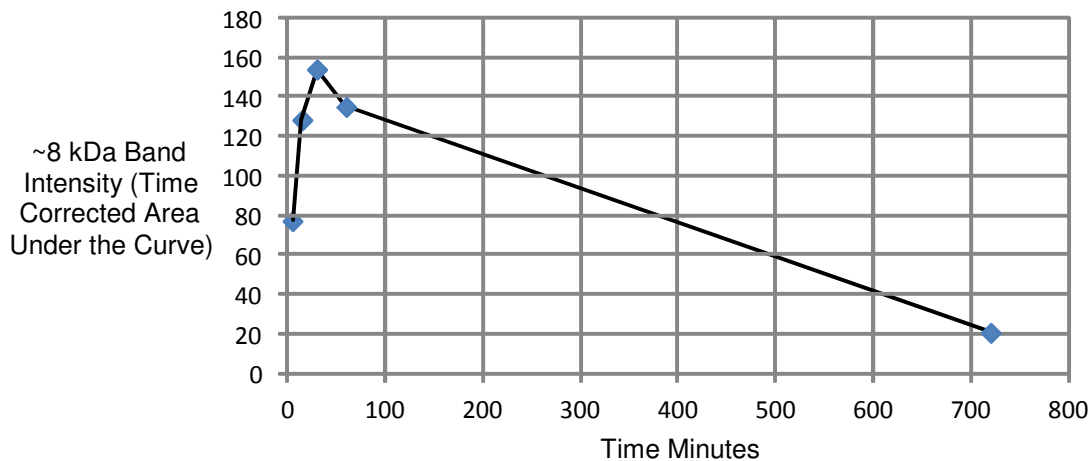
Figure 32 Digestion Time Course ~12 kDa Band



Sample protein:trypsin ratio = 20 ug:10 ng

N.B. It is important to note that during the course of digestion while the major observable blood proteins Alb (~71 kDa) and Hb (~12.5 kDa) decline with increasing time, other bands begin to appear but eventually disappear. These bands (see e.g. the ~8 kDa band at 5, 15, 30 and 60 min in **Figure 30**) are likely intermediate digestion products that are completely digest after overnight incubations (see lane 10 **Figure 30**). **Figure 33** shows the increase and decline of the ~8 kDa band quantitatively.

Figure 33 Digestion Time Course ~8 kDa Band



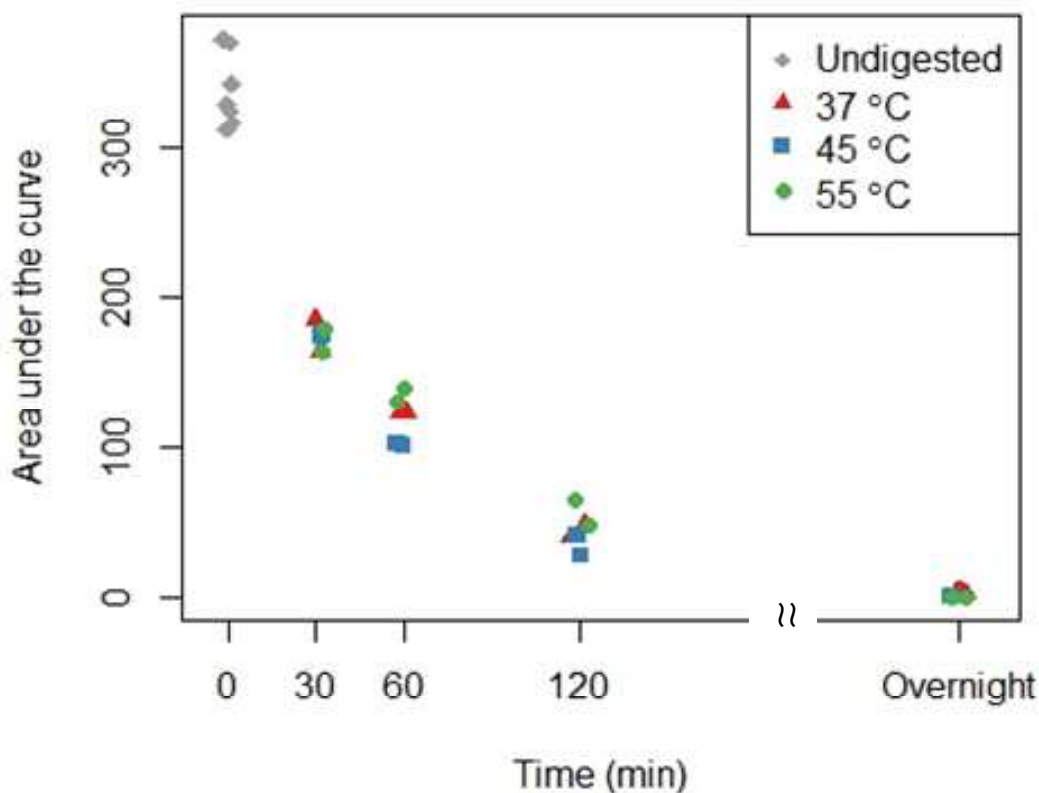
Sample protein:trypsin ratio = 20 ug:10 ng

In subsequent experiments we focus on the digestion of the 71 kDa (Alb) and 12.5 kDa (Hb) bands over relatively short time intervals to monitor the effects of microwaves, organics, etc. on trypsin activity. The appearance of intermediate digestion products (e.g. bands between 5 and 10 kDa) are not monitored as they are eventually fully digested. We chose *not* to perform these experiments on a standard protein, e.g. bovine serum albumin, as such pure proteins do not represent the types of complex protein mixtures we will encounter during casework.

EXPERIMENT 2 – EVALUATION OF TRYPSIN DIGESTION TEMPERATURE OVER TIME

To identify optimal digestion temperature, blood samples were incubated at the 2,000:1 ($\mu\text{g}/\mu\text{g}$) sample protein to trypsin ratio at 37°C, 45°C and 55°C and followed over time. Samples were run on MiGE and the 12.5 kDa Hb band quantified by time corrected area under the curve from resulting electropherograms. **Figure 34** shows the continuous digestion the Hb band at 30, 60, 120 min and overnight. Undigested control (gray diamonds) is a sample with trypsin at zero time quenched with PMSF on ice. Results show little difference of the effects of incubation temperature on digestion.

Figure 34 Comparison of Digestion of the 12/5 kDa Hb Band at 37°C, 45°C, and 55°C Over Time



EXPERIMENT 3 – EVALUATING THE EFFECTS OF ORGANIC SOLVENTS ON TRYPSIN DIGESTION

• **METHODS:** Blood samples were homogenized in a 1:10 (v/v) ratio with RIPA buffer (25mM Tris•HCl pH 7.6, 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, Pierce Rockford, IL). Samples were then centrifuged at 18,000 g for 30 min at 4°C. Supernatant protein was precipitated with a 1:6 (v/v) of acetone, respectively, and the pellet resuspended in 20 µl of 500 mM ammonium bicarbonate (**ABC**). Concentration of samples was measured by bicinchoninic acid assay (**BCA**) and Bradford Assay (Sigma). Reduction was performed using one part dithiothreitol (**DTT**) to ten parts of sample and incubated at 37°C for one hr. Sample was alkylated with a final concentration of 15mM iodoacetamide at room temperature for 30 minutes. A 2,000:1 (µg:µg) ratio of sample protein to trypsin was used. Positive controls were treated at the same sample:trypsin µg ratio but digested overnight in a 37°C incubator (current standard) unless otherwise indicated. Negative controls were treated the same as positive controls except an equal volume of 50 mM ABC was added in place of trypsin. Three µl of 100 mM PMSF was added to all reactions after microwave or incubation and samples immediately placed at -20°C. Samples were thawed and kept on ice until assayed. The extent of trypsin digestion was measured by MiGE electropherograms quantified by time corrected area under the curve (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA, using Agilent Protein 80 Kits).

Five different solvents were tested at varying concentrations to identify the optimal ratio of solvent to ABC for maximum digestion. Digestions were performed with a 2,000:1 (µg/µg) ratio of sample protein to trypsin in a 37°C incubator for 20 min. Table 18 shows the range of solvents tested and the percentages at which maximum digestion occurred (green shading). For acetonitrile 5% and 10% gave similar results. All experiments were performed with an ABC only (i.e. no solvent added) positive control, and an ABC only no trypsin negative control. Once optimal organic solvent concentrations were determined, direct comparison experiments between microwave and incubator were performed for each solvent (see below).

Table 18 Organic Solvents and Percent Range Tested

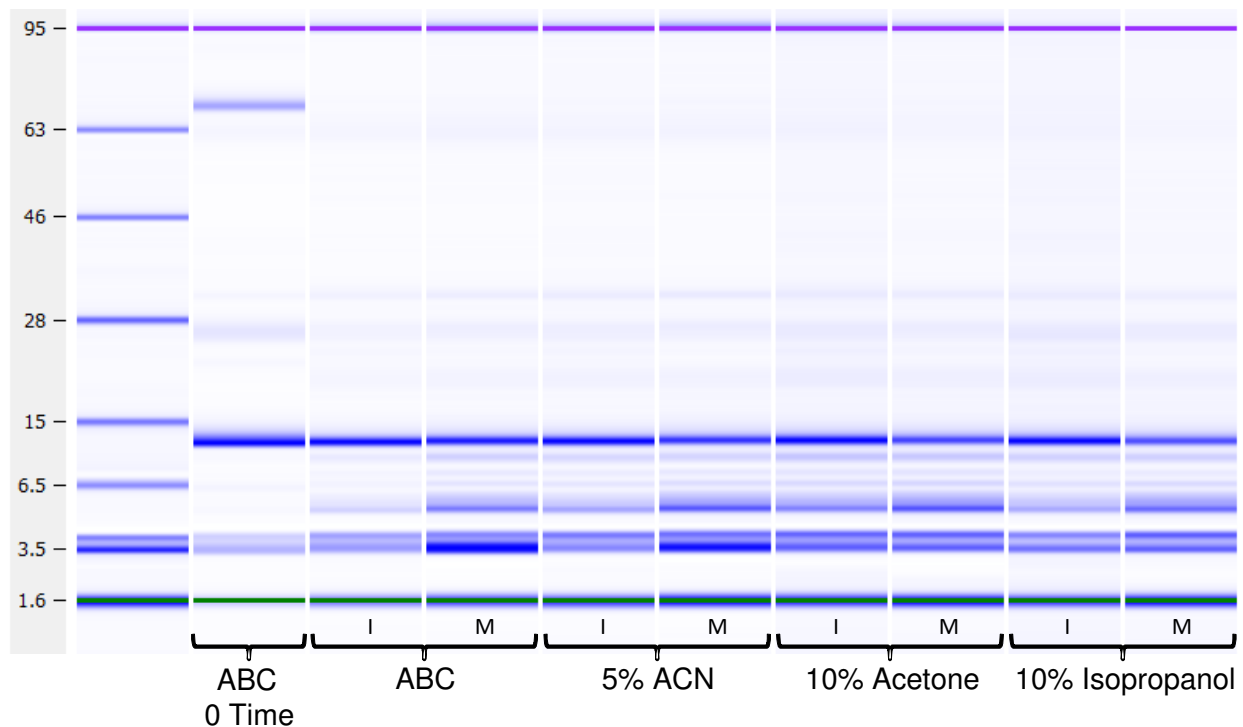
Organic	Percent in ABC						
Acetone	5	10	15	20	25	30	35
Acetonitrile	2	5	10	15	20	25	-
Dimethyl Sulfoxide	2	5	10	15	20	25	-
Isopropanol	2	5	10	15	-	-	-
Methanol	2	5	10	20	30	40	50

Digestion in a 37°C incubator for 20 min with a 2,000:1 (µg/µg) ratio of sample protein to trypsin. Green shadings are percentages that gave the best digestion results.

EXPERIMENT 4 – EVALUATING MICROWAVE-ASSISTED ENZYMATIC Digestion

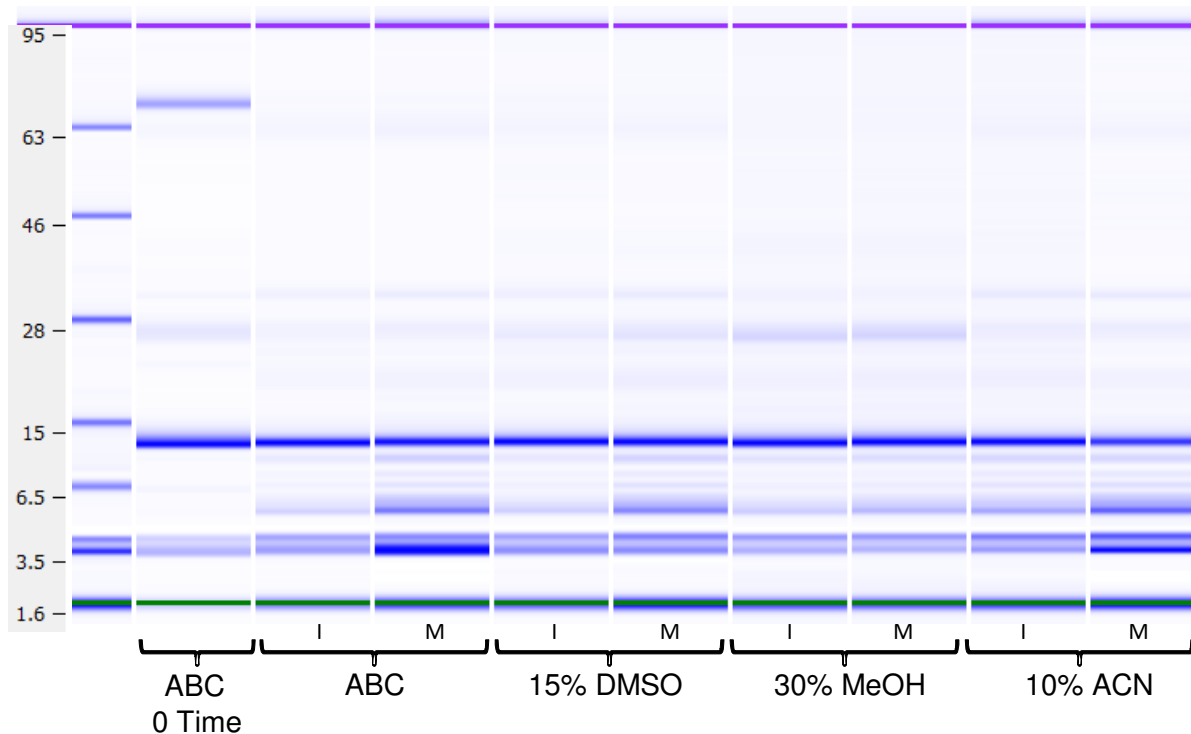
- *CEM MARS6 MICROWAVE*: Enzymatic digestions in the CEM MARS6 microwave are performed in 1.5 ml microfuge tubes. The tubes are held in a small closed container with 25 ml of water below the tubes that does not touch the tubes. This is to ensure that sample tubes are directly exposed microwaves. The 25 ml water bath below the tubes acts as a heat sink capturing microwave energy until the desired temperature is reached above the water bath in the closed container. Temperature is continuously monitored by a probe inserted into a small hole in the cap of a microfuge tube filled with the same volume of buffer as in the digestion tubes. The temperature monitoring tube is also above the water. As temperature and power are dependent variables, microwave input (watts) automatically turns on and off to maintain desired set temperature. Consequently, sample exposure to microwave radiation varies as temperature is maintained. Setting a constant temperature with different power inputs exposes the sample to greater microwave energy but for shorter amounts of time.
- *EXPERIMENTAL DESIGN*: A direct comparison of incubation vs microwave-assisted digestion was performed at the optimal solvent concentrations for the five organic described above (*Experiment 2*). Again, sample protein to trypsin ratio ($\mu\text{g}/\mu\text{g}$) was 2,000:1. Incubation was at 37°C; microwave temperature was held constant at 37°C at 50 watts. Digestion time was 20 minutes. Positive controls were performed with ABC alone (i.e. no solvent). Negative controls included the microwaving/incubating of samples in ABC and all solvents without trypsin for the same amount of time. Under these conditions, no change in the 12.5 kDa (Hb) band was observed in the negative controls (data not shown).
- *RESULTS*: As may be seen in **Figures 35 and 36**, microwave-assisted trypsin digested samples (M) appear to show less intense 12.5 kDa (Hb) bands (particularly in ACN, acetone and isopropanol) than non-microwave treated samples, suggesting increased digestion due to microwaves. Increases in the number and intensity of intermediate digestion products between 3.5 kDa and 10 kDa in the microwave treated samples also suggest improved digestion. This was confirmed by analyzing electropherogram time corrected area under the curve for Hb digestion in each sample.

Figure 35 MiGE Comparison of Incubator and Microwave-Assisted Trypsin Digestion of Blood in Different Organic Solvents at 37°C (50 Watts) for 20 Minutes



I = incubator, M = microwave, ABC = ammonium bicarbonate, ACN = acetonitrile, Iso = isopropanol.
Sample to trypsin ratio 2,000:1

Figure 36 MiGE Comparison of Incubator and Microwave-Assisted Trypsin Digestion of Blood in Different Organic Solvents at 37°C (50 Watts) for 20 Minutes

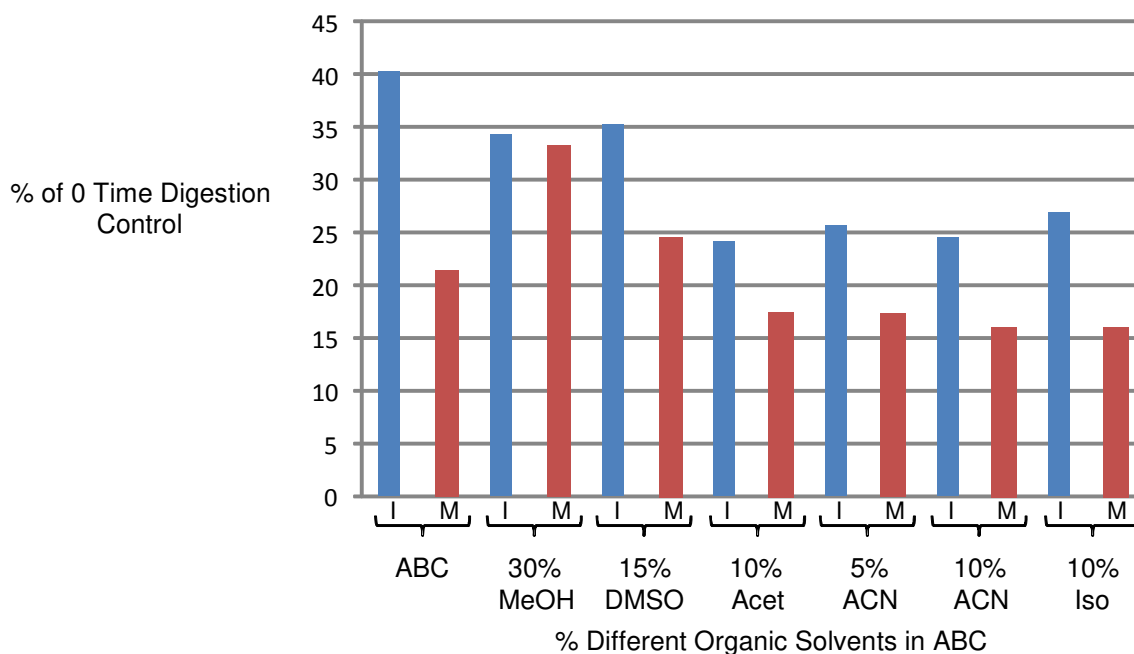


I= incubator, M = microwave, ABC = ammonium bicarbonate, DMSO = dimethyl sulfoxide, MeOH = Methanol, ACN = acetonitrile. Sample to trypsin ratio 2,000:1

Figure 37 shows the percent of the ~12.5 kDa band (Hb) remaining after 20 min digestion at 37°C compared to the zero time negative control (also containing trypsin) for all samples in each organic with and without microwave treatment. First, with respect to the effects of organic solvents on digestion of incubator samples alone (blue bars), all organics demonstrated some improvement with the data segregating into two main groups. The first, 30% MeOH and 15% DMSO reduced the Hb band to about 35% of the starting material (negative control), compared to ABC alone which reduced the Hb band to ~40% of starting material - about a 12.5% improvement. The second group of organics, 10% acetone, 5% ACN, 10% ACN and 10% isopropanol reduced the Hb band to about 25% of the ABC control - approximately a 37.5% improvement.

With respect to the effects of microwaves on trypsin digestion (red bars), with the exception of methanol, microwaves improved digestion efficiency in all cases. Most importantly in the ABC no organic control which showed an improvement of nearly 50% digestion. However, in the 10% acetone, 5% ACN, 10% ACN and 10% isopropanol samples, microwaves demonstrated an approximate 34% improvement compared to their non-microwave counterparts.

Figure 37 Comparison of Incubator and Microwave-Assisted Trypsin Digestion of ~12.5 kDa Band in Different Organic Solvents at 37°C for 20 Minutes

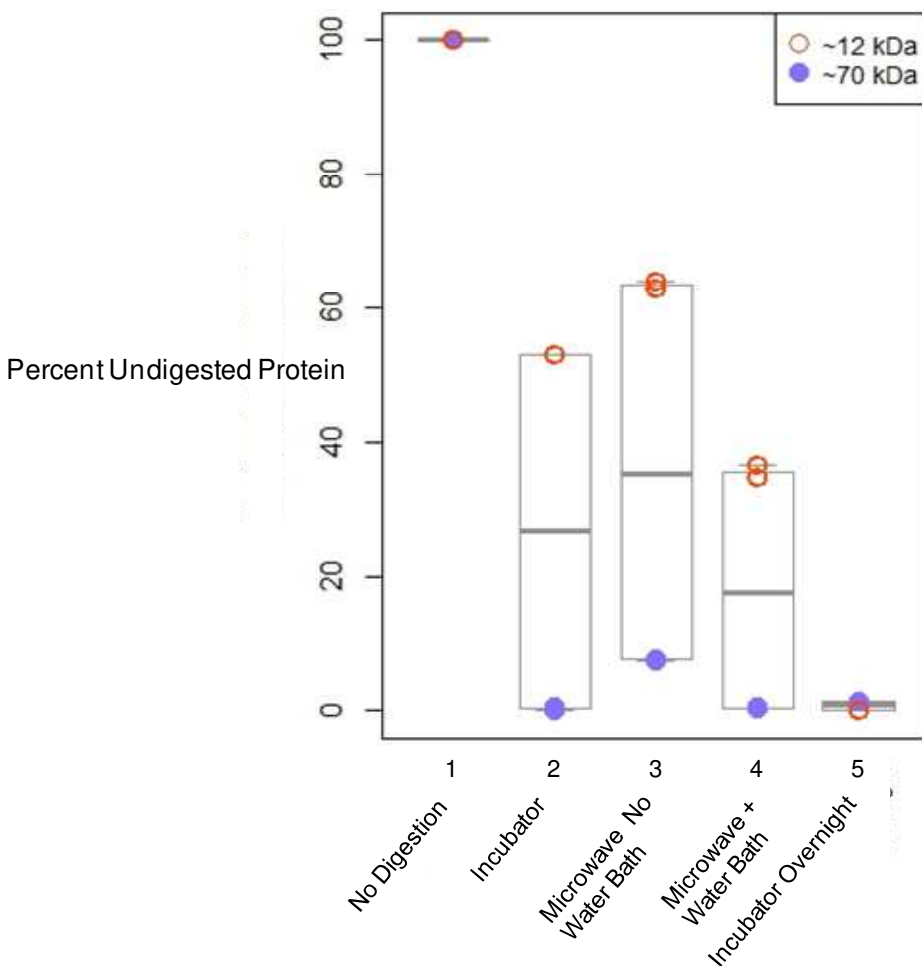


I= incubator, M = microwave, ABC = ammonium bicarbonate, MeOH = Methanol, DMSO = dimethyl sulfoxide, Acet= acetone, ACN = acetonitrile, Iso= isopropanol

EXPERIMENT 5 – EVALUATING MICROWAVE POWER & NEED FOR A WATER BATH

To determine if the level of microwave power has an effect on digestion rate, blood samples were again assayed at 37°C, with a sample protein to trypsin ratio of 2,000:1 but using microwaves at 850 watts. Lanes 2 (incubator) and 4 (microwave) of **Figure 38** demonstrate that the use of microwaves improves protein digestion in a manner similar to what was observed using 50 watts at 37°C (**Figure 37**). To determine the effect of microwaves sample digestion in the absence of the water bath (see *CEM MARS6 Microwave* discussion above) a similar sample was run without the water bath present. Lane 3 demonstrates that microwaving in the absence of a water bath is not as effective as the incubator alone. This is likely due to the instrument trying to maintain a constant 37°C temperature in a small digestion volume without the water “sink” which can modulate the temperature more evenly.

Figure 38 Microwave –Assisted Trypsin Digestion at 37°C and 850 Watts for 15 Minutes in the Presence and Absence of a Water Bath



EXPERIMENT 6 – COMBINED DIGESTION CONDITIONS AT STANDARD TRYPSIN CONCENTRATION

The ultimate goal of these experiments was to identify trypsin digestion conditions that can substantially reduce digestion time. Consequently, we returned to the standard ratio of sample protein to trypsin (i.e. 20 µg sample protein to 1 µg trypsin) and repeated the digestion. As may be seen in Table 19, nearly complete Hb and Alb digestion can be obtained at 37°C in 10% isopropanol within 30 min. Increasing temperature to 45°C or 55°C had little effect (as shown earlier) and digestion at 60 and 120 min made only a modest difference.

Table 19 % Protein Remaining at 20:1 (µg/µg)
 Sample Trypsin Ratio

Time	Temperature		
	37°C	45°C	55°C
30 min	0.58	0.42	0.32
1 hr	0.96	0.69	0.44
2 hrs	0.32	0.62	0.39
Overnight	0.09		

In conclusion, the combination of organic solvents and use of microwaves can significantly reduce the time for Hb and albumin digestion. Near overnight digestion levels can be achieved in 10% isopropanol at 37°C in approximately two hours using the 2,000:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio. Using a 20:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio under these same conditions, these same levels can be achieved in approximately 30 minutes.

SECTION 3 MENSTRUAL BLOOD AND VAGINAL FLUID MARKER STABILITY

One of the criteria for selecting a body fluid assay (DNA, RNA or protein) is the stability of its markers over time and their resistance to environment insults. Thus, we created a series of mock forensic samples of menstrual blood and vaginal fluid that were allowed to age and tested them over a period of 18 months. Our goal was to test the stability of menstrual blood and vaginal fluid marker proteins over time.

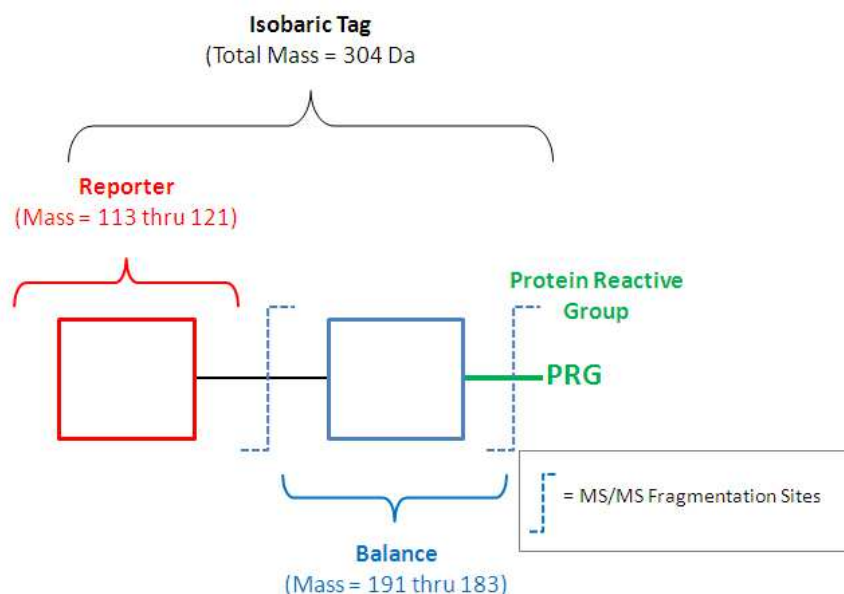
- **MATERIALS & METHODS:** Menstrual blood and vaginal fluid mock samples were prepared by depositing liquid samples from five volunteers onto cotton swabs. All menstrual blood samples were taken from day two of a women's period. To ensure sufficient protein would be available for CPLC treatment (which requires 10 mg of sample), 50 μl of each sample were spotted on cotton Q-tips and stored for 1 week, and 1, 3, 6, 12 and 18 months at room temperature protected from direct sunlight. When samples come of age, the head of the Q-tip was removed, placed in a 2 ml MP Biomedicals screw cap tube (Cleveland, OH) with premeasured amounts of Lysing Matrix D (1.4 mm ceramic beads) and 500 μl extraction buffer (0.5% NP-40 in **PBS**). Samples were homogenized on a FastPrep 24 (MP Biomedicals, Cleveland, OH) reciprocating homogenizer set to maximum speed (6 m/sec) for 40 sec. Samples were then spun at 12,000 g for 30 min at 4°C.

Menstrual blood samples were run on CPLC columns in order to reduce protein dynamic range and increase detection of lower abundance markers. The CPLC procedure is described in detail in Section 1. In brief, CPLC columns were washed twice with PBS, loaded with 10 mg of sample protein, and incubated for 2 hrs at room temperature. After washing with PBS and ddH₂O, protein was eluted with 2% SDS in PBS at 90°C for 15 min. Concentration of eluted protein was also measured by the BCA assay using BSA as the standard. Twenty micrograms of protein per sample were reduced and alkylated using 5 mM TCEP and 15 mM iodoacetamide. Samples were digested with 1 μg of trypsin overnight at 37°C.

In order to quantitatively assess marker protein change with age, isobaric tags were used to label different aged samples from the same volunteer. The isobaric barcode used (iTRAQ from

SCIEX, Framingham, MA) is schematically represented in **Figure 32**. It is composed of three functional groups: i) the *Protein Reactive Group* which allows the tag to covalently react with primary amines from N-terminal or lysine on each peptide, ii) the *Reporter* which is the actual barcode composed of eight different masses (113–119 Da plus 121 Da), and iii) the *Balance* which, in combination with the reporter always results in an isobaric mass of 304 Da for all iTRAQ labels (e.g. Reporter 113 + Balance 191, Reporter 114 + Balance 190, etc.). The importance of keeping the mass of the *Reporter* plus *Balance* constant (isobaric) is to ensure that the same marker peptide from different samples (e.g. samples from one individual aged for different period of time) migrates together in the mass spectrometer which separates and

Figure 39 Schematic Representation of iTRAQ Isobaric Tag



identifies peptides based on their mass. Subsequent fragmentation of the co-migrating peptides in the mass spec's collision chamber releases the *Reporters* (barcodes) which are then detected. A comparison of each barcode's peak height (intensity) shows the relative quantity of the marker peptide present in the different samples. The relative intensities of these ions are proportional to the amount of this peptide in each volunteer's sample. These relative peptide quantities are reported by Protein Pilot software (AB Sciex) as ratios relative to a labeled standard in the run. Protein ratios are obtained using a weighted average of the ratios of all the peptides contributing to the protein identification.

Twenty micrograms of digested peptides from each sample were individually labeled with isobaric tags following the manufacturer's instructions. Two micrograms of each sample of digested peptides were combined and separated by reverse phase C-18 nano-HPLC. HPLC runs were performed on Ultimate 3000 system (Dionex, CA). Tryptic peptides were desalted using an inline C18 trap column (300 μm x 5 mm, 5 μm , 100 \AA , Dionex). Separation was performed on a PepMap100 RP C18 column (75 μm x 250 mm; 3 μm , 100 \AA , Dionex) using a 5–45% linear ACN gradient for 3 hrs at flow rate of 300 nl/min (solvent A 2% acetonitrile, 0.1% TFA; and solvent B 98% acetonitrile, 0.1% TFA). A Dionex Probot was used to spot

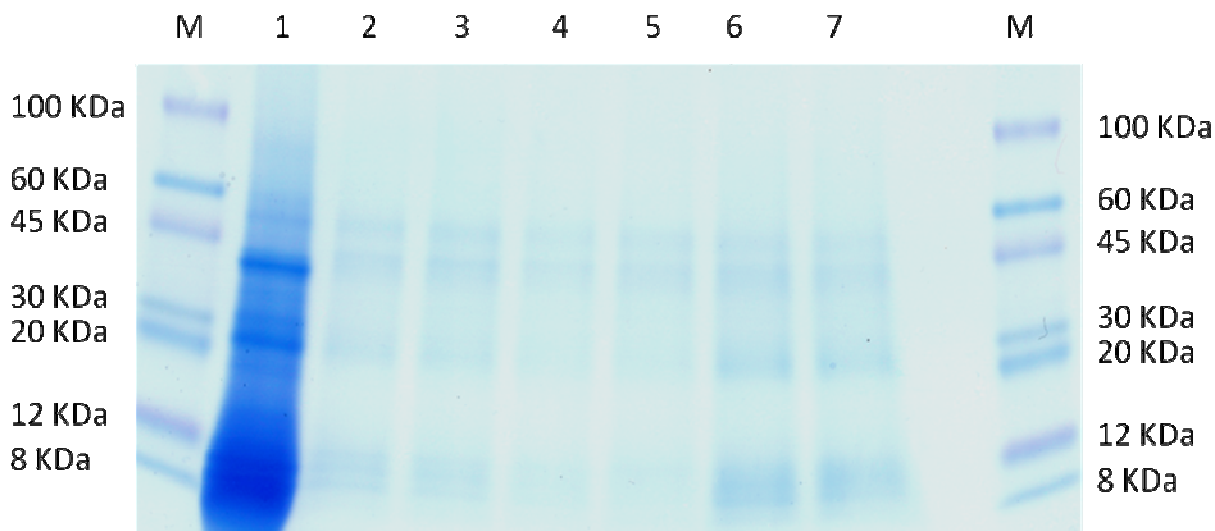
approximately 1 μ l (1:1 volumes eluate:matrix) on a 384 SCIEX MALDI plate (matrix: α -cyano-4-hydroxycinnamic acid at a concentration of 5 mg/ml in 60% ACN, 40% water, 0.1% TFA).

Peptides were analyzed on a Sciex 4800 MALDI TOF/TOF instrument. MS data was acquired at a laser repetition rate of 200 Hz with 600 laser shots/spectrum (50 laser shots/sub-spectrum). MS/MS data was acquired at 200 Hz in 1 kV MS/MS mode with 3200 laser shots/spectrum (40 laser shots/sub-spectrum) with the following TOF/TOF Series Explorer Stop Conditions: Maximum shots per spectrum 3200; minimum shots per spectrum 1000; number of MS/MS fragments 8; S/N of each fragment 75. Top 30 strongest peaks were selected to do MS/MS analysis. Peptides were identified and quantified via automated database searching on raw data using ProteinPilot 4.5 (SCIEX, CA) against Swiss-Prot Human protein database (40,670 Protein Entries, April 2009).

• RESULTS

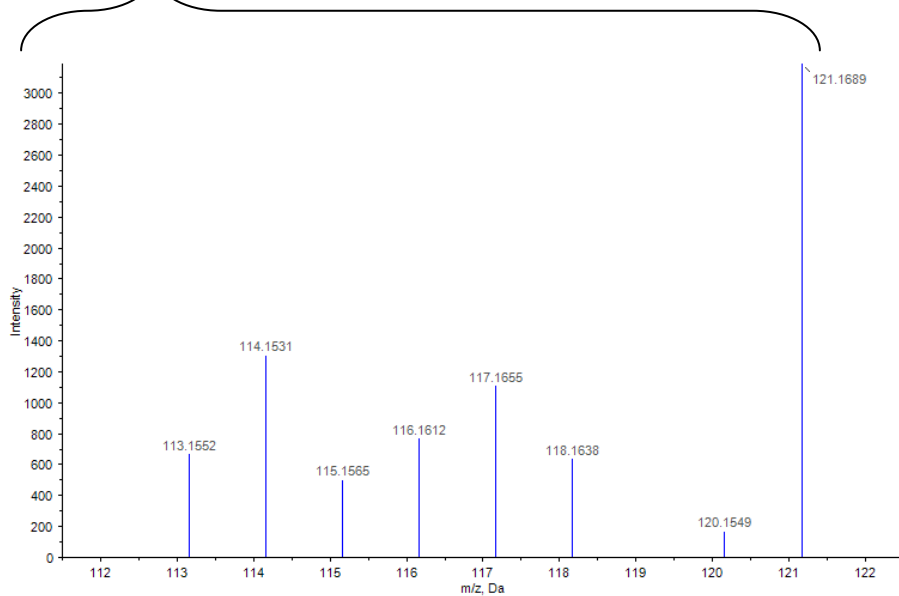
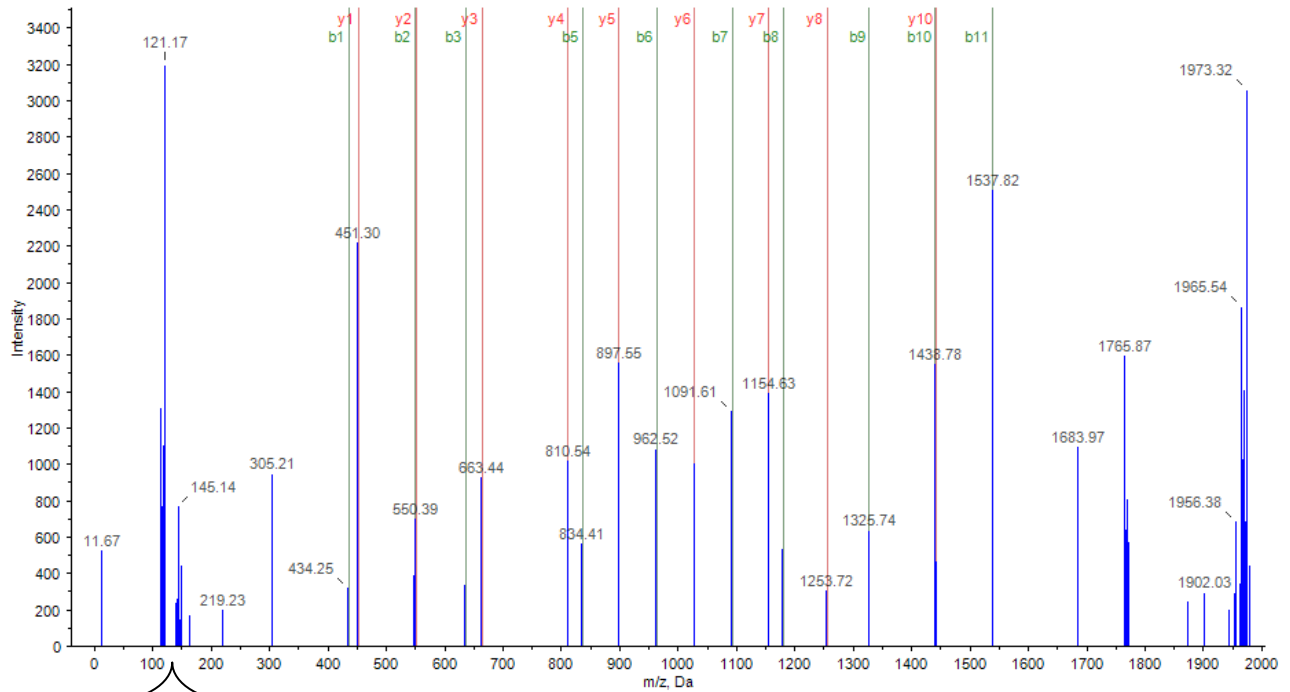
I) VAGINAL FLUID: SDS PAGE gels were used to observe protein profile changes of different aged vaginal fluid samples. **Figure 40** shows protein profiles from the same volume of vaginal fluid which corresponds to \sim 0.5 μ l of original vaginal fluid. Compared with collected liquid samples, proteins extracted from cotton swab were reduced (Lane 1 vs Lanes 2-7). This may be caused by protein loss during the extraction process from the cotton swab. However protein profiles from samples aged 1 week to 1.5 years (Lanes 2-7) were similar. This was confirmed by LC-MALDI results (**Figure 41**). Isobaric labeling allows for direct comparison of peptide signal intensity at each different time point compared with the original "fresh" sample. **Figure 41** shows the spectra of a commonly detected peptide from vaginal fluid marker alpha-2 macroglobulin-like protein (A2ML1). Reporter ions 113, 114, 115, 116, 117, and 118 identify 1 week, 1 month, 3 months, 6 months, 12 months, 18 months aged samples, respectively. Reporter ion 121 is fresh, liquid sample from the same individual. Little change is seen among ions intensities of these labels for marker protein A2ML1 for the dried samples from 1 week to 18 months. The data seen for A2ML1 are typical of most other proteins examined, demonstrating vaginal fluid protein stability over the 18 months tested. **Figure 42** shows the ratio of reporter ion intensity for 4 vaginal fluid marker proteins SPRR3, small proline-rich protein 3 (SPRR3); transmembrane protease serine 11B (TM11B), and transmembrane protease serine 11D (TM11D), as well as A2ML1, from dried, aged samples relative to the same fresh, liquid sample from the same individual. The ion intensities of these markers in dried, aged samples range from approximately a sixth to a third of the signal from the fresh sample, however signals from dried samples do not decrease appreciably over the course of 18 months.

Figure 40 SDS PAGE Gel analysis on different aged vaginal fluid samples.



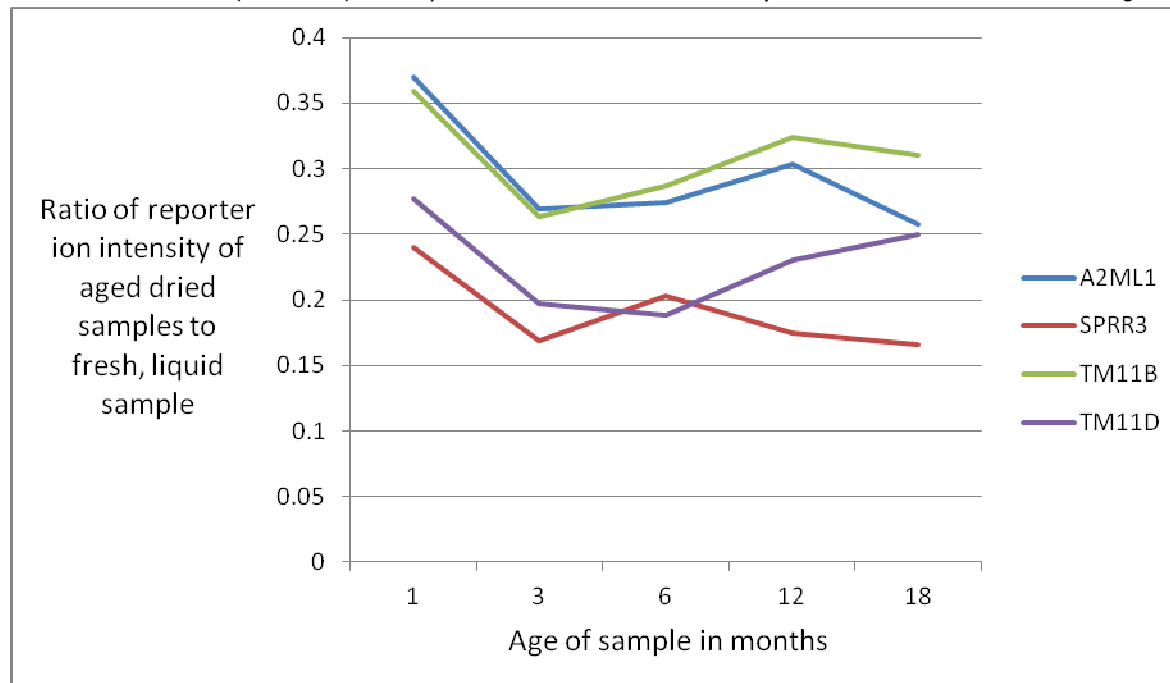
Lane M = Marker; Lane 1 = liquid vaginal fluid samples; Lane 2 = 1 week aged vaginal fluid; Lane 3 = 1 month aged vaginal fluid; Lane 4 = 3 months aged vaginal fluid; Lane 5 = 6 months aged vaginal fluid; Lane 6 = 1 year aged vaginal fluid; Lane 7 = 1.5 year aged vaginal fluid

Figure 41 MS/MS spectrum of vaginal fluid marker protein alpha-2 macroglobulin-like protein (A2ML1) peptide ELSTVQESFLVK from iTRAQ labeled aged vaginal fluid



Reporter ions 113, 114, 115, 116, 117, and 118 identify 1 week, 1 month, 3 months, 6 months, 12 months, 18 months aged samples, respectively. Reporter ion 121 is fresh, liquid sample from the same individual.

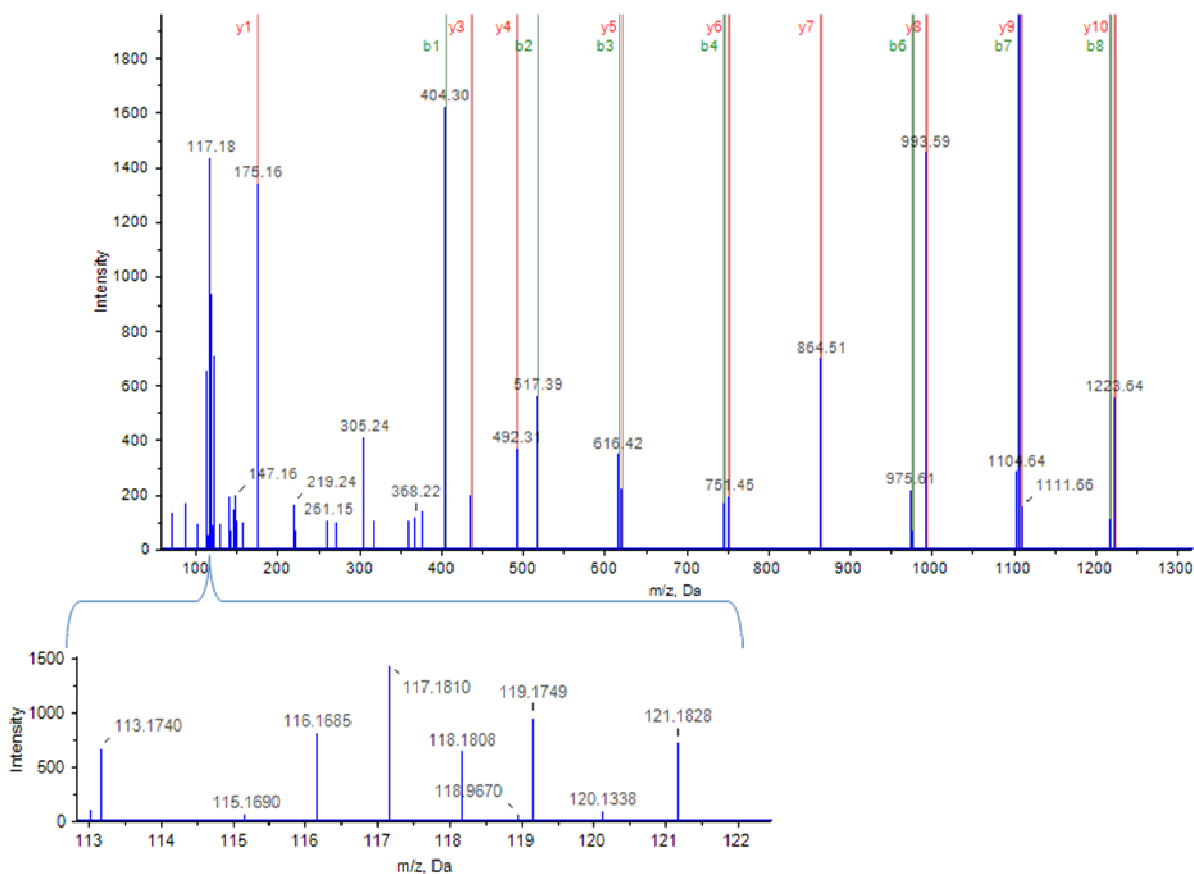
Figure 42 Ratio of reporter ion intensity of dried, aged samples to fresh, liquid sample from the same individual (533D19) for 4 protein markers that are specific to, or enriched in, vaginal fluid



A2ML1: alpha-2 macroglobulin like protein; SPRR3, small proline-rich protein 3; TM11B, transmembrane protease serine 11B; TM11D, transmembrane protease serine 11D

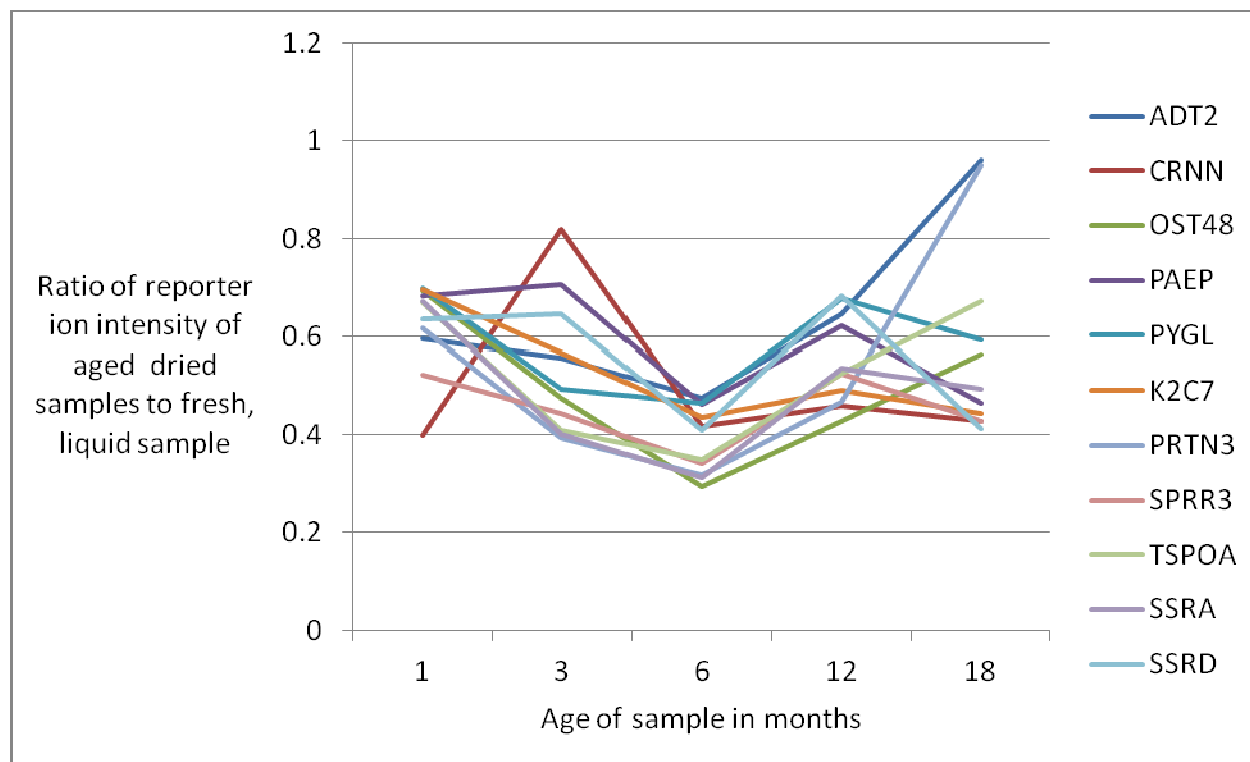
II) MENSTRUAL BLOOD: Due to the limited volume of samples eluted from CPLC columns, menstrual blood samples were not run on SDS PAGE gels. **Figures 43 and 44** show results of HPLC-MALDI analysis of isobarically tagged menstrual blood samples over the course of 18 months. **Figure 43** is the spectra of a commonly detected peptide from menstrual blood marker glycodelin (PAEP); reporter ions 113, 116, 117, 118, 119 and 121 identify 1 month, 3 months, 6 months, 12 months, and 18 months aged, and non-aged samples, respectively. Little change is seen among ions intensities of these labels for the glycodelin peptide, indicating stability over the 18 months tested. **Figure 44** shows the ratio of reporter ion intensity for 9 protein markers that are specific to or enriched in menstrual blood from dried, aged samples relative to the same fresh, liquid sample from the same individual. The ion intensities of these markers in dried, aged samples range from approximately a third to nearly 100% of the signal from the fresh sample, and signals from dried samples do not decrease over the course of 18 months.

Figure 43 MS/MS spectrum of glycodelin peptide VLVEDDEIMQGFIR from iTRAQ labeled aged menstrual blood.



Reporter ions 113, 116, 117, 118, 119 and 121 identify 1 month, 3 months, 6 months, 12 months, and 18 months aged, and non-aged samples, respectively.

Figure 44 Ratio of reporter ion intensity of dried, aged samples to fresh, liquid sample from the same individual (118D2) for 11 protein markers that are specific to, or enriched in, menstrual blood



ADT2, ADP/ATP translocase; CRNN, cornulin; OST48, dolichyl-diphosphooligosaccharide protein glycosyltransferase 48 kDa subunit; PAEP, glycodelin; PYGL, glycogen phosphorylase; K2C7, Keratin type II cytoskeletal 7; PRTN3, myeloblastin; SPRR3, small proline rich protein 3; TSPOA, translocator protein; SSRA, translocon-associated protein subunit alpha; SSRD, translocon-associated protein subunit delta

EFFECTS OF REAGENTS BLUESTAR AND LUMINOL ON VAGINAL FLUID MARKERS: Bluestar Forensic and luminol are common crime scene reagents used to identify bloodstains; both function at alkaline pH. There is a risk of DNA degradation with prolonged exposure to alkaline conditions. Our goal was to determine if Bluestar or luminol interfere with MALDI MS detection of vaginal fluid proteins over a period of 30 days following exposure.

In our previous study we demonstrated little effect of Bluestar and luminol on the stability of blood, saliva and semen marker proteins at 1, 9 and 29 days (NIJ grant 2008-DN-BX-K011). We initially planned to test the effect of Bluestar and luminol on both menstrual blood and vaginal fluid. However, detection of menstrual blood specific markers require CPLC treatment (see Section 1) to reduce protein dynamic range, which means at least 10 mg of menstrual blood protein (~50 μ l) is needed. Since menstrual blood samples did not have sufficient protein remaining, we performed tested the effects of Bluestar and luminol on vaginal fluid only.

• **MATERIALS & METHODS:** Vaginal fluid stains were made by applying 50 μ l of vaginal fluid to a cotton swab. Due to the limited amount of sample remaining, we used swabs that had been spotted with vaginal fluid more than two years prior to treatment with Bluestar or luminol. Ten micrograms of luminol or Bluestar were added to the stains. Stain swatches were further dried

and kept at room temperature for 1 day, 9 days and 29 days before extraction. Vaginal fluids stains made without luminol or Bluestar were positive controls. Clean cotton swabs spotted with 10 µl of luminol or Bluestar were used as negative controls.

Proteins were extracted from vaginal stains using 200 µl of 0.5% NP40 in PBS in a FastPrep24 reciprocating homogenizer using Lysing Matrix D and set to the maximum speed of 6 m/s for 30 sec. Samples were centrifuged and supernatants collected. After spinning at 18,000 g for 30 min, protein concentration was measured with BCA using BSA as standard.

Twenty micrograms of protein per sample were reduced and alkylated using 5 mM TCEP and 15 mM iodoacetamide. Samples were digested with 1 µg of trypsin overnight at 37°C. Two micrograms of digested peptides were separated by reverse phase C-18 nano-HPLC using a 5% - 45% linear gradient of acetonitrile over 3 hours. HPLC runs were performed on Ultimate 3000 system (Dionex, CA). Tryptic peptides were desalted using an inline C18 trap column (300 µm x 5 mm, 5 µm, 100 Å, Dionex). Separation was performed on a PepMap100 RP C18 column (75 µm x 250 mm; 3 µm, 100 Å, Dionex) using a 5–45% linear ACN gradient for 160 min at flow rate of 300 nl/min (solvent A 2% acetonitrile, 0.1% TFA; and solvent B 98% acetonitrile, 0.1% TFA). A Dionex Probot was used to spot approximately 1 µl (1:1 volumes eluate:matrix) on a 384 SCIEX MALDI plate (matrix: α-cyano-4-hydroxycinnamic acid at a concentration of 5 mg/ml in 60% ACN, 40% water, 0.1% TFA).

Peptides were analyzed on a Sciex 4800 MALDI TOF/TOF instrument. MS data was acquired at a laser repetition rate of 200 Hz with 600 laser shots/spectrum (50 laser shots/sub-spectrum). MS/MS data was acquired at 200 Hz in 1 kV MS/MS mode with 3200 laser shots/spectrum (40 laser shots/sub-spectrum) with the following TOF/TOF Series Explorer Stop Conditions: Maximum shots per spectrum 3200; minimum shots per spectrum 1000; number of MS/MS fragments 8; S/N of each fragment 75. Top 30 strongest peaks were selected to do MS/MS analysis. Peptides were identified and quantified via automated database searching on raw data using ProteinPilot 4.5 (SCIEX, CA) against Swiss-Prot Human protein database (40,670 Protein Entries, April 2009).

- **RESULTS:** Due to sample limitations, very old vaginal fluid mock samples (27 months) were used for this study. Compared with our previous study (NIJ Grant DN-BX-K192), the number of proteins identified in these samples was low. However, we were still able to detect at least one, and in most cases more than one vaginal fluid specific marker in all treated samples and positive controls. Treatment of dried vaginal fluid stains with Bluestar or luminol did not affect the overall number of proteins detected (Table 20) or the number of specific markers detected (Table 21). This is consistent with our previous study which showed Bluestar and luminol had no effect on blood, saliva and semen marker protein detection by LC-MALDI.

Table 20 Number of proteins detected in extract from vaginal fluid stains.

Sample	Day	Treatment	# proteins at 95% conf.
P1	1	none	66
P2	1	none	65
B1	1	Blue Star	42
B2	1	Blue Star	34

L1	1	Luminol	65
L2	1	Luminol	38
P3	9	none	65
B3	9	Blue Star	75
B4	9	Blue Star	47
L3	9	Luminol	10
L4	9	Luminol	61
B5	29	Blue Star	56
B6	29	Blue Star	49
L5	29	Luminol	44
L6	29	Luminol	45

Table 21 Vaginal fluid-specific protein markers detected in extract from vaginal fluid stains.

Sample Day Treatment	P1	P2	B1	B2	L1	L2	P3	B3	B4	L3	L4	B5	B6	L5	L6
	1	1	1	1	1	1	9	9	9	9	9	29	29	29	29
	none	none	BS	BS	L	L	none	BS	BS	L	L	BS	BS	L	L
Marker Protein															
SPRR3	8	7	12	3	11	8	14	25	13	1	19	16	15	16	14
A2ML1	2	5	12	9	5	1	26	24	5	0	9	16	15	6	3
TMPRSS11D	0	2	4	0	4	0	0	0	0	0	10	5	4	4	0
TMPRSS11B	0	0	0	0	0	0	0	0	2	0	1	4	2	0	0

SPRR3, small proline-rich protein 3; A2ML1: alpha-2 macroglobulin like protein; TMPRSS11D, transmembrane protease, serine 11D; TMPRSS11B, transmembrane protease, serine 11B; BS, Bluestar; L, luminol

THE VAGINAL FLUID MICROBIOME: In order to identify microbial proteins 96 samples of vaginal fluid were taken from 48 women at days 13 and 19 during the menstrual cycle (~ one week prior to and one week following menses). Samples were extracted, digested with trypsin and processed by HPLC-MALDI TOF/TOF as described above.

Identified proteins were searched against all microbial protein database (Swiss-Prot, 2013). Proteins from *Lactobacillus*, which have been shown to be present in the human vagina tract [1,2] were found in all samples. The two most abundant *Lactobacillus* proteins identified were glyceraldehyde-3-phosphate dehydrogenase (**GAPDH**) and enolase (**ENO**). GAPDH were detected in 46 women at day 13 (96%) and 45 women at day 19 (94%); ENO was detected in 36 women at day 13 (75%) and 38 women day 19 (79%). Other proteins identified from *Lactobacillus* include glucose-6-phosphate isomerase, triosephosphate isomerase, D-lactate dehydrogenase and phosphate acyltransferase. While these were not detected in as many samples, detection of any *Lactobacillus* proteins is sufficient to confirm its presence in a sample.

- **CONCLUSION:** Menstrual blood and vaginal fluid protein markers are stable under the conditions tested for at least 18 months – the end of the trial period. Bluestar and luminol have no noteworthy effect on protein stability for the 30 days over which these samples were tested.

Finally, we confirmed that many *Lactobacillus* proteins that have been used to identify the vaginal fluid from the vaginal fluid microbiome can also be detected by MS of vaginal fluid samples. However, we believe that the identification of human specific vaginal fluid marker proteins are a more definitive test for the presence of vaginal fluid.

SECTION 4 SIMULTANEOUS PROTEIN & DNA EXTRACTION & PURIFICATION– [Part of this section was presented by Kaylee Hershfeld in partial fulfillment of her master’s program at Towson University, Towson, MD.]

When limiting amounts biological samples are available, there may be insufficient material for both DNA and protein extractions and analyses. Here we evaluate the simultaneous extraction of DNA and proteins from a single sample to determine the feasibility of concurrent STR and body fluid identifications.

- **EXPERIMENTAL DESIGN - OVERVIEW:** The AllPrep DNA/RNA/Protein purification kit from Qiagen (Valencia, CA) was used for simultaneous protein and DNA extractions from neat and diluted samples of blood and saliva. To compare protein and DNA recoveries from the AllPrep kit to protein and DNA recoveries from single method extractions, results from the AllPrep kit were compared to methods currently used in the lab: for protein the routine OCME method and that of Mann et al. [8] and for DNA, DNA IQ method (see details below).
- **MATERIALS & METHODS:** All the samples used in this study consisted of human blood and saliva obtained by a single volunteer for each body fluid. Comparison of methods was done using neat, 10x and 100x dilutions of the blood and saliva samples. Each of the methods was done three times per type of sample (i.e. neat, 10x, 100x) for a total of 18 samples per body fluid. Samples were diluted to 10x and 100x using deionized water.

I) ALLPREP DNA/RNA/PROTEIN MINI KIT METHOD: AllPrep chemistry works by solubilizing proteins and nucleic acids in a single lysis buffer. Samples are loaded onto a micro-spin filter to which DNA binds and is subsequently washed and eluted. The flow through contains proteins and RNA. RNA may be isolated by the addition of ethanol and binding to a second micro-spin filter. Here protein would be in the final flow through. However, as RNA isolation is unnecessary, the protocol was modified (see below) so that the original lysate flow through was directly processed for protein.

Thirty ul aliquots of all blood samples (neat and diluted) were extracted using 670 ul of lysis buffer (Qiagen **RLT** buffer with β -mercaptoethanol). A 200 ul aliquot of all saliva samples were extracted with 500 ul of lysis buffer. Samples were homogenized at room temperature using FastPrep-24 reciprocating homogenizer (MP Biomedicals, Cleveland, OH) for 40 seconds at speed 6 m/s using Lysing Matrix D (1.4mm ceramic beads). The homogenized lysate was transferred to an AllPrep DNA spin column where DNA and RNA collection were carried out via the Qiagen protocol. After RNA collection, Qiagen APP buffer was added to the remaining solution to precipitate proteins. The protein pellet was re-solubilized in guanidine hydrochloride (GnHCl). An aliquot of this solution was used for protein quantitative (BCA) and qualitative analyses (SDS MiGE).

On-filter trypsin digestion of samples was performed using the filter-assisted sample preparation (FASP) method of Winiewski et al. [8] as described below (II. FASP METHOD). Samples were loaded onto filters as follow: 20 µg of blood protein (neat and 10x diluted) were applied to 10 kDa MWCO micro-spin filters (Pierce Protein Concentrators PES, ThermoFisher, Rockford, IL). Because of the high protein concentration in these samples sufficient water was added to ensure that the filter was completely covered. The 100 x diluted blood samples and all saliva samples had sufficient volume to cover the filters. Filter units were centrifuged at 15,000 x g for 2-10 min (depending on volume to avoid dryness) at room temperature. Sample alkylation and trypsin digestion are described below.

DNA elution from the DNA spin filter was carried out in accordance with the Qiagen AllPrep protocol and then stored at -20°C until use.

II) FASP METHOD: All blood samples (45 µl aliquots) were extracted with 455 µl of lysis buffer (4% SDS in 50 mM Tris/HCl pH 7.0) All saliva samples (300 µl aliquot) were extracted with 300 µl of lysis buffer. Samples were homogenized at room temperature using FastPrep-24 for 40 seconds at a speed of 6 m/s using Lysing Matrix D (MP Biomedical). One hundred µl aliquots of whole homogenates were saved for protein analysis. The remaining samples were incubated at 70°C for 5 minutes and spun at 16,000 x g for 5 minutes at 20°C. One hundred µl of resulting supernatant was taken for protein concentration measurement and MiGE qualitative analysis. Remaining supernatant was transferred to a clean tube and brought to a final DTT concentration of 5 mM.

To remove SDS from samples (for downstream MALDI processing) 6 µl of neat blood lysate and 200 µl of 8M urea in 0.1 M Tris/HCl pH 8.5 (prepared fresh daily) were mixed and added to the filter units. For the diluted blood (200 µl) and all saliva samples (250 µl) no urea buffer was added. Filter units were centrifuged at 15,000 x g for 7-8 minutes (not to dryness) at 20°C and the flow through discarded. Filters were washed twice with 200 µl of urea buffer.

Samples were alkylated on the filters by adding 100 µl of 50 mM iodoacetamide in 8 M urea, 0.1 M Tris/HCl pH 8.5, mixing at 600 rpm for 1 min and incubating in the dark for 20 min at RT. Filters were centrifuged at 15,000 x g for 7-8 min (not to dryness) at RT and flow through discarded. Filters were then washed two additional times with 8 M urea buffer and three times with 50 mM ammonium bicarbonate (ABC).

Samples were digested on the filters by adding 40 µl of ABC with a sample protein:trypsin ratio of 20:1 (µg/µg). Filter units were capped, mixed at 600 rpm for 1 min and incubated overnight at 37°C. Filter units were centrifuged units at 15,000 x g for 10 min. An additional 40 µl ABC was added and filters recentrifuged. Samples were acidified with 5% formic acid, dried down in a SpeedVac and stored at -20°C until use.

III) NYC OCME METHOD: All blood samples (45 µl) were extracted with 455 µl of 8M urea. Three hundred µl of neat and diluted saliva samples were extracted with 300 µl and 600 µl of 8M urea, respectively. Samples were homogenized at room temperature using the FastPrep-24 for 40 seconds at speed of 6m/s using Lysing Matrix D. One hundred µl aliquots of homogenized samples were saved for protein analysis. The remainder of the samples were transferred to a clean microfuge tube and spun at 18,000 x g for 30 min at 4°C. Protein concentrations of resulting supernatants were assayed.

Prior to digestion, supernatants were reduced with 5 mM TCEP and alkylated with 15 mM IAA; 15 mM DTT was added after alkylation to stop the activity of the IAA. For digestion, a 20:1 ratio of sample protein to trypsin was added and samples were digested overnight at 37°C. Samples were acidified using 5% formic acid, dried down in a SpeedVac and stored at -20°C.

- RESULTS:

I) QUANTITATIVE AND QUALITATIVE PROTEIN ANALYSIS: Quantitative protein analysis was performed by BCA assay using BSA as standard.

Figures 45 and 46 show the amount of protein recovered at each dilution by each method for blood and saliva respectively. The AllPrep neat blood protein concentrations ranged from 48.2-74.2 µg/µl, 10x diluted blood from 3.5-4.0ug/µl, and 100x diluted blood from 0.10-0.15 µg/µl. FASP neat blood ranged from 114.1-132.2 µg/µl, 10 x diluted blood from 8.7-10.5 µg/µl, and 100x diluted blood from 0.99-2.4ug/µl. NYC OCME neat blood ranged from 132.3-140.6 µg/µl, 10x diluted blood from 12.1- 13.4 µg/µl, and 100x diluted blood from 1.3-1.6 µg/µl. Protein concentrations of neat and diluted saliva sample processed with AllPrep were too low to be accurately measured by BCA assay and therefore no data was available for comparison. FASP neat saliva ranged from 1.4 - 2.1 µg/µl, 10x diluted saliva ranged from 0.11-0.14 µg/µl, and 100x diluted saliva from 0.014-0.021 µg/µl. NYC OCME neat saliva ranged from 0.88-1.2 µg/µl, 10x diluted saliva from 0.11-0.17 µg/µl, and 100x diluted saliva from 0.006-0.008 µg/µl.

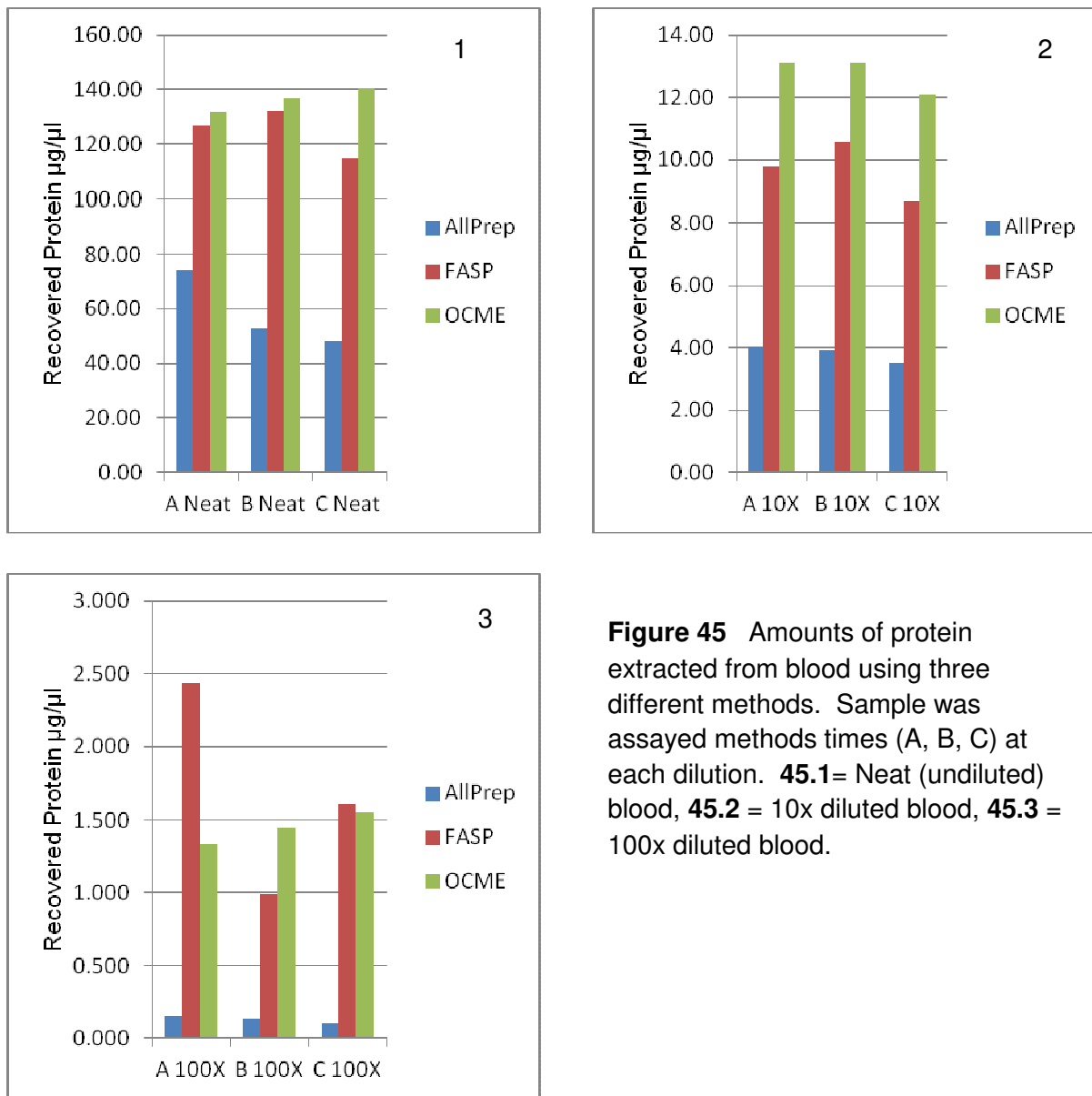


Figure 45 Amounts of protein extracted from blood using three different methods. Sample was assayed methods times (A, B, C) at each dilution. **45.1**= Neat (undiluted) blood, **45.2** = 10x diluted blood, **45.3** = 100x diluted blood.

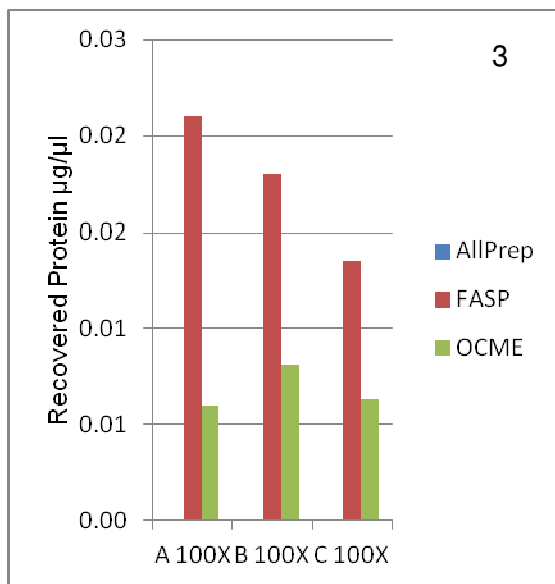
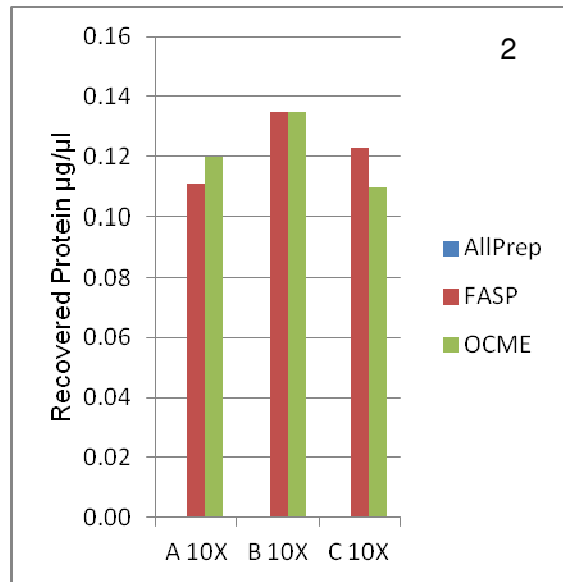
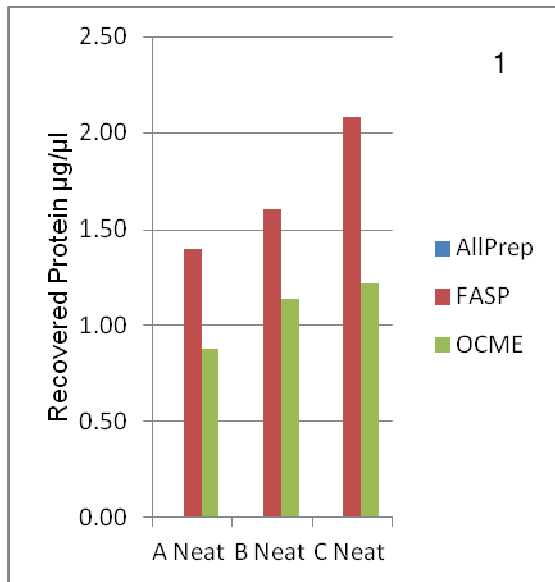


Figure 46 Amount of protein extracted from saliva using three different methods. Sample was assayed methods times (A, B, C) at each dilution. **46.1**= Neat (undiluted) saliva, **46.2** = 10x diluted saliva, **46.3** = 100x diluted saliva.

These data suggest that the standard OCME method extracts the greatest amount of protein from blood with the FASP method a close second. The FASP method appears extract more proteins from saliva (except 10x diluted sample) than the OCME method. AllPrep extracts the lowest amount of protein from both blood and saliva.

Qualitative protein analysis was performed by SDS MiGE (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA) using Agilent Protein 80 Kits. **Figure 47** shows a digital gel image comparing protein extraction for the same neat blood sample using the three methods. The qualitative results compare well with the quantitative results shown in **Figure 45**.

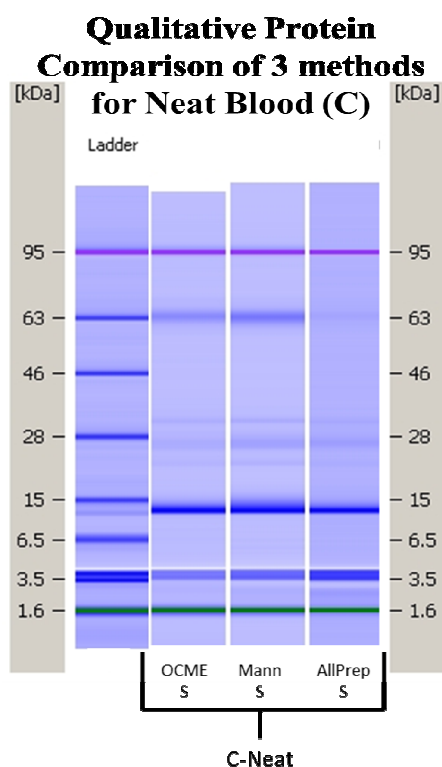


Figure 47 SDS MiGE digital gel comparison of supernatant fraction from all three protein extraction methods for a neat blood sample.

High performance liquid chromatography was performed on an Ultimate 3000 system (Dionex, Sunnyvale, CA) using a 30 minute elution gradient of 2-40% ACN containing 0.1% TFA at a flow rate of 300 nl/min. Eluted peptides were mixed in-line with equal volumes of HCCA matrix (described above) and 96 aliquots spotted on a MALDI plate. MS and MS/MS data were collected on a Sciex 4800 MALDI TOF/TOF using ProteinPilot software (Sciex, Framingham, MA). All analyses were based on peptide identifications with >95% confidence. Table 22 summarizes the average number of peptides, proteins and mass spectra from each of the AllPrep samples analyzed. Due to instrument constraints 10x blood AllPrep and 100x saliva AllPrep were not processed.

Table 22 Average number of peptides, proteins, and spectra identified for each type of samples extracted with AllPrep.

	Proteins Detected	Proteins Before Grouping	Distinct Peptides	Spectra Identified	% Total Spectra
Neat Blood	50	84	638	1313	81.0
100X Blood	52	103	559	1214	88.9
Neat Saliva	103	195	865	1395	82.3
10X Saliva	19	40	73	90	75.1

Blood-specific protein markers are listed in Table 23, with the number of peptides identified from each protein for neat and diluted blood samples extracted with AllPrep. The relatively small number of proteins identified in blood appears consistent with the MiGE gel image **Figure 40**, which shows a strong Hb band extracted by AllPrep. If Hb is preferentially extracted by AllPrep it could reduce MS/MS blood proteins identified due to dynamic range suppression of minor protein components. Saliva-specific protein markers are listed in Table 24, with the number of peptides identified from each protein for neat and diluted saliva samples extracted with AllPrep.

Table 23 Blood-specific protein markers found in blood samples extracted with AllPrep

	No. peptides at >95% confidence					
	Neat A	Neat B	Neat C	100x A	100x B	100x C
Hemoglobin subunit beta (HBB)	91	125	121	56	108	105
Hemoglobin subunit alpha (HBA1/HBA2)	72	73	85	36	58	59
Band 3 anion transport protein (SLC4A1)	3	4	5	-	-	-
Spectrin alpha chain, erythrocytic 1 (SPTA1)	1	-	-	-	-	-

- No peptides identified

A, B, C = separate extractions of the same blood sample

Table 24 Saliva-specific protein markers found in saliva samples extracted with AllPrep

	No. peptides at >95% confidence		
	Neat D	Neat E	Neat F
Alpha-amylase 1 (AMY1A)	48	44	52
Cystatin-SA (CST2)	24	15	17
Protein LEG1 homolog (LEG1/C6orf58)	1	1	1

A, B, C = separate extractions of the same blood sample

II) QUANTITATIVE AND QUALITATIVE DNA ANALYSIS: Blood samples were extracted using the DNA IQ – Database Protocol (Promega, TB297). Saliva samples were extracted with the DNA IQ System – Small Sample Casework Protocol (Promega, TB296). AllPrep DNA extraction followed the manufacturer’s protocol as described above. Extracted DNA samples were

quantified using Plexor HY System (Promega, Madison, WI) on the Rotor-Gene Q (Qiagen, Valencia, CA). Twenty-five μl of blood were used as the initial sample volume. Forty μl of saliva as the initial sample volume and 50 μl of the elution buffer were used. **Figure 48** shows DNA quantitation results from blood and saliva prepared by AllPrep and DNA IQ.

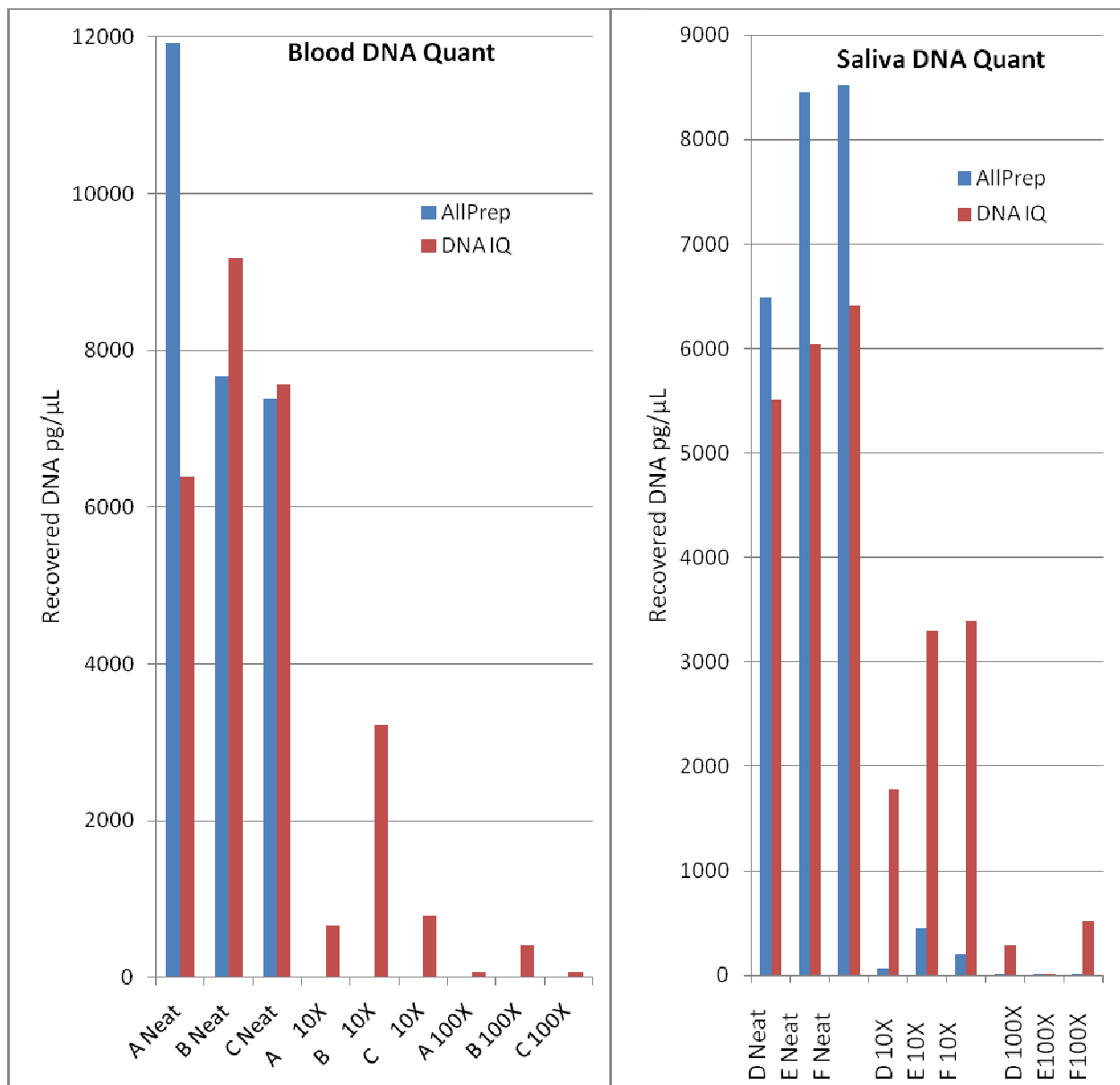


Figure 48 Amounts of DNA extracted from neat and diluted blood saliva samples using AllPrep NS DNA IQ. Left = Blood, right = Saliva. A, B, C are separate extractions of the same sample.




STR analysis was performed using PowerPlex 16 HS System (Promega, Madison, WI) on a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA). DNA was

analyzed and genetic profiles generated using GeneMapper Software (Applied Biosystems, Foster City, CA).

For each sample, verification of true allelic peaks was made by examining any baseline noise, or obvious artifact peaks, as well as ensuring that there was no contamination (i.e. only single source profiles expected). A DNA profile with true allelic peaks at 7 of 13 STR loci was considered a usable profile. Table 25 indicates which AllPrep processed samples produced useable DNA profiles in comparison with the DNA IQ processed samples.

Table 25 Determination of whether or not a usable DNA profile was obtained from each of the AllPrep processed samples in comparison to the DNA IQ processed samples.

	Neat Blood A	Neat Blood B	Neat Blood C	10x Blood A	10x Blood B	10x Blood C	100x Blood A	100x Blood B	100x Blood C	Neat Saliva D	Neat Saliva E	Neat Saliva F	10x Saliva D	10x Saliva E	10x Saliva F	100x Saliva D	100x Saliva E	100x Saliva F
All Prep	Usable Profile	Usable Profile	Usable Profile	Not Useable Profile	Not Useable Profile	Not Useable Profile	Not Useable Profile	Not Useable Profile	Not Useable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Not Useable Profile	Not Useable Profile	Not Useable Profile
DNA IQ	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Usable Profile	Failed to amplify due to an instrumentation error	Usable Profile

	Usable Profile
	Not Useable Profile
	Failed to amplify due to an instrumentation error

- **DISCUSSION:** Both protein-only extraction and purification methods recovered considerably more protein than the AllPrep kit for both blood and saliva samples. The FASP method and NYC OCME method both recovered twice as much protein as the AllPrep kit for neat blood. For 10x diluted blood, FASP recovered twice as much, while the NYC OCME method recovered almost three times as much protein. The 100x diluted protein results were variable and both FASP and NYC OCME methods recovered 6-25 times more protein. AllPrep saliva protein concentrations were too low to be measured by the BCA assay, meaning there was minimal recovery of protein. Qualitative analysis using MiGE, supported the quantitative results; as seen in **Figure 40**, the AllPrep gel image shows lower band intensity than the other two methods. The difference in recovery by these methods is likely due to the different chemistries and the number of steps involved in the processing methods. The AllPrep kit required multiple reagents and multiple steps in order to get the final protein fraction for digestion suggesting that there was more opportunity for protein loss than in the other protocols. The simplest method for processing the samples was the NYC OCME method that did not require additional filtering for digestion like FASP and AllPrep required. The advantage of using filters in the FASP digestion is for digesting larger sample sizes at once and allows for more purified proteins prior to the digestion allowing for a more concentrated final protein fraction of digested peptides.

HPLC-MALDI TOF/TOF analysis identified blood-specific protein markers hemoglobin subunit beta (HBB) and hemoglobin subunit alpha (HBA1/HBA2) in both undiluted and 100x diluted blood samples extracted with AllPrep. Additional blood-specific markers band 3 anion transport protein (SLC4A1) and spectrin alpha chain, erythrocytic 1 (SPTA1) were found in some undiluted samples. Saliva-specific protein markers α -amylase 1 (AMY1A), cystatin-SA

(CST2), and protein LEG1 homolog (LEG1/C6orf58) were found in all undiluted saliva samples extracted with AllPrep, however no saliva-specific markers were found in the 10x diluted saliva samples.

Almost all the DNA IQ extracted samples produced full, usable DNA profiles. Neat blood, neat saliva, and 10x diluted saliva were the only AllPrep samples that produced useable profiles. One possible cause is the AllPrep reagent present in the purified DNA sample interfering with the DNA analysis kit.

- **CONCLUSION:** Although separate protein and DNA extraction and purification methods yielded the best results, the results of dual protein/DNA extraction with the AllPrep kit were good enough to be used in forensic casework when testing options are limited. The MS proteomic analysis was more successful than the DNA analysis in that body fluid specific protein markers were identified for neat and diluted blood samples and neat saliva samples. Useable DNA profiles were obtained for the neat blood and saliva samples as well as for the 10x saliva samples. Different simultaneous protein/DNA extraction kits, or altering instrument parameters, may ameliorate the AllPrep results and produce useable profiles from more diluted samples. Our results show that AllPrep dual extraction could be an effective tool for obtaining both proteomic and genetic information from a single sample, especially when dealing with small evidence samples that may only allow for one extraction to be conducted. However, there is opportunity for improvement.

CONCLUSIONS AND IMPLICATIONS FOR POLICY, PRACTICE AND FUTURE RESEARCH

CONCLUSIONS: The overall goal of this work was to assess methods that could aid in reducing body fluid sample preparation time prior to mass spectrometry analysis (which is rapid) in order to improve throughput. This goal was achieved through the evaluation of CPLC for dynamic range reduction, and the use of organic solvents and microwaves to reduce trypsin digestion time which has traditionally been done overnight. A second goal was to determine the effects of long term aging and the chemicals luminol and Bluestar (commonly used forensic reagents that can be caustic) on the stability of menstrual blood and vaginal fluid protein markers. This aim also included an evaluation of the mass spectrometry method to simultaneously identify vaginal fluid microbiome markers along with our identified vaginal fluid human markers. The final goal was to evaluate the ability to simultaneously extract proteins and DNA from a sample in order to reduce sample consumption for body fluid and STR identify when sample is limiting. Success of each of these objectives is discussed sequentially below.

1. REDUCTION OF SAMPLE PROCESSING TIME – CPLC DYNAMIC RANGE REDUCTION AND USE OF ORGANIC SOLVENTS AND MICROWAVES TO REDUCE TRYPSIN DIGESTION TIME

Results demonstrated that CPLC treatment is effective in reducing the dynamic range of complex biological samples and increasing the number of proteins identified. This was particularly true for menstrual blood. However, while CPLC treatment of mixed body fluids (blood, saliva and semen) also showed a protein dynamic range reduction, in these mixtures marker proteins detection did not improve. For saliva in a mixture with blood, it appears that amylase is out-competed for binding on the CPLC column as it is found in the flow through fraction. Similarly, semenogelin does not appear in the CPLC eluates. Why these marker

proteins cannot find high affinity CPLC aptamers to bind to may be due to similarities in their structure/function – e.g. both form large complexes with other proteins in their respective body fluids. Amylase, along with specific markers proline-rich proteins, statherin and histatin, has been shown to complex with highly glycosylated mucin 5 to form complexes $> 40 \times 10^6$ Da [4,5]. Similarly semenogelins 1 and 2 (glycosylated) are cross-linked by cystine bond and transglutaminases [6,7] forming large complexes. These carbohydrate-coated multimer complexes may sequester amylase and semenogelins, preventing them from finding sufficient numbers high affinity aptamers to bind to and thus effectively reducing their enrichment. Clearly additional research will be required to identify mechanisms. However, for the forensic identification of menstrual blood CPLC appreciably reduced assay time.

Results from experiments aimed at reducing trypsin digestion time demonstrated that the combination of organic solvents and use of microwaves can appreciably reduce the time for Hb and albumin digestion. Near overnight digestion levels were achieved in 10% isopropanol at 37°C in approximately two hours using the 2,000:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio. Using a 20:1 ($\mu\text{g}/\mu\text{g}$) sample protein/trypsin ratio under these same conditions achieved similar results in approximately 30 minutes, making these conditions ideal for high throughput body fluid testing.

2. STABILITY OF MENSTRUAL BLOOD AND VAGINAL FLUID MARKER PROTEINS & IDENTIFICATION OF VAGINAL FLUID MICROBIOME MARKERS

Menstrual blood and vaginal fluid protein markers were shown to be stable for at least 18 months (the end of the trial period) under conditions tested. Bluestar and luminol showed no appreciable effect on protein stability for the 30 days over which these samples were exposed to these reagents. Additionally, we confirmed that many *Lactobacillus* proteins that have been used to identify the vaginal fluid from the vaginal fluid microbiome can also be detected by mass spectrometry of vaginal fluid samples. However, we believe that the identification of human specific vaginal fluid marker proteins is a more definitive test for the presence of vaginal fluid.

3. SIMULTANEOUS EXTRACTION OF PROTEINS AND DNA FROM A SINGLE SAMPLE

Results demonstrated that protein and DNA can be simultaneously extracted from a single sample and analyzed for body fluid and STR identity markers. Although separate protein and DNA extraction and purification methods yielded the best results, the results of dual protein/DNA extraction with the Qiagen AllPrep kit were good enough to be used in forensic casework when testing options are limited. The MS proteomic analysis was more successful than the DNA analysis in that body fluid specific protein markers were identified for neat and diluted blood samples and neat saliva samples. Useable DNA profiles were obtained for the neat blood and saliva samples as well as for the 10x saliva samples. Different simultaneous protein/DNA extraction kits, or altering instrument parameters, may ameliorate the AllPrep results and produce useable profiles from more diluted samples. Our results show that AllPrep dual extraction could be an effective tool for obtaining both proteomic and genetic information from a single sample, especially when dealing with small evidence samples that may only allow for one extraction to be conducted. However, there is opportunity for improvement to this method with additional research needed.

IMPLICATIONS FOR POLICY AND PRACTICE:

Current methods used for body fluid identification encompass a wide range of technologies including chemical, biochemical, immunochemical, enzymatic, spectrophotometric and others. Some rely on technologies that are a decades old, some are sensitive, other not, some rapid, other laborious hardly any are confirmatory. This assortment of techniques presents practical laboratory as well as policy problems as courts and juries come to expect test results that meet the high standards set by DNA testing.

The practical laboratory problems are many; some tests require biological activity, meaning that sample degradation in the field or in storage limit their usefulness. The need to perform multiple tests on an unknown sample to find out which body fluid is present is both time and sample consuming and requires multiple, and often expensive, instruments, as well as trained personnel to operate them. There are currently no regularly performed test for menstrual blood and vaginal fluid, and species identification is rarely performed because of expense and time.

Recent progress in the field of proteomics has demonstrated that a single methodology, mass spectrometry, can rapidly, and cost effectively, identify all body fluids in a sample simultaneously and confirmatively. Use of this technology will have implications for both the policy and practice of forensic science. Results from the work performed in this application can unify all body fluid identifications into a single sensitive method and will establish tests for menstrual blood, vaginal fluid and body fluids mixtures. The 2009 National Research Council's report *Strengthening Forensic Science in the United States: A Path Forward* has identified the need to bring all fields of forensics to the level of quality routinely provided by DNA testing, and policies to establish common standards and practices for body fluid testing are inevitable. Because mass spectrometry identifies the amino acid sequences of body fluid specific peptides, it offers a high level of certainty for results. Also, as the use of MS becomes more prevalent, its methods more streamlined and user friendly, sample extraction and preparations "kits" may become available and in our opinion mass spectrometry may become, in practice, the gold standard of body fluid testing.

We believe adoption of MS for body fluid detection would result in a single, uniform assay, using the same extraction, purification and detections methods for all body fluids, eliminating multiple techniques, saving time, money and sample, and giving the same confirmatory results for all body fluids. Above all, it would allow for the establishment of common, uniform standards throughout the forensic community in a manner similar to what is now in place for DNA testing. Its adoption would be beneficial not only to forensic science, but the criminal justice system, as the source of DNA (testing of which is approaching the level of a few cells) will be able to be identified – meaning that a jury will be more confident that a person's DNA was not the results of shed epithelial cells or transfer from a touched object, but from a specific body fluid – e.g. blood, saliva, semen, menstrual blood or vaginal fluid. This knowledge can only improve the quality of justice.

We have identified multiple unique and enriched protein markers in blood, saliva and semen. The strength of these protein markers rests in the fact that they perform specific biological tasks necessary for these body fluids to perform their functions (e.g. hemoglobin in blood, semenogelin in semen and amylase in saliva) and consequently these markers are well established in their respective body fluids in the fields of biology, medicine and forensics. Further, we have identified multiple proteins in each body fluid and multiple peptides from each

protein. The presence of multiple proteins and peptides imparts greater confidence to the accuracy of the assay, especially when trying to deconvolute body fluid mixtures. In addition, we have also identified multiple functional markers in menstrual blood and vaginal fluid which will allow confident identification of these body fluids. Indeed, because of our ability to identify menstrual blood specific markers, we can distinguish a mixture of venous blood and vaginal fluid from menstrual blood. Such tests may prove decisive in criminal cases, especially sexual assaults.

This work has focused on making these markers useful to the forensic practitioner in a way that is confirmatory, fast and inexpensive. Microwave digesting significantly reduced protein digestion time from overnight to less than one hour, and CPLC reduces menstrual blood processing time from day to hours. We believe that further work, described in the conclusion of the body fluid mixtures section, will improve mixture deconvolution detection levels even more than our current assay's ability.

The identification of multiple *functional* proteins unique to or enriched in these five body fluids will directly contribute to both the policy and practice of forensic investigations. With regard to policy, we believe that the demonstration of a panel of *functional* proteins markers will bring solid scientific approach to serology analysis, making it a reliable confirmatory test. With respect to practice, the NYC OCME, the largest public forensic laboratory in the nation, is moving forward with confirmatory proteomic serology test because it is accurate, reliable, fast, amenable to high throughput and less expensive than the other molecular alternatives currently available.

IMPLICATIONS FOR FUTURE RESEARCH – While mass spectrometry is already serving as a sensitive and powerful method with clinical applications, its use in biological forensic science is just beginning. As mentioned, incorporation of MS body fluid identification would be levels of confidence in these results to those of DNA STR testing and consequently improving confidence in the criminal justice system. While ready for use in biological forensic science, more research needs to be done. Many more samples and body fluid types and mixtures need to be tested. Improvement in sample preparation methods examined and standards set. Additionally, as mass spectrometry analysis automatically gives species identification, examination into the standards for reporting out these data also need to be established.

REFERENCES

1. Akutsu T, Motani H, Watanabe K, Iwase H, Sakurada K. Detection of bacterial 16S ribosomal RNA genes for forensic identification of vaginal fluid. *Leg Med (Tokyo)* 2012; **14**:160–2.
2. Benschop CC, Quaak FC, Boon ME, Sijen T, Kuiper I. Vaginal microbial flora analysis by next generation sequencing and microarrays; can microbes indicate vaginal origin in a forensic context? *International journal of legal medicine* 2012; **126**:303–310.
3. Yang H, Zhou B, Deng H, Prinz M, Siegel D. Body fluid identification by mass spectrometry. *International journal of legal medicine* 2013; **127**:1065–1077.
4. Iontcheva I, Oppenheim F, Troxler R. Human salivary mucin MG1 selectively forms heterotypic complexes with amylase, proline-rich proteins, statherin, and histatins. *Journal of dental research* 1997; **76**:734–743.
5. Thornton DJ, Khan N, Mehrotra R, *et al.* Salivary mucin MG1 is comprised almost entirely of different glycosylated forms of the MUC5B gene product. *Glycobiology* 1999; **9**:293–302.
6. Lilja H, Laurell C-B. The predominant protein in human seminal coagulate. *Scandinavian journal of clinical and laboratory investigation* 1985; **45**:635–641.
7. Peter A, Lilja H, Lundwall Å, Malm J. Semenogelin I and semenogelin II, the major gel-forming proteins in human semen, are substrates for transglutaminase. *European Journal of Biochemistry* 1998; **252**:216–221.
8. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nature methods* 2009; **6**:359.
9. Bauer M, Patzelt D. Identification of menstrual blood by real time RT-PCR: technical improvements and the practical value of negative test results. *Forensic science international* 2008; **174**:55–59.
10. Fleming RI, Harbison S. The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids. *Forensic Science International: Genetics* 2010; **4**:244–256.
11. Haas C, Klessner B, Maake C, Bär W, Kratzer A. mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR. *Forensic Science International: Genetics* 2009; **3**:80–88.
12. Juusola J, Ballantyne J. Multiplex mRNA profiling for the identification of body fluids. *Forensic Science International* 2005; **152**:1–12.
13. Nussbaumer C, Gharehbaghi-Schnell E, Korschineck I. Messenger RNA profiling: a novel method for body fluid identification by real-time PCR. *Forensic science international* 2006; **157**:181–186.
14. Richard MLL, Harper KA, Craig RL, Onorato AJ, Robertson JM, Donfack J. Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis. *Forensic Science International: Genetics* 2012; **6**:452–460.

15. Haas C, Klessner B, Kratzer A, Bär W. mRNA profiling for body fluid identification. *Forensic Science International: Genetics Supplement Series* 2008; **1**:37–38.
16. Hanson EK, Lubenow H, Ballantyne J. Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. *Analytical biochemistry* 2009; **387**:303–314.
17. Ibberson D, Benes V, Muckenthaler MU, Castoldi M. RNA degradation compromises the reliability of microRNA expression profiling. *BMC biotechnology* 2009; **9**:1.
18. Jung M, Schaefer A, Steiner I, *et al.* Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. *Clinical chemistry* 2010; **56**:998–1006.
19. Peirson SN, Butler JN. RNA extraction from mammalian tissues. *Circadian Rhythms: Methods and Protocols* 2007:315–327.
20. Brown T, Mackey K, Du T. Analysis of RNA by Northern and slot blot hybridization. *Current protocols in molecular biology* 2004:4–9.
21. Gilman M. Preparation of cytoplasmic RNA from tissue culture cells. *Current protocols in molecular biology* 2002:4–1.
22. Blank A, Dekker C, Schieven G, Sugiyama R, Thelen M. Human body fluid ribonucleases: detection, interrelationships and significance. In: *Nucleic acids symposium series.*; 1980:203–209.
23. Inoue H, Kimura A, Tuji T. Degradation profile of mRNA in a dead rat body: basic semi-quantification study. *Forensic science international* 2002; **130**:127–132.
24. Tanner M, Berk L, Felten D, Blidy A, Bit S, Ruff D. Substantial changes in gene expression level due to the storage temperature and storage duration of human whole blood. *Clinical & Laboratory Haematology* 2002; **24**:337–341.
25. Shaw JL, Diamandis EP. Distribution of 15 human kallikreins in tissues and biological fluids. *Clinical Chemistry* 2007; **53**:1423–1432.
26. Wang Z, Luo H, Pan X, Liao M, Hou Y. A model for data analysis of microRNA expression in forensic body fluid identification. *Forensic Science International: Genetics* 2012; **6**:419–423.
27. Zubakov D, Boersma AW, Choi Y, van Kuijk PF, Wiemer EA, Kayser M. MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. *International journal of legal medicine* 2010; **124**:217–226.
28. Shomron N, Golan D, Hornstein E. An evolutionary perspective of animal microRNAs and their targets. *BioMed Research International* 2009; **2009**.
29. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *cell* 2005; **120**:15–20.

30. Lim LP, Lau NC, Garrett-Engele P, *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005; **433**:769–773.
31. Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH, Cuppen E. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* 2005; **120**:21–24.
32. Pigati L, Yaddanapudi SC, Iyengar R, *et al.* Selective release of microRNA species from normal and malignant mammary epithelial cells. *PloS one* 2010; **5**:e13515.
33. Saito Y, Liang G, Egger G, *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer cell* 2006; **9**:435–443.
34. Scott GK, Mattie MD, Berger CE, Benz SC, Benz CC. Rapid alteration of microRNA levels by histone deacetylase inhibition. *Cancer research* 2006; **66**:1277–1281.
35. Mor E, Cabilly Y, Goldshmit Y, *et al.* Species-specific microRNA roles elucidated following astrocyte activation. *Nucleic acids research* 2011:gkq1325.
36. De Bustos C, Ramos E, Young JM, *et al.* Tissue-specific variation in DNA methylation levels along human chromosome 1. *Epigenetics & chromatin* 2009; **2**:1.
37. Eckhardt F, Lewin J, Cortese R, *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature genetics* 2006; **38**:1378–1385.
38. Hong S-J, Kang M-I, Oh J-H, *et al.* DNA methylation and expression patterns of key tissue-specific genes in adult stem cells and stomach tissues. *Journal of Korean medical science* 2009; **24**:918–929.
39. Frumkin D, Wasserstrom A, Budowle B, Davidson A. DNA methylation-based forensic tissue identification. *Forensic Science International: Genetics* 2011; **5**:517–524.
40. Grönniger E, Weber B, Heil O, *et al.* Aging and chronic sun exposure cause distinct epigenetic changes in human skin. *PLoS Genet* 2010; **6**:e1000971.
41. El-Maarri O, Becker T, Junen J, *et al.* Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. *Human genetics* 2007; **122**:505–514.
42. Adkins RM, Krushkal J, Tylavsky FA, Thomas F. Racial differences in gene-specific DNA methylation levels are present at birth. *Birth Defects Research Part A: Clinical and Molecular Teratology* 2011; **91**:728–736.
43. Andraos C, Koorsen G, Knight JC, Bornman L. Vitamin D receptor gene methylation is associated with ethnicity, tuberculosis, and TaqI polymorphism. *Human immunology* 2011; **72**:262–268.
44. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004; **429**:457–463.

45. Vercelli D. Genetics, epigenetics, and the environment: switching, buffering, releasing. *Journal of Allergy and Clinical Immunology* 2004; **113**:381–386.
46. Choi S-W, Friso S. Epigenetics: a new bridge between nutrition and health. *Advances in Nutrition: An International Review Journal* 2010; **1**:8–16.
47. Ehrich M, Zoll S, Sur S, Van Den Boom D. A new method for accurate assessment of DNA quality after bisulfite treatment. *Nucleic acids research* 2007; **35**:e29.
48. Grunau C, Clark S, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic acids research* 2001; **29**:e65–e65.
49. Lill JR, Ingle ES, Liu PS, Pham V, Sandoval WN. Microwave-assisted proteomics. *Mass spectrometry reviews* 2007; **26**:657–671.
50. Brewis IA, Gadella BM. Sperm surface proteomics: from protein lists to biological function. *Molecular human reproduction* 2010; **16**:68–79.
51. Issaq HJ, Waybright TJ, Veenstra TD. Cancer biomarker discovery: opportunities and pitfalls in analytical methods. *Electrophoresis* 2011; **32**:967–975.
52. Righetti PG, Fasoli E, Boschetti E. Combinatorial peptide ligand libraries: the conquest of the “hidden proteome” advances at great strides. *Electrophoresis* 2011; **32**:960–6.
53. Sparbier K, Wenzel T, Dihazi H, *et al.* Immuno-MALDI-TOF MS: New perspectives for clinical applications of mass spectrometry. *Proteomics* 2009; **9**:1442–1450.
54. Remily-Wood ER, Liu RZ, Xiang Y, *et al.* A database of reaction monitoring mass spectrometry assays for elucidating therapeutic response in cancer. *PROTEOMICS-Clinical Applications* 2011; **5**:383–396.
55. Hachani J, Duban-Deweere S, Pottiez G, Renom G, Flahaut C, Périni J-M. MALDI-TOF MS profiling as the first-tier screen for sickle cell disease in neonates: Matching throughput to objectives. *PROTEOMICS-Clinical Applications* 2011; **5**:405–414.
56. Al-Tarawneh SK, Border MB, Dibble CF, Bencharit S. Defining salivary biomarkers using mass spectrometry-based proteomics: a systematic review. *Omics: a journal of integrative biology* 2011; **15**:353–361.
57. Thacker S, Yadav SP, Sharma RK, *et al.* Evaluation of sperm proteins in infertile men: a proteomic approach. *Fertility and sterility* 2011; **95**:2745–2748.
58. Zegels G, Van Raemdonck GA, Coen EP, Tjalma WA, Van Ostade XW. Comprehensive proteomic analysis of human cervical-vaginal fluid using colposcopy samples. *Proteome science* 2009; **7**:17.
59. Hortin GL, Sviridov D. The dynamic range problem in the analysis of the plasma proteome. *J Proteomics* 2010; **73**:629–36.

60. Boschetti E, Righetti PG. The art of observing rare protein species in proteomes with peptide ligand libraries. *Proteomics* 2009; **9**:1492–510.
61. Ringrose JH, van Solinge WW, Mohammed S, *et al.* Highly efficient depletion strategy for the two most abundant erythrocyte soluble proteins improves proteome coverage dramatically. *Journal of proteome research* 2008; **7**:3060–3063.
62. Wang H, Chang-Wong T, Tang H-Y, Speicher DW. Comparison of extensive protein fractionation and repetitive LC-MS/MS analyses on depth of analysis for complex proteomes. *Journal of proteome research* 2010; **9**:1032–1040.
63. Sijen T. Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Sci Int Genet.* 2015, **18**: 21-32
64. Hanson EK, Mirza M, Rekab K, Ballantyne J. The identification of menstrual blood in forensic samples by logistic regression modeling of miRNA expression. *Electrophoresis.* 2014, **35**: 3087-3095
65. Li Z, Bai P, Peng D, Long B, Zhang L, Liang W. Screening and confirmation of microRNA markers for distinguishing between menstrual and peripheral blood. *Forensic Sci Int Genet.* 2015, **5**: e353-e355
66. Haas C, Hanson E, Anjos M, Ballantyne K, Banemann R, *et al.* RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: Results of a fourth and fifth collaborative EDNAP exercise. *Forensic Sci Int Genet.* 2014;**8**:203-212
67. Jakubowska J, Maciejewska A, Bielawski K, Pawłowski R. mRNA heptaplex protocol for distinguishing between menstrual and peripheral blood. *Forensic Sci Int Genet.* 2014;**13**:53-60
68. Jakubowska J, Maciejewska A, Pawłowski R. mRNA Profiling for Vaginal Fluid and Menstrual Blood Identification. *Methods Mol Biol.* 2016;**1420**:33-42
69. Richard ML, Harper KA, Craig RL, Onorato AJ, Robertson JM, *et al.* Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis. *Forensic Sci Int Genet.* 2011.
70. Zubakov D, Kokmeijer I, Ralf A, Rajagopalan N, Calandro L, Wootton S, Langit R, Chang C, Lagace R, Kayser M. Towards simultaneous individual and tissue identification: A proof-of-principle study on parallel sequencing of STRs, amelogenin, and mRNAs with the Ion Torrent PGM. *Forensic Sci Int Genet.* 2015, **17**: 122-128
71. Ambion TB Ambion Technical Bulletin.
72. Roeder AD, Haas C. Body Fluid Identification Using mRNA Profiling. *Methods Mol Biol.* 2016;**1420**:13-31

73. Wang Z, Luo H, Pan X, Liao M, Hou Y. A model for data analysis of microRNA expression in forensic body fluid identification. *Forensic Sci Int Genet.* 2012, 6: 419-423
74. Sauer E, Reinke AK, Courts C. Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. *Forensic Sci Int Genet.* 2016, 22: 89-99
75. Sirker M, Fimmers R, Schneider PM, Gomes I. Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification. *Forensic Sci Int Genet.* 2017, 27: 41-49
76. Lee HY, An JH, Jung S, Oh Y, Lee EY, Choi A, Yang WI, Shin K. Genome-wide methylation profiling and a multiplex construction for the identification of body fluids using epigenetic markers. *Forensic Sci Int Genet.* 2015, 17: 17-24
77. Leontiou CA, Hadjidaniel MD, Petros M, Antoniou P, Ioannides M, Patsalis PC. Bisulfite Conversion of DNA: Performance Comparison of Different Kits and Methylation Quantitation of Epigenetic Biomarkers that Have the Potential to Be Used in Non-Invasive Prenatal Testing. *PLoS One.* 2015, 10(8):e0135058.
78. Silva D, Antunes J, Balamurugan K, Duncan G, Alh CS, McCord B. Developmental Validation studies of epigenetic DNA methylation markers for the detection of blood, semen and saliva samples. *Forensic Sci Int Genet.* 2016, 23: 55-63
79. An, J.H., et al., DNA methylation-specific multiplex assays for body fluid identification. *Int J Legal Med,* 2013. 127(1): p. 35-43.
80. Bernstein, B.E., A. Meissner, and E.S. Lander, The mammalian epigenome. *Cell,* 2007. 128(4): 669-81.
81. Forat S, Huettel B, Reinhardt R, Fimmers R, Haidl G, Denschlag D, Olek K. Methylation Markers for the Identification of Body Fluids and Tissue from Forensic Trace Evidence. *PLoS One.* 2016, 11(2): e0147973
82. Gršković B, Zrnec D, Vicković S, Popović M, Mršić G. DNA methylation: the future of crime scene investigation? *Mol Biol Rep.* 2013, 40(7):4349-60.
83. Kader, F. and M. Ghai, DNA methylation and application in forensic sciences. *Forensic Sci Int,* 2015. 249: p. 255-265.
84. Lee HY, Lee SD, Shin KJ. Forensic DNA methylation profiling from evidence material for investigative leads. *BMB Rep.* 2016, 49 (7): 359-369
85. Lee HY, Jung SE, Lee EH, Yang WI, Shin KJ. DNA methylation profiling for a confirmatory test for blood, saliva, semen, vaginal fluid and menstrual blood. *Forensic Sci Int Genet.* 2016 ,24:75-82

86. Lin YC, Tsai LC, Lee JC, Su CW, Tzen JT, Linacre A, Hsieh HM. Novel identification of biofluids using a multiplex methylation sensitive restriction enzyme-PCR system. *Forensic Sci Int Genet.* 2016,25:157-165
87. Lin YC, Tsai LC, Lee JC, Liu KL, Tzen JT, Linacre A, Hsieh HM. Novel identification of biofluids using a multiplex methylation-specific PCR combined with single-base extension system. *Forensic Sci Med Pathol.* 2016,(2):128-38
88. Sauer E, Reinke AK, Courts C. Validation of forensic body fluid identification based on empirically normalized miRNA expression data. *Forensic Sci Int Genet Suppl Series.* 2015, 5: e464-e464

DISSEMINATION OF RESEARCH FINDINGS

• Publications

1. Proteomic Analysis of Menstrual Blood, Yang H, Zhou B, Prinz M, Siegel D was published in *Molecular and Cellular Proteomics 11*: 1024-1035 (2012).
2. Body Fluid Identification by Mass Spectrometry, Yang H, Zhou B, Deng, H, Prinz M, Siegel D *International Journal of Legal Medicine 127(6)*:1065-77 (2013)

• Presentations

1. Evaluation of Simultaneous Protein & DNA Extraction and Purification, Presented by Kaylee Hershfeld in partial fulfillment of her master's program at Towson University, Towson, MD, fall 2013.
2. Molecular Serology Rapid, Confirmatory, Proteomic Identification of Body Fluids, Tissues and People, Presented by Donald Siegel, Ph.D. at the 2014 annual Green Mountain DNA Conference.
3. Identification of Marker Proteins in Menstrual Blood Using LC-MALDI Mass Spectrometry, Presented by Matthew Goldstein, B.S., graduate student, Virginia Commonwealth University Department of Forensic Science, at the 2014 American Academy of Forensic Sciences Annual Meeting, Seattle, WA