Novel MS–based Approaches (MRM and iMALDI) for Biomarker Discovery and Validation

Christoph H. Borchers
The mandate of the University of Victoria Genome British Columbia Proteomics Centre is to provide exceptional service and support for proteomics research in the areas of protein identification and characterization, and quantitative proteomics for biomarker discovery and validation. Our research interests are focused on novel technology development in structural proteomics, metabolomics, clinical proteomics, and protein imaging with the ultimate goal of applying our technologies to customer research projects.

**Featured Services**
- Biomarker Discovery and Validation
- MALDI Imaging
- Quantitative Proteomics
- Structural Proteomics
- Peptide Synthesis
- Protein Characterization
- Protein Identification
- Bioinformatics
- Metabolomics

**Publications**
- Quantitative iTRAQ proteome and comparative transcriptome analysis of elicitor-induced Norway spruce (Picea abies) cells reveals elements of calcium signaling in the early conifer defense response - February 6, 2009
- Membrane protease degradomics: proteomic identification and quantification of cell surface protease substrates. - January 26, 2009
- Identification of serum biomarkers for lung cancer by proteomic analysis - January 20, 2009

www.proteincentre.com
• Fee-for-service

• Large-scale collaborative projects
  • Industry partnerships
  • Supported by Genome Canada and Genome BC

• Products
  • Peptide Standards
  • Crosslinkers
• **Quantitative proteomics**
  - iTRAQ, multiple reaction monitoring, iMALDI

• **Comprehensive protein characterization**
  - analysis of protein expression & modification levels

• **Phospho–proteomics**
  - signaling pathways

• **Metabolomics for biomarker discovery**

• **Structural proteomics**
  - for drug development & design

• **MALDI tissue imaging**
Directed MS based quantitation

- target protein of interest must be known

- analysis is rapid
  - hours not days
  - higher throughput than “shotgun” MS approaches

- accurate quantitation
  - highly reproducible (CV’s <10%)
  - absolute concentration determination

- robust mass spec instrumentation
  - triple quadrupoles/MALDI–MS
Multiple Reaction Monitoring (MRM)

Application of the HPP SIS peptide cocktail
Multiple Reaction Monitoring

Collision chamber

- tandem MS (MS/MS) scan mode
  - low noise & high sensitivity
  - unique to QqQ
- broadlinear range ($10^3$ – $10^4$)
- well established in small molecule quantitation
GOAL: To create a library of isotopically labeled, standard peptides for absolute concentration determination using MRM assays

Simple sample preparation
- no depletion/fractionation
- direct analysis of trypsin digested plasma

Rapid analysis
- 30 – 60 min per analysis
- duplicate analyses
Applications

- Multiplexed protein expression profiling
  - analogous to cDNA microarrays
  - reveal disease associated expression patterns
  - high throughput potential

- Development of MRM assays
  - quantitation
  - potential for translation to clinical settings

- Plasma trypsin digestion kinetics
  - analytical comparison of denaturants & solvent systems

- Applicable to any biological system
Plasma Protein Assays by MRM:
45 moderate–high abundance proteins

<table>
<thead>
<tr>
<th>Protein Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum Albumin</td>
</tr>
<tr>
<td>Transferrin</td>
</tr>
<tr>
<td>Apolipoprotein A–I</td>
</tr>
<tr>
<td>Alpha–1–acid Glycoprotein 1</td>
</tr>
<tr>
<td>Apolipoprotein A–II Precursor</td>
</tr>
<tr>
<td>Haptoglobin Beta Chain</td>
</tr>
<tr>
<td>Transthyretin</td>
</tr>
<tr>
<td>Hemopexin</td>
</tr>
<tr>
<td>Fibrinogen Gamma Chain</td>
</tr>
<tr>
<td>Fibrinogen Alpha Chain</td>
</tr>
<tr>
<td>Fibrinogen Beta Chain</td>
</tr>
<tr>
<td>Apolipoprotein C–III</td>
</tr>
<tr>
<td>Alpha–2–macroglobulin</td>
</tr>
<tr>
<td>Apolipoprotein C–I Lipoprotein</td>
</tr>
<tr>
<td>Alpha–1–antichymotrypsin</td>
</tr>
<tr>
<td>Complement C3</td>
</tr>
<tr>
<td>Vitronectin</td>
</tr>
<tr>
<td>Inter–alpha–trypsin Inhibitor Heavy Chain H1</td>
</tr>
<tr>
<td>Vitamin D–binding Protein</td>
</tr>
<tr>
<td>Alpha–1B–glycoprotein</td>
</tr>
<tr>
<td>Antithrombin–III</td>
</tr>
<tr>
<td>Apolipoprotein A–IV</td>
</tr>
<tr>
<td>Complement Factor H</td>
</tr>
<tr>
<td>Complement B Factor</td>
</tr>
<tr>
<td>Gelsolin, isoform 1</td>
</tr>
<tr>
<td>Ceruloplasmin</td>
</tr>
<tr>
<td>Complement Component C9</td>
</tr>
<tr>
<td>Clusterin</td>
</tr>
<tr>
<td>Zinc–alpha–2–glycoprotein</td>
</tr>
<tr>
<td>Serum amyloid P–component</td>
</tr>
<tr>
<td>Heparin Cofactor II</td>
</tr>
<tr>
<td>Prothrombin</td>
</tr>
<tr>
<td>Plasma Retinol–binding Protein Precursor</td>
</tr>
<tr>
<td>Complement C4 Beta Chain</td>
</tr>
<tr>
<td>Complement C4 Gamma Chain</td>
</tr>
<tr>
<td>Apolipoprotein B–100</td>
</tr>
<tr>
<td>Alpha–2–antiplasmin</td>
</tr>
<tr>
<td>Kininogen–1</td>
</tr>
<tr>
<td>Plasminogen</td>
</tr>
<tr>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>Coagulation Factor XIIa Heavy Chain</td>
</tr>
<tr>
<td>Afamin</td>
</tr>
<tr>
<td>Beta–2–glycoprotein I</td>
</tr>
<tr>
<td>Angiotensinogen</td>
</tr>
<tr>
<td>L–selectin</td>
</tr>
</tbody>
</table>

All 45 isotopically peptides have been synthesized, purified and optimized.
Panel of standards is expandable

addition based on CVD association

~120 proteins with a circulating concentration >L-selectin

- Adiponectin
- Aldolase C
- Angiotensin-converting Enzyme
- Apolipoprotein D
- Apolipoprotein L1
- Aspartate Aminotransferase
- CD105 (Endoglin)
- Creatine Kinase-b
- Coagulation Factor XIII (A Chain)
- Fibrinopeptide A
- Fibronecin
- Histidine rich glycoprotein
- Thrombospondin-1
- Gliarial Fibrillary Acidic Protein (GFAP)
- GPIIB Soluble
- Tropomyosin 1 Alpha Chain
- Complement C1 Inactivator
- Von Willebrand Factor
- Alpha-1-antitrypsin
- A2–hs–glycoprotein

20 new peptides synthesized, MRM optimization & assay characterization underway

<table>
<thead>
<tr>
<th>Name</th>
<th>Accession</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>coagulation Factor XIII</td>
<td>P00488</td>
<td>coagulation</td>
</tr>
<tr>
<td>paraoxonase (PON1, 2, 3)</td>
<td>(P27169, Q15165, Q15166)</td>
<td>increased levels in chronic arterial obstructive disease (Limminello, 1994 #1433)</td>
</tr>
<tr>
<td>platelet-derived growth factor (PDGF)</td>
<td>P14063</td>
<td>strongly related to cardiovascular risk (Koehn, 2003 #1657)</td>
</tr>
<tr>
<td>fibrinogen</td>
<td>P02671</td>
<td>specific biochemical marker of myocardial injury (Av, 2002 #1418)</td>
</tr>
<tr>
<td>creatine kinase-MB</td>
<td>P12277, P06732</td>
<td>increased levels in chronic arterial obstructive disease (Limminello, 1994 #1433)</td>
</tr>
<tr>
<td>myosin heavy chain, cardiac</td>
<td>P13533, P02792</td>
<td>increased levels in chronic arterial obstructive disease (Limminello, 1994 #1433)</td>
</tr>
<tr>
<td>albumin</td>
<td>P02768</td>
<td>increased levels with increased risk of cardiovascular mortality (Shaper, 2004 #1833)</td>
</tr>
<tr>
<td>fibronectin</td>
<td>P02761</td>
<td>increased in patients with ACS and is associated with adverse outcome (Otani, 2001 #617)</td>
</tr>
<tr>
<td>alpha-2 macroglobulin (AAT)</td>
<td>P01024</td>
<td>Alpha phase reactant</td>
</tr>
<tr>
<td>alpha-2 macroglobulin</td>
<td>(P01009, P02656)</td>
<td>C3 is more strongly associated with previous myocardial infarction than other risk factors (Muscati, 2000 #1731)</td>
</tr>
<tr>
<td>alantinase</td>
<td>P02656</td>
<td>Major plasma protease inhibitor marker of CHD independent of cholesterol (Koen, 2003 #1456)</td>
</tr>
<tr>
<td>thrombospondin-1</td>
<td>P04464</td>
<td>Might function as an alternative substrate for thrombus formation (Jurek, 2003)</td>
</tr>
<tr>
<td>von Willebrand Factor</td>
<td>P04464</td>
<td>A repressor for rapid inhibition of activated protein C by plasminogen activator inhibitor-1 (Gschom, 1995 #1465)</td>
</tr>
<tr>
<td>smooth muscle myosin heavy chain</td>
<td>P35749</td>
<td>Intracoronary level may be a biochemical marker for the prediction of restenosis (Tsukou, 2000 #982)</td>
</tr>
<tr>
<td>haptoglobin</td>
<td>P00737</td>
<td>Subjects with H2-2 had significantly higher serum total and free cholesterol concentration (Brazenor, 1999 #982)</td>
</tr>
<tr>
<td>apolipoprotein A-II</td>
<td>P02652</td>
<td>Major plasma protease inhibitor</td>
</tr>
<tr>
<td>alpha 2 macroglobulin</td>
<td>(alpha 2 M)</td>
<td>May be predictive of recurrent ischemia (Scharf, 1998 #1184)</td>
</tr>
<tr>
<td>fibrinogen B beta-1-42</td>
<td>P02675</td>
<td>Low level associated with mortality and myocardial infarction five years after CABG (Skinner, 1998 #983)</td>
</tr>
<tr>
<td>apolipoprotein</td>
<td>P02647</td>
<td>Marker of brain damage (Herrmann, 2000 #1564)</td>
</tr>
<tr>
<td>glial fibrillary acidic protein</td>
<td>(GFAP)</td>
<td>Acute phase protein</td>
</tr>
<tr>
<td>haptoglobin</td>
<td>P14136</td>
<td>Ceruloplasmin was reported to be an independent risk factor for cardiovascular disease (Kem, 2002 #464)</td>
</tr>
<tr>
<td>apolipoprotein D</td>
<td>P00290</td>
<td>Lipoprotein</td>
</tr>
<tr>
<td>albumin</td>
<td>P00450</td>
<td>Cellular fibrinogen may be a marker protein for endothelial cell activation (Kaners, 2001 #1639)</td>
</tr>
<tr>
<td>alpha(1)-antichymotrypsin</td>
<td>P01011</td>
<td>Major plasma protease inhibitor</td>
</tr>
</tbody>
</table>
MS Protein Assays by MRM: Workflow for assay development

Workflow for assay development

Select proteotypic tryptic peptides

MRM optimization by nano-infusion

Co-elution of heavy and natural peptides?

Synthesis of isotopically labeled peptide

YES

Pep balance amount of heavy peptide used vs. natural peptide

YES

Reversed phase HPLC purification

MRM assay

Assay characterization:
- Analytical precision
- Linearity & linear response
- Limit of quantitation (LOQ)

Amino acid analysis
- Capillary zone electrophoresis
- MALDI-TOF MS
Proteotypic peptide selection

ceruloplasmin query of GPMdb

>2,100 observations

- Proteotypic peptides
  - Spectral databases (PeptideAtlas & GPMdb), in house MS/MS data
  - complete cleavage products
  - avoid residues susceptible to chemical mods
    - Cys, Met, Trp, N-terminal Gln
Peptide synthesis

- Isotopically labeled peptides synthesized in house
  - 6 peptides in parallel (5 µmol scale)

- $^{13}$C labeled Lys or Arg
  - 6 – 10 amu shift

- Concentration determination by AAA

- CZE & MALDI–TOF purity assessment
  - 96.1% average purity

CZE analysis of ceruloplasmin EYTDASFTNR

MALDI–TOF analysis of ceruloplasmin EYTDASFTNR
Q1/Q3 ion pair optimization

- syringe infusion (300 nL/min) of pure peptides
- peptide ionization optimization
- fragment ion selection
- **co-elution** with SIS peptide to verify natural peptide signal
  - $R_t - 3^{rd}$ analyte specific criteria

- **11.4-fold** average signal increase
  - $>70$ peptides optimized to date

- **Maximizing sensitivity & specificity**
  - multiple ion pairs
  - relative signal intensity of multiple MRM ion pairs (pure vs. matrix)
Concentration Balancing

- approximate 1:1 endogenous levels
- maintain a high quality SIS signal

Haptoglobin beta chain
Benefits of internal standards: Analyte specificity

Fibrinogen gamma chain

Gelsolin isoform I

2.0e4

29 30 31

27 28 29
SIS Peptides Reduce Variation

- 12 replicate injections
  - 14 nL plasma equiv.
  - peak areas of 45 peptides

- IS normalization
  - minimize variation between analyses

- 37/45 peptides <6% CV
Assay Linearity: Calibration Curve & Analyte Response

- **Albumin, serum (575.3 / 937)**
  - Analyte Name: Albumin, serum
  - Calibration Equation: \( y = 0.86729 x + 0.09843 \) \( r = 0.95 \)

- **Haptoglobin beta chain (490)**
  - Analyte Name: Haptoglobin beta chain
  - Calibration Equation: \( y = 0.95562 x + 0.00398 \) \( r = 0.95 \)

- **Fibrinogen gamma chain (40)**
  - Analyte Name: Fibrinogen gamma chain
  - Calibration Equation: \( y = 0.99431 x + 7.21992 \times 10^{-4} \) \( r = 0.95 \)

- **Alpha-1-antichymotrypsin (531.3 / 633.4)**
  - Analyte Name: Alpha-1-antichymotrypsin
  - Calibration Equation: \( y = 0.97762 x + 5.18127 \times 10^{-4} \) \( r = 0.99988 \)

- **Transferrin (815.4 / 693.4)**
  - Analyte Name: Transferrin
  - Calibration Equation: \( y = 1.15681 x - 0.03839 \) \( r = 0.95 \)

- **Complement C3 (501.8 / 731.4)**
  - Analyte Name: Complement C3
  - Calibration Equation: \( y = 0.96274 x + 9.82389 \times 10^{-4} \) \( r = 0.99962 \)

- **r > 0.99 for 42 peptides**
MS Protein Assays by MRM: Quantitation cocktail

Multiplexed quantitation of 45 proteins

Human plasma trypsin digest
1 µg (14 nL) per analysis

IS conc. balanced to match natural abundance

- haptoglobin β IS
- albumin 0.2x
- fibrinogen β
- Apo A-IV
- inter-α-trypsin inhibitor
- natural form
- hemopexin
- transferrin
Plasma Protein Assays by MRM:
Summary

• simultaneous multiplexed protein quant.
  • expandable (>>100 proteins possible)
  • highly reproducible
    • analytical variation <10% for 44/45 peptides

• rapid & sensitive
  • 60 minute analyses per sample
  • Attomole detection sensitivity

• untapped potential applications in protein expression profiling
iMALDI

immuno-MALDI
iMALDI – workflow

1. Serum
2. Trypsin Digest
3. Labeled Peptide (PSA Fragment)
4. MS
5. MS²
6. Bind
7. Anti-PSA Peptide Affinity Beads
8. Elute & analyze
iMALDI – Automation

MALDI spotting

Image:
- Spot On Last Acquired Image
- Spot On This Image
- Path: C:\Program Files\2010\Images\Bruker spotting\2.bmp

Spotting:
- Matrix Reservoir: MCT 1.5mL Left
- Target Plate: Bruker
- Spotting Volume: 100 μL

Wash Dispense Nozzle after Spotting:
- Wash Reagent Reservoir: MCT 0.6mL Left
- Wash Volume: 200 μL
- Times to Wash: 3 times
iMALDI: Signaling pathway assay
iMALDI:
Signaling pathway assay: EGFR

EGFR signal transduction in tumor cells

Tumors with elevated EGFR expression

- Head & neck: 80–100%
- Renal cell: 50–90%
- NSCLC: 40–80%
- Ovarian: 35–70%
- Pancreatic: 30–50%
- Colorectal: 25–77%
- Breast: 14–91%
- Gastric: 40%
iMALDI:
Signaling pathway assay: EGFR

EGFR peptide selection for antibody production

963MHLPSPTDNSFYR^975
m/z = 1564.7

EGFR tryptic peptide
aa 963-975
iMALDI: Signaling pathway assay: EGFR

*in vitro* sensitivity:

- SUM102 cell line
- < 25 cells
**iMALDI:** Signaling pathway assay: EGFR

**in vivo sensitivity:**

- <1/25 of biopsy sample
- from a basal-like primary human tumor BR97-0137b
iMALDI:
Signaling pathway assay: EGFR

Specificity:
MS/MS of EGFR peptide from ~25 SUM102 cells

MS/MS of EGFR peptide from ~25 SUM102 cells
## iMALDI – Capture of EGFR isoforms

Capture of both endogenous isoform peptide and spiked-in natural peptide from 1 ug of U87MGdeltaEGFR cell line lysate

<table>
<thead>
<tr>
<th>Mass (m/z)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1340 - 1430</td>
<td>Mass range</td>
</tr>
<tr>
<td>1349.6047</td>
<td>NYVVTVDHGSCVR - Spiked in natural peptide at 100 fmol, captured by the same antibody in the same experiment</td>
</tr>
<tr>
<td>1406.6235</td>
<td>GNYVVTDHGSCVR - Endogenous isoform peptide captured by iMALDI method</td>
</tr>
</tbody>
</table>

**Graph:**
- **Final - Shots 400 - 6162-02 Started October 1; Run #31; Label**
- **Mass (m/z):** 1340 - 1430
- **Peptide:** GNYVVTDHGSCVR (endogenous), NYVVTVDHGSCVR (spiked-in)

**Legend:**
- **Red Arrow:** Mass 1406.6235
- **Blue Arrow:** Mass 1349.6047
iMALDI Assay

PSA kit

- Labeled Peptide (PSA Fragment)
- Anti-PSA Peptide Affinity Beads
- diaPAC

- Liquid Handling system, MALDI–TOF
  - HT, MS-only, small, “cheap”, diagnostic device
  - high sensitivity & dynamic range
- System
  - Software for clinics not research labs

www.proteincentre.com
iMALDI Assay: Summary

- Only 1 mAb: cheap assay development!
- No dilution series
- Absolute quantitation
- Attomole sensitivity
- MS provides molecular specificity
- Protein isoforms specificity!
- Allows peptide surrogate marker quantification
- Multiplexing with multiple Abs
Acknowledgments

University of North Carolina

- Jiang Jian
- Carol Parker