Application of Proteomics to Cancer Biomarker Discovery

O. John Semmes

Director, Center for Biomedical Proteomics
Department of Microbiology and Molecular Cell Biology
Virginia Prostate Center
Eastern Virginia Medical School Norfolk, VA

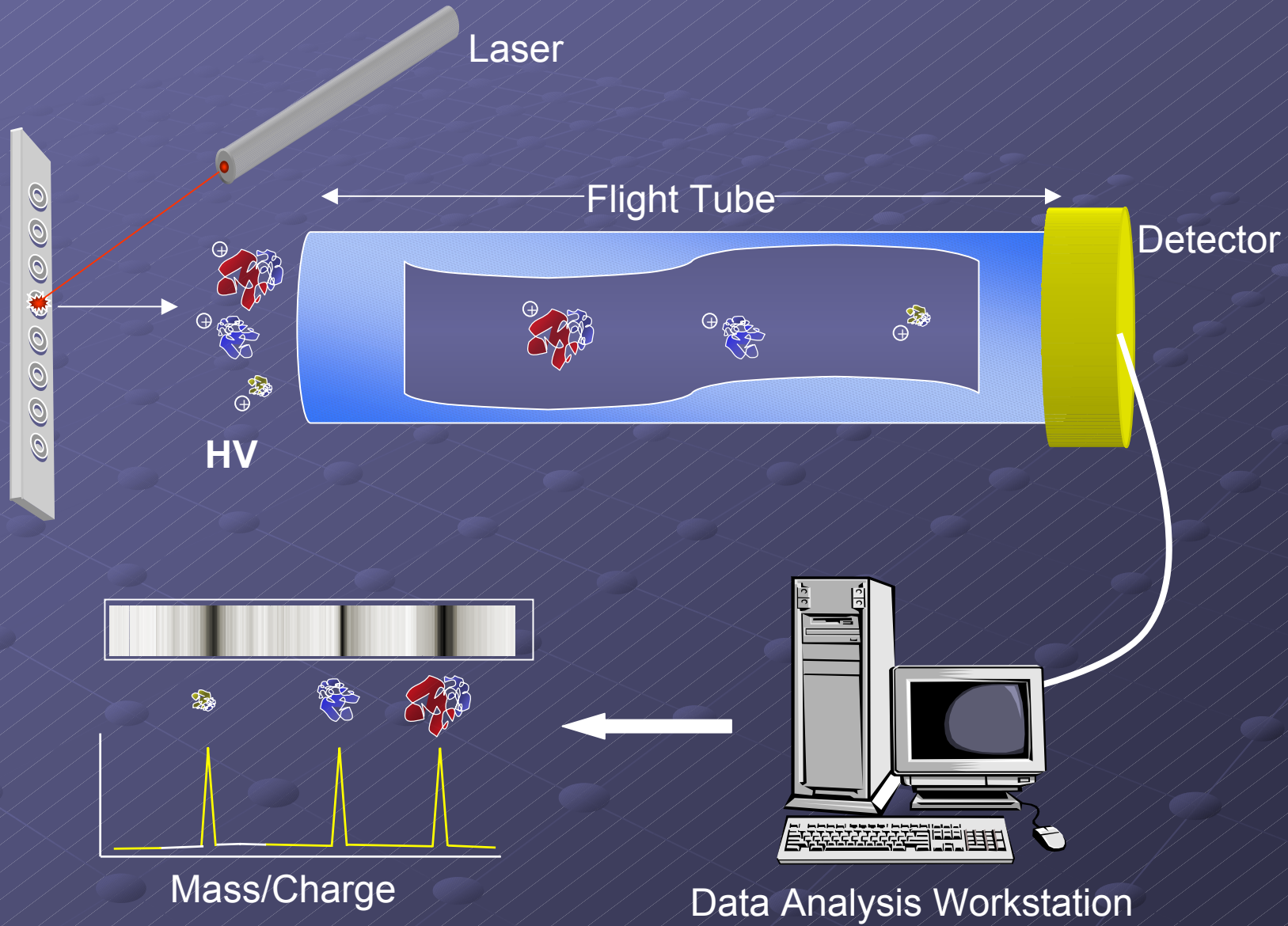
Eastern Virginia Medical School
Established 1973

Eastern Virginia Medical School
Discovery Laboratory
Early Detection Research Network

VIRGINIA PROSTATE CENTER

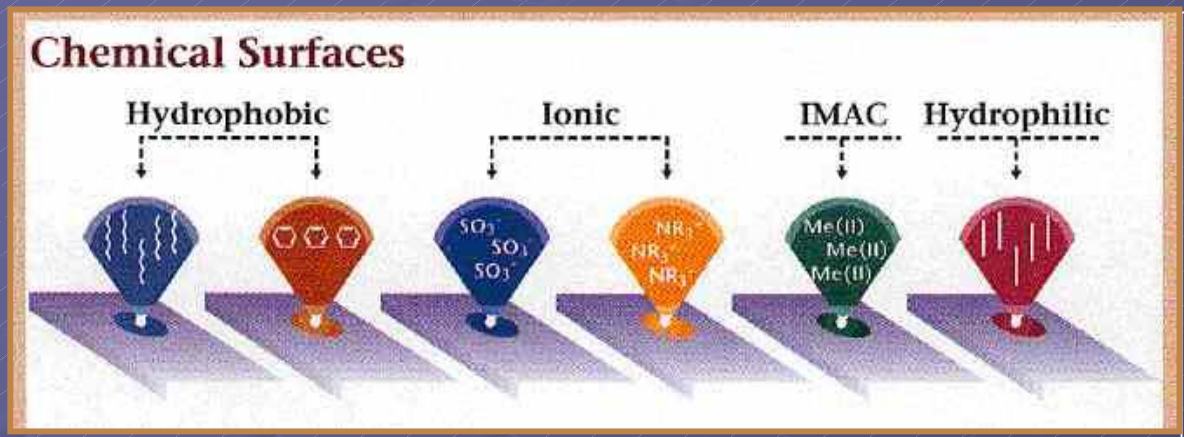
Center For Biomedical Proteomics
Eastern Virginia Medical School

Surface Enhanced Laser Desorption/Ionization (SELDI) Time-Of-Flight (TOF) Mass Spectroscopy



Proteomics Using SELDI Technology

Surface Enhanced Laser Desorption



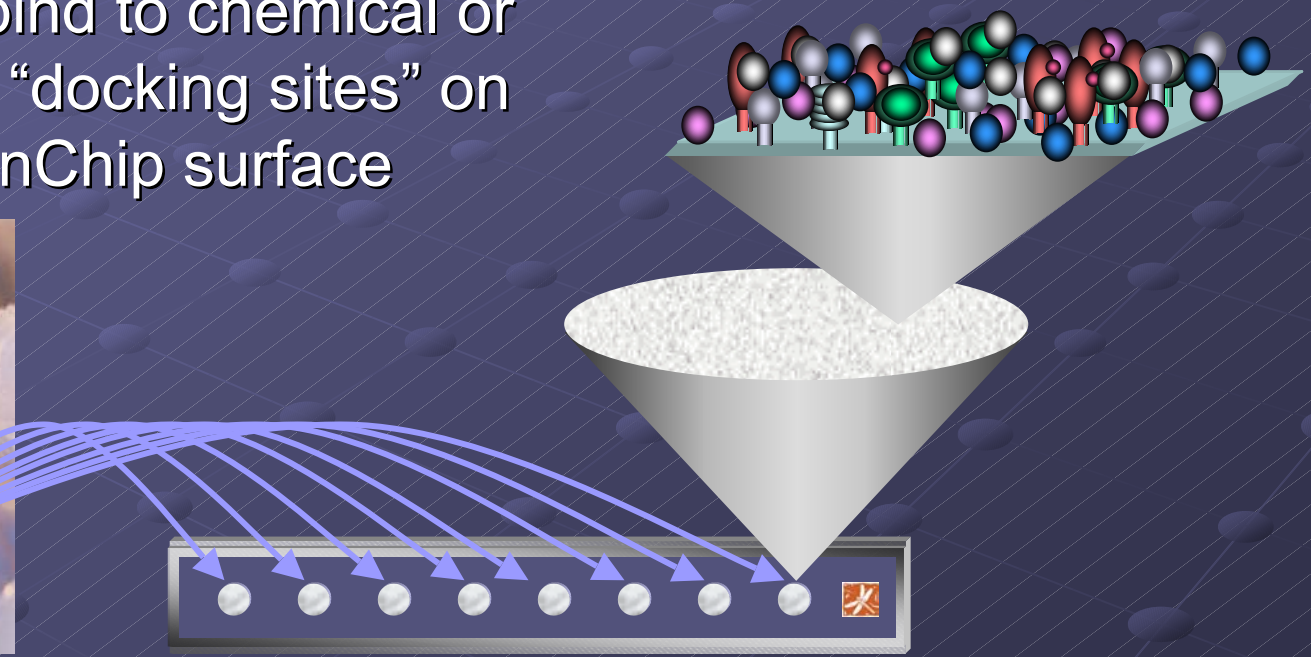
←Surface Chemistries Each chip binds a specific set of proteins based on the chromatographic surface of the ProteinChip®.

←Protein Chips

Each spot on the chip will contain sample from a control or diseased/treated source. The spots are analyzed separately and a mass spectra is created for each spot representing the proteins bound to the chip surface.

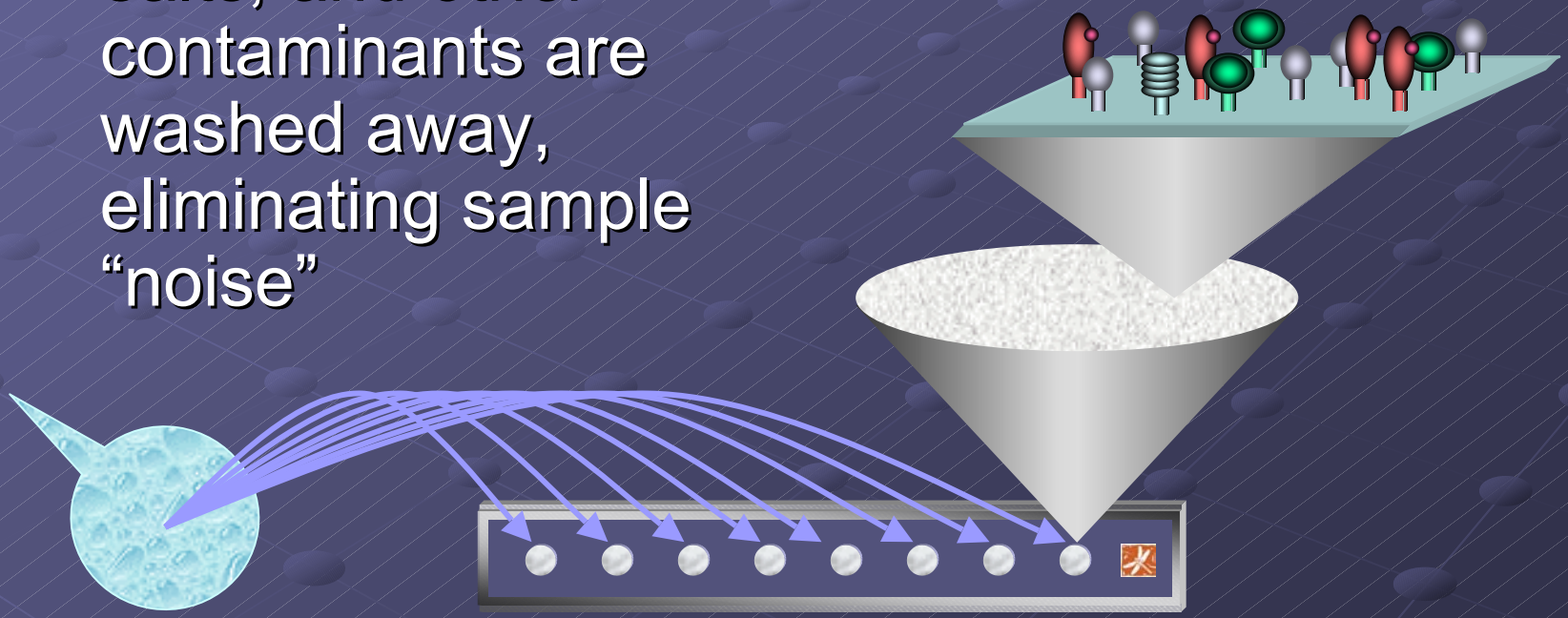
ProteinChip Technology: **Protein Binding**

- Crude sample is placed (and processed) on a ProteinChip Array
- Proteins bind to chemical or biological “docking sites” on the ProteinChip surface

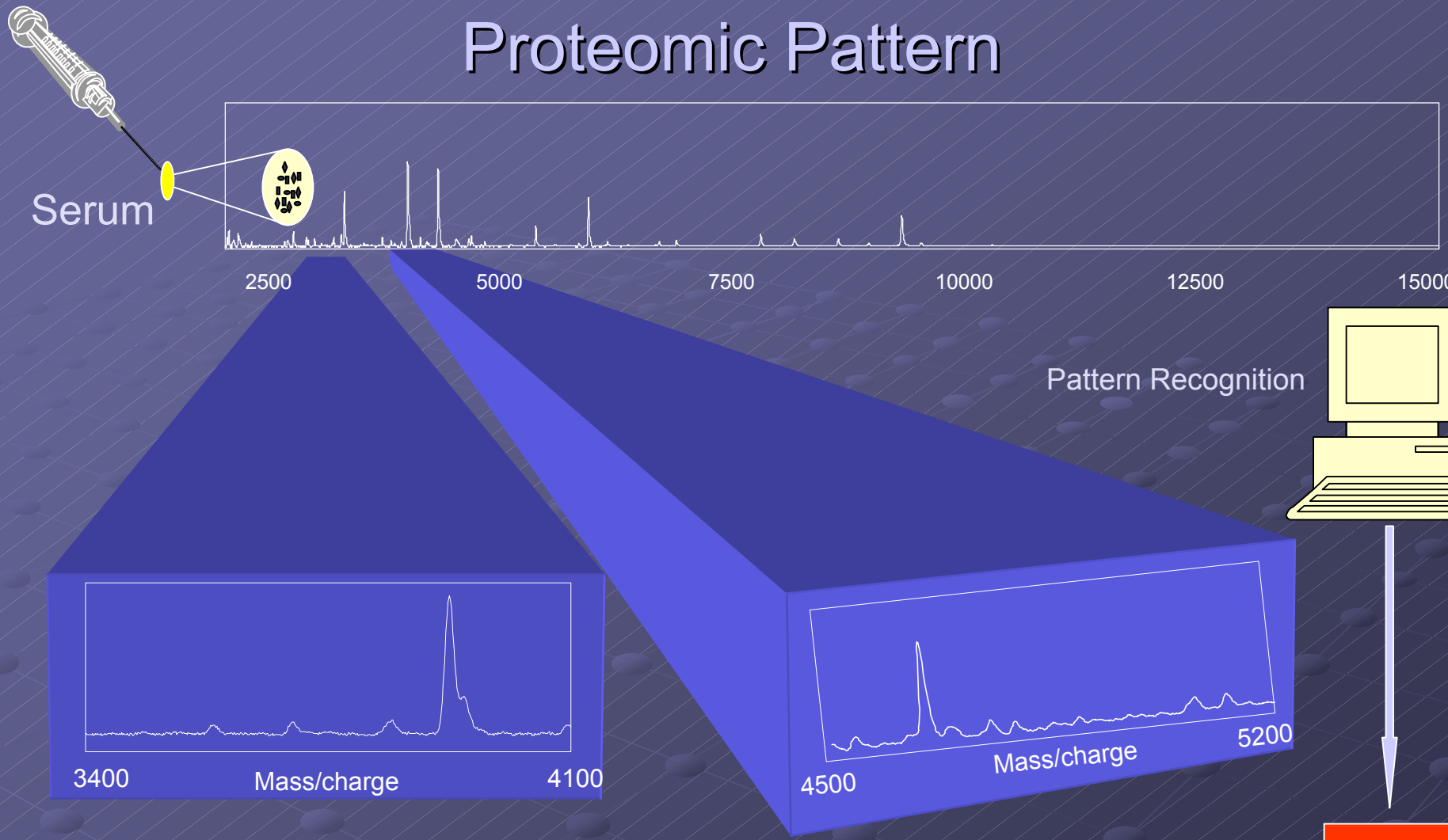


ProteinChip Technology: **Washing Reduces Non-Specific Binding**

- Non-binding proteins, salts, and other contaminants are washed away, eliminating sample "noise"

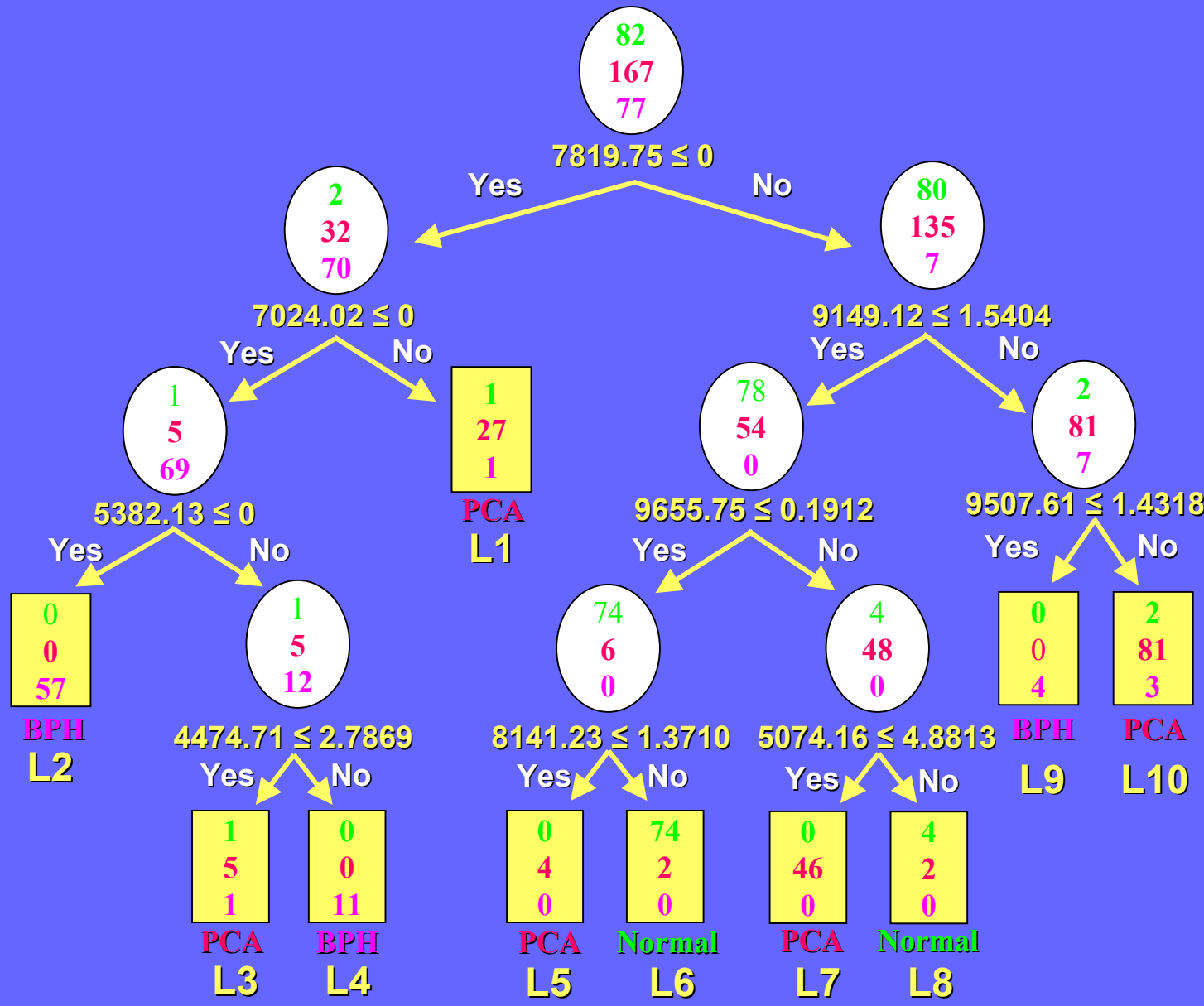


Proteomic Pattern



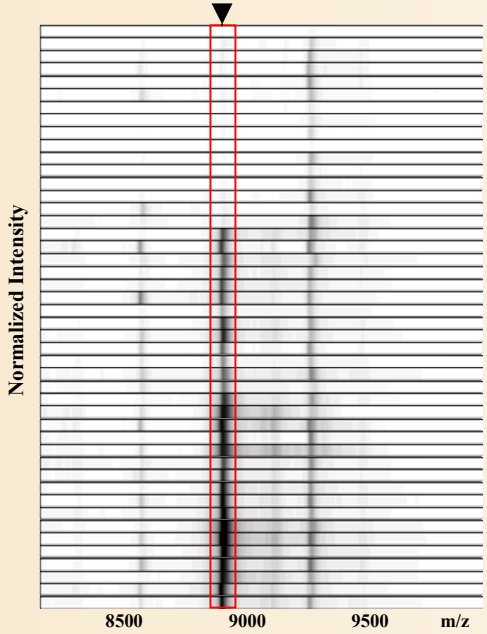
The discriminating pattern formed by a subset of proteins or peptides buried among the entire repertoire of thousands of proteins represented in the serum sample.

DECISION TREE CLASSIFICATION

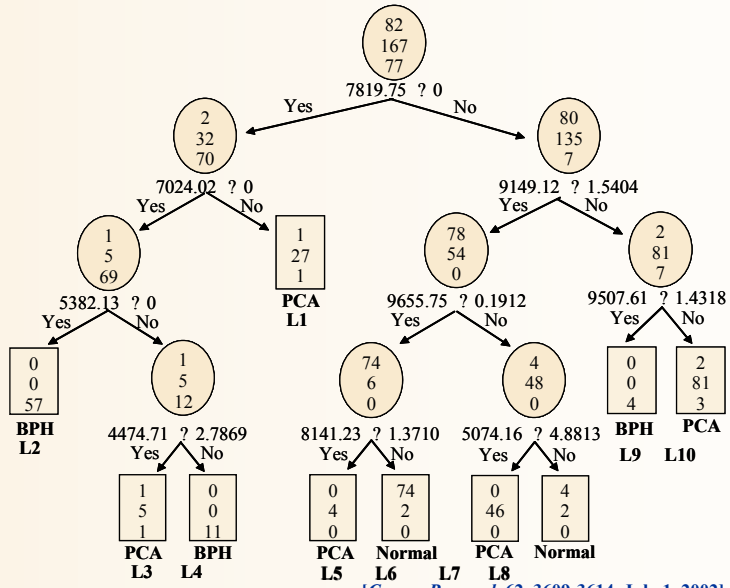


Summary of Biomarker Discovery and Identification

SELDI-TOF

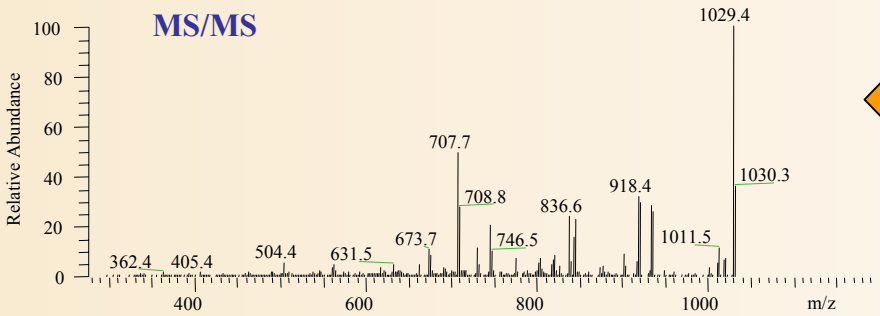


Classification and Regression Tree Analysis

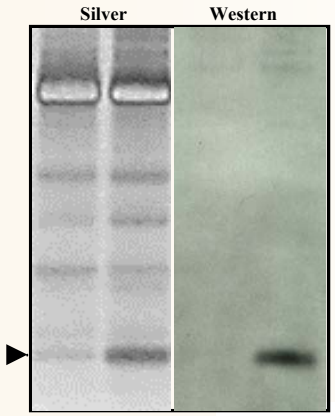


[*Cancer Research* 62, 3609-3614, July 1, 2002]

Identification



Purification

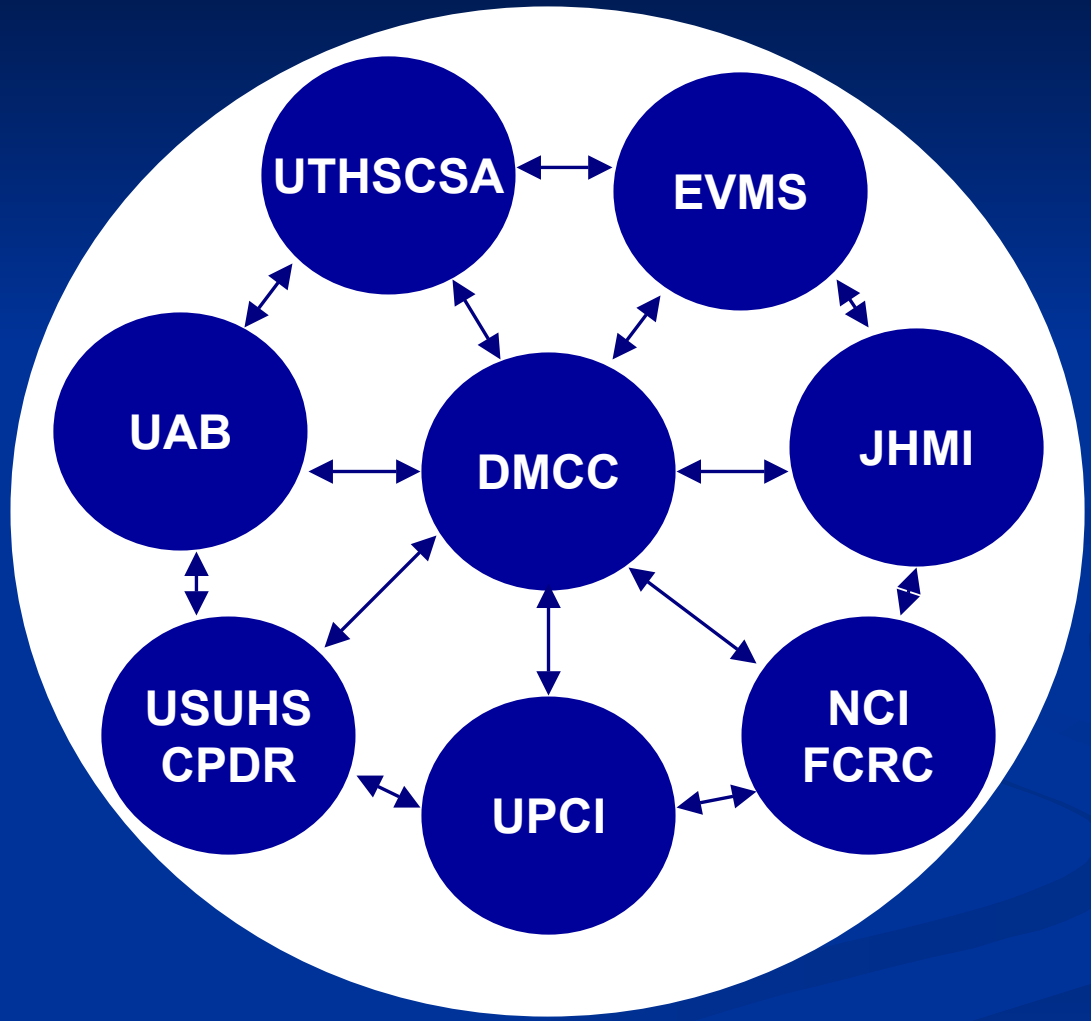


CENTER FOR BIOMEDICAL PROTEOMICS
EASTERN VIRGINIA MEDICAL SCHOOL

Decision Tree Classification of the Blinded Test Set

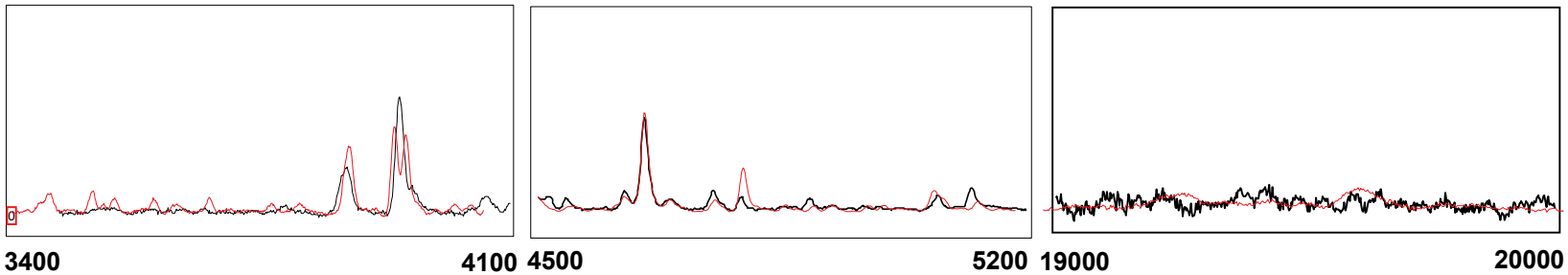
Sample	Normal		BPH		PCA		Misclassified Rate	
Normal (N=15)	15	100.00%	0	0.00%	0	0.00%	0	0.00%
BPH (N=15)	0	0.00%	14	93.33%	1	6.67%	1	6.67%
PCA Stage T1,T2 (N=15)	3	6.67%	0	0.00%	12	80.00%	3	20.00%
PCA Stage T3, T4 (N=15)	1	6.67%	1	6.67%	13	86.67%	2	13.33%
Total Samples (N=60)							6	10.00%

Validation Team

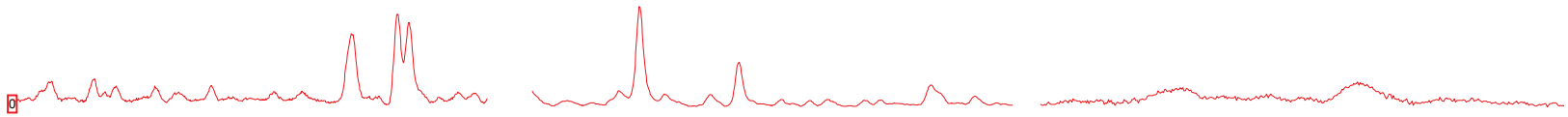


Proteomic Pattern of Sera from Patient

Class 1 Profile

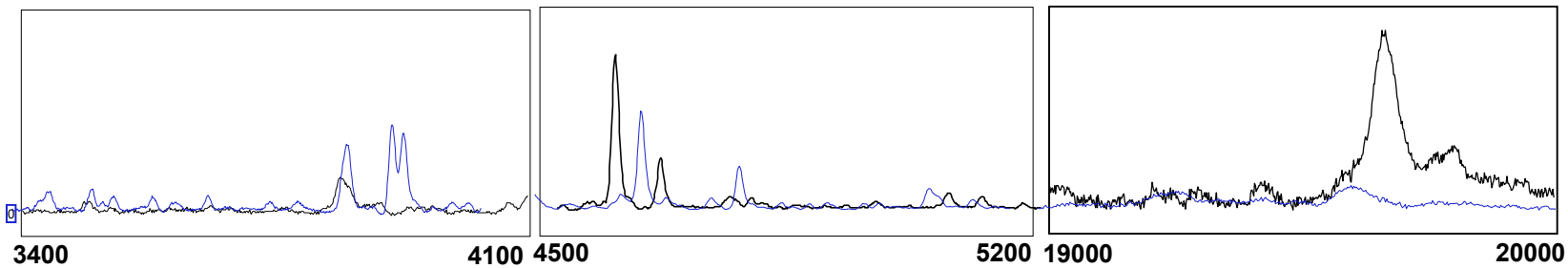


Sample

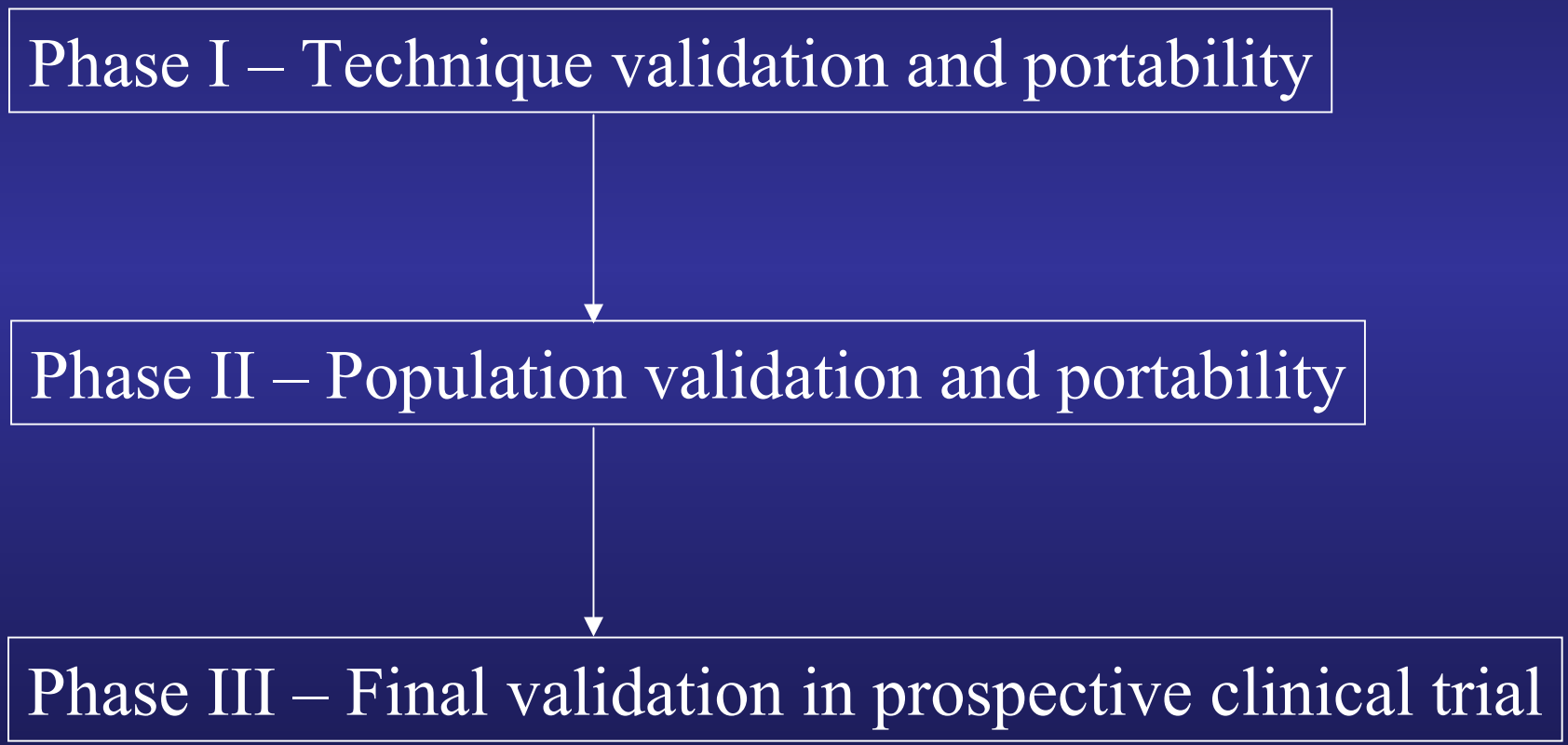


not a C2 pattern

Class 2 Profile



The Prostate SELDI Design



The Prostate SELDI Design

Phase I – Technique validation and portability

Prove we can execute the technique at a number of edrn sites and replicate the discrimination with the algorithm.

Phase II – Population validation and portability

Prove we can 'get the same results' for prostate cancer diagnosis, again at several edrn sites.

Phase III – Final validation in prospective clinical trial

Validate the study in PCPT

The Prostate SELDI Design

Phase I – Technique validation and portability

Synchronize the output of all six SELDI instruments

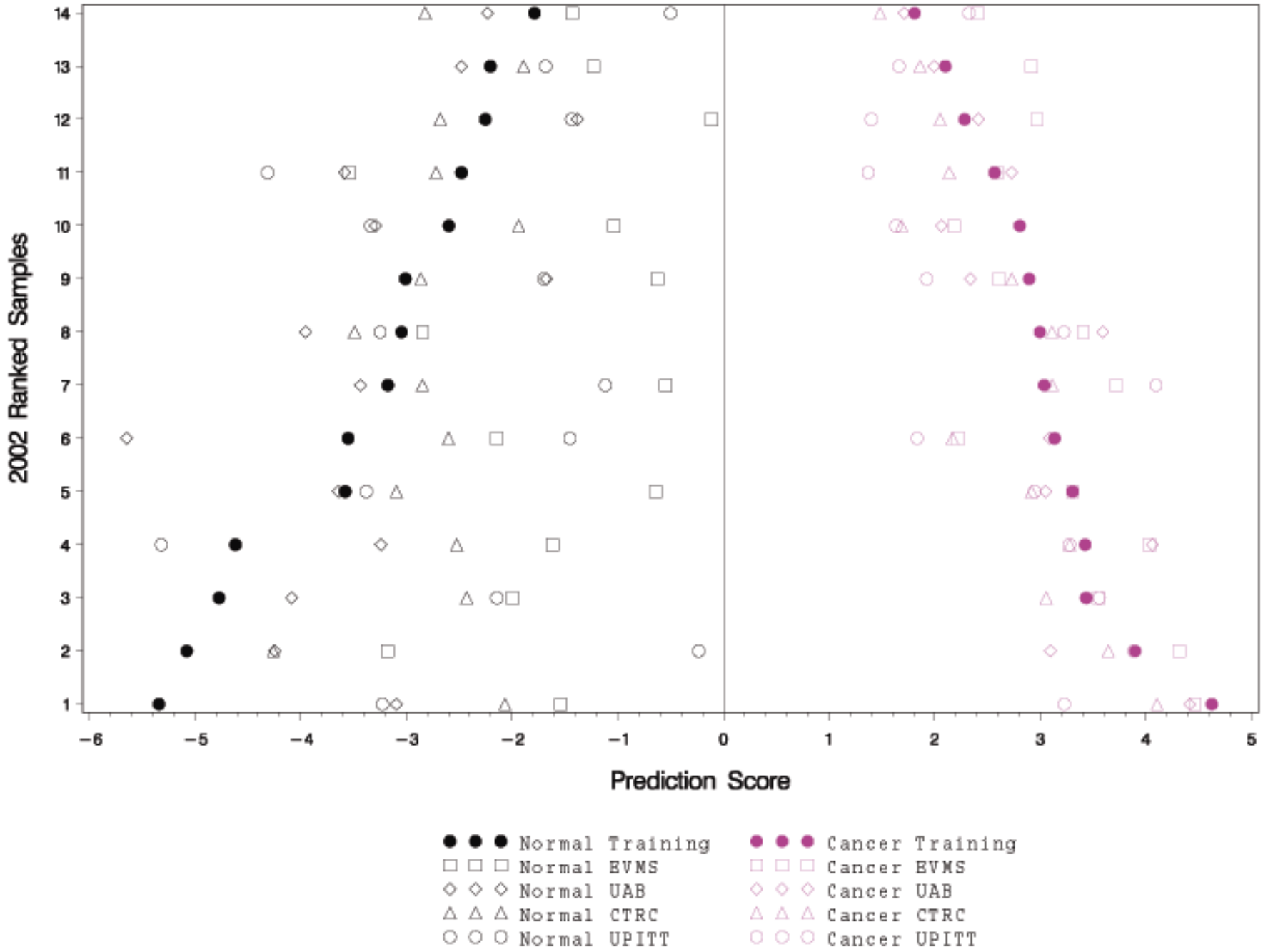
Challenge an existing classification algorithm with case/control data from each site

Data set was collected in 2002

14 PCa and 14 healthy men; must correctly classify 12/14

All sites submitting data (6/6) passed 100% perfect classification

Challenge with case/control collected from five geographically distinct sites


2002 Ranked Samples
14
13
12
11
10
9
8
7
6
5
4
3
2
1
−6
−5
−4
−3
−2
−1
0
1
2
3
4
5
Prediction Score
Normal Training
Normal EVMS
Normal UAB
Normal CTRC
Normal UPITT
Cancer Training
Cancer EVMS
Cancer UAB
Cancer CTRC
Cancer UPITT

EDRN VALIADTION COLLABOARATIVE

Investigators

John Semmes, Ph.D. EVMS BDL
Ian Thompson UTHSSA EDRN CEL
William Grizzle UAB EDRN Validation laboratory
Alan Partin JHU EDRN CEL
Ziding Feng, FHCC EDRN DMCC
William Bigbee, UPCC EDRN BDL
Shiv Srivastava, USUHS
Tim Veenstra, NCI
Bao-Ling Adam, UGA

Eastern Virginia Medical School
Biomarker Discovery Laboratory

Investigators

John Semmes, Ph.D.
John Davis, M.D.
Jose Diaz, M.D., Ph.D.
Rick Drake, Ph.D.
Christine Laronga, M.D.
Paul Schellhammer, M.D.
Jeffery T. Wadsworth, M.D.

Fellows

Alberto Corica, M.D.
Daniel Holterman, Ph.D.
Gunjan Malik, Ph.D.
Lining Qi, Ph.D.

Staff

Diane Brassil
Lisa Cazares
MaryAnn Clements
Tarek Kendil
Brian Main
Michelle Moody
Michael Ward

E
V
M
S

Biostatistics/Computation

WMRI
William Cooke, Ph.D.
Dasha Malaryncko, Ph.D.
Dennis Manos, Ph.D.
Michael Trossett, Ph.D.
Eugene Tracy, Ph.D.

The William & Mary Team

Dariya Malyarenko

Dennis Manos

Michael Trosset

Eugene Tracy

Haijian Chen

Pete Harris

Christine Hopkins

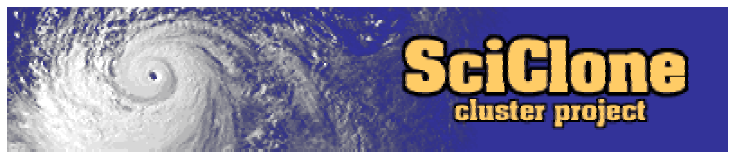
Natalie Pearcy

Amy Wilkerson

Data management and analysis software by

INCOGEN

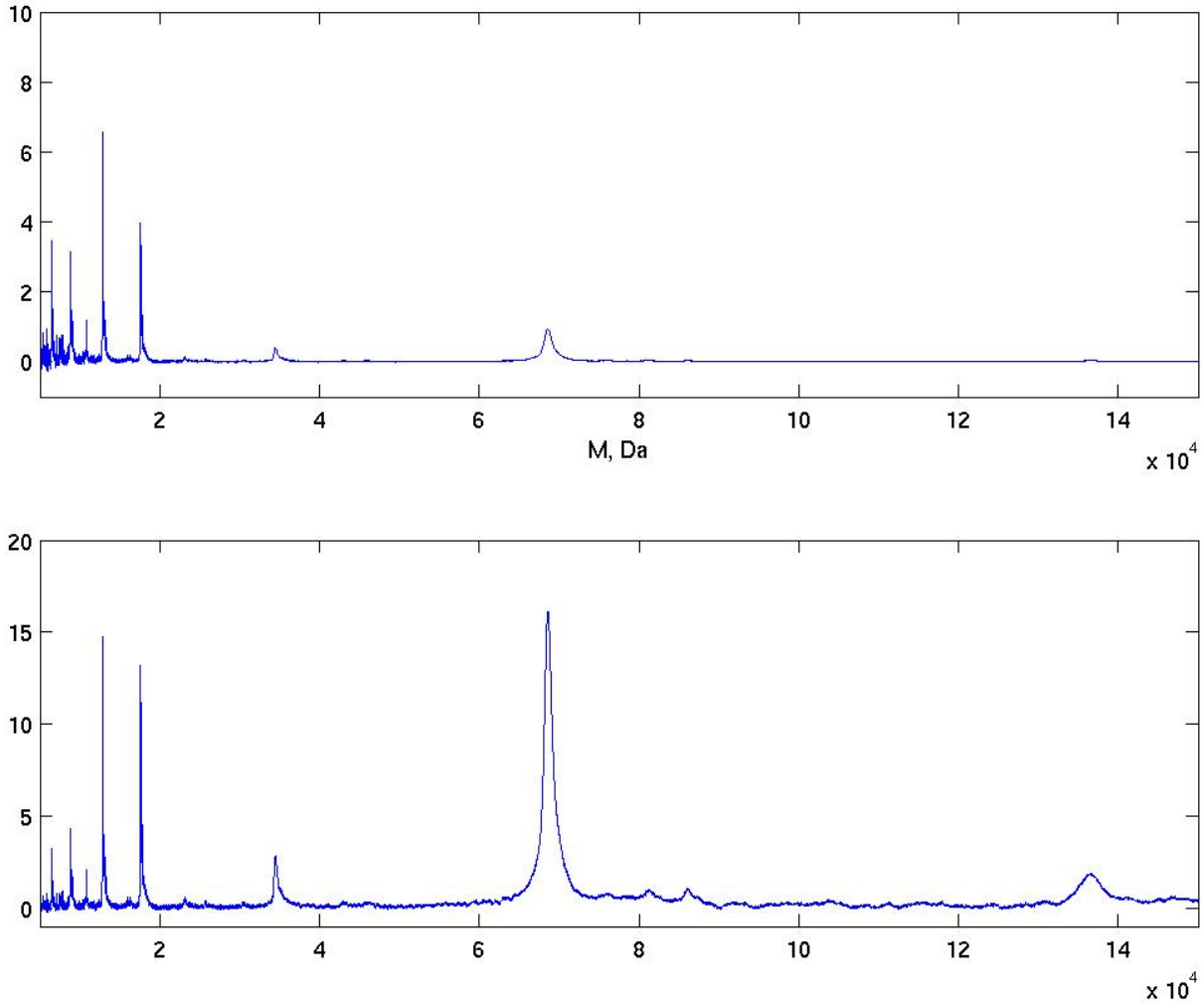
Powered by



William & Mary Goals

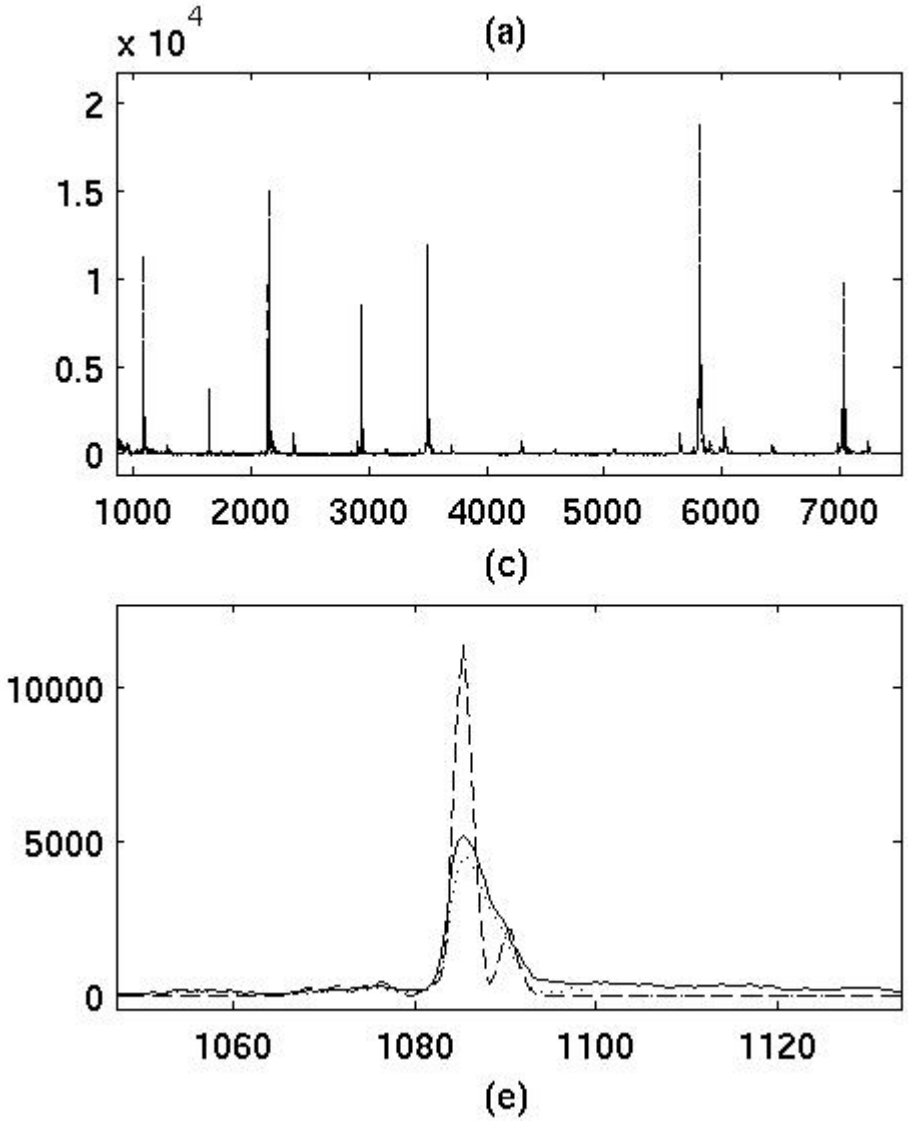
- Enhance Signal/Noise
- Explore classification schemes
- Connect markers to biological processes

Data processing



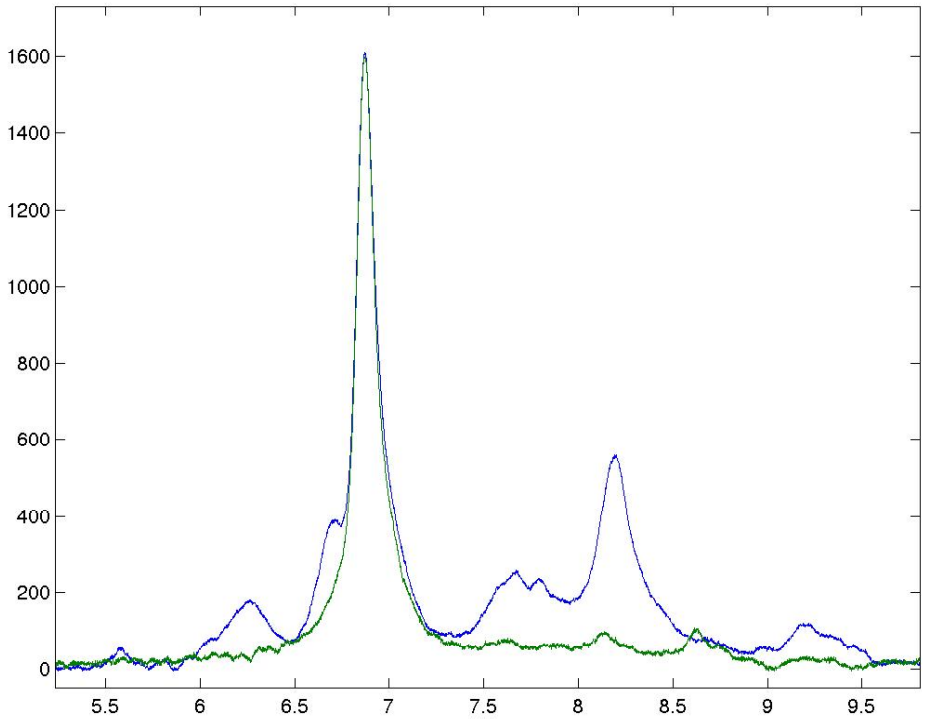
- Rescaling improves peak identification

Filtering



- Filtering resolves hidden peaks

Consider human albumin



- Why so broad?
- Why is there structure?

Limits to resolution

- Time variation

$$\frac{\Delta M}{M} \approx \frac{\Delta t}{T} \xrightarrow[M \to \infty]{} 0$$

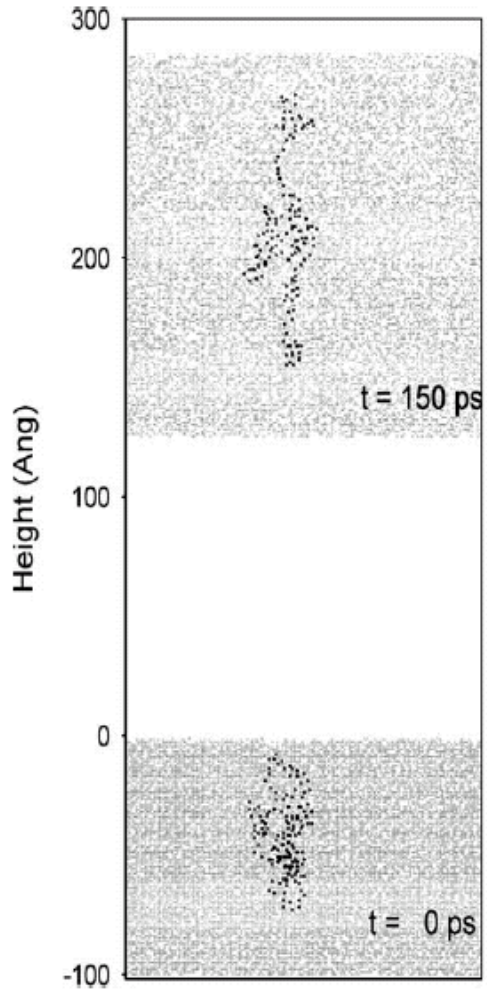
- Length variation (independent of mass)

$$\frac{\Delta M}{M} \approx \frac{\Delta L}{L}$$

Limits to resolution

- Initial velocity

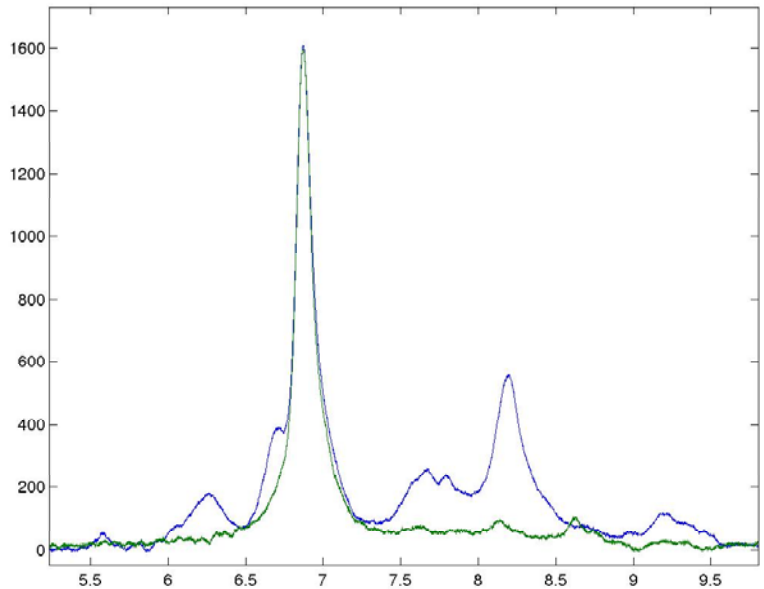
$$\frac{\Delta M}{M} \approx \frac{v_i}{v} < \frac{100\ \text{m/s}}{3000\ \text{m/s}}$$



B.J. Garrison, A. Delcorte, L.V. Zhigilei, T.E. Itina, K.D. Krantzman, Y.G. Yingling, C.M. McQuaw, E. J. Smiley, and N. Winograd, *Appl. Surf. Sci*. 203-204, (2003).

Limits to resolution

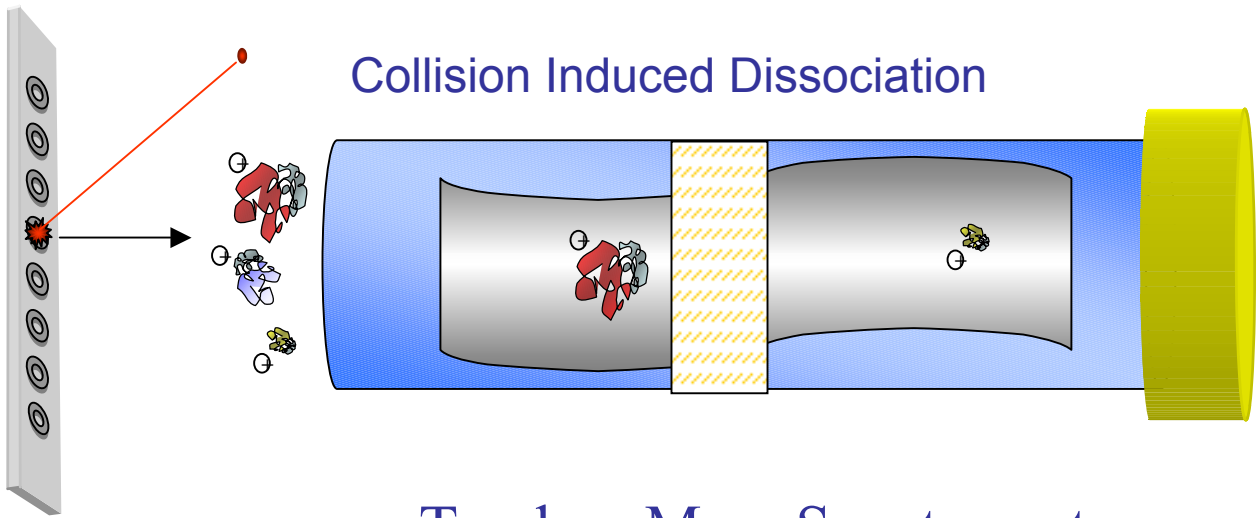
- Biology! Proteins come with modifications:
  Phosphorylated
  Acetylated
  Methylated
  Sulfated



How can you tell if a shoulder is a peak or a problem?

MS/MS to the rescue

- Can we distinguish biologically different molecules with nearly the same mass?



Optical Ionization/Dissociation would be better

(Compared to Collision-Induced Dissociation)

- Wavelength tunability and intensity variation can enhance biological selectivity
- Specifics of fragmentation are unimportant, as long as there is a difference!
- Photons do not change the ion speed.

The upgraded FEL is the perfect post-ionizer

- High rep rate makes it quasi continuous

$$\Delta x = \left(3000\ \text{m/s}\right)\left(53\ \text{ns}\right) = 170\,\mu\text{m}$$

As long as the laser waist is no larger than 170 microns, every ion will see a pulse.

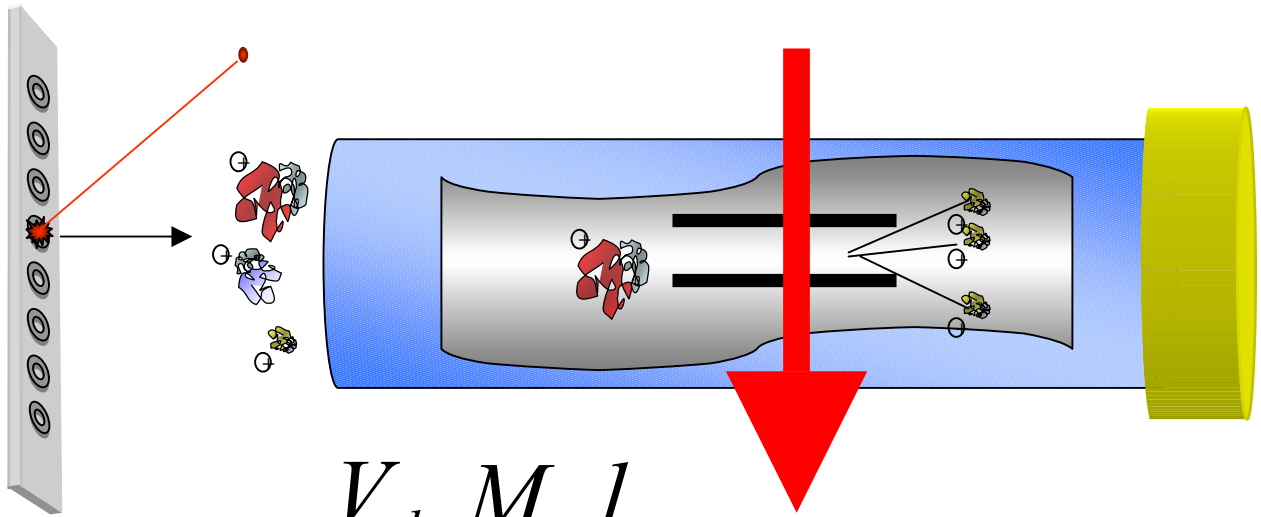
The upgraded FEL is the perfect post-ionizer

- High peak power enhances multi-photon ionization or fragmentation

$$I = \frac{100\,\mu\mathrm{J}}{(1\ \mathrm{mm})^2\,(0.5\ \mathrm{ps})} = 20\ \mathrm{GW/cm^2}$$

The upgraded FEL is the perfect post-ionizer

- Simultaneously obtain survey MS ***with*** fragmentation. The data density goes up!



$$\Delta\Phi \approx \frac{V_d}{V_a}\frac{M}{m}\frac{l}{d}$$

Proteomics for Cancer Biomarker Discovery

- EVMS methodology is already working.
- Improved S/N is within sight.
- Data lives in a world with very high dimensions.
- The upgraded FEL could significantly advance the connection of biomarkers to biological processes.