



Comparative Analysis of the Human and Mouse Serum Proteomes

Brian L. Hood

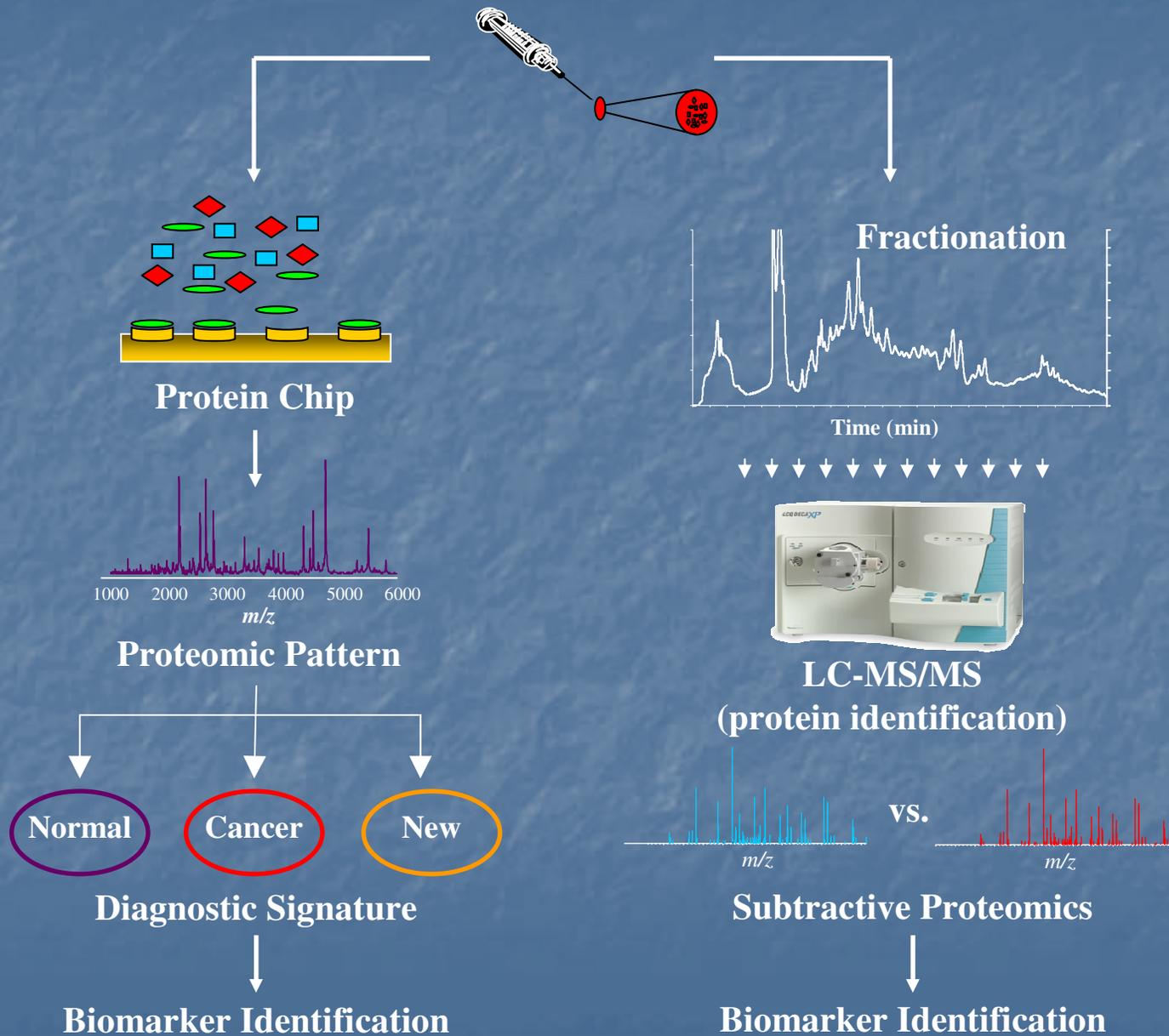
Laboratory of Proteomics and Analytical Technologies

SAIC- Frederick, Inc.

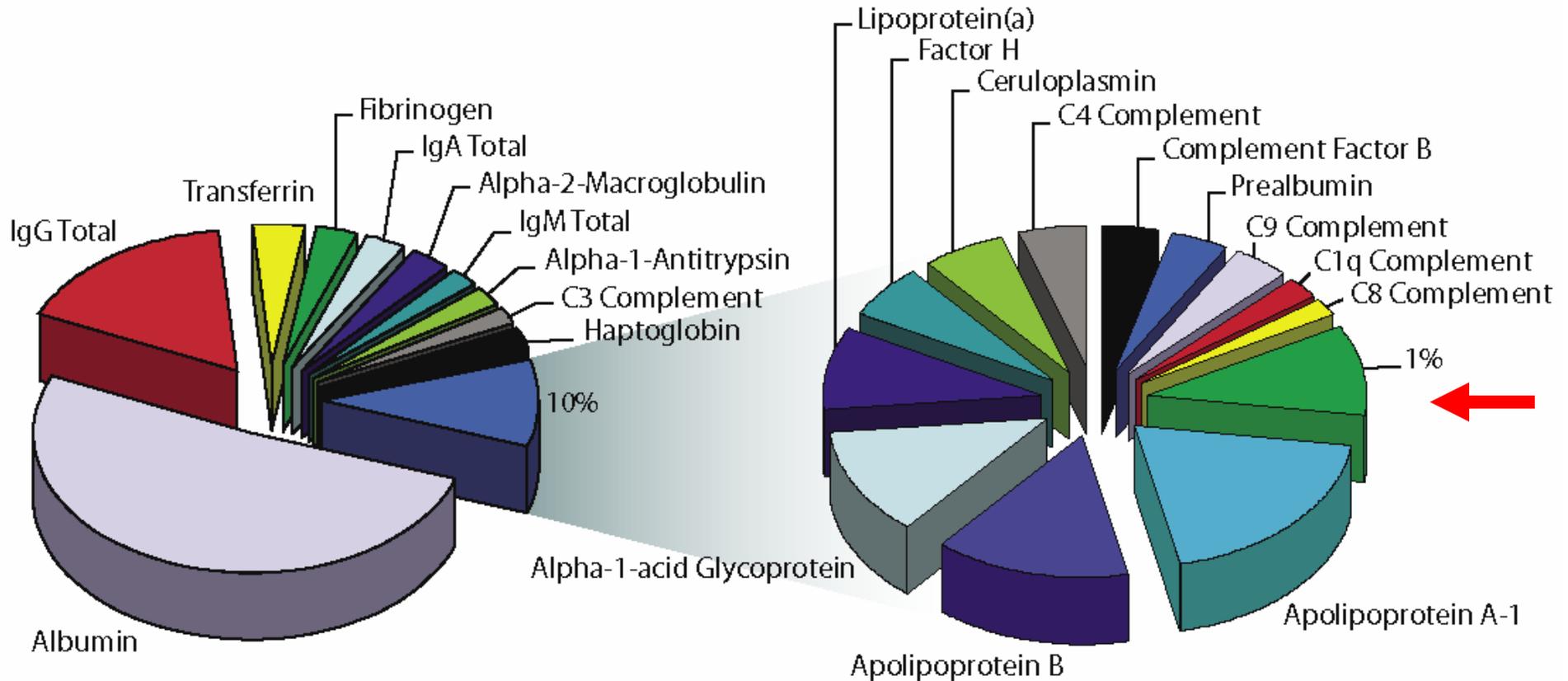
National Cancer Institute - Frederick



Different Paths To Biomarker Discovery



22 Proteins Comprise 99% Of The Protein Mass In Serum



90%

10%

Human Serum Proteomic Investigations

- Global human serum proteome survey

Can we account for the presence of disease and cellular process-related proteins in serum?

- Low molecular weight protein/peptide proteome

Can we deplete the high molecular weight fraction for more effective interrogation of the source of the diagnostic information?

- Investigation of bound peptides to high abundant serum proteins

Is there histopathological content bound to the highly abundant carrier proteins, such as albumin?

Human Serum Proteomic Investigations

- Global human serum proteome survey

Can we account for the presence of disease and cellular process-related proteins in serum?

- Low molecular weight protein/peptide proteome

Can we deplete the high molecular weight fraction for more effective interrogation of the source of the diagnostic information?

- Investigation of bound peptides to high abundant serum proteins

Is there histopathological content bound to the highly abundant carrier proteins, such as albumin?

Global Human Serum Proteomic Analysis Strategy



200 μ L Serum



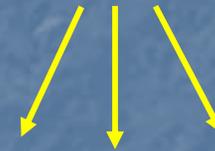
Digest Proteins
with Trypsin



...KHIYTPLINHGFRITG...

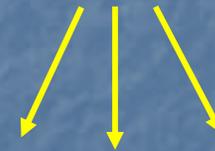


...K + HIYTPLINHGFR + ITG...



Peptide Isoelectric
Focusing

20 Fractions



Strong Cation
Exchange HPLC

140 Fractions

Reversed-phase microcapillary liquid chromatography
online with tandem mass spectrometry



Analysis Of The Global Human Serum Proteome

IEF

473 Proteins
957 Unique Peptides

IEF/SCX

1143 Proteins
2071 Unique Peptides

Total Proteins and Peptides Identified

1444 Unique Proteins
2646 Unique Peptides

Analysis of the Human Serum Proteome

King C. Chan, David A. Lucas, Denise Hise, Carl F. Schaefer, Zhen Xiao, George M. Janini, Kenneth H. Buetow, Haleem J. Issaq, Timothy D. Veenstra and Thomas P. Conrads *Clinical Proteomics* (2004) 1, 1-225.

Human Serum Proteomic Investigations

- Global human serum proteome survey

Can we account for the presence of disease and cellular process-related proteins in serum?

- Low molecular weight protein/peptide proteome

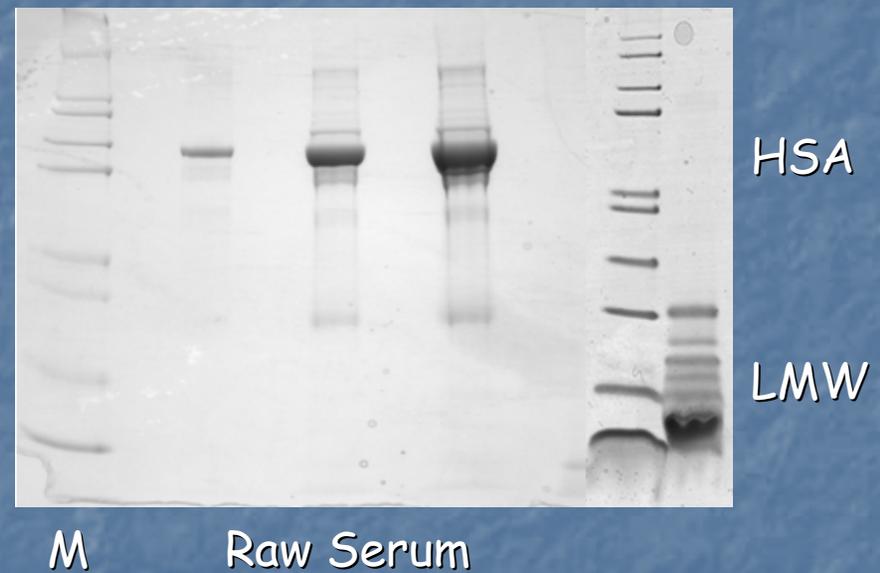
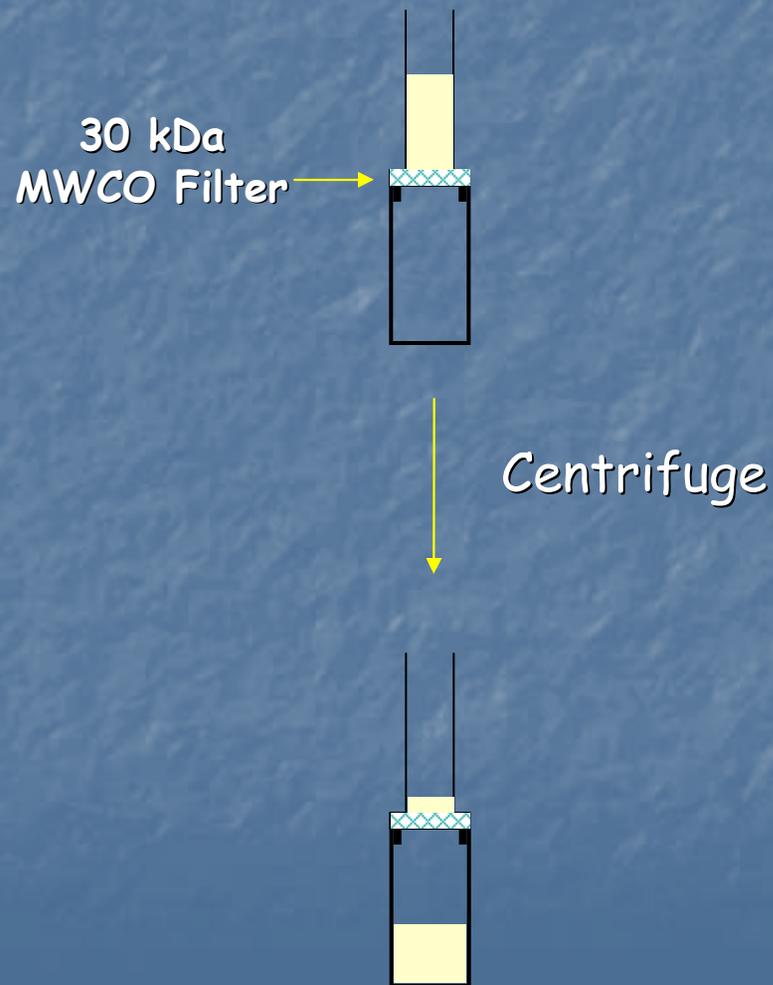
Can we deplete the high molecular weight fraction for more effective interrogation of the source of the diagnostic information?

- Investigation of bound peptides to high abundant serum proteins

Is there histopathological content bound to the highly abundant carrier proteins, such as albumin?

High Molecular Weight Protein Depletion Using Ultrafiltration

Dilute raw serum 1:5 in
25 mM NH_4HCO_3 , pH 8.2 / 20%
acetonitrile



Analysis Of The LMW Human Serum Proteome

LMW

340 Proteins
808 unique peptides

High molecular weight proteins were identified -
serum is comprised of a variety of proteolytic
fragments from larger proteins

Characterization of the Low Molecular Weight Human Serum Proteome

Radhakrishna S. Tirumalai, King C. Chan, DaRue A. Prieto, Haleem J. Issaq,

Thomas P. Conrads and Timothy D. Veenstra *Molecular and Cellular Proteomics* (2003) 2.10, 1096-1103.

Human Serum Proteomic Investigations

- Global human serum proteome survey

Can we account for the presence of disease and cellular process-related proteins in serum?

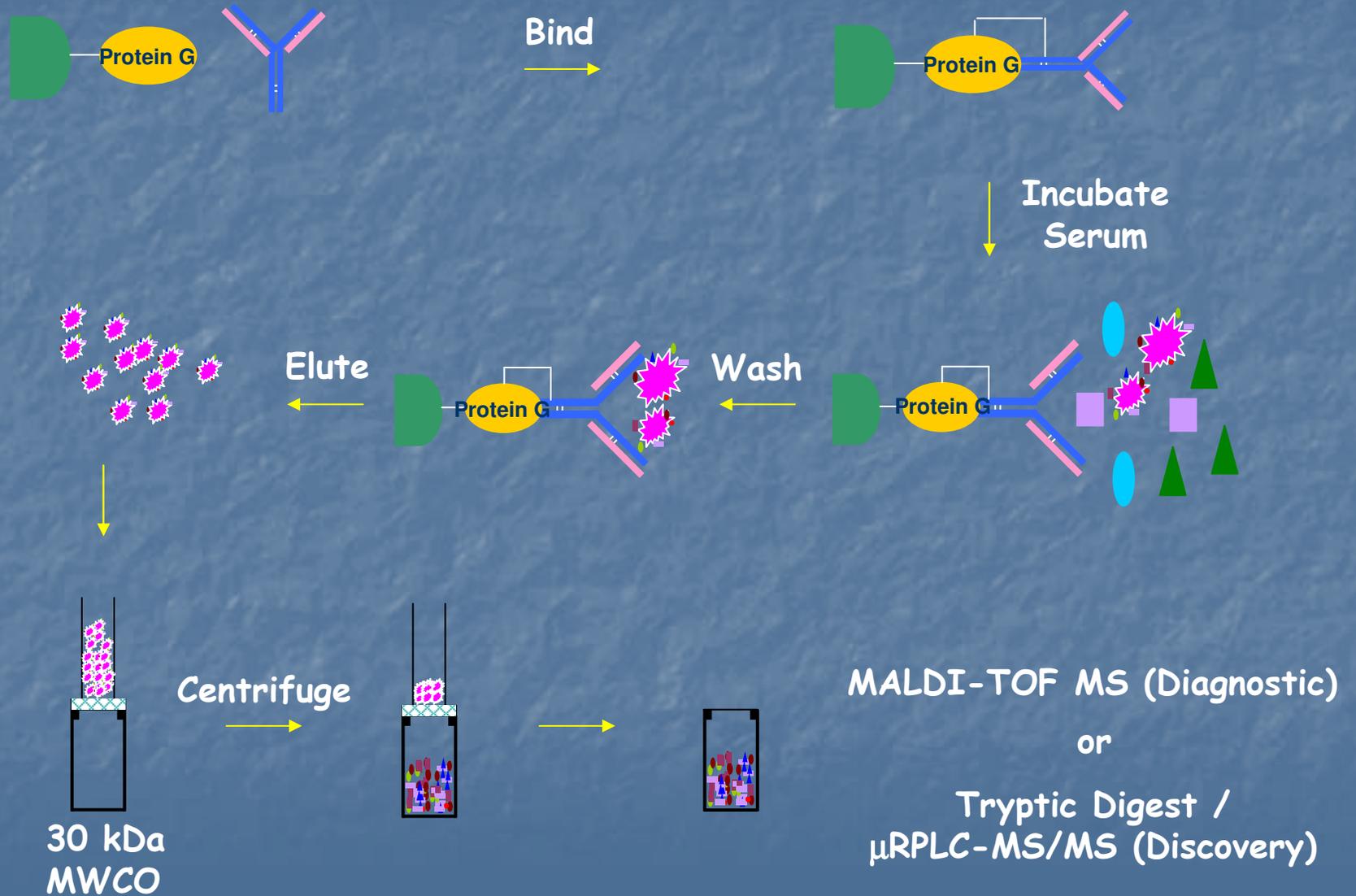
- Low molecular weight protein/peptide proteome

Can we deplete the high molecular weight fraction for more effective interrogation of the source of the diagnostic information?

- Investigation of bound peptides to high abundant serum proteins

Is there histopathological content bound to the highly abundant carrier proteins, such as albumin?

Human Serum 'Interactome' Investigation



Human Serum 'Interactome' Investigation

<u>Protein</u>	<u>Accession</u>	<u>Association</u>	<u>Mr (kDa)</u>
Glycosylasparaginase	P20933	Albumin	37.2
Paraneoplastic antigen MA1	O95144	IgA	39.8
Meningioma-expressed antigen 6/11	O15320	Albumin	90.9
Dihydropteridine reductase	P09417	Apolipoprotein	25.8
Coagulation factor VII precursor	P08709	Albumin	51.6
ASM-like phosphodiesterase 3a	Q92484	Apolipoprotein	51.3
Prostate transglutaminase	P49221	IgA, IgG	77.1
Pregnancy-plasma protein-A	Q13219	IgG, IgM	181.0
Hsc70-interacting protein	P50502	Albumin	41.3
Ryanodine receptor 2	Q92736	Albumin	564.6
Bone morphogenetic protein 3b	P55107	IgA	53.1
Prostate-specific antigen	P07288	Albumin, IgG	28.7

High abundant serum proteins may act as 'molecular sponges'

Human Serum 'Interactome' Analysis

IP

210 Proteins
378 unique peptides

67% not observed in the global human serum proteome
73% not observed in LMW serum proteome

Twelve proteins currently used as clinical biomarkers

An Investigation into the Human Serum “Interactome”

Ming Zxou, David A. Lucas, King C. Chan, Haleem J. Issaq, Emanuel F. Petricoin III,
Lance A. Liota, Timothy D. Veenstra and Thomas P. Conrads *Electrophoresis* (2004) 25, 1289-1298.

Global Analysis Of The Serum Proteome

Identification of proteins from all functional classes and cellular locations.

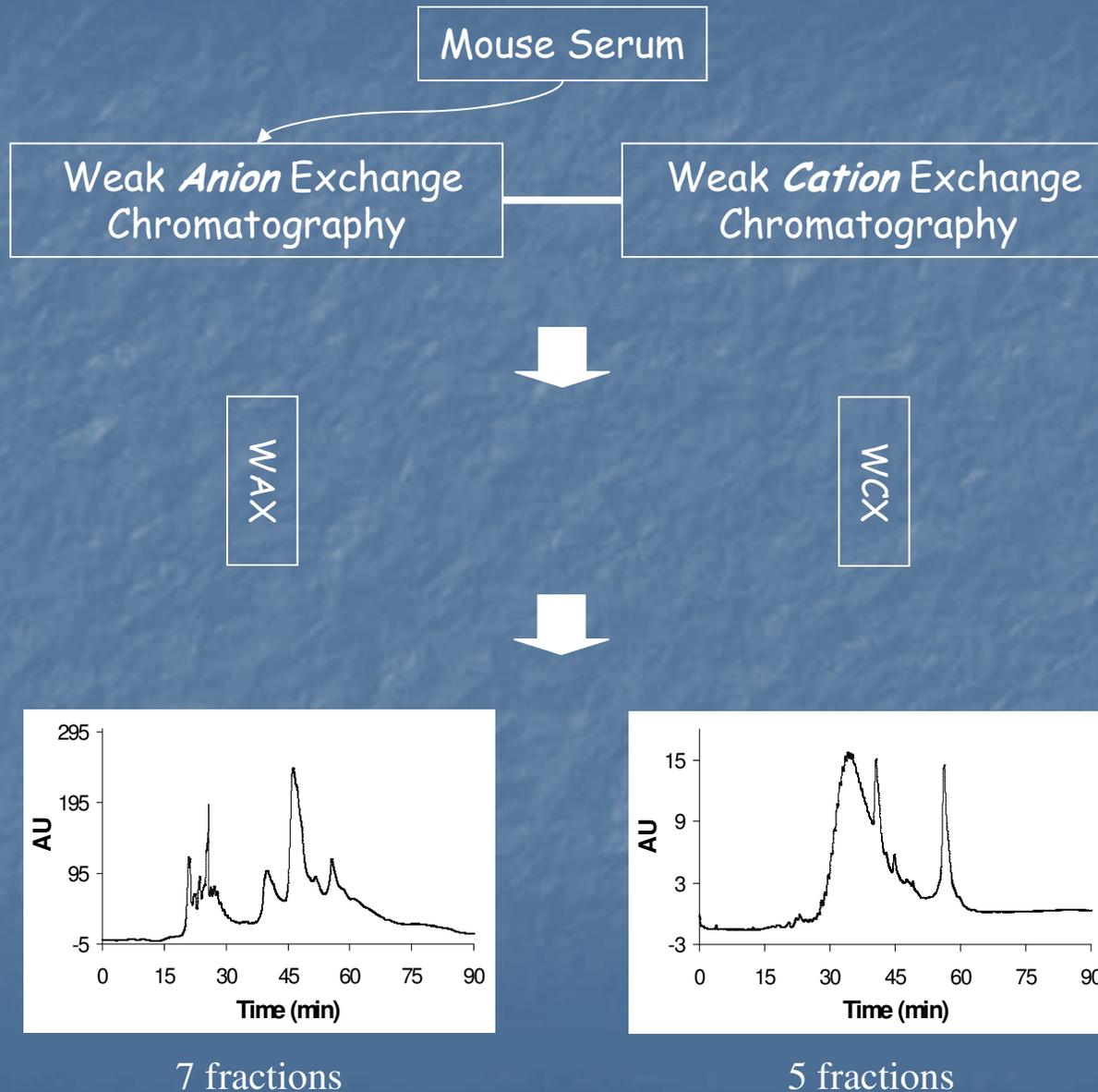
Clearly an archive of diagnostic information exists within serum.

Can we use serum proteome analyses as a tool for the identification of clinically relevant biomarkers or for the timely diagnosis of diseased states? Where do we start?

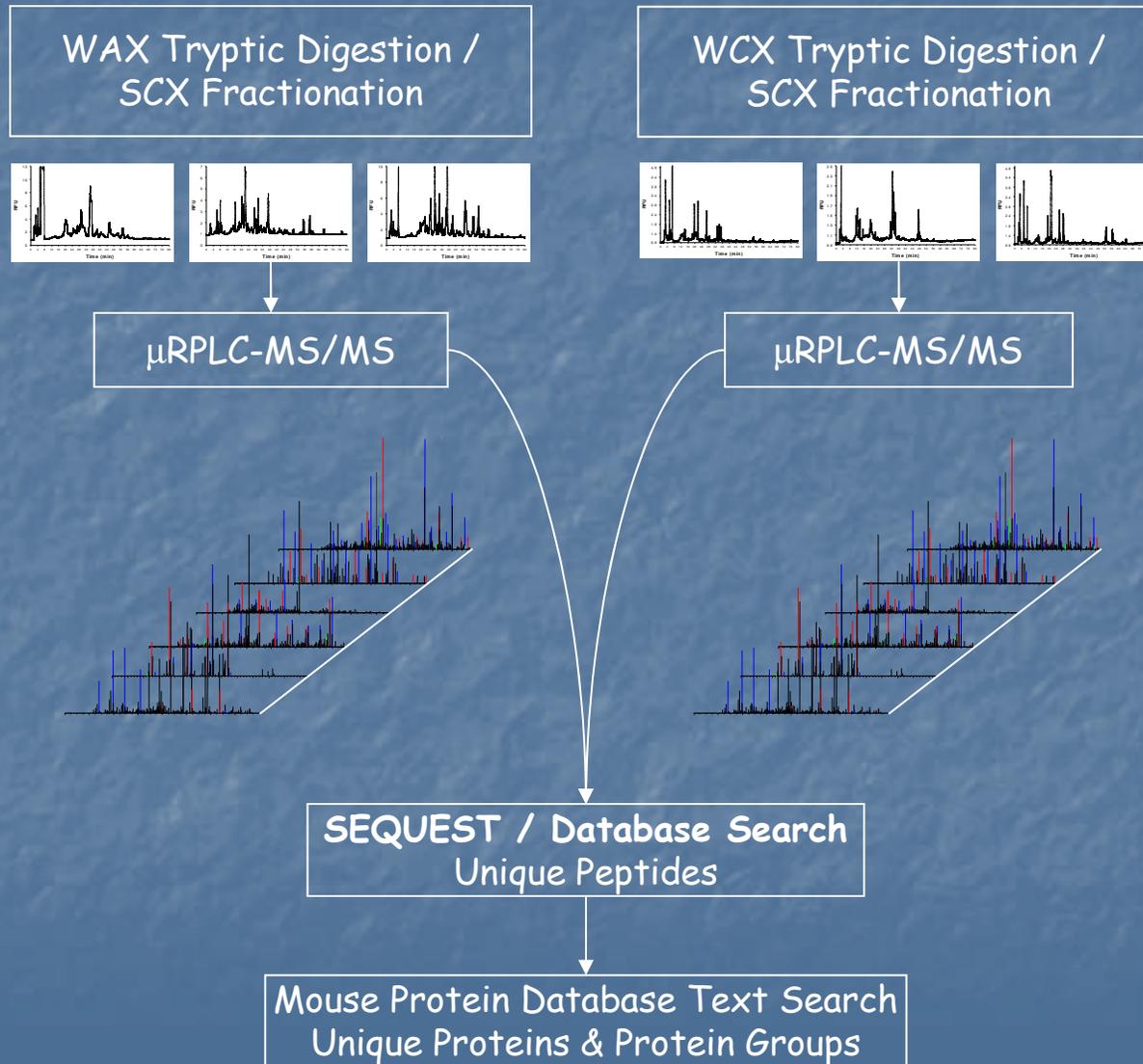
Human Versus Mouse Serum Proteomes

- Mouse models are extensively used for studying human disease
 - Genetic variability can be controlled
 - Reduce/minimize experimental artifacts
 - Fundamental differences in tumorigenesis
- Mouse serum proteome is largely uncharacterized
 - Mouse genome is 14% smaller than the human genome
 - Very similar in content and chromosomal organization
 - 80% of mouse coding regions have a direct orthologue in the human genome
- Evaluation of these similarities is **fundamental** to the comparative question
- Global characterization of mouse proteome using a multidimensional fractionation approach on the protein and peptide level without depletion of highly abundant proteins

Global Mouse Serum Proteomic Analysis Strategy



Global Mouse Serum Proteomic Analysis Strategy



Global Mouse Serum Proteome

4,565 Unique Proteins
12,384 Unique Peptides

2,909 Protein Groups

1,399 containing
unique protein

1,510 without
unique protein

ABCC Peptide Exact Match Form

Address <http://wwwdev.abcc.ncifcrf.gov/app/htdocs/proteomics/plainText.php?file=/abcc/users/web/95023.fasta.exact.peptides&outputMode=HTML&> Go Links >>

List of protein hits by peptide:

pep1	Q8K1F0	←	Unique protein identifier - SPIKES - Unique Proteins
pep2	HV56_MOUSE	←	
pep3	HV17_MOUSE	←	
pep3	HV18_MOUSE	←	Non-unique protein identifier - Protein Groups
pep3	HV19_MOUSE	←	
pep3	HV20_MOUSE	←	
pep3	HV21_MOUSE	←	
pep3	HV22_MOUSE	←	
pep3	HV23_MOUSE	←	
pep3	HV24_MOUSE	←	
pep3	HV25_MOUSE	←	
pep4	Q920E7	←	
pep5	MRTB_MOUSE	←	
pep6	APM1_MOUSE	←	
pep7	I10R_MOUSE		
pep8	Q8C6Q3		
pep9	PDX6_MOUSE		
pep10	GTP1_MOUSE		
pep11	HBB1_MOUSE		
pep11	HBB2_MOUSE		
pep11	Q9ROS6		
pep11	Q91V86		
pep12	HBA_MOUSE		
pep12	Q91VB8		
pep13	Q811P8		
pep14	K6PP_MOUSE		
pep15	Q8K0E8		
nen16	NMF4_MOUSE		

Done Internet

Global Mouse Serum Proteome

~~6,075~~ Unique Proteins
12,384 Unique Peptides

2,909 Protein Groups

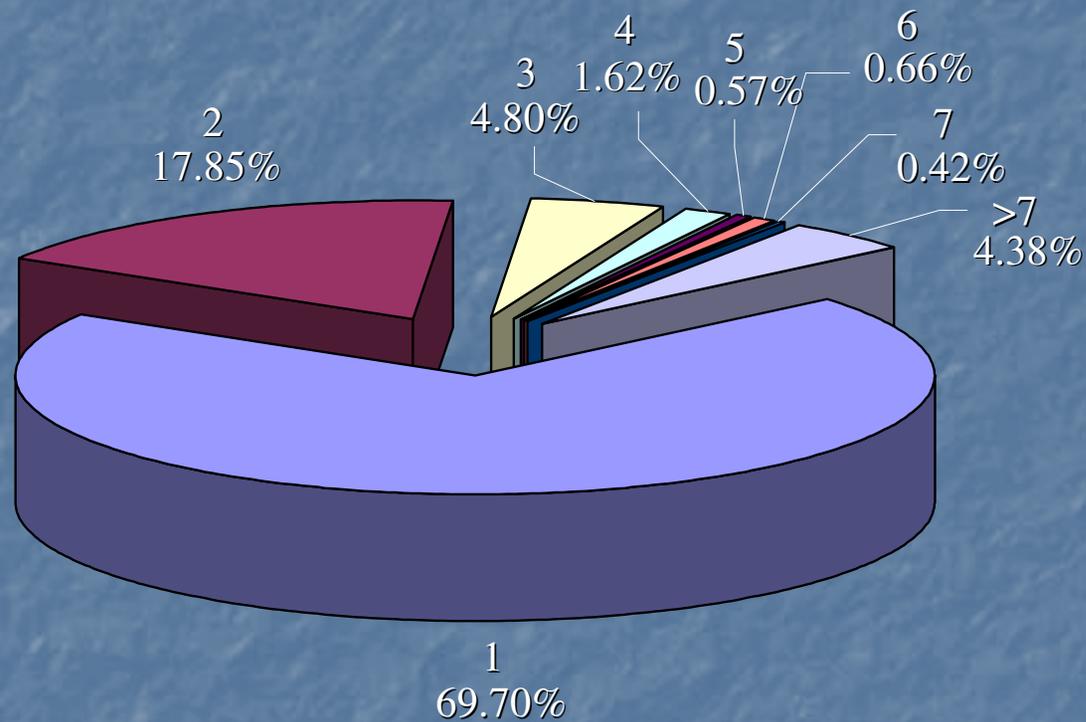
1,399 containing
unique protein

1,510 without
unique protein

Increased confidence?

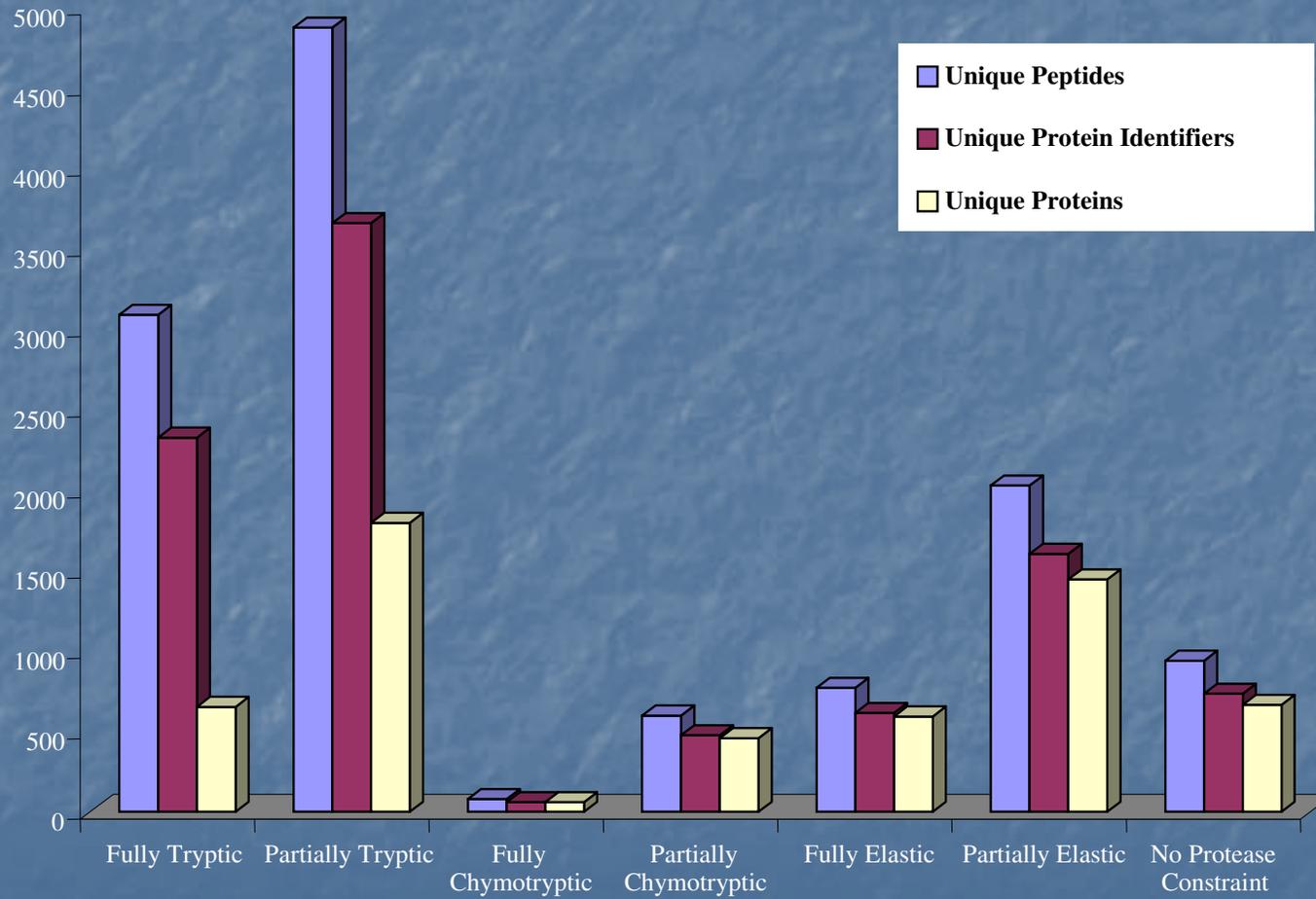


Global Mouse Serum Proteome



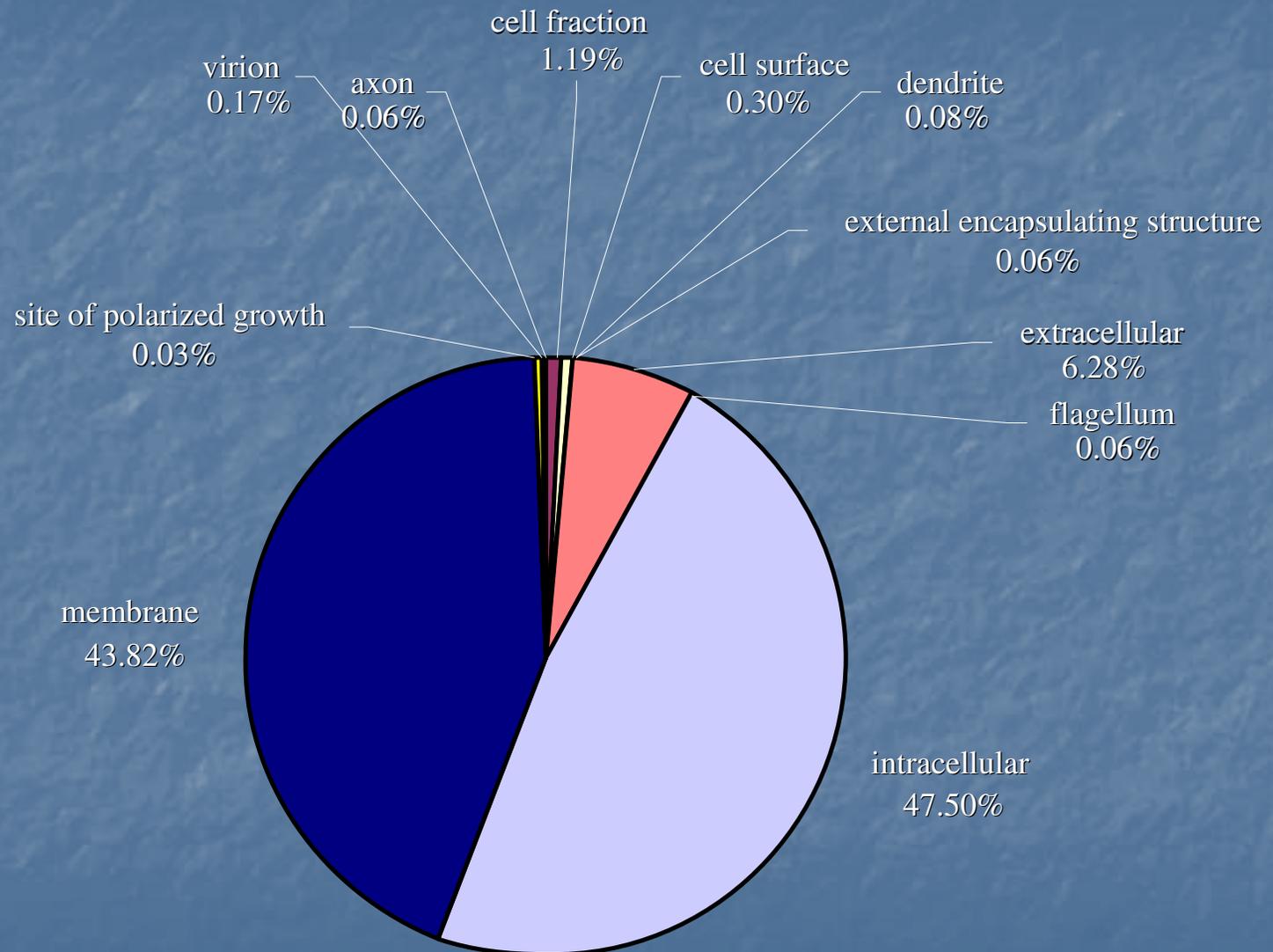
Number of Unique Peptides per Unique Protein

Global Mouse Serum Proteome



Identifications based on Enzymatic Constraint

Global Mouse Serum Proteome



Transmembrane Prediction In Serum Proteins

- TMHMM - Model for the prediction of transmembrane helices

COX2_BACSU

len=278

ExpAA=68.69

First60=39.89

PredHel=3

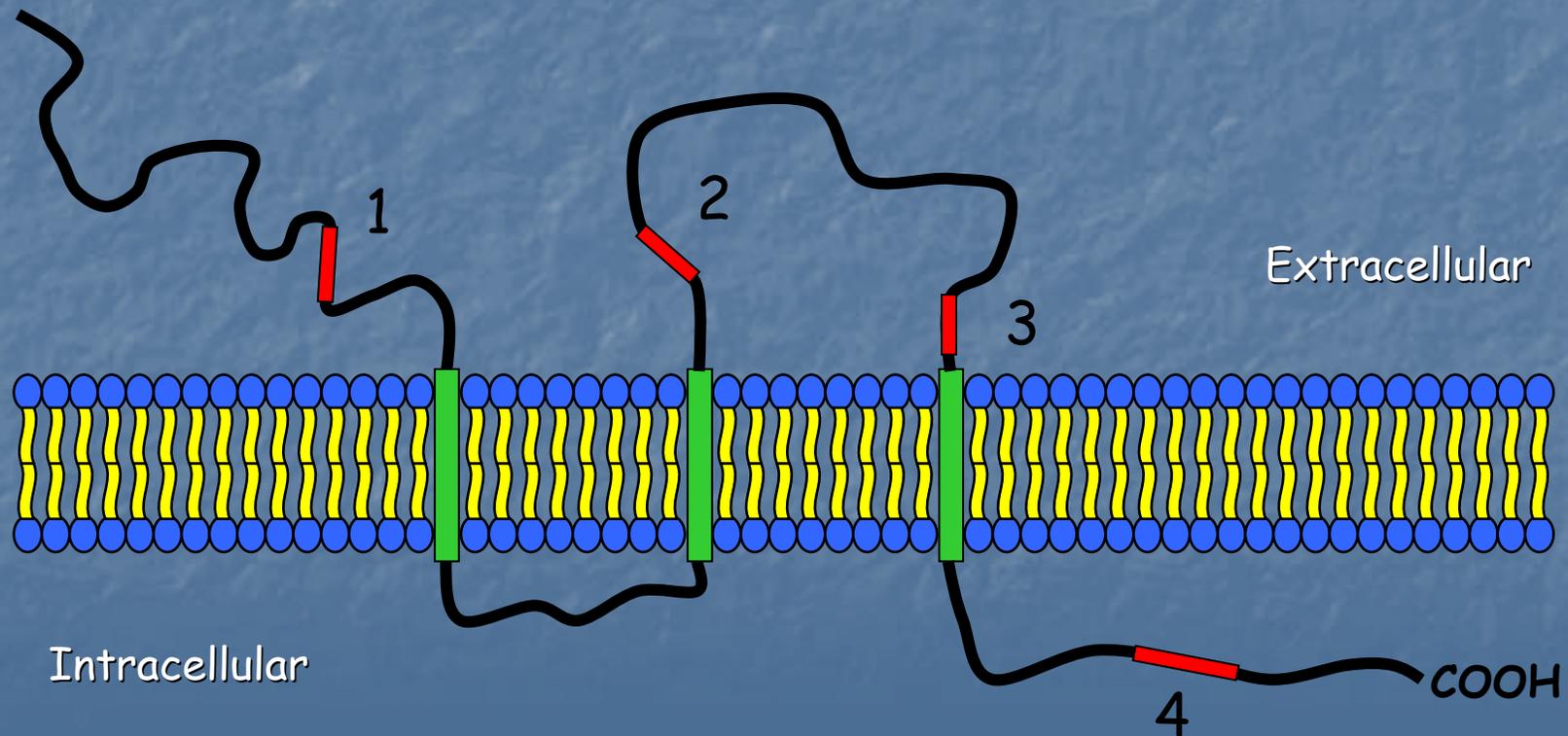
Topology=i7-29o44-66i87-109o

- N-terminal predicted TM helices may be signal peptides
PHOBIUS - predicts TM helices and signal peptides
- Does not account for beta-sheet membrane protein structures
- 66% of unique peptides from proteins predicted to contain transmembrane helices were extracellular in nature

Global Mouse Serum Proteome

Clipped or shed proteins from the surface of cells proximal to the circulatory system contribute significantly to the makeup of this biofluid proteome

Glutamate receptor KA-2



Comparative Mouse Serum Proteome Analysis

- Largest mouse serum proteome analysis to date
 - Foundation for future mouse serum experiments
 - Ultimately web-accessible resource
 - Comparison to HUPO results vital
 - Impractical for routine analysis

- Global analysis of control and Lewis lung carcinoma mouse serum proteomes

Can we see differences in this xenograft mouse model?

Comparative Mouse Serum Proteome Analysis



Control Serum



Tryptic Digest



SCXLC Fractionation



Analyze by μ RPLC-ESI-MS/MS



Lewis Lung Carcinoma Serum



Tryptic Digest

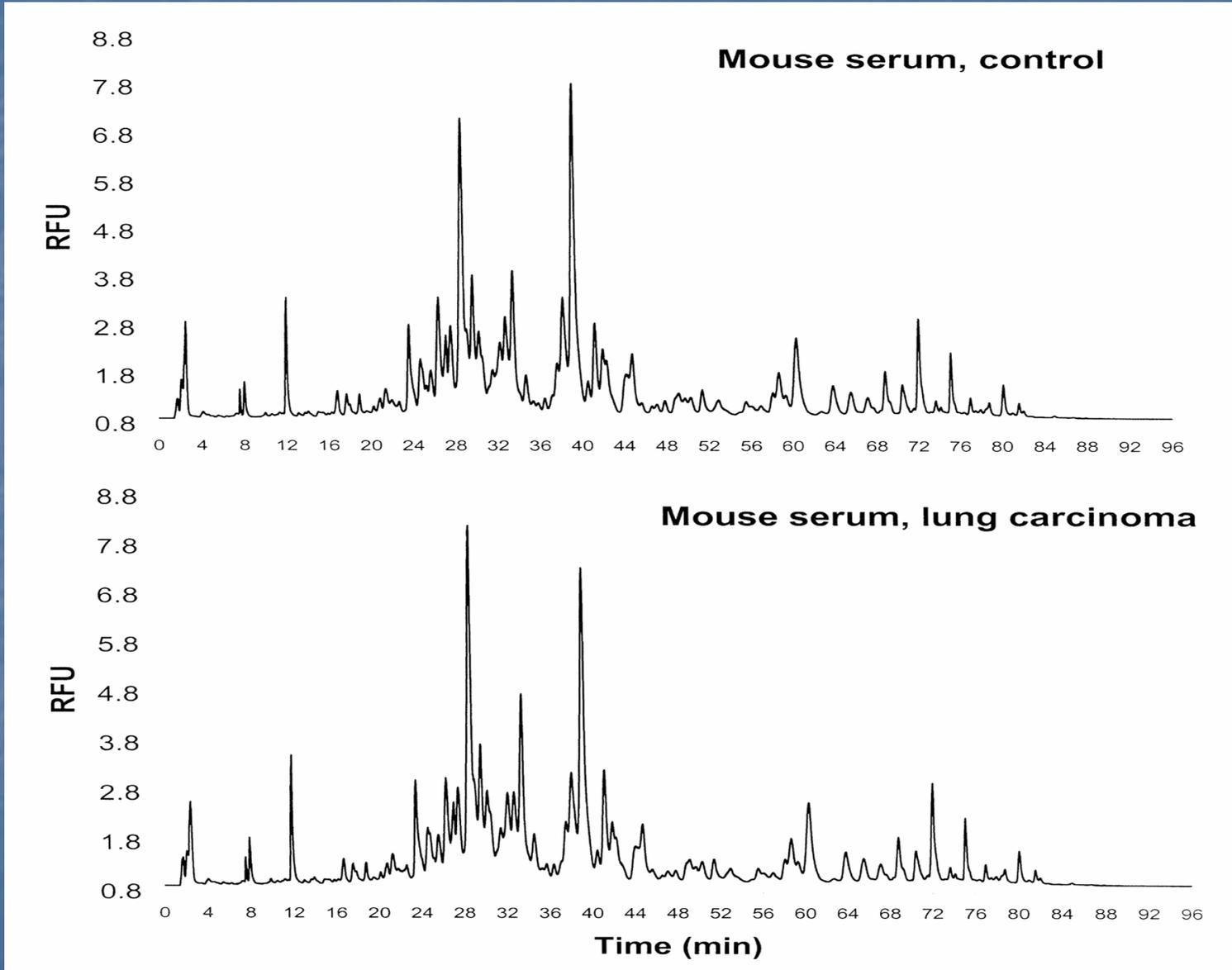


SCXLC Fractionation



Analyze by μ RPLC-ESI-MS/MS

Comparative Mouse Serum Proteome SCX



Comparative Mouse Serum Proteome Analysis

Control mouse serum proteome

1,474 unique proteins

4,432 unique peptides

370 identified by ≥ 2 unique peptides

LLC mouse serum proteome

1,608 unique proteins

5,258 unique peptides

380 identified by ≥ 2 unique peptides

576 common proteins

214 common (≥ 2 unique peptides)

Serum proteome comparison

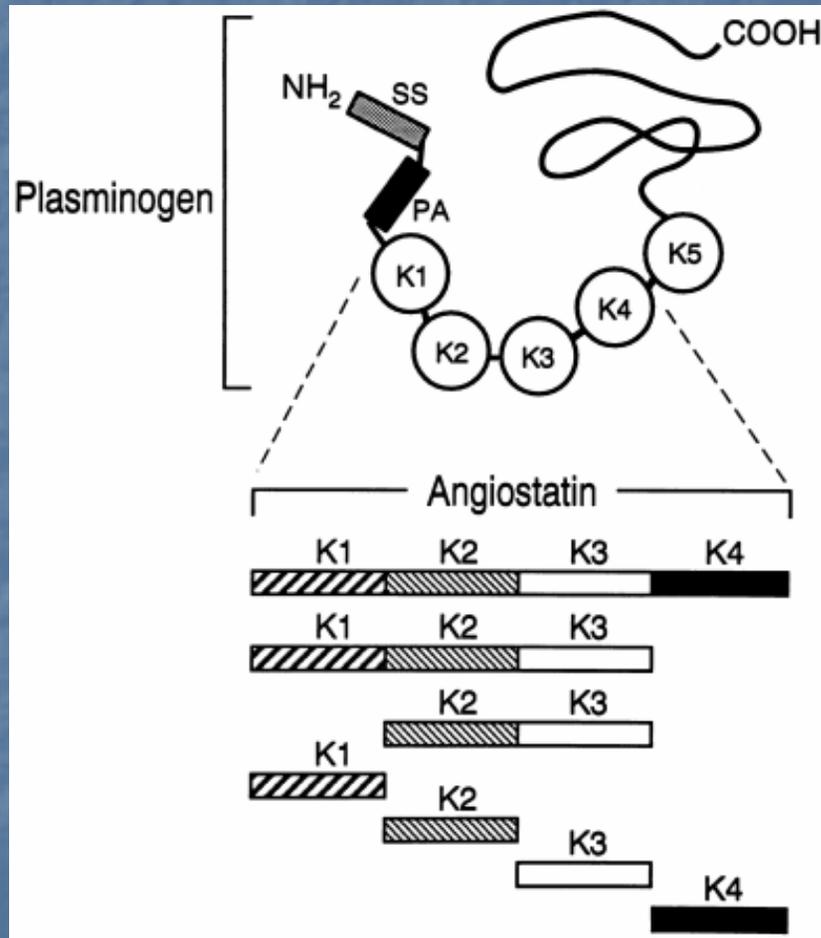
Individual unique peptides

Total # of unique peptides

Comparative Mouse Serum Proteome Analysis

Reference	Ctrl Spikes	LC Spikes	Spikes Diff	Ctrl Total Spikes	LC Total Spikes	Total Spikes Diff
ALBU_MOUSE	125	184	59	2157	4483	2326
TRFE_MOUSE	68	104	36	394	776	382
HEMO_MOUSE	24	46	22	125	264	139
PLMN_MOUSE	14	35	21	58	83	25
FINC_MOUSE	60	78	18	122	228	106
CFAH_MOUSE	6	22	16	13	38	25
Q9DBD0	15	31	16	38	87	49
Q8K0E8	45	61	16	174	248	74
Q99K47	61	77	16	224	257	33
KNG_MOUSE	18	32	14	55	126	71
HPT_MOUSE	7	20	13	19	89	70
CERU_MOUSE	21	34	13	40	69	29
APOH_MOUSE	5	17	12	12	46	34
THRB_MOUSE	11	23	12	21	45	24
CO5_MOUSE	32	44	12	41	70	29
A2MG_MOUSE	113	124	11	418	685	267
ITH1_MOUSE	16	25	9	41	37	-4
A2HS_MOUSE	19	28	9	44	106	62
FLNA_MOUSE	54	62	8	97	99	2
EGFR_MOUSE	7	14	7	19	38	19
ITH3_MOUSE	7	14	7	21	40	19

Plasminogen, Angiostatin, And Cancer



- First described in Lewis lung carcinoma
- Consists of first four kringle domains
- Triple-loop, disulfide-linked structures
- Angiogenesis inhibitor
- Not likely derived from the tumor, lack detectable mRNA for plasminogen
- Combinations of kringles result in varying inhibitory effects
- Found in a variety of other proteins

Comparative Mouse Serum Proteome Analysis

Regional peptide map of unique plasminogen peptides observed in control serum

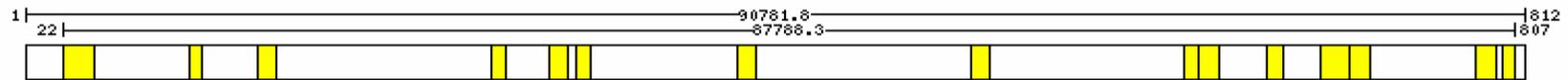
Reference: PLMN_MOUSE

Database: new_mouse_121803.fasta

Header: PLMN_MOUSE (P20918) Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin]

Avg Mass: 90781.8

Coverage: 155/812 = 19% by amino acid count



1 MDHKEVILLF LLLLKPGQGD SLDGYISTQG ASLFLTKKQ LAAGVSDCL AKCEGETDFV CRSFQYHSKE QOCVIMAENS

81 KTSSIIIRMD VILFEKRVYL SECKTGIGNG YRGTMSRTKS GVACQKWGAT FPHVPNYSPTS THPNEGLEEN YCRNPDNDEQ

161 GPWCYTTDPD KRYDYCNIFE CEEECMYCSG EKYEKISKST MSGLDCAWD SQSPHAHGYI PAKFPSKNLK MNYCRNPDGE

241 PRPWCFTTDP TKRWEYCDIP RCTTPPPPPS PTYQCLKGRG ENYRGTVSVT VSGKTCQRWS EQTPHRHNRT PENFPCKNLE

321 ENYCRNPDGE TAPWCYTTDS QLRWEYCEIP SCESSASPDQ SDSSVPPEEQ TPVVQECYQS DGQSYRGTSS TTITGKKCQS

401 WAAMFPHRHS KTPENFPDAG LEMNYCRNPD GDKGPWCYTT DPSVRWEYCN LKRCSETGGS VVELPTVSQE PSGPSDSETD

481 CMYGNKDYR GKTAVTAAGT PCQGWAQEP HRHSIFTPQT NPRAGLEKNY CRNPDGDVNG PWCYTTNPRK LYDYCDIPLC

561 ASASSFECGK PQVEPKKCPG RVVGGCVANP HSWPWQISLR TRFTGQHFCG GTLIAPEWVL TAAHCLEKSS RPEFYKVI LG

641 AHEEYIRGSD VQEISVAKLI LEPNNRDIAL LKLSRPATIT DKVIPACLPS PNYMVADRTI CYITGWGETQ GTFGAGRLKE

721 AQLPVIENKV CNRVEYLNRR VKSTELCAGQ LAGGVDSCQG DSGGPLVCFE KDKYILQGV T SWGLGCARP N KPGVYVRVSR

801 FVDWIEREMR NN

Comparative Mouse Serum Proteome Analysis

Regional peptide map of unique plasminogen peptides observed in LLC serum

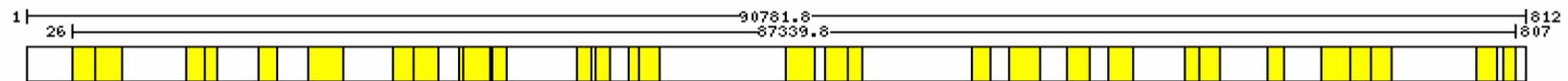
Reference: PLMN_MOUSE

Database: new_mouse_121803.fasta

Header: PLMN_MOUSE (P20918) Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin]

Avg Mass: 90781.8

Coverage: 331/812 = 41% by amino acid count



1 MDHKEVILLF LLLLKPGQGD SLDGYISTQG ASLFSLTKKQ LAAGGVSDCL AKCEGETDFV CRSFQYHSKE QQCHEMAENS

81 KTSSIIRMRD VILFEKRVYL SECKTGIGNG YRGTMSRTKS GVACQKWGAT FPHVPNYSPTS THPNEGLEEN YCRNPDNDEQ

161 GPWCYTTPDP KRYDYCNIFE CEEECMYCSG EKYEGKISKI MSGLDQAWD SQSPHAHGYI PAKFPKSNLK MNYCRNPDGE

241 PRPWCFTTDP TKRWEYCDIP RCTTPPPPPS PTYQCLKGRG ENYRGTVSVT VSGKTCQRWS EQTPHRHNRT PENFPCKNLE

321 ENYCRNPDGE TAPWCYTTPDS QLRWEYCEIP SCESSASPDQ SDSSVPPEEQ TPVVQECYQS DGQSYRGTS TTITGKKCQS

401 WAAMFPHRHS KTPENFPDAG LEMNYCRNPD GDKGPWCYTT DPSVRWEYCN LKRCSETGGS VVELPTVSQE PSGPSDSETD

481 CMYGNKDYR GKTAVTAAGT PCQGWAAQEP HRHSIFTPQT NPRAGLEKNY CRNPDGDVNG PWCYTTNPRK LYDYCDIPLC

561 ASASSFECGK PQVEPKKCPG RVVGGCVANP HSWPWQISLR TRFTGQHFCG GTLIAPEWVL TAAHCLEKSS RPEFYKVILG

641 AHEEYIRGSD VQEISVAKLI LEPNNRDIAL LKLSRPATIT DKVIPACLPS PNYMVADRTI CYITGWGETQ GTFGAGRLKE

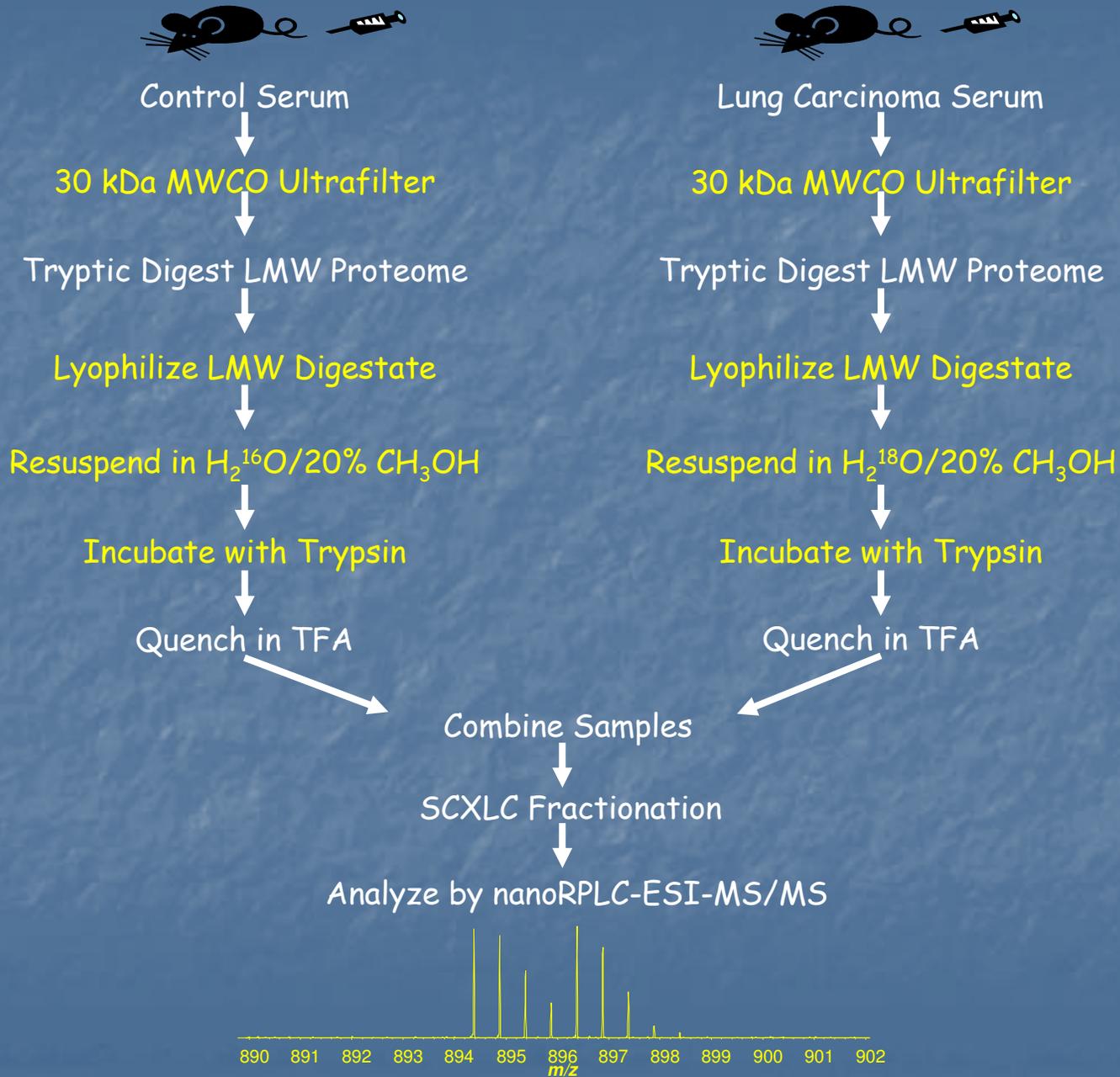
721 AQLPVIENKV CNRVEYLNRR VKSTELCAGQ LAGGVDSCQG DSGGPLVCFE KDKYILQGVT SWGLGCARPV KPGVYVVRVSR

801 FVDWIEREMR NN

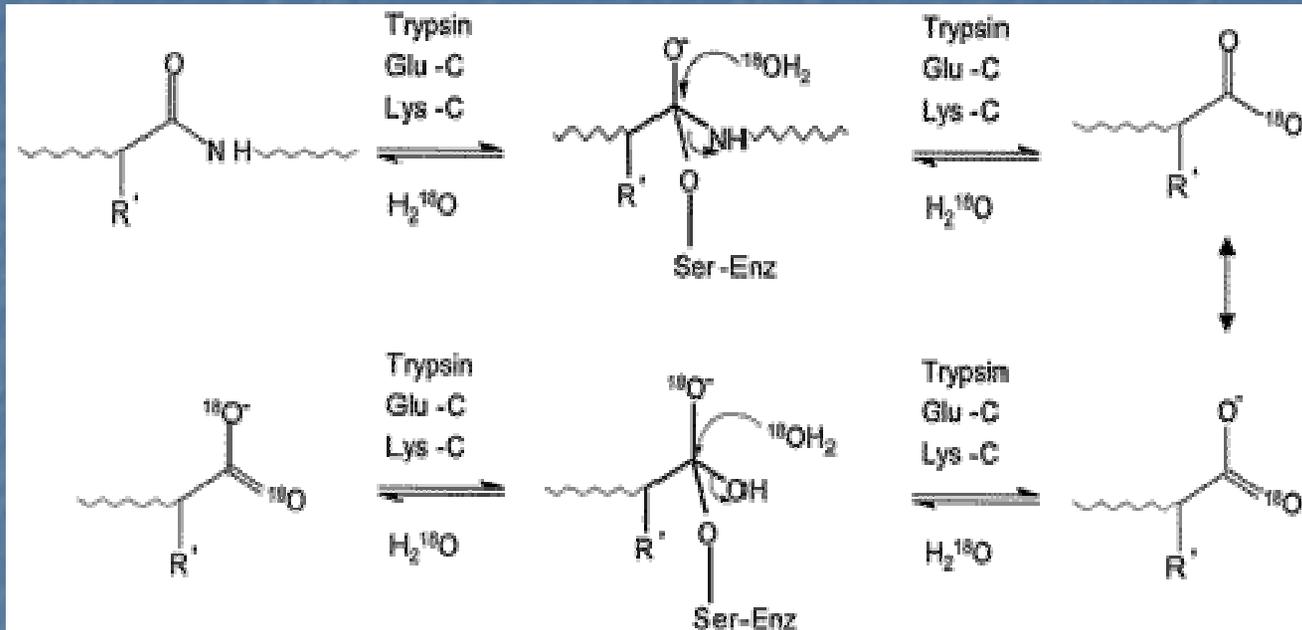
Isotope Labeling Of LMW Mouse Serum Proteome

- Why not use isotope labeling to detect abundance differences?
Serum is notoriously difficult to label
- Reduce complexity - LMW preparation of serum samples
- Enzyme-mediated ^{18}O labeling of tryptic peptides
- LC-MS/MS analysis using LTQ linear ion trap
Dual detectors and faster scan speed
Trap more ions than the LCQ Deca
Data dependent acquisition - top 5 molecular ions for
MS/MS

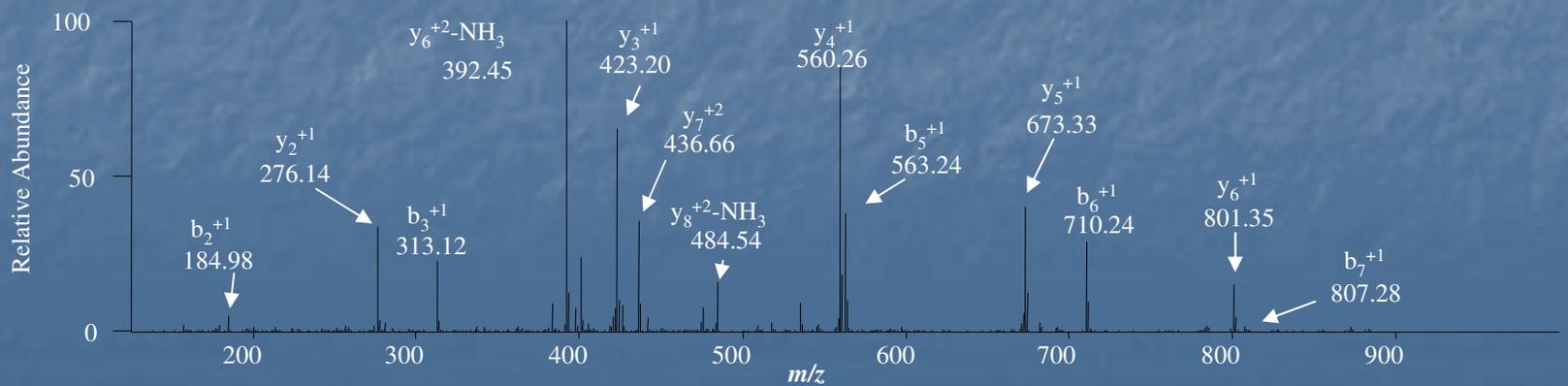
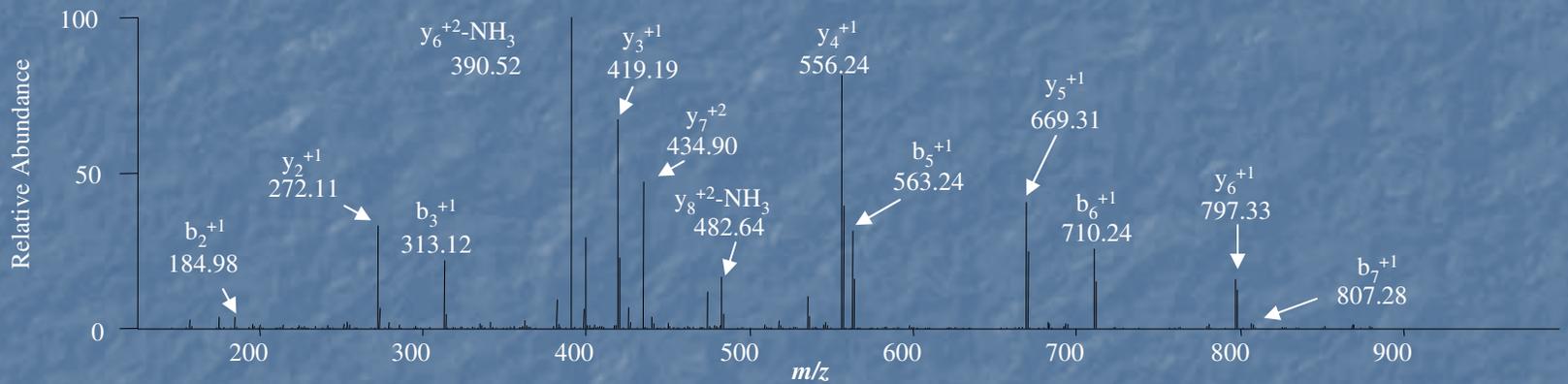
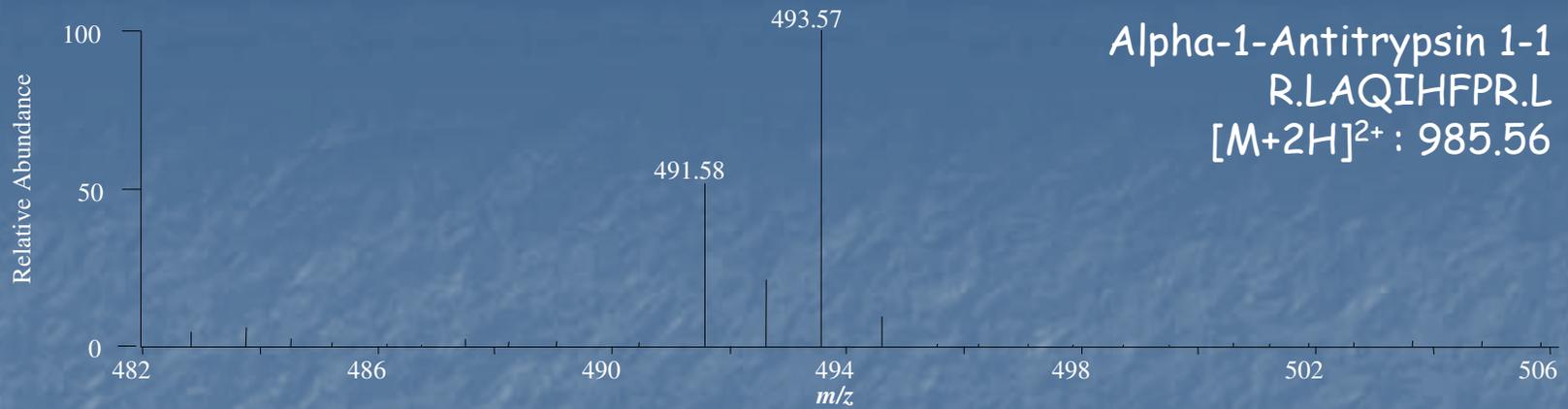
^{18}O Labeling Of LMW Mouse Serum Proteome



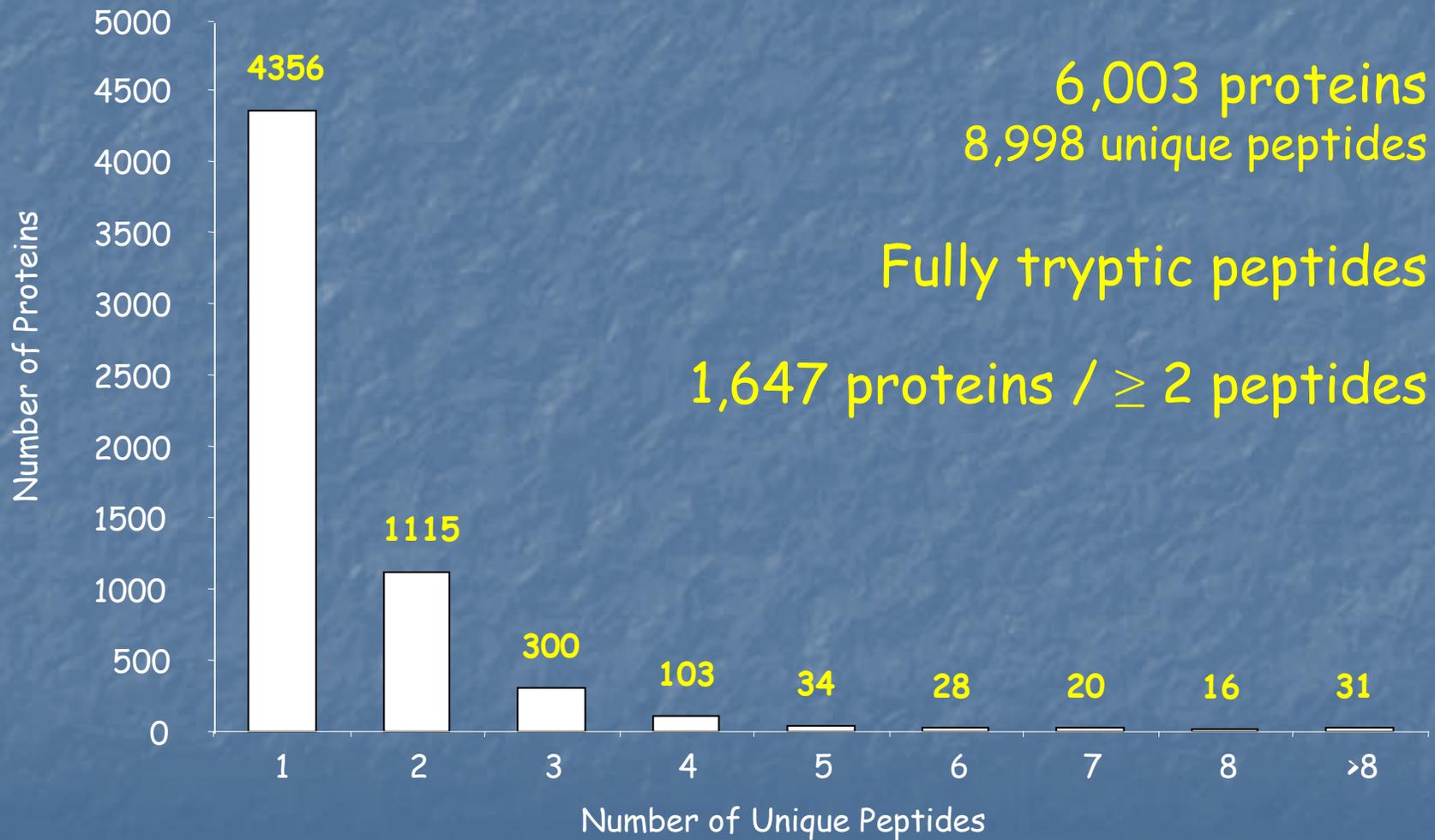
Enzyme Mediated ^{18}O Labeling



Complete incorporation results in a 4 Da mass shift



¹⁸O-labeled LMW Mouse Serum Proteome



Comparative Proteomics In Mouse Serum

- Can we use these analyses to help us identify and validate specific targets?
- VEGF/VEGF-receptor pathway involved in angiogenesis
 - VEGF-R - receptor tyrosine kinase family
 - VEGF homodimers / VEGF-PLGF heterodimers
 - VEGF induced cell proliferation is blocked by angiostatin
- Comparative mouse serum proteome analysis:
 - VEGF and PLGF unique peptides found in LLC but not control
- ¹⁸O-labeled LMW mouse serum proteome analysis:
 - VEGF-R increased 7-fold in LLC LMW serum
- EGF-receptor is up in global mouse serum proteome analysis

Uniques:	Ctrl - 7	LLC - 14
Total uniques;	Ctrl - 19	LLC - 38

Mouse Serum Proteomics

- *Global mouse proteome has been investigated*
First proteomic investigation in this significant animal model
Database of information that can be found in serum
Foundation for future proteomic investigations of cancer and disease
- *Comparative global and LMW mouse serum proteome analyses*
Observe relevant changes in serum between two states
Target proteins with significant differences between normal and cancerous states based on peptide abundances
Isotope labeling for evaluation of protein changes in serum
Validation is necessary to confirm the results from these analyses

Laboratory of Proteomics and Analytical Technologies

Dr. Timothy D. Veenstra (Director)

Separations Technology Lab

Dr. Haleem J. Issaq
Dr. King Chan
Dr. Nick Stewart
Mr. Stephen Fox

Mass Spectrometry Lab

Dr. Thomas P. Conrads
Dr. Ming Zhou
Dr. Josip Blonder
Dr. Li-Rong Yu
Dr. Xia Xu
Dr. Brian L. Hood
Dr. Zhaojing Meng
Dr. Zhen Xiao
Dr. Nic Stewart
Mr. Aaron Lucas
Ms. DaRue Prieto
Mr. John Roman
Mr. Al Richter
Mr. Tim Waybright
Ms. Grace Kim

NMR Lab

Dr. Gwendolyn Chmurny
Dr. Que Van
Mr. John Klose

ABCC

Bob Stephens