

Mass Spectrometry of Proteins (Biomolecules)

Martin Hubálek, PhD

**Institute of Organic Chemistry and Biochemistry AS CR,
v.v.i.**

Biomolecules

- Proteins, glycoproteins, phosphoproteins, other PTM
- Polysaccharides, glycans
- Nucleic Acids, Oligonucleotides

Soft ionization techniques:

MALDI

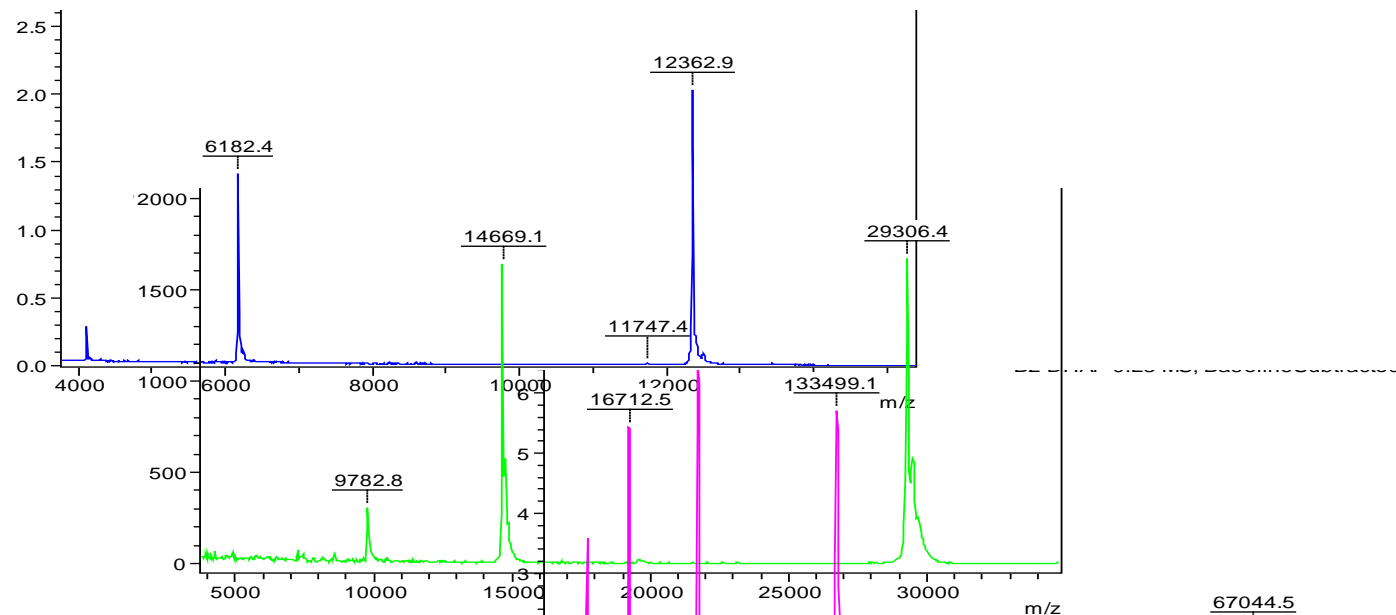
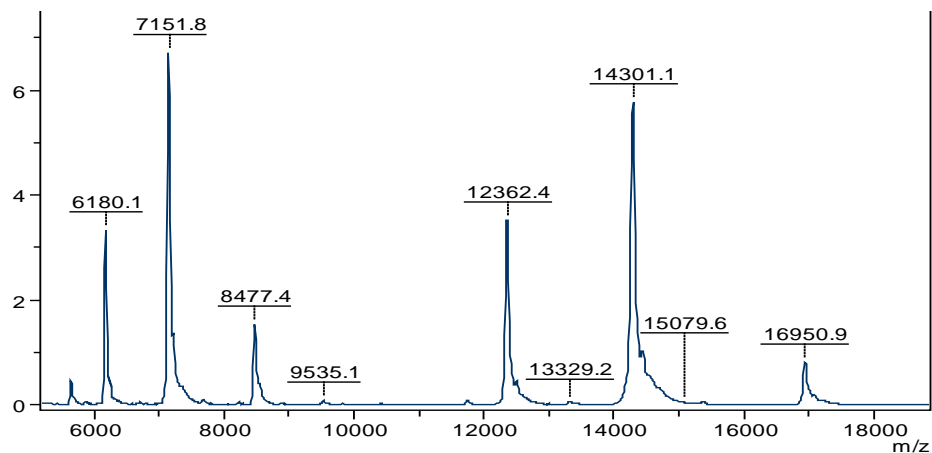
ESI

Proteins

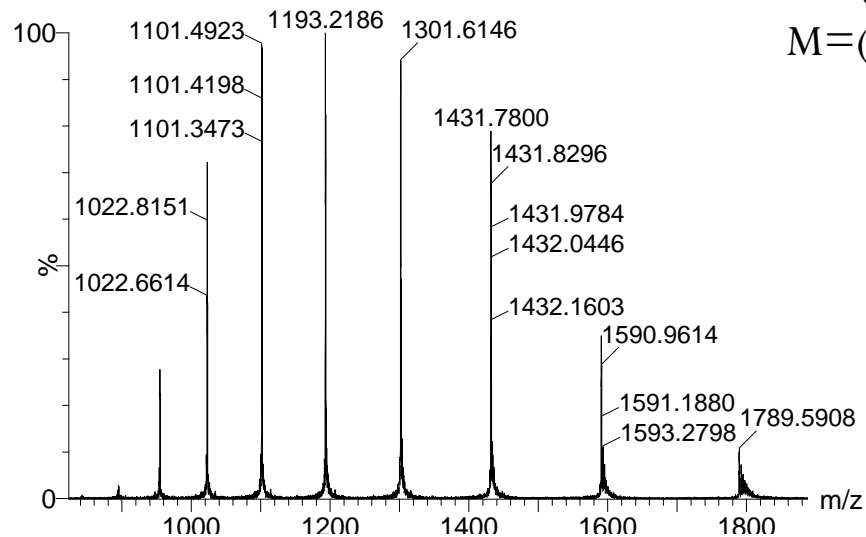
- What can MS tell about proteins?
 - Intact mass - **MS**
 - Confirmation of right molecular weight
 - Quality control of purified proteins, recombinant proteins, synthetic peptides
 - Analysis of modification
 - Fragmentation of ions – **MS/MS**
 - Identification of proteins
 - Intact protein molecule – N-term and C-term sequencing
 - In source decay – MALDI-TOF
 - Protein digestion – separation of peptides on NanoLC – tandem MS of peptides – search against selected database
 - Quantitation

Mass of Intact protein

- MALDI - TOF

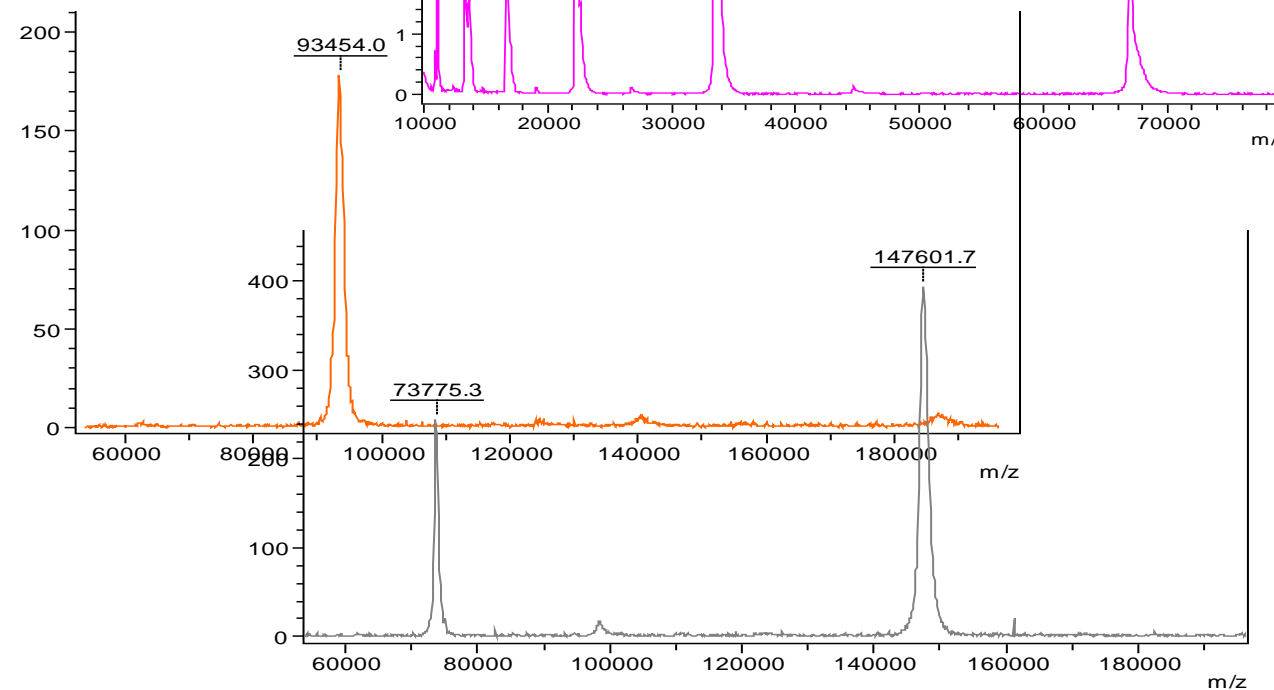


- ESI – (TOF, IT, QTOF etc)



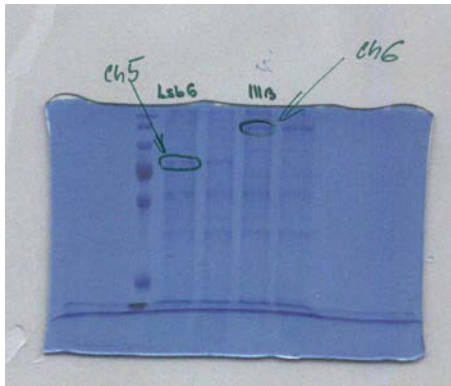
$$z = (y-1)/(x-y)$$

$$M = (x*z) - z$$

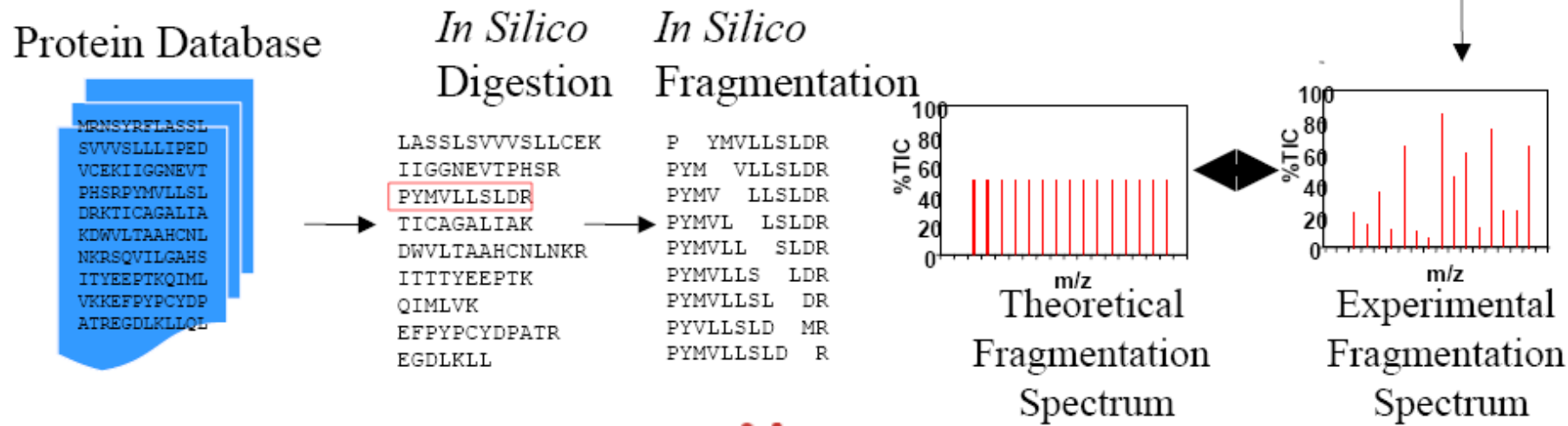
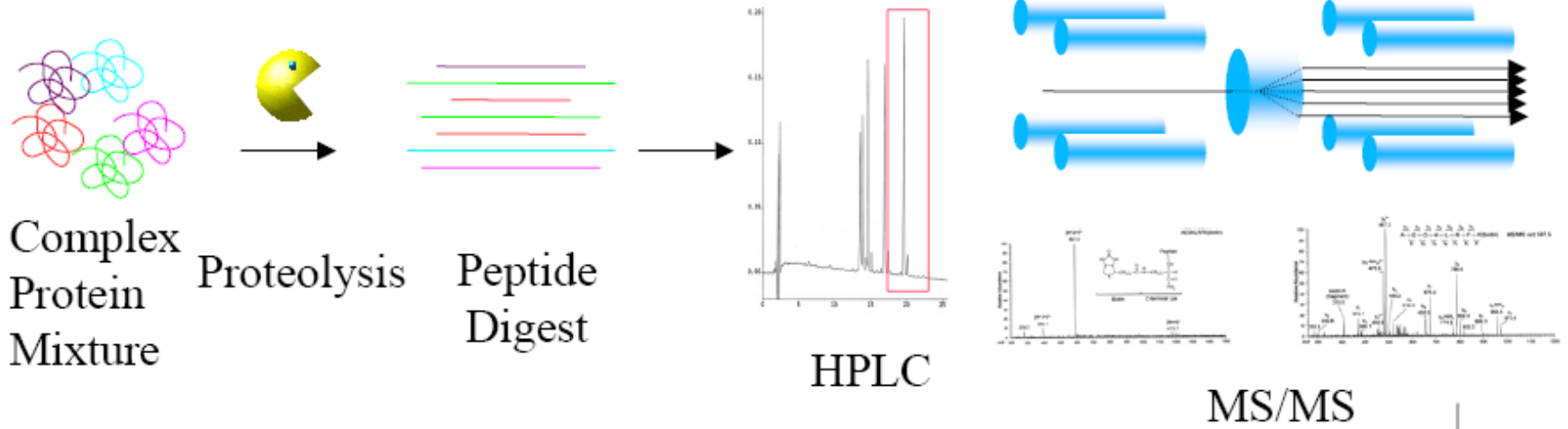


Identification of protein by Tandem MS

- Gel

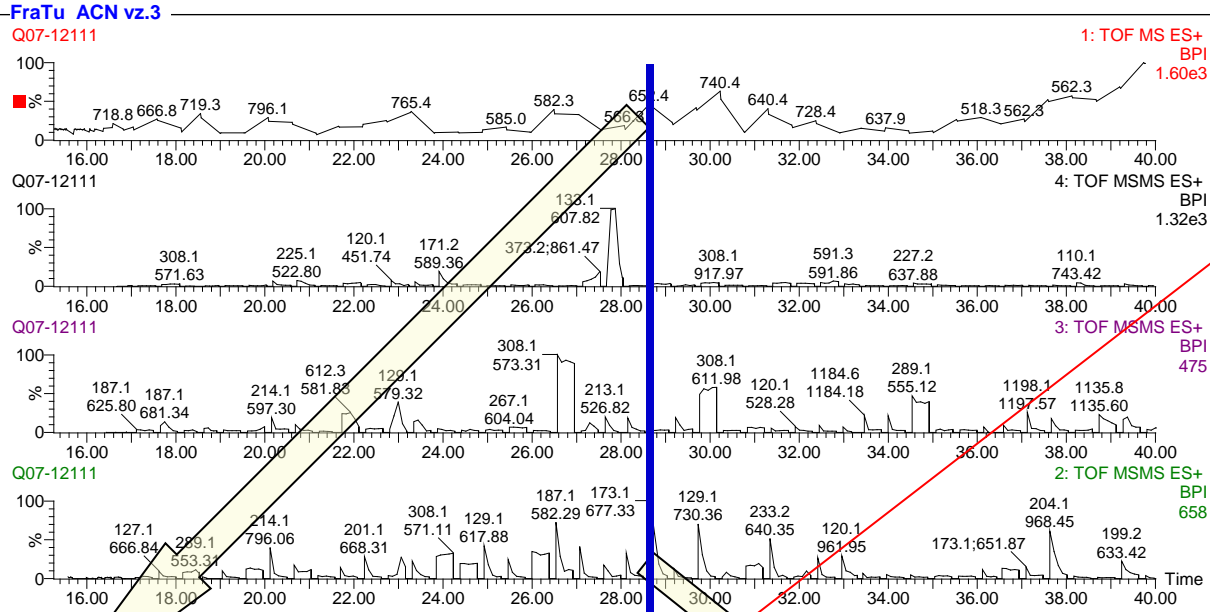


- Protein(s) in solution

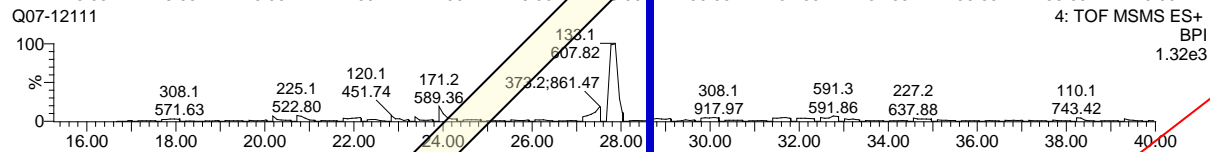


- Sample preparation
 - Compatible with enzymatic digestion, LC separation and ESI (the most intolerant step – no detergent, very low salt concentration)
 - Reduction and alkylation of -S-S-
 - Basic buffer - generally
- Digestion
 - Enzymatic – **Trypsin**, Chymotrypsin, AspN, GluC, LysC
 - Desalting off-line and on-line
 - Optimal final buffer – 0,1 % Formic acid
- LC-MS/MS
 - Nano LC – 300 nl/min, gradient 20 min – 3 hours (depends on sample complexity)
 - MS/MS – data dependent acquisition
- Database search
 - Appropriate search engine and right database!

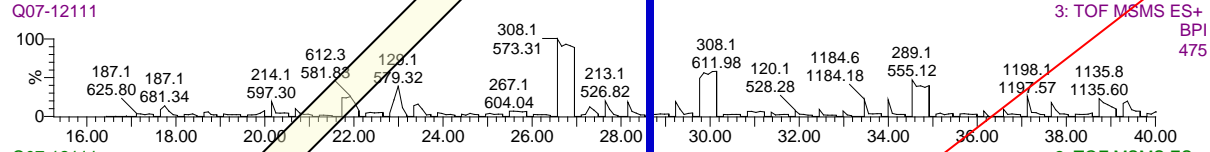
MS



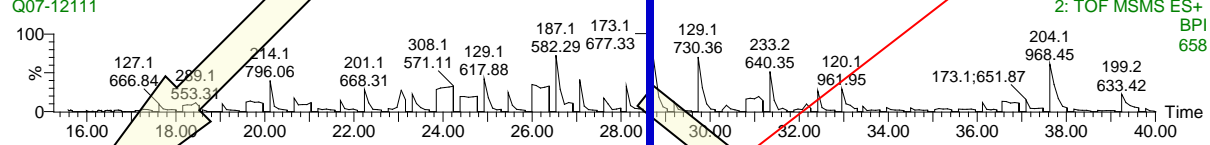
MSMS 1



MSMS 2

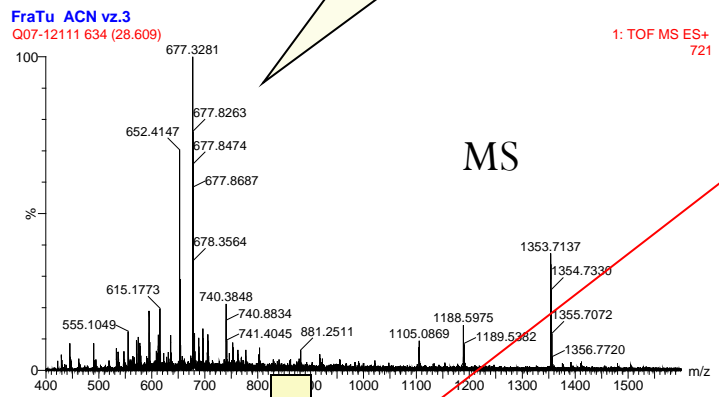


MSMS 3

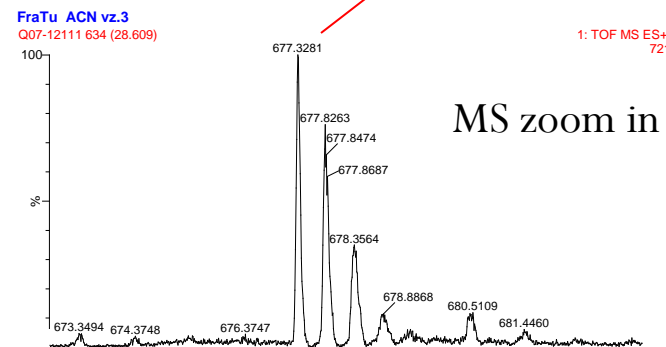
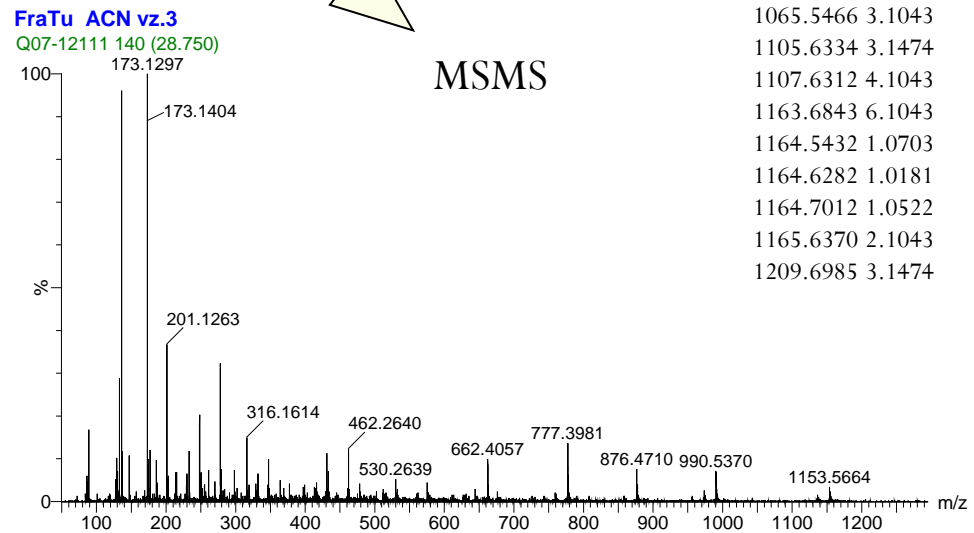


677.3281, 721, 2

- 70.0332 1.0794
- 70.0671 8.1202
- 72.0876 11.3469
- 73.5504 2.1043
- 74.0656 3.1383
- 83.0607 2.1134
- 84.0828 25.2925
- 85.0589 1.0522
- 85.0898 3.0612
- 86.1002 48.8639
- 96.0747 4.1043
- 98.0287 3.1927
- 98.0612 2.2358
- 98.1019 5.9909
- ...
- 283.2370 2.1927
- 284.1254 11.0181
- 284.1629 37.8798
- 284.2126 9.7188
- 285.0384 4.7891
- ..
- 1065.5466 3.1043
- 1105.6334 3.1474
- 1107.6312 4.1043
- 1163.6843 6.1043
- 1164.5432 1.0703
- 1164.6282 1.0181
- 1164.7012 1.0522
- 1165.6370 2.1043
- 1209.6985 3.1474



MS



MSMS scoring function

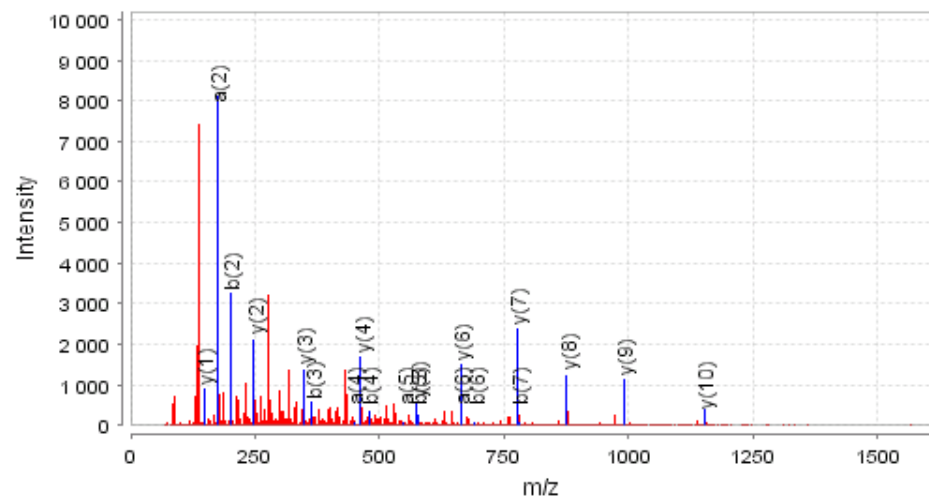
The comparison of theoretical and experimental MS/MS spectra is performed by **scoring function** and the score (ideally complemented by **p-value**) is used to recognize the correct peptide from the database. Reliable **peptide identification** can be then considered for **protein identification**.

Value Types: Percentage of most intense peaks: 70<x<100 50<x<70 30<x<50 10<x<30 <10

	1	2	3	4	5	6	7	8	9	10	11	12
a		-0.006		-0.018	-0.026	-0.059						
a-H ₂ O		-0.006	-0.006	0.025	0.024							
a-NH ₃				-0.029	-0.028	-0.041						
b		-0.008	-0.014	-0.019	-0.025	-0.027	-0.006					
b-H ₂ O		-0.007	-0.016	-0.018	-0.011							
b-NH ₃				-0.032	-0.032	-0.04			-0.112			
	S	L	Y	N	V	D	S	I	D	V	T	K
y			-0.013	-0.033	-0.034	-0.026	-0.026	-0.023	-0.017	-0.012	-0.008	-0.006
y-H ₂ O					-0.018	-0.039	0.029		-0.012	-0.01	-0.008	
y-NH ₃			-0.044	-0.047								
	12	11	10	9	8	7	6	5	4	3	2	1

Match of an experimental spectrum with a peptide sequence

- Intense peaks should match
- As many as possible peaks should match
- Series of contiguous matches



MSMS scoring function

Score - dependent on the search engine

E-value - statistical expectation value

- expected number of peptides with score better or equal than observed score under the assumption that peptides are matching the experimental spectrum by **random chance**
- Invariant under different scoring methods, clearer interpretation of goodness of match

- Correlation function – **Sequest**
- Shared fragment count and dot blot
 - **MASCOT, TANDEM, OMSSA**
- Empirically observed rules – **SpectrumMill**
- Statistically derived fragmentation frequencies – **Phenyx**
- sequence tag algorithm, allowing the degree of implication by an MS/MS spectrum of each region of a database to be determined on a continuum – **Paragon (Protein Pilot)**

Neither the best match nor a high search score (or low E-value) are reliable indicators of true match – ***discriminating true from false is critical step***

Modification

- Fixed modification
 - Cystein alkylation by Iodoacetamide - +57mu
- Variable
 - Met oxidation
 - Pyroglutamate
 - Phosphorylation

The more modification allowed the more time and space taken to for search

Statistical assessment of peptide assignments in large-scale datasets

- *discriminating true from false*
 - Probabilistic approaches that provide statistical measure of confidences and estimates of error rates

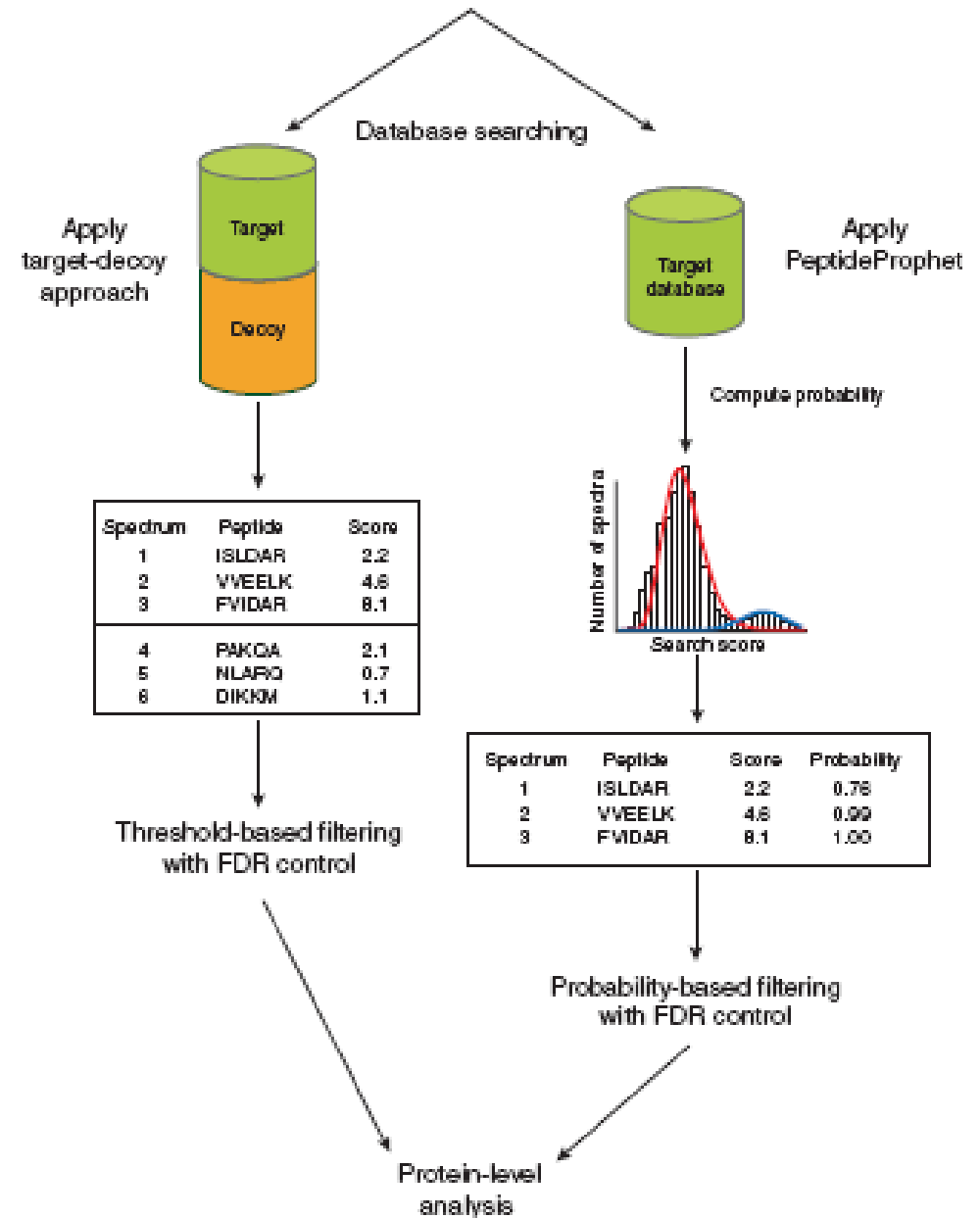
1. Target-decoy searching

- Spectra searched against decoy database (database augmented with reverse, or randomized or shuffled sequence of the same db)
- FDR cut off
 - False discovery rate – e.g. 5%
 -
 - double the searching time,
 - is the random sequence really random?

2. Empirical bayes approach

Peptide prophet

- Two component mixture of distributions representing correct and incorrect identification



Can be combined together

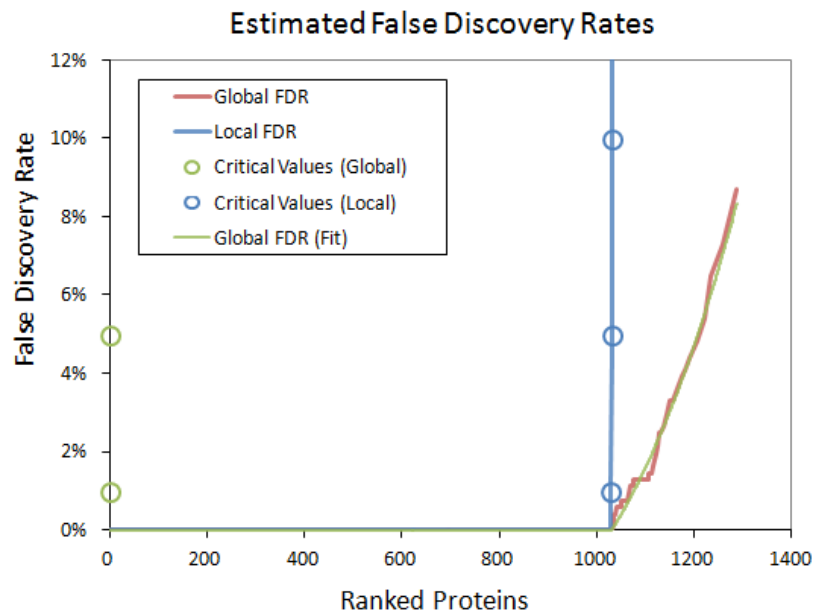
Protein Level FDR Analysis

Proteins Identified at Critical False Discovery Rates

Number of Proteins Detected

Critical FDR	Local FDR	Global FDR	Global FDR from Fit
1.0%	<i>1029</i>	<i>1067</i>	1076
5.0%	1031	<i>1205</i>	<i>1208</i>
10.0%	1031		

** It is recommended you use numbers in bold and avoid using numbers in italics.*



Correspondence between FDR Levels and ProteinPilot Reported Confidences

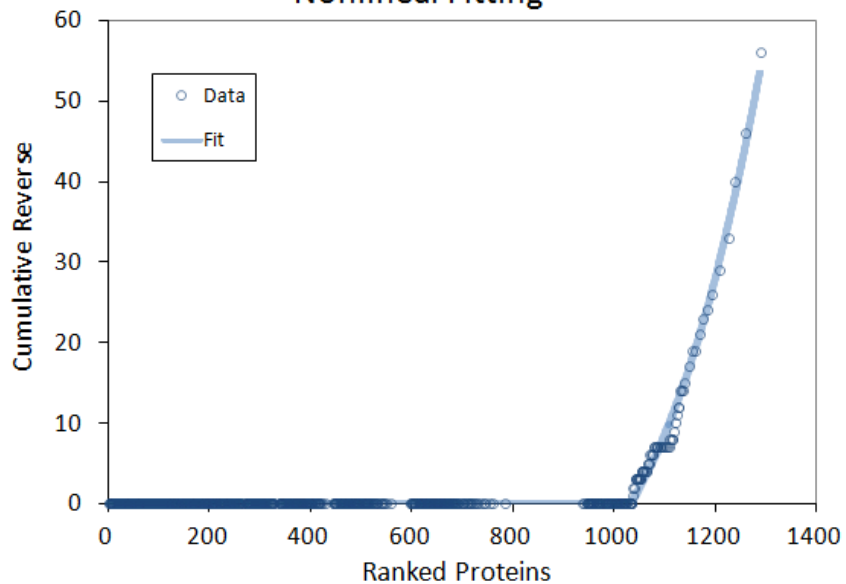
Corresponding ProteinPilot Confidence

Critical FDR	Local FDR	Global FDR	Global FDR from Fit
1.0%	<i>85.5%</i>	<i>68.4%</i>	63.7%
5.0%	84.2%	<i>18.7%</i>	<i>16.8%</i>
10.0%	84.2%		

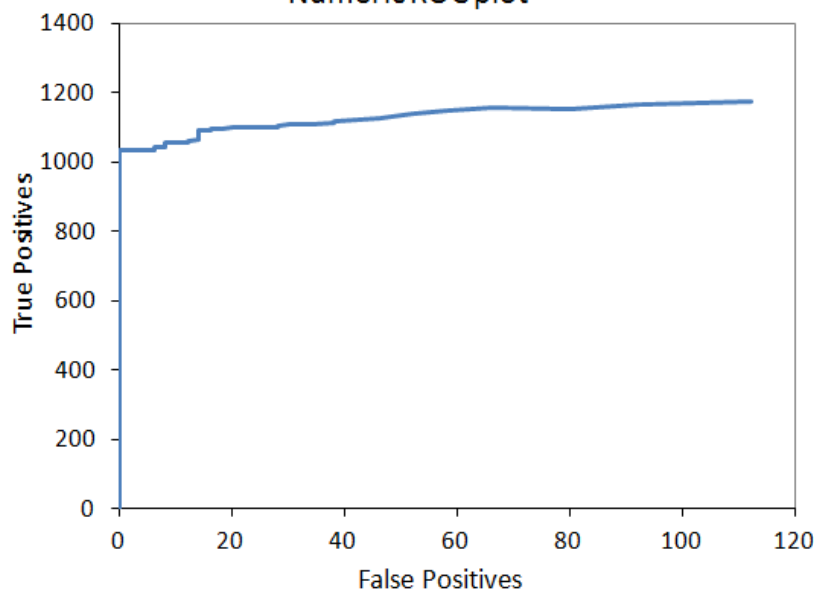
Corresponding Unused ProtScore

Critical FDR	Local FDR	Global FDR	Global FDR from Fit
1.0%	<i>0.840</i>	<i>0.500</i>	0.440
5.0%	0.800	<i>0.090</i>	<i>0.080</i>
10.0%	0.800		

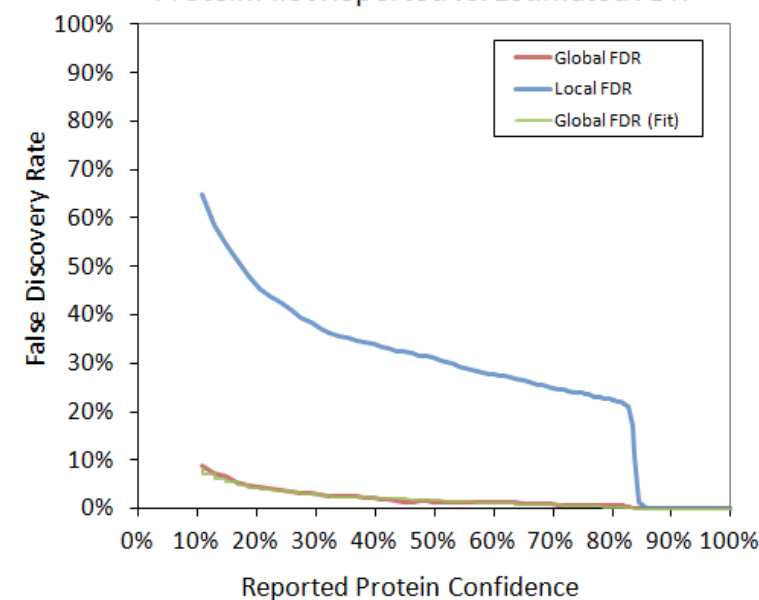
Nonlinear Fitting



Numeric ROC plot

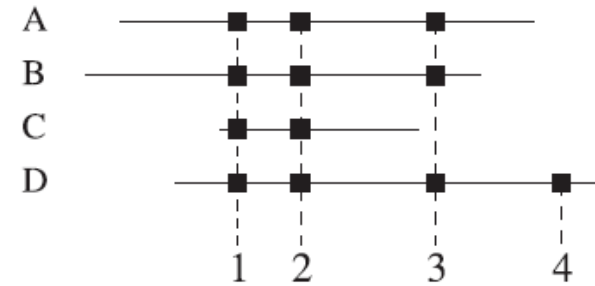


ProteinPilot Reported vs. Estimated FDR



Protein identification

- Combination of peptide ID
- Score = sum of the highest score for each peptide (standard)
- Sources of errors
 - peptides shared by several proteins
 - Homologous proteins
 - Splice variants
 - Redundant entries in the db
 - correct peptides tend to group into a relatively small number of proteins – e.g. human serum
 - Dominant peptides from few proteins only
 - Incorrect spectral identification match randomly to the much larger number of proteins – increase in FDR on protein level

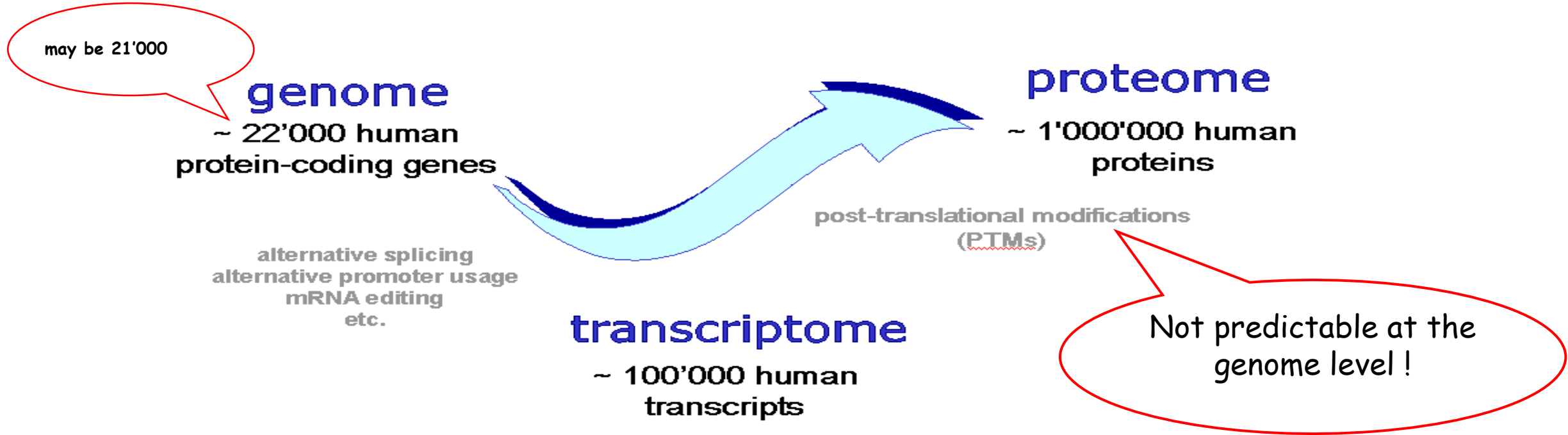


Parsimony principle (Occam's razor)

- Determining the smallest number of proteins that can account for all observed peptides
- Scoring functions
 - Bayes rule-based – **Protein Prophet**
 - Concept of number of sibling peptides
 - Poisson distribution-based statistics
 - Protein length, model of protein abundance etc.

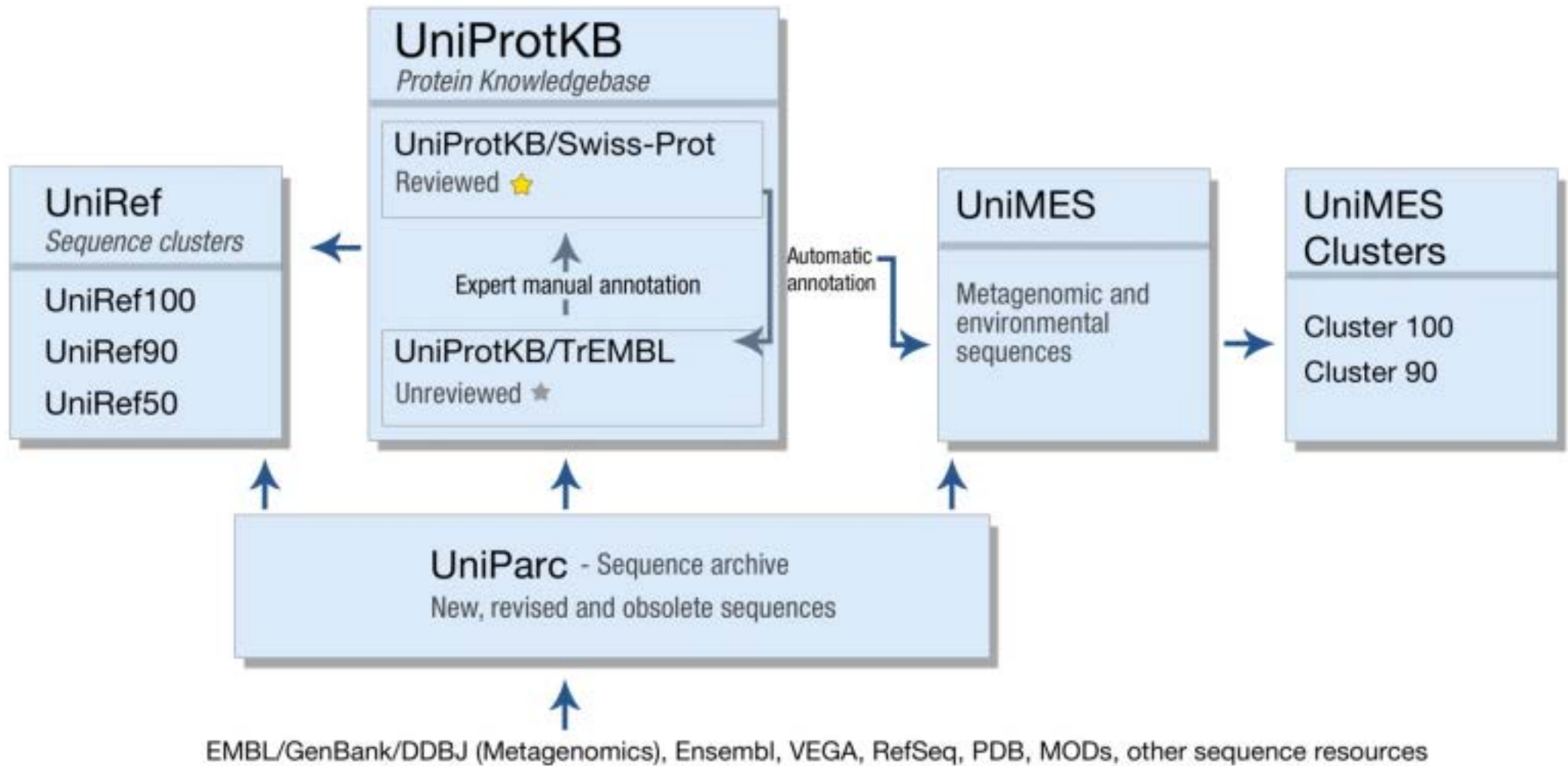
Proteome complexity

Example: human



(Jensen O.N., Curr. Opin. Chem. Biol., 2004, 8, 33-41, PMID: 15036154).

UniProt databases



How to understand the results?

ProteinPilot™ Software - [Result - C:\AB SCIE\X\ProteinPilot Data\Data\2013\2013_November\131114_gely\131114_B2.group]

File Configure Window Help

Workflow Tasks

- Identify Proteins
- LC...
- Spot-Based (MS only)...
- Spot-Based (MS and MS/MS)...
- View
- Analysis Log...
- Result...
- Export
- Peptide Summary...
- Protein Summary...
- MGF Peaklist(s)...

Protein ID Spectra Summary Statistics

Proteins Detected

N	Unused	Total	% Cov	Accessio...	Name	Species	Peptides(96%)	Biological Processes	Molecular Functions	PANTHER ID
1	132.27	132.27	75.0	jb l CaLo...	confirm C terminus, putative OS=Jakub Benysek...	IXOSC	166			
2	57.56	57.56	61.9	cont 000...	pdh1 FNI_A A Chain A, Crystal Structure Of Por...	Sus scrofa (contaminant)	66			
3	42.05	42.05	46.8	cont 000...	cra hCP181205 keratin 10(epidermolytic hype...	Homo sapiens (contaminant)	25			
4	36.10	36.10	45.8	cont 000...	cra hCP1809934.2 keratin 1(epidermolytic hyp...	Homo sapiens (contaminant)	22			
5	25.86	25.86	20.6	cont 000...	gl 114939 sp P00722 BGAL_ECOLI Beta-galact...	Escherichia coli (contaminant)	15			
6	9.15	9.15	35.9	cont 000...	gl 7441526 pir A69068 beta-casein variant OrH...	Bos taurus (contaminant)	6			
7	8.00	8.00	26.2	cont 000...	gl 115646 sp P02662 CAS1_BOVIN Alpha-S1-c...	Bos taurus (contaminant)	5			
8	7.01	7.11	14.4	cont 000...	sp P48666 Keratin, typeII cytoskeletal 6C (Cyt...	Homo sapiens (contaminant)	4			
9	4.01	4.01	43.6	cont 000...	sp P81605 Dermcidin precursor (Preproteolys...	Homo sapiens (contaminant)	2			
10	2.25	2.25	17.6	cont 000...	gl 27806963 ref NP_776953.1 casein alpha-S2...	Bos taurus (contaminant)	2			
11	2.00	2.00	25.8	cont 000...	gl 27881412 ref NP_776719.1 casein kappa[Bo...	Bos taurus (contaminant)	1			

Protein Group 1 - confirm C terminus, putative OS=Jakub Benysek GN=IscW_ISCW007008 PE=3 SV=1

Proteins in Group						Peptides in Group										
N	Unused	Total	Accessio...	Name	Species	Con...	Conf	Sequence	Modifications	Cleavages	ΔMass	Prec MW	z	Sc	Spectrum	Type
1	132.27	132.27	jb l CaLo...	confirm C terminus,...	IXOSC	2.00	99	AFTFYSNGVYDFPK			-0.0001	1654.76...	2	23	1.1.1.4914.4	Win...
						2.00	99	AHGLASDEDYGY		cleaved Y-...	-0.0047	1367.55...	2	16	1.1.1.4685.4	Win...
						2.00	99	AHGLASDEDYGYLGQDGI		cleaved I-C...	0.0004	1950.85...	3	26	1.1.1.4870.2	Win...
						2.00	99	AHGLASDEDYGYLGQDGI...	Carbamidomethyl(C)@20		0.0038	2592.12...	4	27	1.1.1.4719.8	Win...
						2.00	99	ALANVGPSVSDAA		cleaved A-...	-0.0003	1382.74...	2	14	1.1.1.4933.3	Win...
						2.00	99	ALANVGPSVSDAAVK			-0.0027	1609.90...	2	16	1.1.1.4924.6	Win...
						2.00	99	APTNALFFPK			-0.0055	1054.57...	2	15	1.1.1.4720.3	Win...
						2.00	99	ASDEDYGYLGQDGI...	Carbamidomethyl(C)@16	cleaved L-...	0.0099	2213.92...	3	21	1.1.1.4727.2	Win...
						2.00	99	AVGYGT LQGEFYWL I K		cleaved L-...	-0.0015	1793.93...	2	24	1.1.1.4940.4	Win...
						2.00	99	AVLPDVTSPFVR		cleaved Q-...	-0.0007	1478.77...	2	13	1.1.1.4990.4	Win...
						2.00	99	DGTVL I SQK		cleaved N-...	0.0002	1021.54...	2	15	1.1.1.4708.3	Win...
						2.00	99	DNNGVASQGT YVE LHHH	Carbamidomethyl(C)@4	cleaved H-...	-0.0011	2036.87...	4	19	1.1.1.4578.3	Win...
						2.00	99	DNNGVASQGT YVE LHHH	Carbamidomethyl(C)@4	cleaved H-...	-0.0012	2173.93...	4	19	1.1.1.4524.2	Win...

Protein Sequence Coverage - confirm C terminus, putative OS=Jakub Benysek GN=IscW_ISCW007008 PE=3 SV=1

```

I I S V Q L K G G S D L N K L C I L L A V V A F V G L S L G R S Q L P E W Q D V Y K V G V L Y P Y A E I R E P T G Y F D A S Q N L S R I D Y R G M V Q T I Q V A P D A E G R S E A P Y G A N Y K I A M P O K E T W Q P R T C F L V N G T I N N T V P L Q A V L E D V T S F E F V R Q E S C W E G T T A V D E P A E R A S C E R F Q L A V T N E H R V S K Y T L W V S R D G G R A V P R Y L M G Y N
T L L G S H F D K Y E L I Y H G F S R K P L P S S V E D I T H L I N E T C R Q P P G P G A E H L A L H S P M A E F M D G H D A H M S A F D K F K E D H S R D Y G H T H E H R R R D I F R Q N L R F I H S T N R A N G Y T I V E N H L A D R S E E L G Y L R G L H S A P T N A L P F P K D R F S S D L P D Y I D W R L Y G A V T P V K D Q A V C G S C W S P G T V G E L B A A L E R K T G R L T R L S E Q Q L V
D C S W N Q N N C D G G E D P R A Y E Y T R A H G L A S D E D Y G A Y L G Q D G I C H D T K R A T V T T I R N Y I N V T D K E S L Q K A L A N V G P S V S I D A A V K A P T F Y S N G V Y D F P K R N D T D G L D H A V L A V G Y G T L Q G E P Y W L I K N S W S T Y W N D G V L I S Q R D N N C G V A S Q G T Y V E L H H H H H H
    
```

Protein ID from gel band
1D ELFO

Protein Group 1 - confirm C terminus, putative OS=Jakub Benysek GN=IscW_ISCW007008 PE=3 SV=1

Proteins in Group						Peptides in Group											
N	Unused	Total	Accessio...	Name	Species	Con...	Conf	Sequence	Modifications	Cleavages	ΔMass	Prec MW	z	Sc	Spectrum	Type	
1	132.27	132.27	jb IrcatL...	confirm C terminus,...	IXOSC	2.00	99	AFTFYSNGVFYDPK			-0.0001	1654.76...	2	23	1.1.1.4914.4	Win...	
						2.00	99	AHGLASDEDYGYAY		cleaved Y-...	-0.0047	1367.55...	2	16	1.1.1.4665.4	Win...	
						2.00	99	AHGLASDEDYGYAYLGQDGI		cleaved I-C...	0.0004	1950.85...	3	26	1.1.1.4870.2	Win...	
						2.00	99	AHGLASDEDYGYAYLGQDGI...	Carbamidomethyl(C)@20		0.0038	2592.12...	4	27	1.1.1.4719.8	Win...	
						2.00	99	ALANVGPSVSVS IDAA		cleaved A-...	-0.0003	1382.74...	2	14	1.1.1.4933.3	Win...	
						2.00	99	ALANVGPSVSVS IDAAVK			-0.0027	1609.90...	2	16	1.1.1.4924.6	Win...	
						2.00	99	APTNALPFPK			-0.0055	1054.57...	2	15	1.1.1.4720.3	Win...	
						2.00	99	ASDEDYGYAYLGQDGI	Carbamidomethyl(C)@16	cleaved L-...	0.0099	2213.92...	3	21	1.1.1.4727.2	Win...	
						2.00	99	AVGYGT LQGE P Y W L I K		cleaved L-...	-0.0015	1793.93...	2	24	1.1.1.4940.4	Win...	
						2.00	99	AVLPDVS T F E F V R		cleaved Q-...	-0.0007	1478.77...	2	13	1.1.1.4990.4	Win...	
						2.00	99	DGYV L I S Q K		cleaved N-...	0.0002	1021.54...	2	15	1.1.1.4708.3	Win...	
						2.00	99	DNNGVASQGT Y V E L H H H	Carbamidomethyl(C)@4	cleaved H-...	-0.0011	2036.87...	4	19	1.1.1.4578.3	Win...	
						2.00	99	DNNGVASQGT Y V E L H H H H	Carbamidomethyl(C)@4	cleaved H-...	-0.0012	2173.93...	4	19	1.1.1.4524.2	Win...	

N	Unused	Total	% Cov	Accessio...	Name	Species	Peptides(95%)	Bi
1	132.27	132.27	75.0	jb IrcatL...	confirm C terminus, putative OS=Jakub Benysek...	IXOSC	166	
2	57.56	57.56	61.9	cont 000...	pdb 1FNI_A A Chain A, Crystal Structure Of Por...	Sus scrofa (contaminant)	66	
3	42.05	42.05	46.8	cont 000...	cra hCP1812051 keratin 10(epidermolytic hype...	Homo sapiens (contaminant)	25	
4	36.10	36.10	45.8	cont 000...	cra hCP1609934.2 keratin 1(epidermolytic hyp...	Homo sapiens (contaminant)	22	
5	25.86	25.86	20.6	cont 000...	gij 114939 sp P00722 BGAL_ECOLI Beta-galact...	Escherichia coli (contaminant)	15	
6	9.15	9.15	35.9	cont 000...	gij 7441526 pir A59068 beta-casein variant CnH...	Bos taurus (contaminant)	6	
7	8.00	8.00	26.2	cont 000...	gij 115646 sp P02662 CAS1_BOVIN Alpha-S1-c...	Bos taurus (contaminant)	5	
8	7.01	7.11	14.4	cont 000...	spt P48666 Keratin, typell cytoskeletal 6C (Cyt...	Homo sapiens (contaminant)	4	
9	4.01	4.01	43.6	cont 000...	spt P81605 Dermcidin precursor(Preproteolysi...	Homo sapiens (contaminant)	2	
10	2.25	2.25	17.6	cont 000...	gij 27806963 ref NP_776953.1 casein alpha-S2...	Bos taurus (contaminant)	2	
11	2.00	2.00	25.8	cont 000...	gij 27881412 ref NP_776719.1 casein kappa[Bo...	Bos taurus (contaminant)	1	

Protein Sequence Coverage - confirm C terminus, putative OS=Jakub Benysek GN=IscW_ISCW007008 PE=3 SV=1

II SVQLKGGSDLNMKLCILLAVVA FVGLSLGRSQELPEWGDVYKVKGVLYLPYAEIREPFTGYFDASQNL SRIDYYRGMVQTIQVAPDAEGRSEAPYGAN YKIA YMPDKE TWQPRRT C FLVNGT INNTVPLQAVLPDVT S FEFV RQESCWEGT TAVDEPAERAS CERFQLAVTNEHRVSKYT LWVSRDGGQRAVPRRYLMMGYN
 TLLGSHFDKYELIYHGFSRKPLPSSVFDITHLINETCRQFPGPGEHLALHSPMAEFMDGHDAHMSAFDKFKEDHSRDYGHHTHEHRRRDI FRQNLRFIHSNTRNRNGYTVENVHLADRSSEELG YLRGLRHSR APTNALPFPKDRFSSDLPDYIDWRLYGAVT PVKQAVCGSCWSFGTVGEL E AALFRKTGKLR LSEQQLV
 DCSWNQGNNGCDGGEDFRAYEYIRAHGLASDEDYGYAYLGQDGI CHDTRV NATVTTIRNYINVTDKESLQKALANVGPSVSVS IDAAVK AFTFYSNGVFYDPKCRNDTDGLDHA V L AVGYGT LQGE P Y W L I R K N S W S T Y W G N D G Y V L I S Q K D N N C G V A S Q G T Y V E L H H H H H H

	A	B	C	D	E	F	G	H	I	J
1	N	Unused	Total	%Cov	%Cov(50)	%Cov(95)	Accession	Name	Species	Peptides(95%)
2	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C	IXOSC	237
3	2	52.1	52.1	62.33	57.4	51.57	cont 0001	pdb 1FNI	Sus scrofa	76
4	2	0	52.1	60.17	55.41	49.78	cont 0001	spt P0076	Sus scrofa	76
5	2	0	52.1	62.33	57.4	51.57	cont 0000	gi 291448	Sus scrofa	76
6	2	0	52.1	62.33	57.4	51.57	cont 0000	gi 494360	Sus scrofa	76
7	3	29.2	29.2	37.73	31.68	31.68	cont 0001	cra hCP1E	Homo sap	17
8	3	0	29.2	37.73	31.68	31.68	cont 0001	rf NP_00E	Homo sap	17
9	3	0	29.2	37.73	31.68	31.68	cont 0001	sp P04264	Homo sap	17
10	4	27.49	27.49	37.66	37.66	36.22	cont 0001	cra hCP1E	Homo sap	17
11	4	0	27.49	40.24	40.24	38.7	cont 0001	trm Q8N1	Homo sap	17
12	4	0	25.15	35.08	35.08	33.56	cont 0001	pir KRHU	Homo sap	16
13	4	0	25.15	35.08	35.08	33.56	cont 0001	spt P1364	Homo sap	16
14	5	23.89	23.89	18.55	16.6	16.6	cont 0001	gi 114939	Escherichi	13
15	6	10	10	26.17	26.17	26.17	cont 0000	gi 115646	Bos taurus	5
16	7	6.58	6.58	32.88	23.42	23.42	cont 0000	gi 278069	Bos taurus	4
17	8	6.33	6.33	38.76	35.89	35.89	cont 0001	gi 744152	Bos taurus	5
18	8	0	6.33	36.16	33.48	33.48	cont 0000	gi 459292	Bos taurus	5
19	8	0	6.33	36.16	33.48	33.48	cont 0000	gi 115660	Bos taurus	5
20	8	0	4.2	27.23	24.55	24.55	cont 0000	gi 162805	Bos taurus	4
21	9	4.75	4.75	21.66	14.71	12.57	tr Q5D575	Actin (Fra	IXORI	3
22	9	0	4.75	21.54	14.63	12.5	tr Q5CAR	Actin OS=	IXORI	3
23	9	0	4.75	24.18	16.42	14.03	tr B7PBG2	Actin, put	IXOSC	3

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	Formula Bar	V	W	X	Y	Z	AA	AB
1	N	Unused	Total	%Cov	%Cov(50)	%Cov(95)	Accession	Names	Used	Annotatio	Contrib	Conf	Sequence	Modificati	Cleavages	dMass	Prec MW	Prec m/z	Theor MW	Theor m/z	Theor z	Sc	Spectrum	Specific	Time	Precursor	Precursor	Elution
2	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	AFTFYSGVGFYD	cleaved D	0.000105	1429.619	715.8168	1429.619	715.8168	2	14	1.1.1.4867	1	44.1906	196.227	44.2223		
3	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	AFTFYSGVGFYD	Ammonia-loss(N)@	-0.00172	1637.739	819.8765	1637.74	819.8774	2	18	1.1.1.4765	1	40.7634	1615.713	40.7791		
4	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	AHGLASDEDYGY	cleaved Y	-0.00459	1367.558	684.7865	1367.563	684.7888	2	16	1.1.1.4501	1	30.1511	1328.73	30.1561		
5	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	AHGLASDEDYGYLGC	cleaved D	-0.00185	1780.752	891.3834	1780.754	891.3843	2	19	1.1.1.4622	1	34.8851	270.4055	34.7998		
6	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	AHGLASDEDYGYLGC	cleaved I	-0.00825	1950.851	976.4329	1950.859	976.4371	2	23	1.1.1.4693	1	37.7562	586.3611	37.7613		
7	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	AHGLASDI	Carbamidomethyl(C	-9.039	2583.08	862.0338	2592.119	865.0468	3	20	1.1.1.4585	1	33.4759	1703.331	33.4439		
8	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	ALANVGPVSVSIDA	cleaved A	-0.00336	1382.737	692.3759	1382.741	692.3776	2	15	1.1.1.4750	1	40.0856	235.0982	40.069		
9	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	ALANVGPVSVSIDA	AVK	-0.0034	1609.901	805.9575	1609.904	805.9592	2	19	1.1.1.4731	1	39.3289	501.4322	39.1705		
10	1	184.43	184.43	82.04	80.14	74.78	jb IrCatLo	confirm C terminus, putative O			2	99	ALANVGPVSVSIDA	A missed K-	0.02065	3246.68	1083.234	3246.66	1083.227	3	20	1.1.1.5207	1	49.8459	315.0218	49.8065		

Protein ID from complex protein mixture – human tissue

ProteinPilot™ Software - [Result - C:\AB SCIE\ProteinPilot Data\Data\2013\2013_August\130819_Saskova\130819_0813_1_group]

File Configure Window Help

Protein ID
Spectra
Summary Statistics

Workflow Tasks

Identify Proteins

LC...

Spot-Based (MS only)...

Spot-Based (MS and MS/MS)...

View

Analysis Log...

Result...

Export

Peptide Summary...

Protein Summary...

MGF Peaklist(s)...

Proteins Detected

N	Unused	Total	% Cov	Accessio...	Name	Species	Peptides(95%)	Biological Processes	Molecular Functions	PANTHER ID
1	55.91	55.91	51.2	trjB4DNT...	cDNA.FLJ54370, highly similar to Heat shock 70...	HUMAN	34			
2	52.58	52.58	51.0	spjP0790...	Heat shock protein HSP90-alpha OS=Homo sa...	HUMAN	29			
3	50.42	50.42	69.6	spjP0673...	Alpha-enolase OS=Homo sapiens GN=ENO1 P...	HUMAN	28			
4	49.86	49.86	40.6	spjP2231...	Ubiquitin-like modifier-activating enzyme 1 OS=...	HUMAN	28			
5	48.98	48.98	71.6	spjP1461...	Pyruvate kinase isozymes M1/M2 OS=Homo sa...	HUMAN	27			
6	45.45	45.45	69.8	trjQ5SU1...	Beta 5-tubulin OS=Homo sapiens GN=TUBB PE...	HUMAN	30			
7	44.82	44.82	48.3	spjP1363...	Elongation factor 2 OS=Homo sapiens GN=EEF...	HUMAN	25			
8	40.72	40.72	39.8	trjH6VRG...	Keratin 1 OS=Homo sapiens GN=KRT1 PE=3 S...	HUMAN	24			
9	35.25	35.25	46.3	spjP1080...	80 kDa heat shock protein, mitochondrial OS=H...	HUMAN	21			
10	35.05	35.05	34.1	spjP1102...	78 kDa glucose-regulated protein OS=Homo sa...	HUMAN	18			
11	34.73	49.19	51.6	trjB4DMA...	cDNA.FLJ54023, highly similar to Heat shock pr...	HUMAN	29			
12	34.52	34.52	66.7	spjP0667...	ATP synthase subunit beta, mitochondrial OS=H...	HUMAN	19			
13	34.09	34.09	42.2	spjP5507...	Transitional endoplasmic reticulum ATPase OS=...	HUMAN	18			
14	33.98	33.98	66.4	spjP1227...	Creatine kinase B-type OS=Homo sapiens GN=...	HUMAN	20			
15	33.09	33.09	60.1	trjB3KT0...	cDNA.FLJ37398 fis, clone BRAMY2027467, hig...	HUMAN	22			
16	33.01	33.01	61.9	spjP0066...	Phosphoglycerate kinase 1 OS=Homo sapiens...	HUMAN	20			
17	32.41	32.41	32.1	spjP0987...	Poly [ADP-ribose] polymerase 1 OS=Homo sapi...	HUMAN	17			
18	32.07	38.14	36.9	spjP1114...	Heat shock cognate 71 kDa protein OS=Homo s...	HUMAN	23			
19	30.66	30.66	41.1	spjP2570...	ATP synthase subunit alpha, mitochondrial OS=...	HUMAN	17			
20	30.18	30.18	57.7	spjP4092...	Malate dehydrogenase, mitochondrial OS=Hom...	HUMAN	15			
21	29.40	33.52	32.8	trjQ5CA...	Tumor rejection antigen (Gp96) 1 OS=Homo sa...	HUMAN	17			
22	29.22	29.22	39.9	spjO4370...	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4...	HUMAN	14			
23	28.63	28.63	31.9	trjB4DH0...	cDNA.FLJ50510, highly similar to Heat shock 70...	HUMAN	14			

Protein Group 1 - cDNA.FLJ54370, highly similar to Heat shock 70 kDa protein 1 OS=Homo sapiens PE=2 SV=1

Proteins in Group						Peptides in Group										
N	Unused	Total	Accessio...	Name	Species	Con...	Conf	Sequence	Modifications	Cleavages	ΔMass	Prec MW	z	Sc	Spectrum	Type
1	55.91	55.91	trjB4DNT...	cDNA.FLJ54370, hig...	HUMAN	2.00	99	AFYPEEISSMVLTK			0.0005	1613.80...	3	20	1.1.1.9096.3	Win...
18	32.07	38.14	spjP1114...	Heat shock cognate...	HUMAN	2.00	99	AQIHDLVLVGGSTR			-0.0007	1464.80...	3	16	1.1.1.8558.4	Win...
29	25.74	27.77	trjB7Z4V...	cDNA.FLJ51907, hig...	HUMAN	2.00	99	ARFEELCSDLFR	Carbamidomethyl(C)@7	missed R-F...	-0.0032	1541.72...	3	13	1.1.1.8896.4	Win...
	0.00	55.85	trjQ59EJ...	Heat shock 70kDa pro...	HUMAN	2.00	99	CQEVISWLDANTLAEKDEP...	Carbamidomethyl(C)@1 Deamidated(Q)@2 Dehydrated(E)@3	missed K-D...	-0.0005	2644.21...	3	19	1.1.1.9379.3	Win...
	0.00	55.85	trjA8K3I0...	Heat shock 70kDa pro...	HUMAN	2.00	99	DAGVIAGLNVLRL			-0.0022	1196.68...	2	20	1.1.1.8931.6	Win...
	0.00	55.85	spjP0810...	Heat shock 70 kDa pr...	HUMAN	2.00	99	DAKLDKAQIHDLVLVGGSTR		missed K-L...	-0.0064	2136.16...	4	13	1.1.1.8802.4	Win...
	0.00	55.85	spjP0810...	Heat shock 70 kDa pr...	HUMAN	2.00	99	EIAEAYLGFVIVAVIVP...			0.0022	3000.48...	3	27	1.1.1.9470.2	Win...

Protein Sequence Coverage - cDNA.FLJ54370, highly similar to Heat shock 70 kDa protein 1 OS=Homo sapiens PE=2 SV=1

MAKAAAIIGDIDGTTPTSYVAFTDTERLIGDAAPKQVALNPQNTVFDAKRLIGRKFQDPPVQSDMKHWPQFVINDGDKPKVQVSYRGDTRKAFYPEEISSMVLTKMKEIAEAYLGYPVNTAVITVPVAFNDSSQQRATKADGVIAGLNVLRIINEPTAAAIAYGLDRTGKGRNVLIFDLGGGTFDVSILTIIDGGIFEVKATAGDTHLGGEDFDNRLVNHFFVEEFKHKHKDDISQNKRAVRRLRACERAKRLLSSSTQASLEIDSLFEGIDFYTITRFRFEELCSDLFRSTLEPVEKALRDAKLDKAQIHDLVLVGGSTRIPKVGQLLQDFFNDRDLNKSINPDEAVYGAQVAAILMGDKSENVQDLLLLDVAPLSLGLTAGGVTALIKRNTIIPKQQTQIFTTYSNDQQGVLIQVYEGERAMTKDNNLLGRFELSGIPAPRGPQIEVTFDIDANGILNVTADKSTGKANKIITINDKGRLSKEEIERMVQAEAKYKAEDEVQRRVSAKNALESYAPNMSAVEDEGLKGRISSEADKVKVLDKQCEVIVSWLDANTLAEKDEFEHKKRKELEQVCNPIISGLYQAGGPGGGFGAQQPKGGSGSGPTIEEVD

Proteins Detected

N	Unused	Total	% Cov	Accessio...	Name	Species	Peptides(95%)	E
1	55.91	55.91	51.2	tr B4DNT...	cDNA FLJ54370, highly similar to Heat shock 70...	HUMAN	34	
2	52.58	52.58	51.0	sp P0790...	Heat shock protein HSP 90-alpha OS=Homo sa...	HUMAN	29	
3	50.42	50.42	69.6	sp P0673...	Alpha-enolase OS=Homo sapiens GN=ENO1 P...	HUMAN	28	
4	49.86	49.86	40.6	sp P2231...	Ubiquitin-like modifier-activating enzyme 1 OS=...	HUMAN	28	
5	48.98	48.98	71.6	sp P1461...	Pyruvate kinase isozymes M1/M2 OS=Homo sa...	HUMAN	27	
6	45.45	45.45	69.8	tr Q5SU1...	Beta 5-tubulin OS=Homo sapiens GN=TUBB PE...	HUMAN	30	
7	44.82	44.82	48.3	sp P1363...	Elongation factor 2 OS=Homo sapiens GN=EEF...	HUMAN	25	

Protein Group 1 - cDNA FLJ54370, highly similar to Heat shock 70 kDa protein 1 OS=Homo sapiens PE=2 SV=1

Proteins in Group					
N	Unused	Total	Accessio...	Name	Species
1	55.91	55.91	tr B4DNT...	cDNA FLJ54370, hig...	HUMAN
18	32.07	38.14	sp P1114...	Heat shock cognate...	HUMAN
29	25.74	27.77	tr B7Z4V...	cDNA FLJ51907, hig...	HUMAN
	0.00	55.85	tr Q59EJ...	Heat shock 70kDa pro...	HUMAN
	0.00	55.85	tr A8K5I0...	Heat shock 70kDa pro...	HUMAN
	0.00	55.85	sp P0810...	Heat shock 70 kDa pr...	HUMAN

Peptides in Group											
Con...	Conf	Sequence	Modifications	Cleavages	ΔMass	Prec MW	z	Sc	Spectrum	Type	
2.00	99	AFYPEEISSMVLTK			0.0005	1613.80...	3	20	1.1.1.9096.3	Win...	
2.00	99	AQIHDLVLVGGSTR			-0.0007	1464.80...	3	16	1.1.1.8558.4	Win...	
2.00	99	ARFEELCSDLFR	Carbamidomethyl(C)@7	missed R-F...	-0.0032	1541.72...	3	13	1.1.1.8896.4	Win...	
2.00	99	CQEVISWLDANTLAEKDEF...	Carbamidomethyl(C)@1 Deamidated(Q)@2 Dehydrated(E)@3	missed K-D...	-0.0005	2644.21...	3	19	1.1.1.9379.3	Win...	
2.00	99	DAGVIAGLNVLK			-0.0022	1196.68...	2	20	1.1.1.8931.6	Win...	
2.00	99	DAKLDKAQIHDLVLVGGSTR		missed K-L...	-0.0064	2135.16...	4	13	1.1.1.8802.4	Win...	
2.00	99	EIAEAYLGYPVTNAVITVP...			0.0022	3000.48...	3	27	1.1.1.9470.2	Win...	

22	29.22	29.22	39.9	sp O4370...	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4...	HUMAN	14	
23	28.63	28.63	31.9	tr B4DH0...	cDNA FLJ50510, highly similar to Heat shock 70...	HUMAN	14	

Request for MS Proteomic Analysis

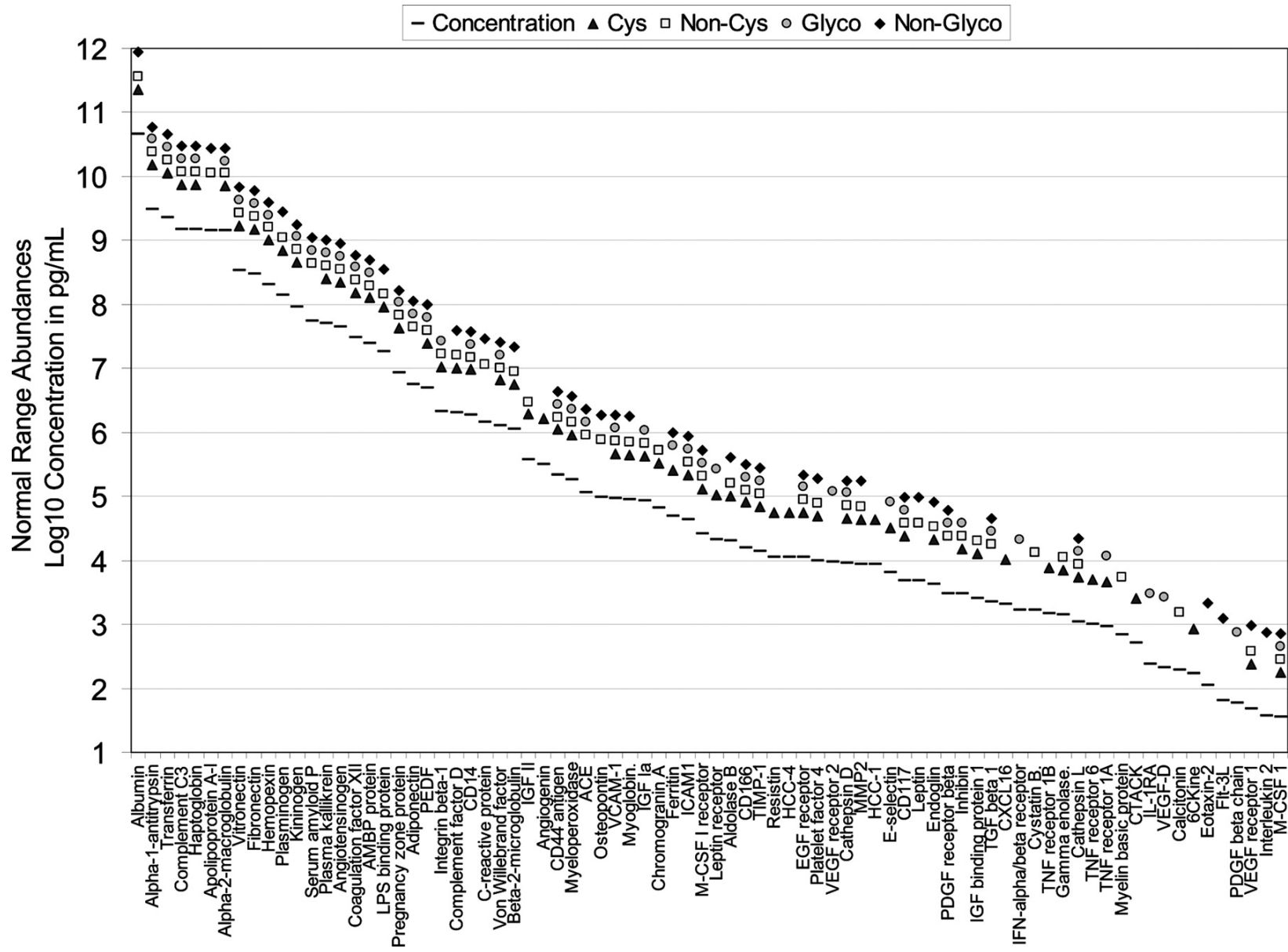
Request for MS Proteomic Analysis	
Date accepted	
From e-mail or phone #	
Team code	

Sample description				
Sample name(s)				
Sample type	Liquid <input type="checkbox"/>		Gel <input type="checkbox"/>	Solid <input type="checkbox"/>
	buffer		stain ing	
volume (μl)				
Estimated total amount of proteins (μg and/or pmol)		Based on eg. BCA, Lowry, Bradford		
Biological source (organism)				
Other relevant information eg. handling, storage				

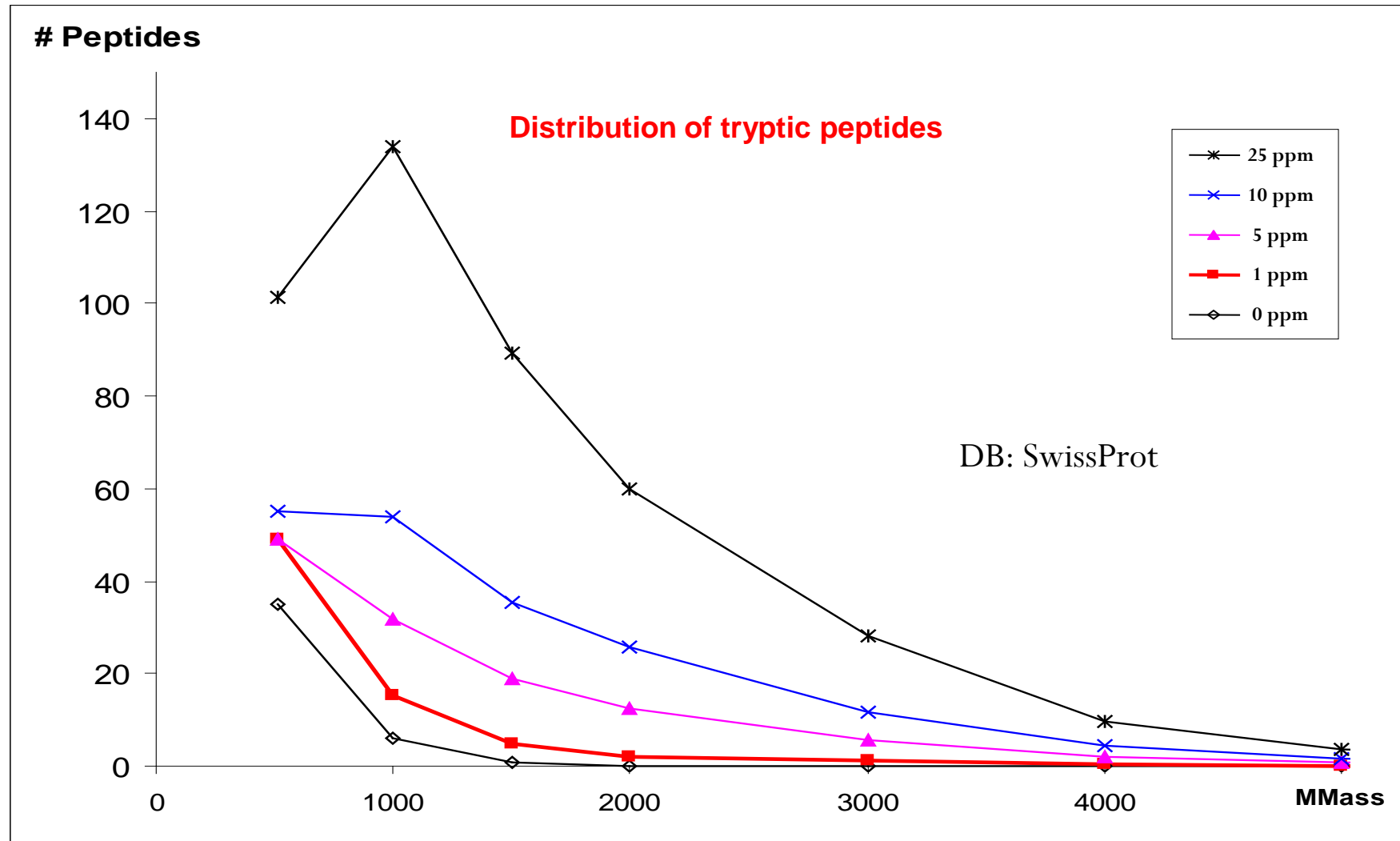
Analysis	
Request specification eg. identification of one or all protein(s)	
Protein sequence (if known)	
Database requested	
Additional information eg. return or storage request, proposed digestion enzyme, proposed measurement queue	

Proteomic team

- Jana Horáková
- Karel Rucker

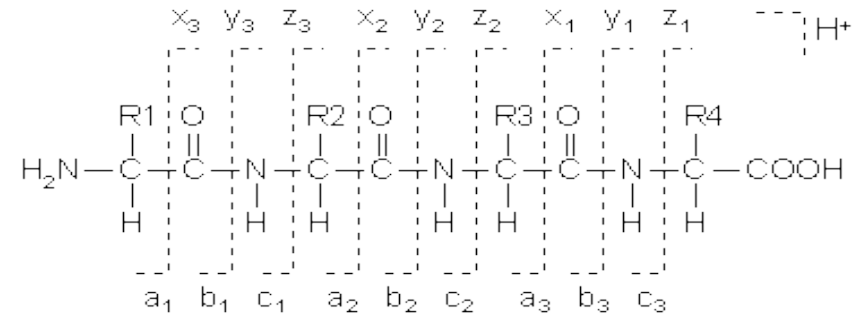
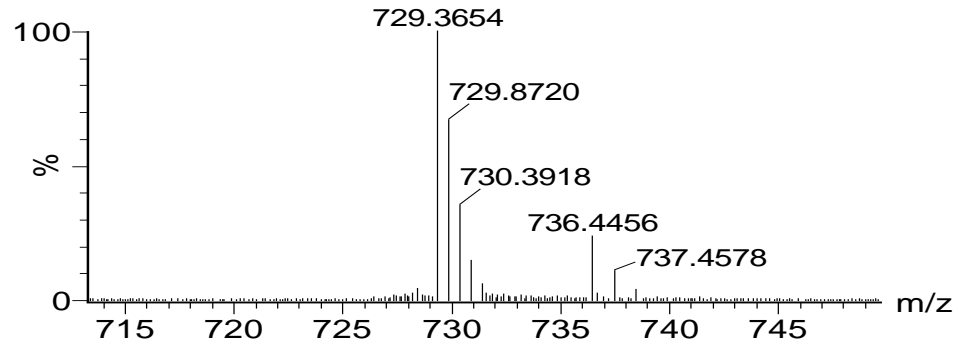


Mass Accuracy and peptide frequency

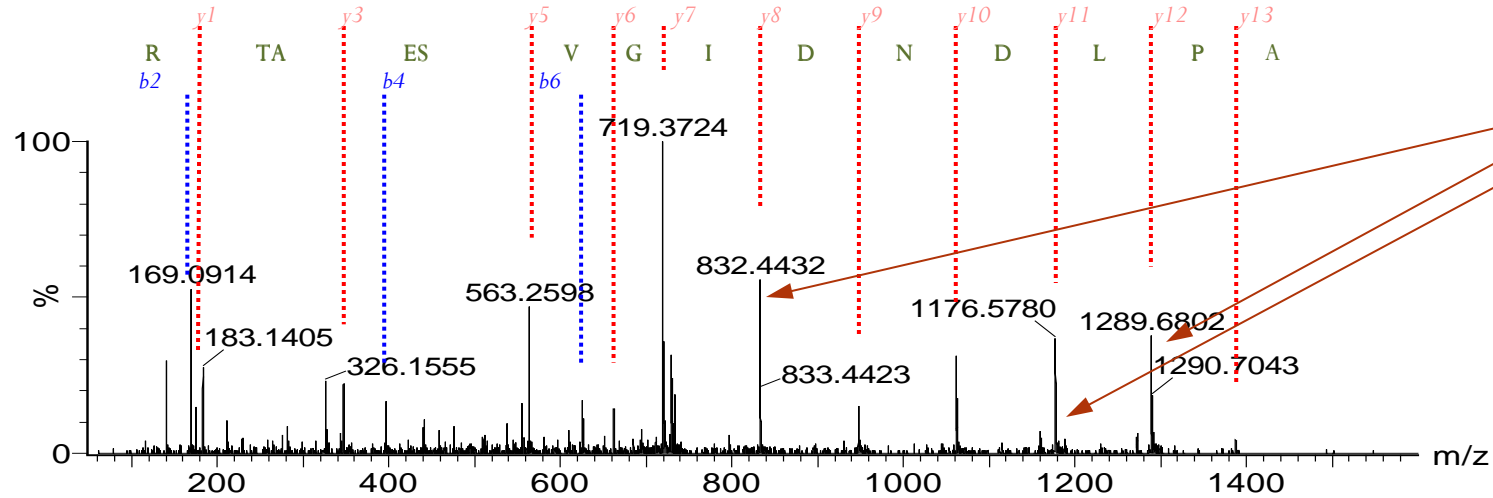


APLDNDIGVSEATR (Beta-galaktosidáza)

MS



MSMS



Seq	#	B	Y	# (+1)
A	1	72.04498	1457.72365	14
P	2	169.09775	1386.68654	13
L	3	282.18181	1289.63378	12
D	4	397.20875	1176.54971	11
N	5	511.25168	1061.52277	10
D	6	626.27862	947.47984	9
I	7	739.36269	832.45290	8
G	8	796.38415	719.36884	7
V	9	895.45256	662.34737	6
S	10	982.48459	563.27896	5
E	11	1111.52719	476.24693	4
A	12	1182.56430	347.20434	3
T	13	1283.61198	276.16722	2
R	14	1439.71309	175.11955	1

Přechody:

ID	Q1	Q3	CE
APLDNDIGVSEATry8	729,4	832,5	42
APLDNDIGVSEATry12	729,4	1289,6	42
APLDNDIGVSEATry11	729,4	1176,6	42

