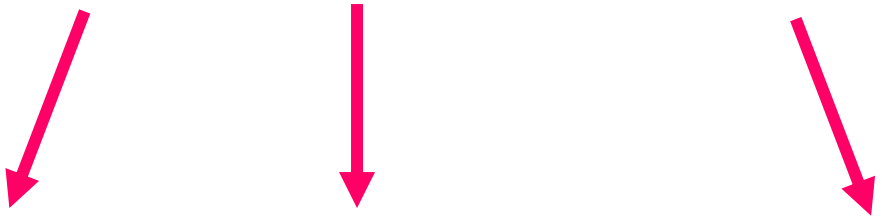
PAF introduction

Protein Analysis Facility
(PAF)

Protein Analysis Facility (PAF) is a core laboratory devoted to the study of proteins and proteomes

The PAF (also called UNIL Proteomics Platform) is a research and service core facility devoted to the high-performance analysis of proteins

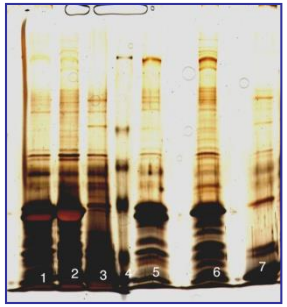


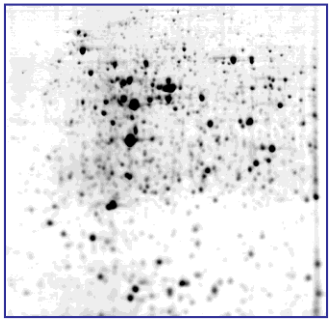
- **Service :**
  - **Identification of gel-separated proteins by MS**
  - **Project discussion**
  - **2D-PAGE and/or other multi-dimensional separation techniques**

- **Research :**
  - **Development of proteomics methods**
  - **Applications to biological problems**

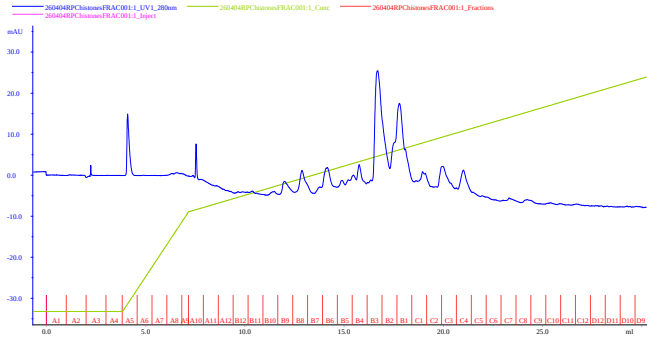
- **Teaching :**
  - **Techniques, possibilities & limitations of proteomics approaches**

PAF technology and instrumentation

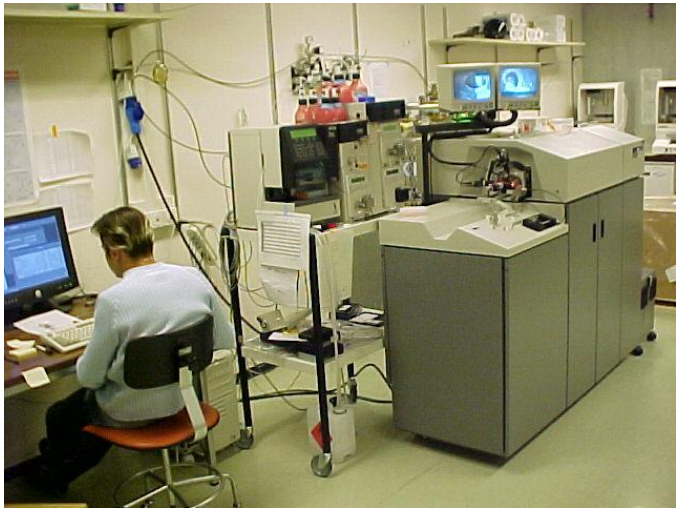




1D-electrophoresis **2D-electrophoresis**



Liquid Chromatography



Nano-HPLC- Quadrupole-time-of-flight Mass spectrometer

Nano-HPLC-triple quadrupole ion trap Mass spectrometer

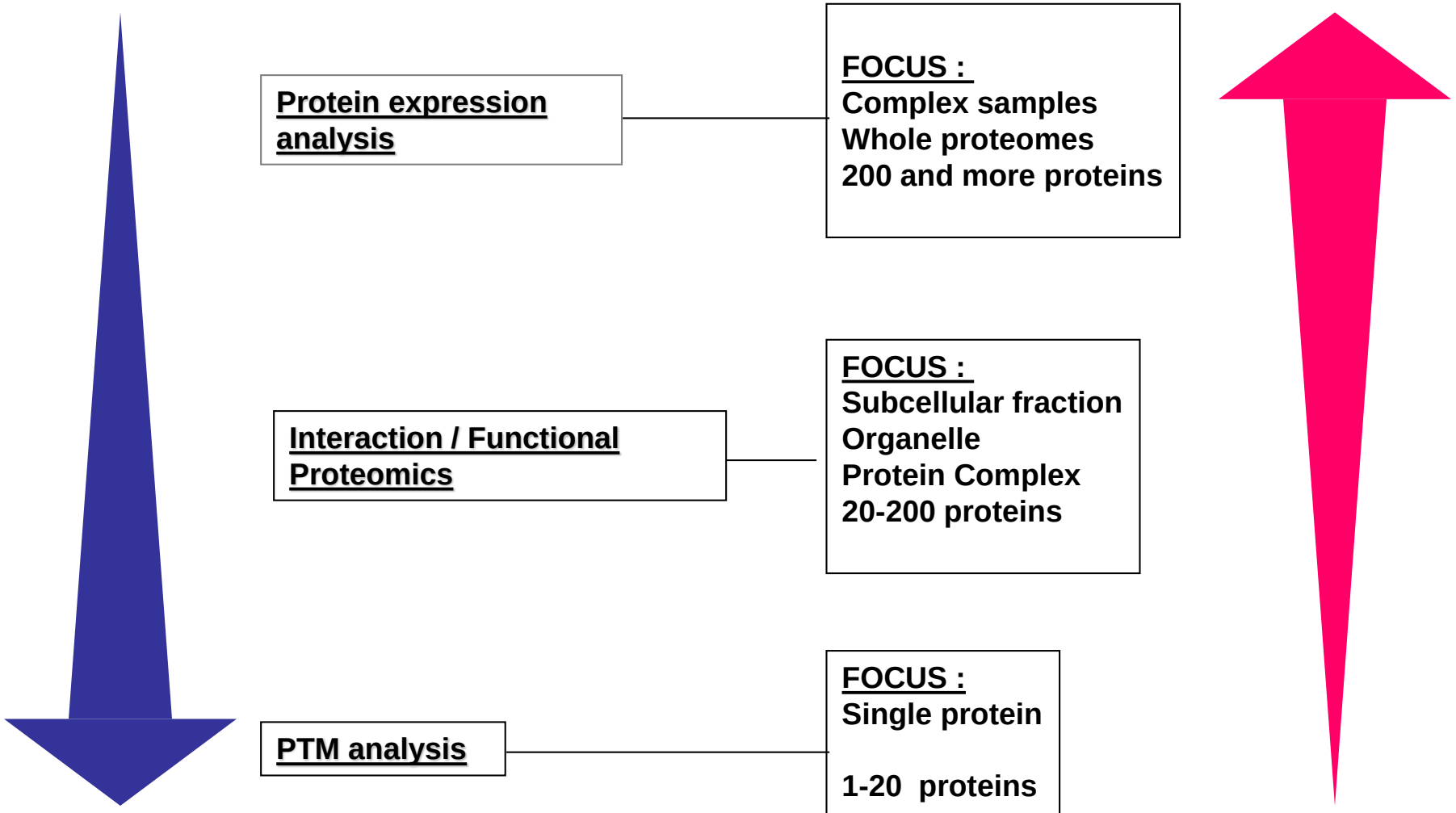
New : access to ABI 4700 TOF/TOF™ (courtesy Dept. Biochemistry)

MALDI - Tandem Time of Flight Mass Spectrometer for High Throughput Protein identification

Challenges of *in vivo* proteomics

- Complexity :
  - 35'000 genes (?) in *H. sapiens,*
  - *15'000 expressed in a single cell ?*
  - *but >50'000 chemically different protein species ?*
- Dynamic range :
  - 10^5 x or 10^6 x between low and high-abundance proteins
- Plasticity :
  - continuous variation in protein expression pattern (every s), PTM's, degradation,…

Overview : classes of proteomics experiments



post-translational modification (*PTM*)

Proteomics

- 2D-PAGE
- Mass Spectrometry for proteomics
- Proteomics workflows and applications

2D-PAGE

IEF: the principle

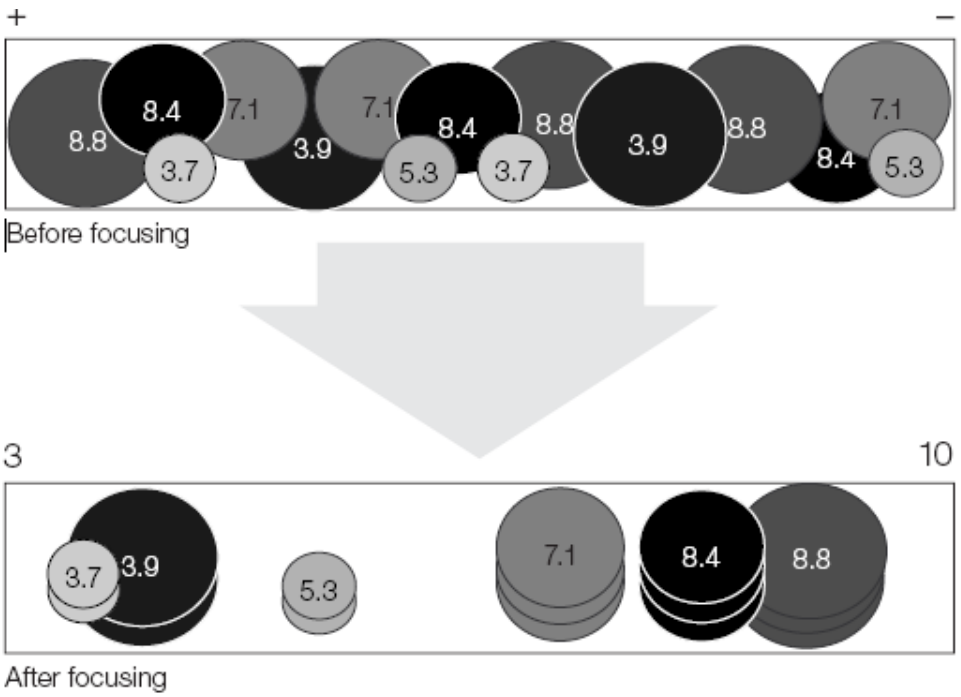
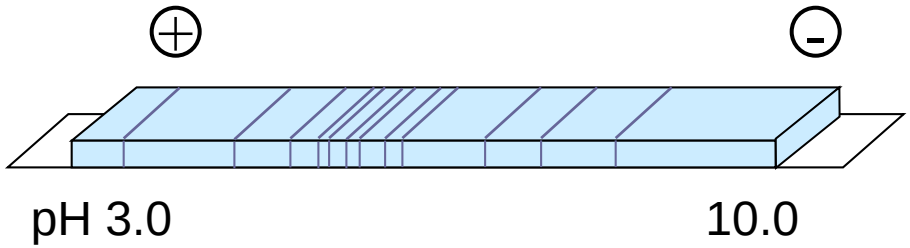


Fig. 3.1. A mixture of proteins is resolved on a pH 3–10 IPG strip according to each protein's pI and independently of its size, as described in the IEF section.

How to create a pH gradient ?

Improvements in 2D-PAGE

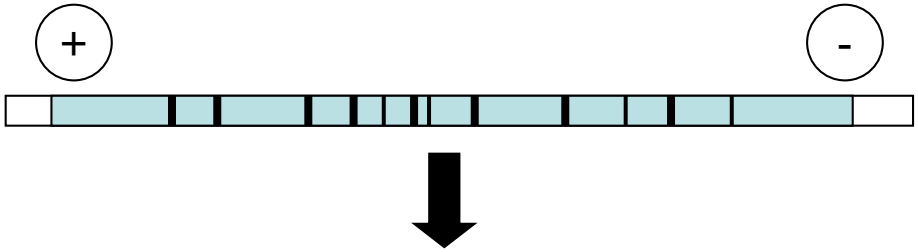


- IPG (Immobilised PH Gradient) strips for the first dimension

  pH-forming chemical groups are grafted onto the polyacrylamide matrix, creating a mechanically stable pH gradient

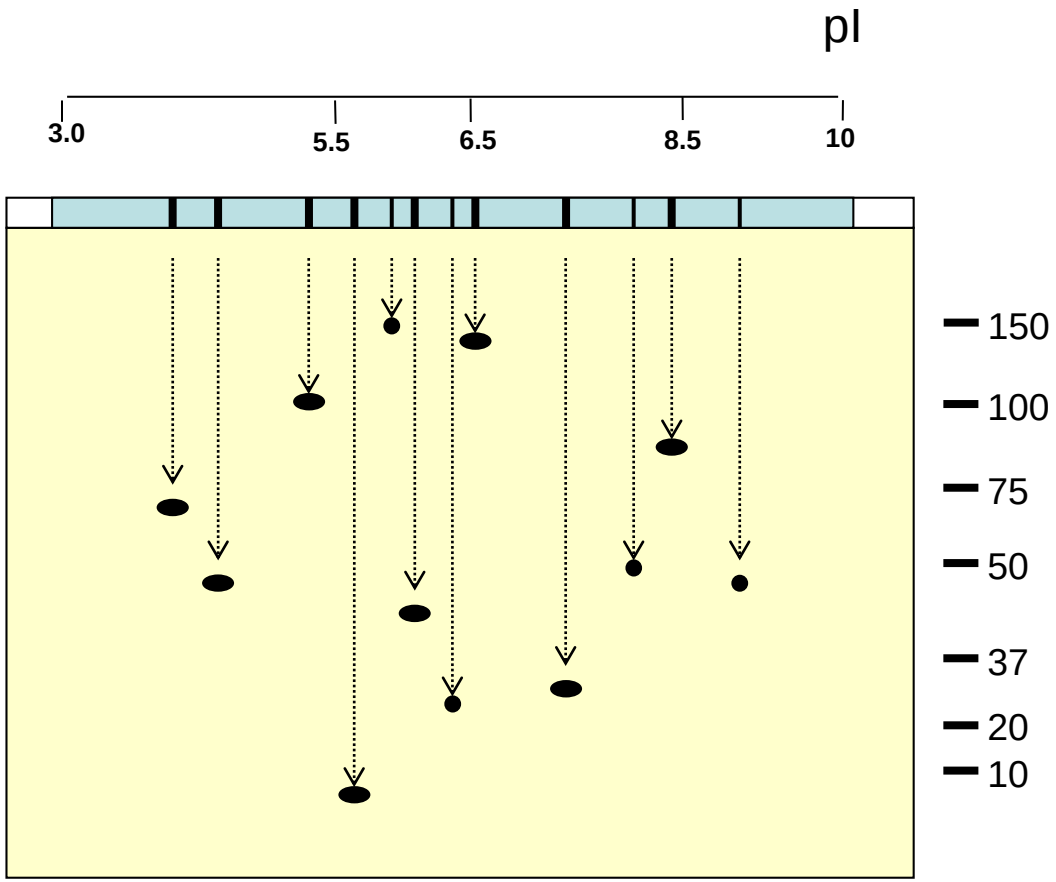
++ mechanical stability
++ reproducibility
++ loadable amounts
++ „zoom“ pI ranges

After IEF : equilibration and 2nd dimension



Equilibration step 1 :
- **SDS**
- **Buffer pH 6.8**
- **DTT (reduce –S-S-)**

Equilibration step 2 :
- **SDS**
- **Buffer pH 6.8**
- **Iodoacetamide (alkylate –S-S-)**



* Dithiothreitol (*DTT*)

One protein → many spots

Post-translational modifications can result in changes in pI and/or MW

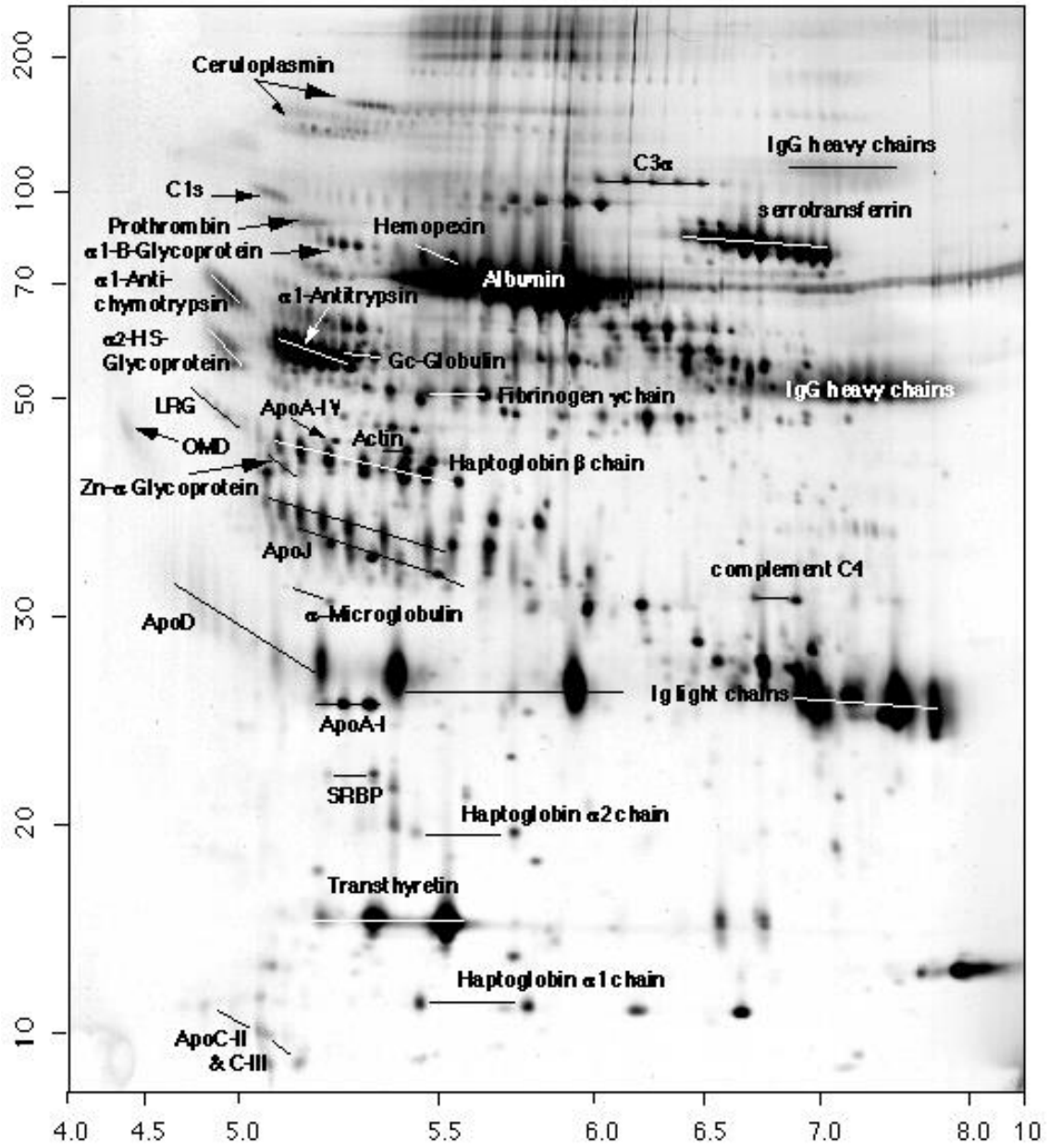
Spot “trains” for extensively modified proteins

Caused by

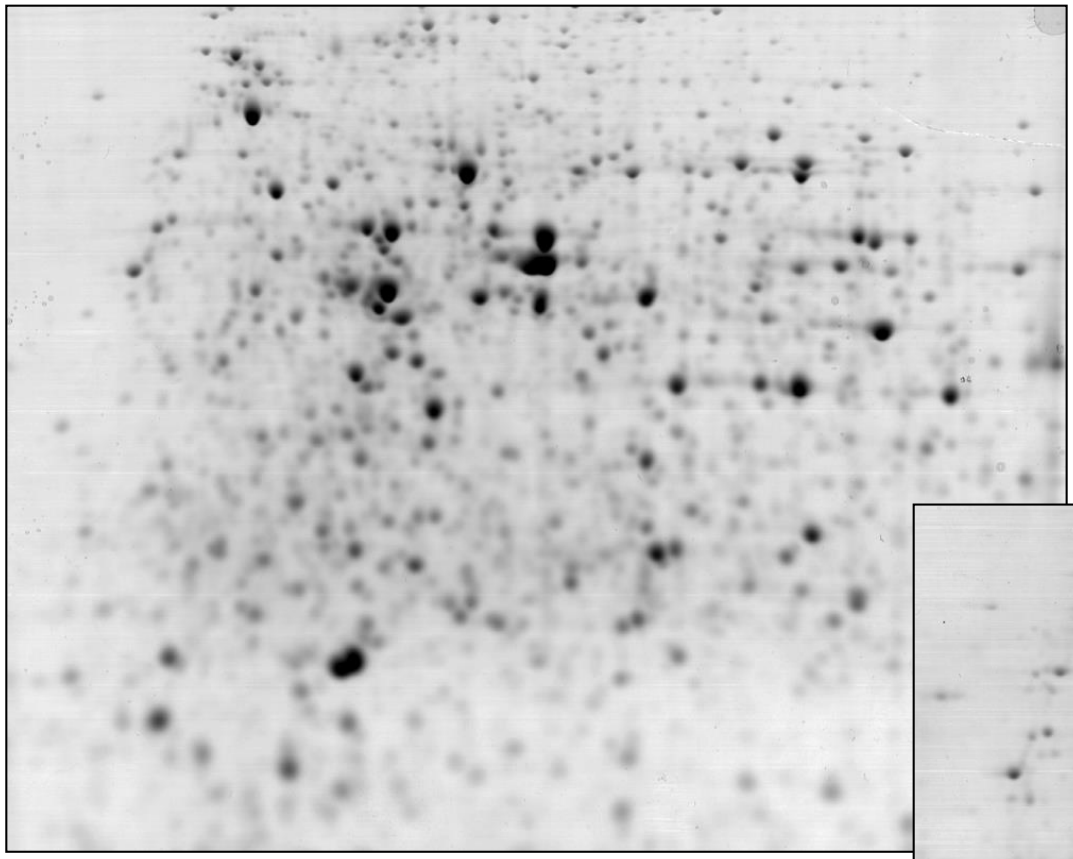
Glycosylation
Phosphorylation
Acetylation (K)
…

Database of 2D images with clickable spots :

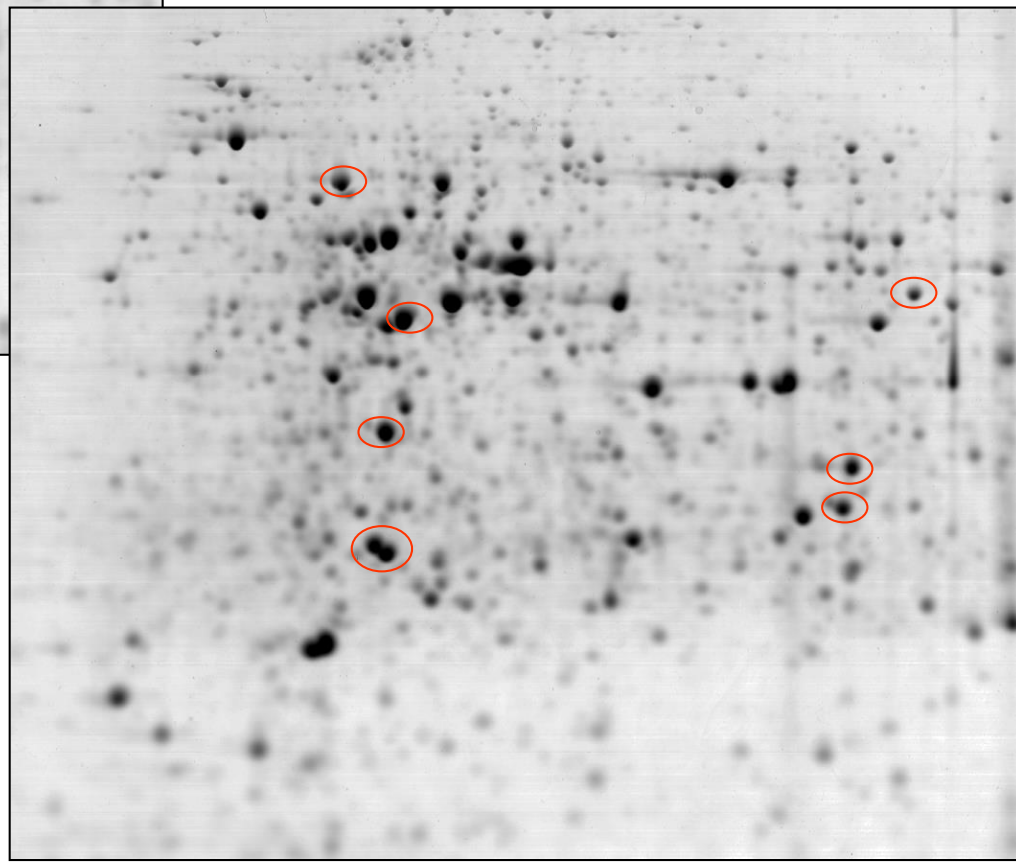
www.expasy.org/ch2d/



2D-PAGE and image analysis are used for studying changes in composition of the proteome



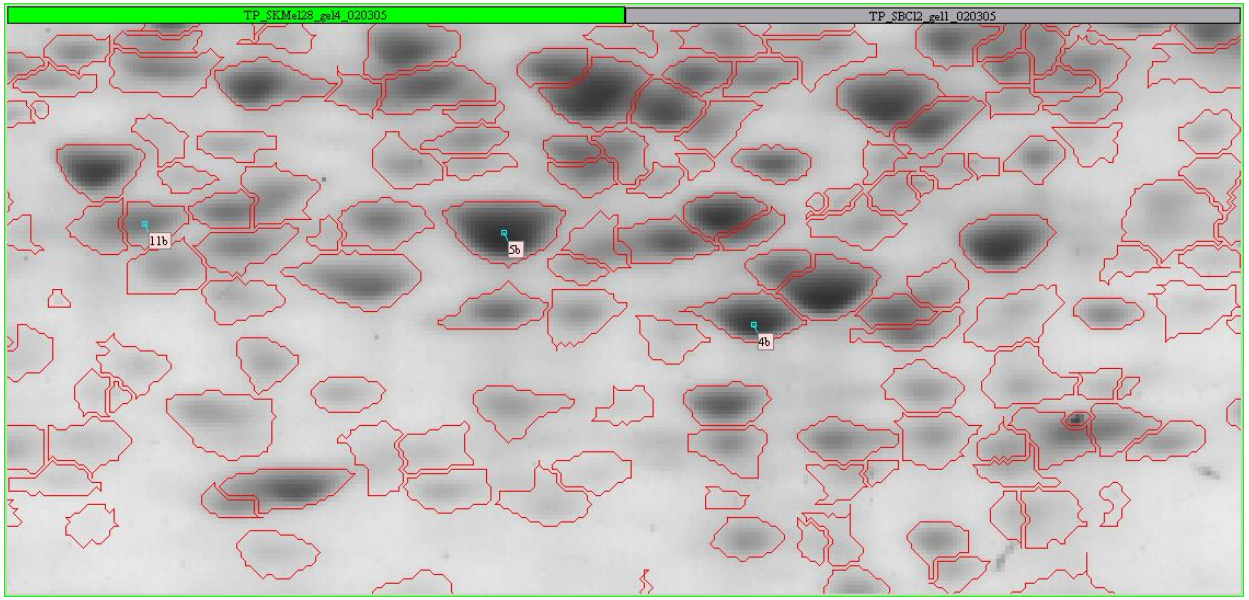
Control

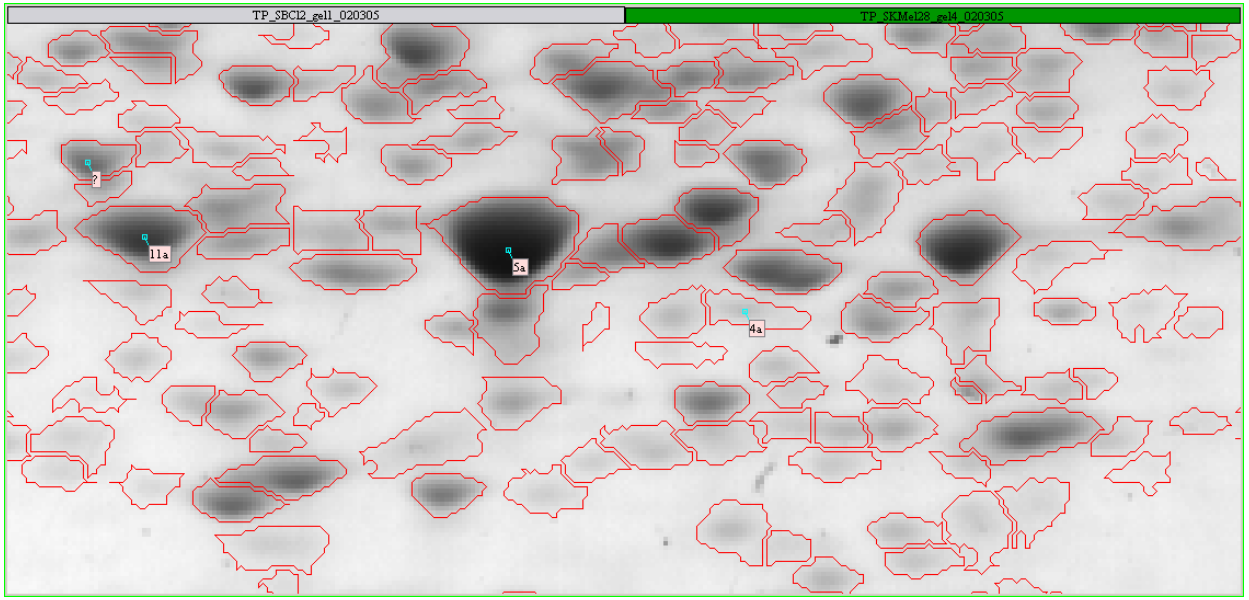


Stimulus applied

Software-based image analysis

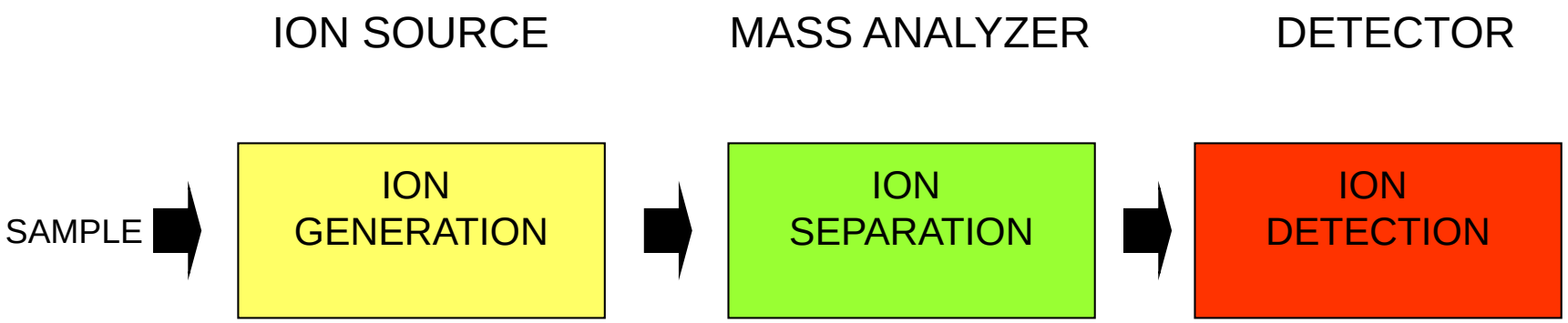
- **Spot detection**
- **Spot quantification**
- **Gel-to-gel**
- **Matching**
- **Presence/absence of spot**
- **Up/down regulation**
- **Statistic analysis**





Mass Spectrometry in proteomics

Mass spectrometry : essential functions



ESI : Electrospray Ionisation

MALDI : Matrix Assisted Laser Disorption/Ionization

Quadrupoles
Ion traps
Time-of-flight with reflectron
TOF/TOF
FT-ICR
(Fourier transform – Ion Cyclotron Resonance)

Faraday cup
Scintillation counter
Electromultiplier
High-energy dynodes with electronmultiplier
Array (detector)
FT-MS

Masses and mass measurements

- All mass spectrometers function measure molecules in their ionized state

- All values determined by MS are relative to the **m/z** assumed by the molecule after the ionization process

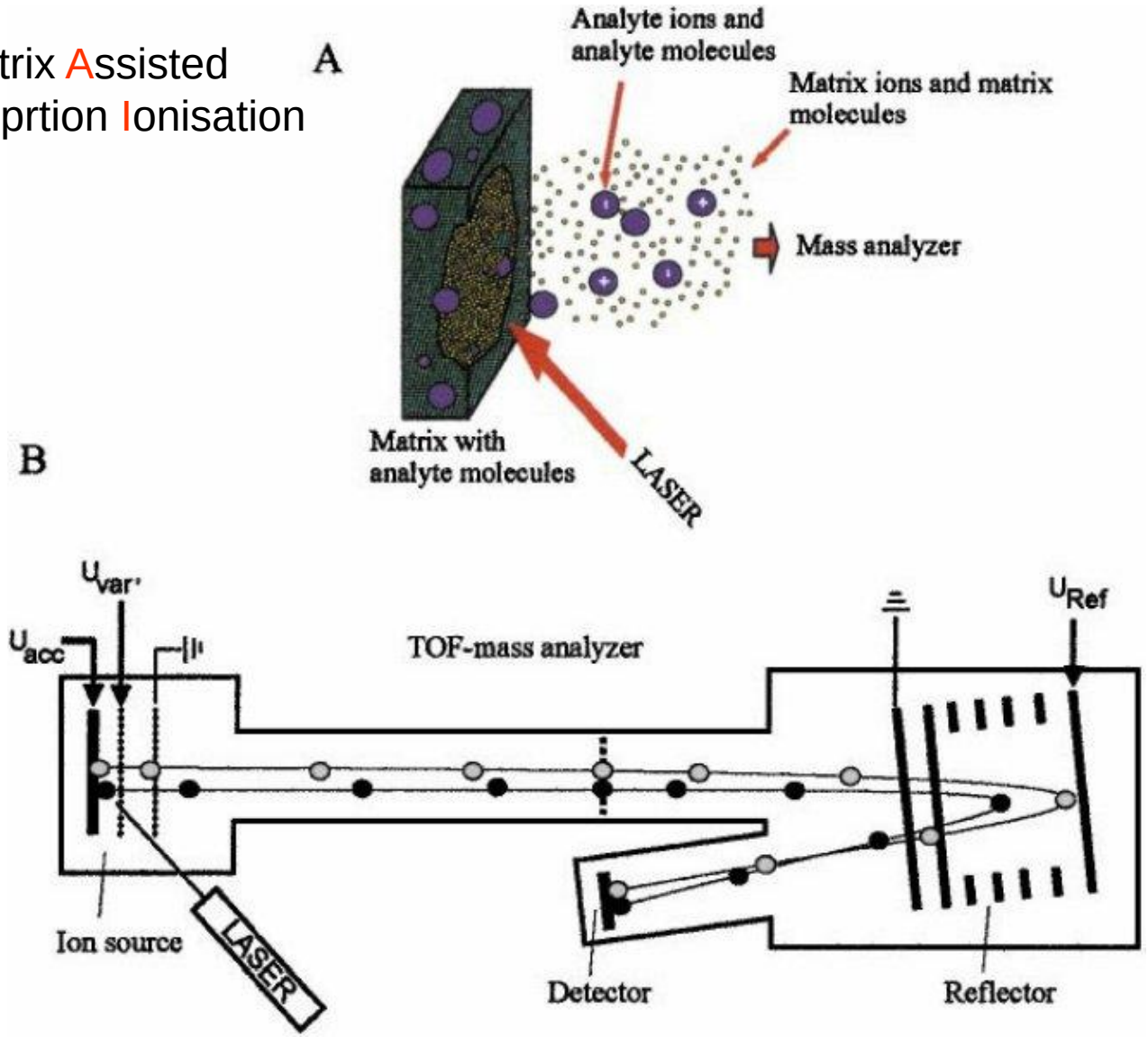
The relationship between the molecular mass (**m**) and the m/z value can be calculated as follows:

$$\mathbf{m/z = (m + (mA * z)) / z}$$

mA is the mass of the adduct responsible for ionization (typically H+ for positive MS mode).

MALDI IONISATION

MALDI (Matrix Assisted Laser Desoprtion Ionisation



MALDI TOF (Time Of Flight)

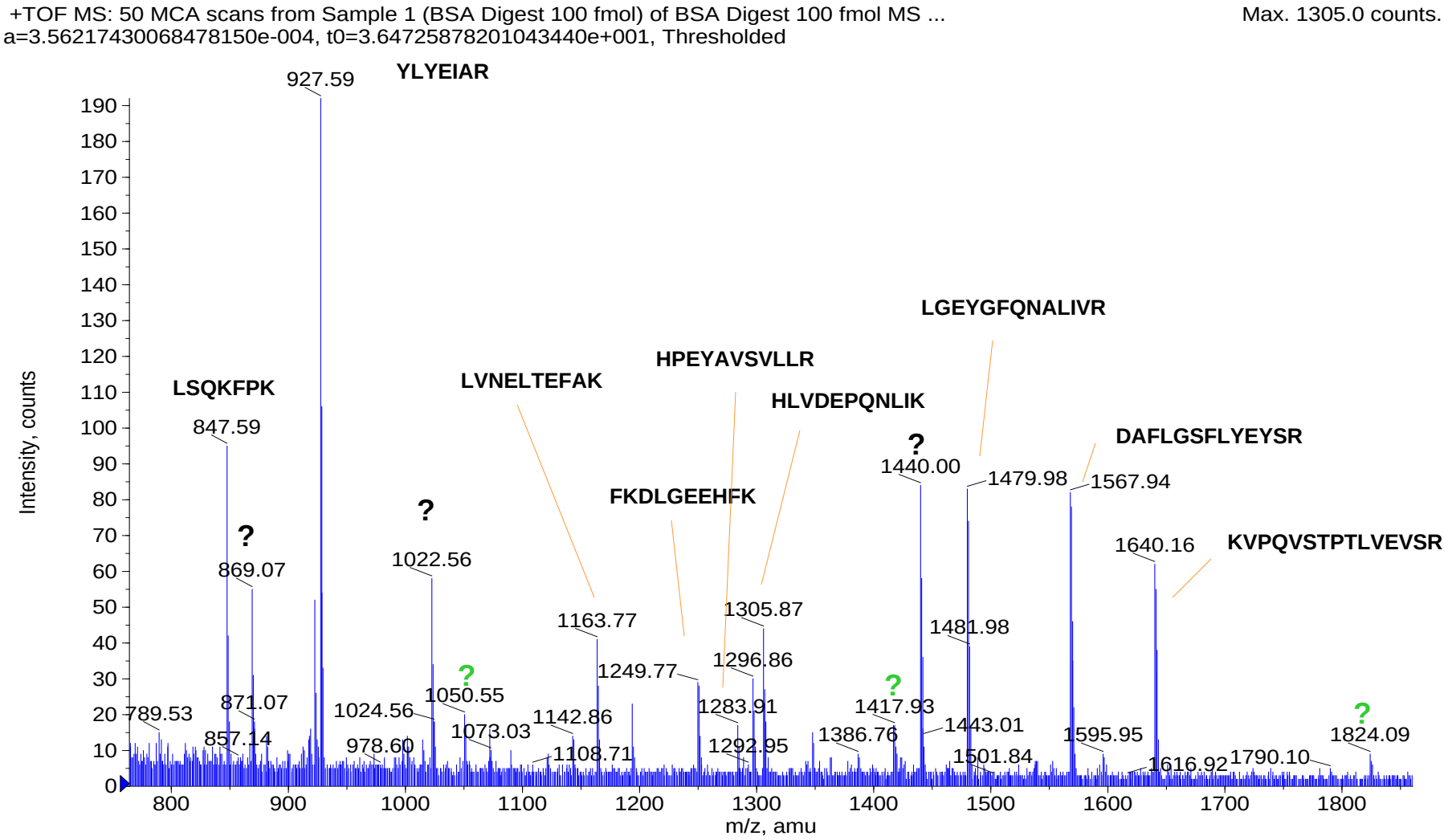
MALDI TOF

- Great for Peptide Mass Fingerprinting
  - Fast
  - Easy to measure
  - Sensitive
  - Salt-tolerant (to some extent)
  - Also good for larger MW (small proteins)
  - Sample on a stable support (no time constraints)
  - 1+ ions → simpler data analysis

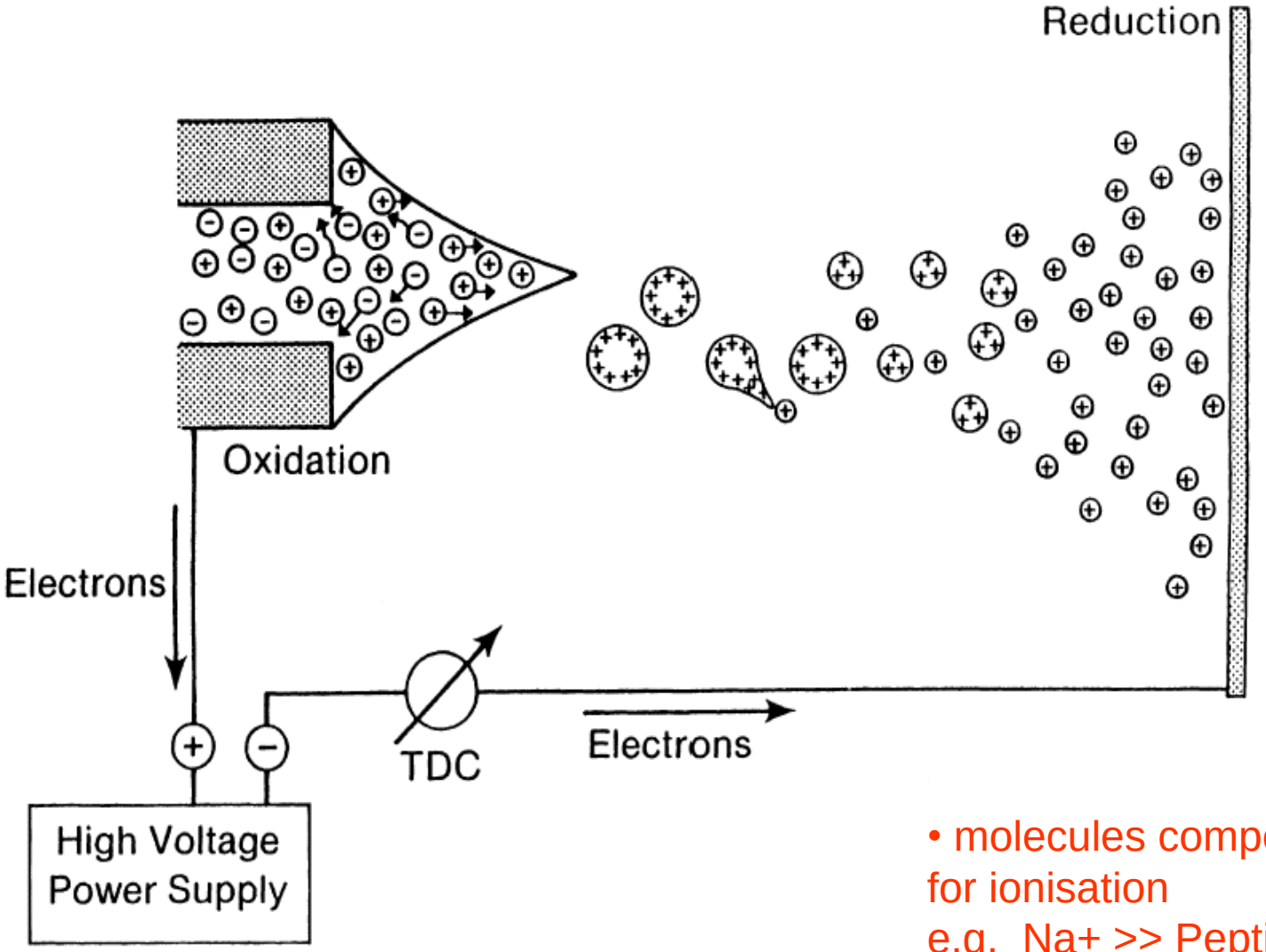
disadvantages ►

- High accuracy needs careful calibration
- Difficult (but possible) to do MS/MS by MALDI
- Signal suppression in complex mixtures
- Crystallisation conditions influence results

MALDI-TOF of a tryptic digest of BSA



ELECTROSPRAY IONISATION



• molecules compete for ionisation
e.g. Na+ >> Peptide

P. Kebarle, M. Peschke / Analytica Chimica Acta 406 (2000) 11–35

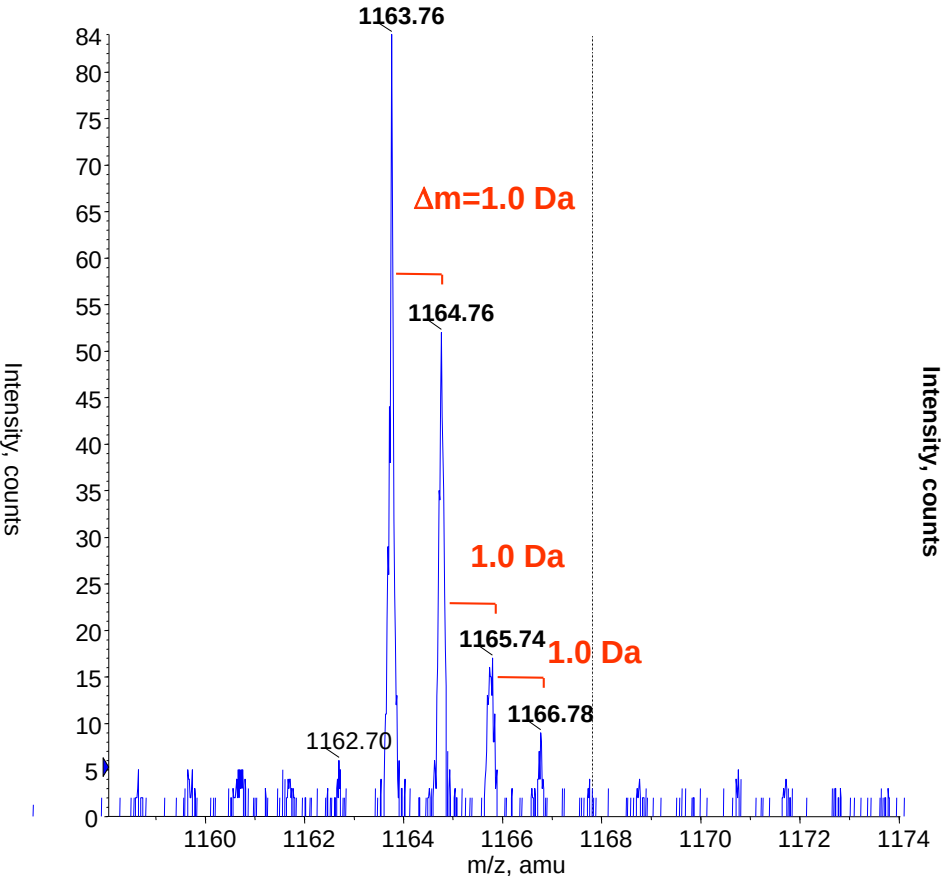
ELECTROSPRAY IONISATION

- Great for MS/MS
  - Can be directly coupled to reversed phase LC (separation !)
  - Sensitive
  - Excellent for MS/MS due to 2+/3+ ions

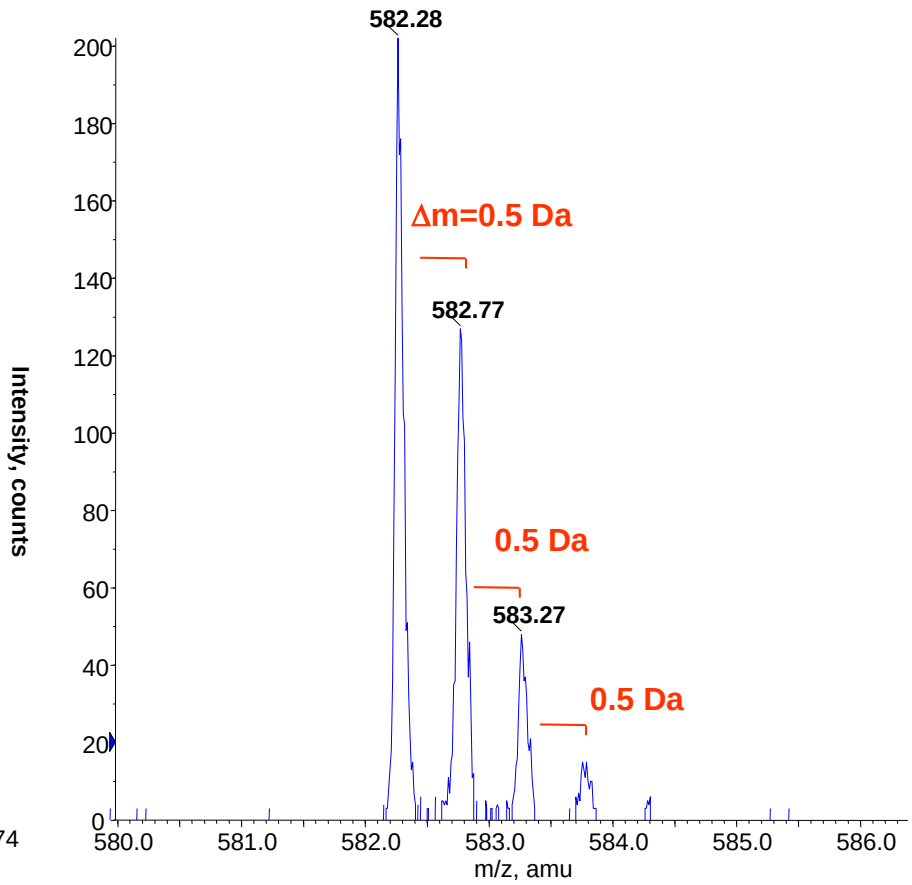
disadvantages ►

- Sample introduction more complex
- Data analysis more difficult (2+/3+ ions)
- one-shot sample analysis (time constraints)
- Very low tolerance to contaminants

1+ versus 2+ ions



1+



2+

Modes of measurement : MS & MS/MS

- Ion production (ionisation)
- Ion separation
- Ion detection

→ MS

- Ion production (ionisation)
- Ion separation – isolation of “parent” ion
- Ion fragmentation (CID)
- Ion separation – separate fragment ions
- Ion detection - measure fragment ions

→ Tandem MS
MS/MS

- CID= collision induced dissociation
- Precursor ion = parent ion : the one being fragmented
- Daughter ions = fragment ions produced by CID
- Tandem mass spectrometry = MS/MS
  - - here : the combination of ion selection / CID / fragment analysis

Tandem MS (MS/MS) facts

- MS/MS results in the acquisition of pseudo-sequence information for multiple (as many as possible) tryptic peptides in addition to intact peptide mass information

- MS/MS needs an instrument able to perform ion isolation and fragmentation, in addition to regular MS

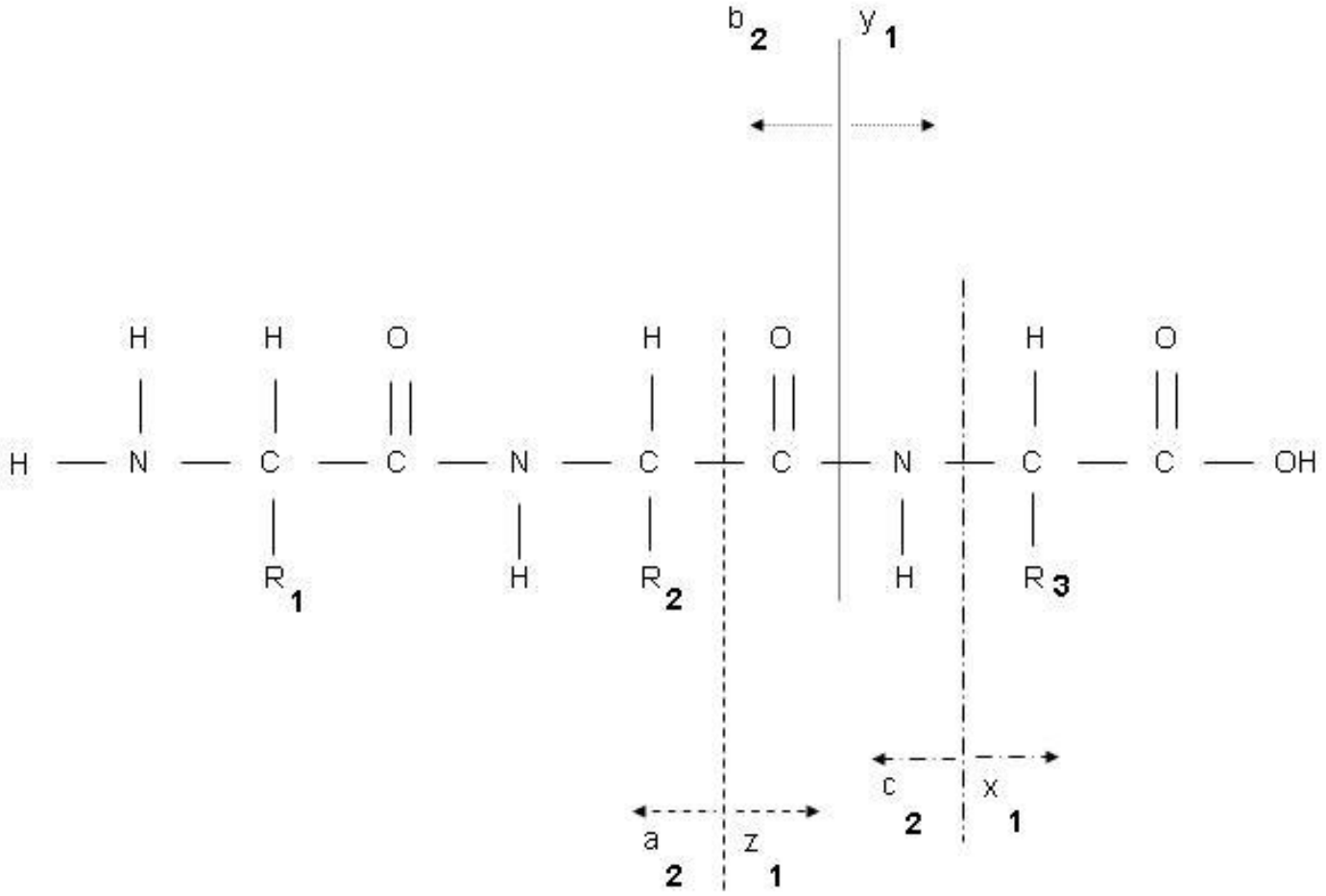
- MS/MS results in the acquisition of multiple orthogonal data files (1 spectrum / peptide)

MS/MS Glossary and facts

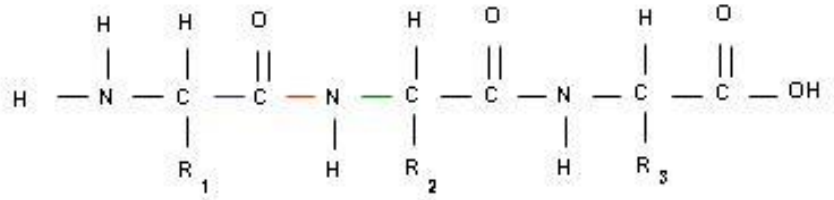
•ESI of tryptic peptides typically generates doubly charged ions due to the presence of Lys or Arg at the C terminal end of the peptides

•y and b-ion series fragments are usually observed in MS/MS fragmentation spectra.

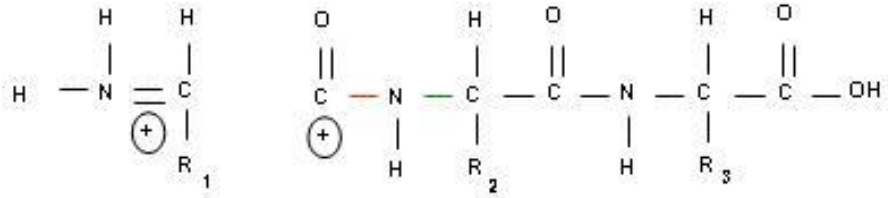
Covalent bonds being broken → ion series



Covalent bonds being broken → ion series

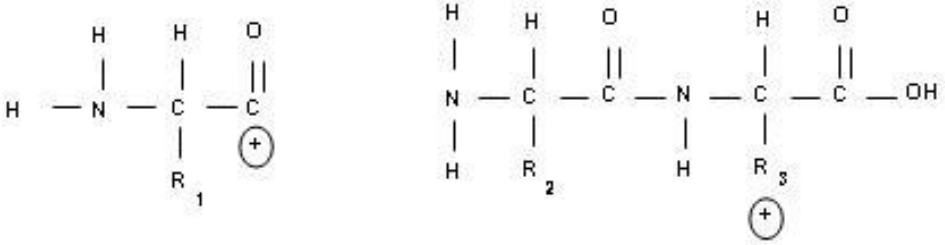


a1



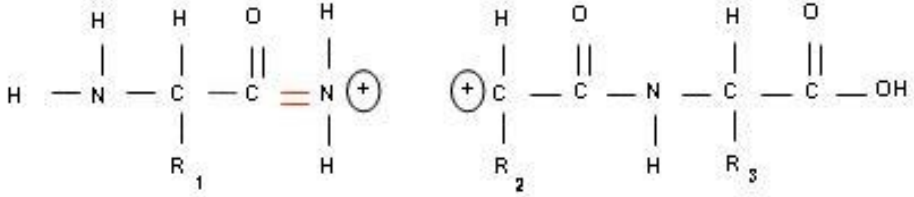
z2

b1



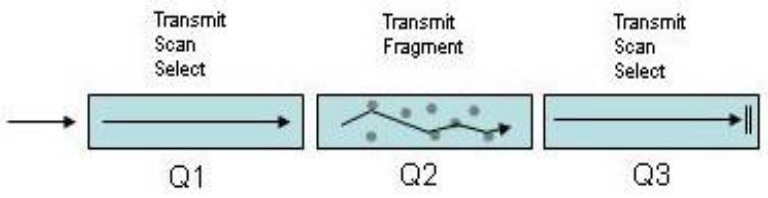
y2

c1

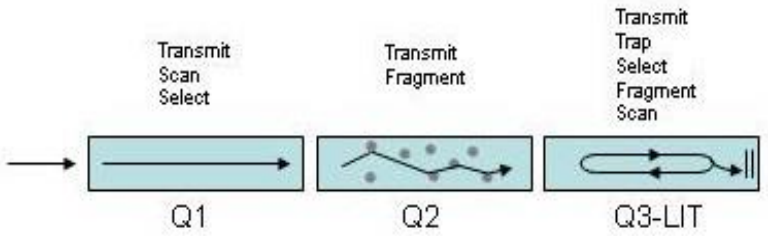


x2

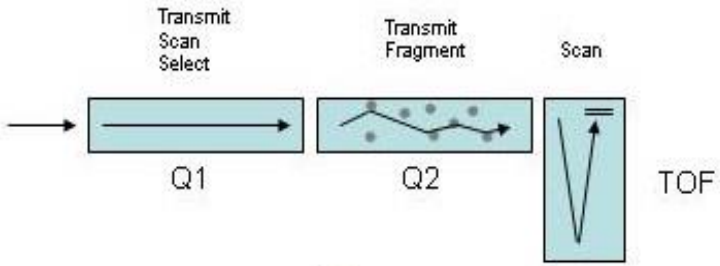
ESI & Quadrupole-based instruments



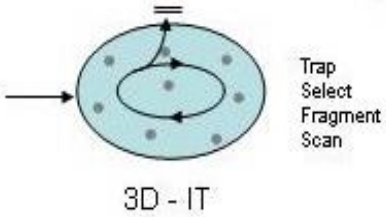
Triple Quadrupole



Triple Quadrupole-
Ion trap



Quadrupole-Quadrupole
TOF

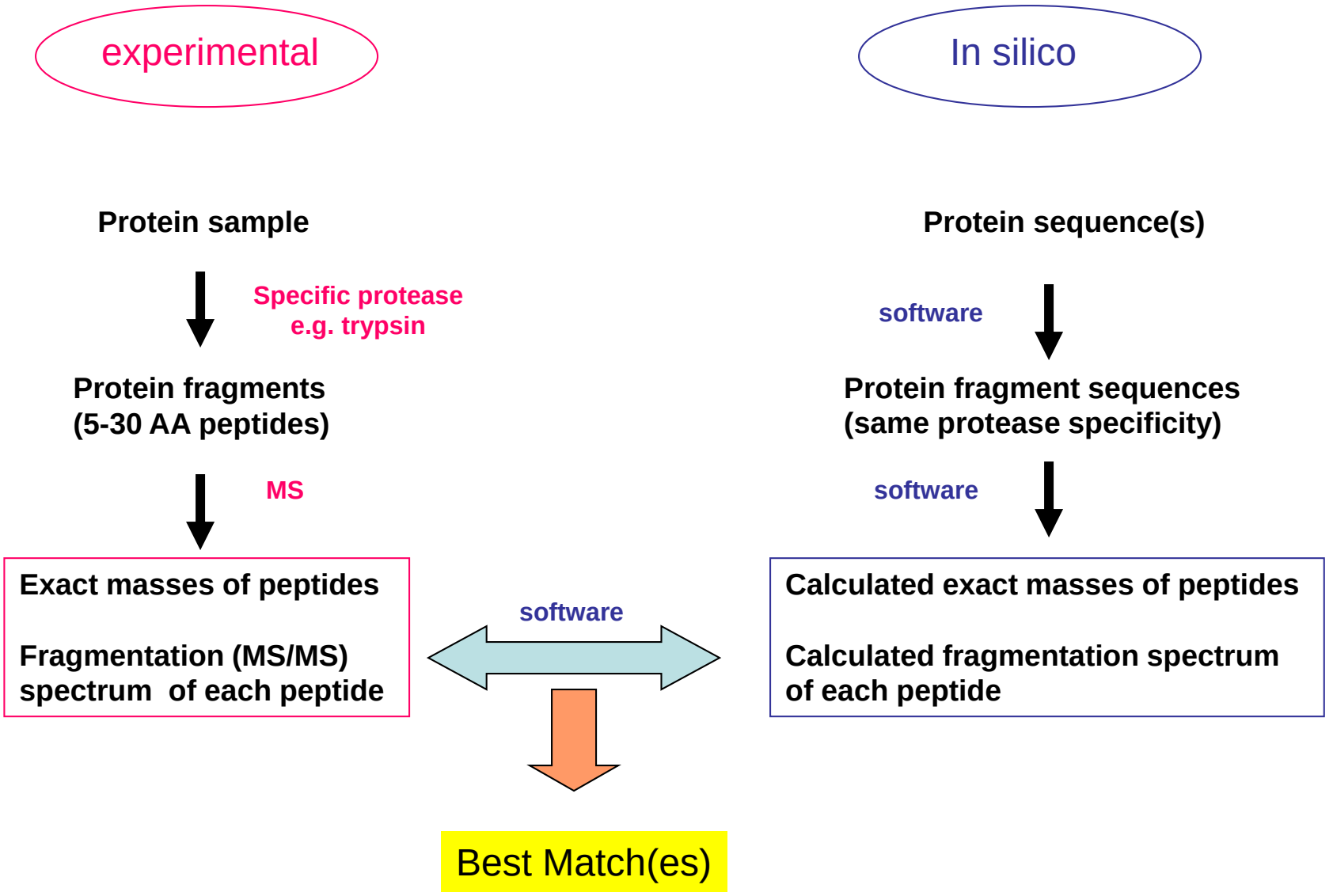


Ion Trap (3D trap)

All these instruments can perform MS/MS fragmentation experiments

- Identifying proteins by mass spectrometry

MS-based protein identification : general concept



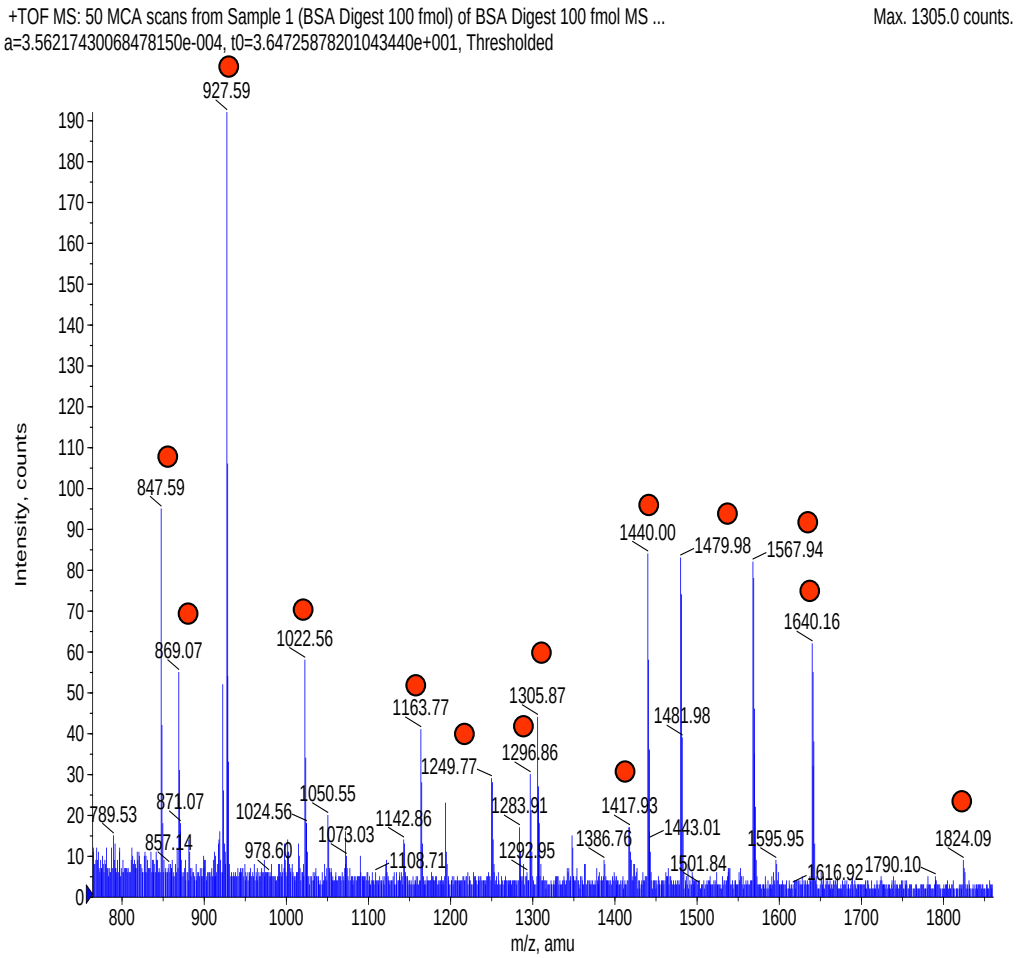
- Protein identification by Peptide Mass Fingerprinting (PMF)

Information contained in MS spectrum

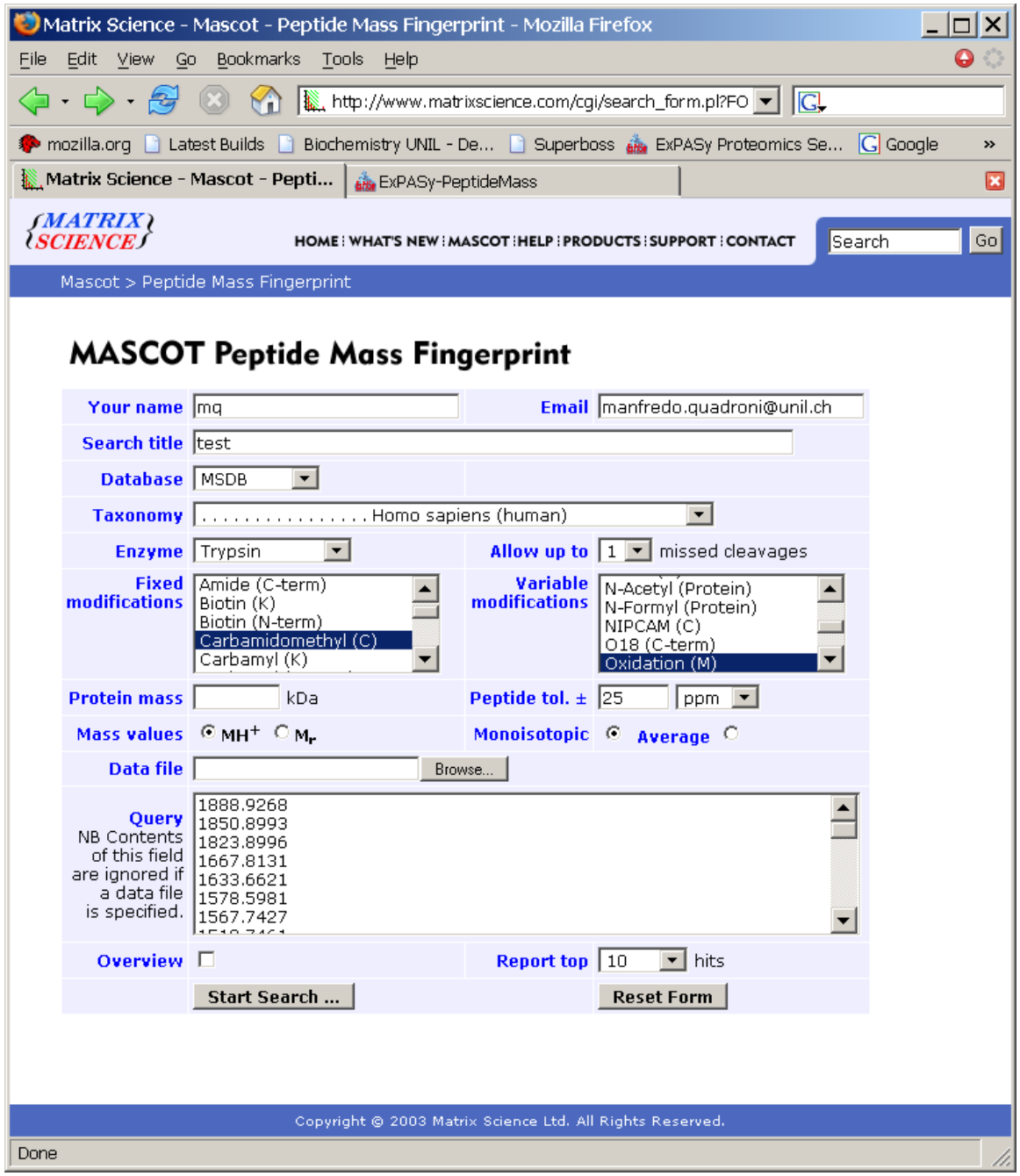
Extracted peak list

m/z

847.5896
869.0722
922.5712
923.5815
927.5904
1022.5551
1050.5533
1163.7695
1164.7531
1193.7393
1249.7705
1250.8103
1296.8556
1297.8499
1305.8668
1416.8929
1440.0008
1479.9773
1482.9583
1567.9417
1640.1635
1824.06
…



Search form…


Matrix Science - Mascot - Peptide Mass Fingerprint - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.matrixscience.com/cgi/search_form.pl?FO
mozilla.org Latest Builds Biochemistry UNIL - De... Superboss ExPASy Proteomics Se... Google
Matrix Science - Mascot - Pepti...
ExPASy-PeptideMass
MATRIX SCIENCE
HOME | WHAT'S NEW | MASCOT | HELP | PRODUCTS | SUPPORT | CONTACT
Search
Go
Mascot > Peptide Mass Fingerprint
MASCOT Peptide Mass Fingerprint
Your name
mq
Email
manfredo.quadroni@unil.ch
Search title
test
Database
MSDB
Taxonomy
. . . . . Homo sapiens (human)
Enzyme
Trypsin
Allow up to
1
missed cleavages
Fixed modifications
Amide (C-term)
Biotin (K)
Biotin (N-term)
Carbamidomethyl (C)
Carbamyl (K)
Variable modifications
N-Acetyl (Protein)
N-Formyl (Protein)
NIPCAM (C)
O18 (C-term)
Oxidation (M)
Protein mass
kDa
Peptide tol. ±
25
ppm
Mass values
MH+
Mr
Monoisotopic
Average
Data file
Browse...
Query
NB Contents of this field are ignored if a data file is specified.
1888.9268
1850.8993
1823.8996
1667.8131
1633.6621
1578.5981
1567.7427
Overview
Report top
10
hits
Start Search ...
Reset Form
Copyright © 2003 Matrix Science Ltd. All Rights Reserved.
Done

Results page…

Concise Summary Report (bsa pmf bad) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://pcib306/mascot/cgi/master_results.pl?file=../data/20050314/F012247.dat

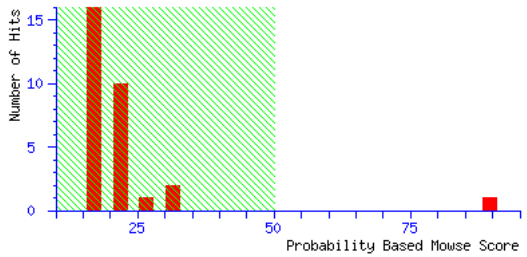
Biochemistry UNIL - De... Superboss ExPASy Proteomics Se... Google NEW MASCOT SearchLog-new Google Scholar

Mascot Search Results

```
User           : mg
Email          :
Search title   : bsa pmf bad
Database       : Sprot 20050217 (172233 sequences; 62615309 residues)
Taxonomy       : Other mammalia (6060 sequences)
Timestamp      : 14 Mar 2005 at 09:31:21 GMT
Top Score      : 89 for ALBU_BOVIN, (P02769) Serum albumin precursor (A
```

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event
Protein scores greater than 50 are significant (p<0.05).



http://pcib306/mascot/cgi/protein_view.pl?file=../data/20050314/F012247.dat&hit=1

Mascot Search Results: Protein View - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://pcib306/mascot/cgi/protein_view.p

Biochemistry UNIL - De... Superboss ExPASy Proteomics Se... Google NEW MASCOT

Protein View

Match to: **ALBU_BOVIN** Score: **89** Expect: **7e-06**
(P02769) Serum albumin precursor (Allergen Bos d 6) (BSA)

Nominal mass (M_r): **71244**; Calculated pI value: **5.82**
NCBI BLAST search of ALBU_BOVIN against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Bos taurus

Fixed modifications: Carbamidomethyl (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **22**
Number of mass values matched: **11**
Sequence Coverage: **19%**

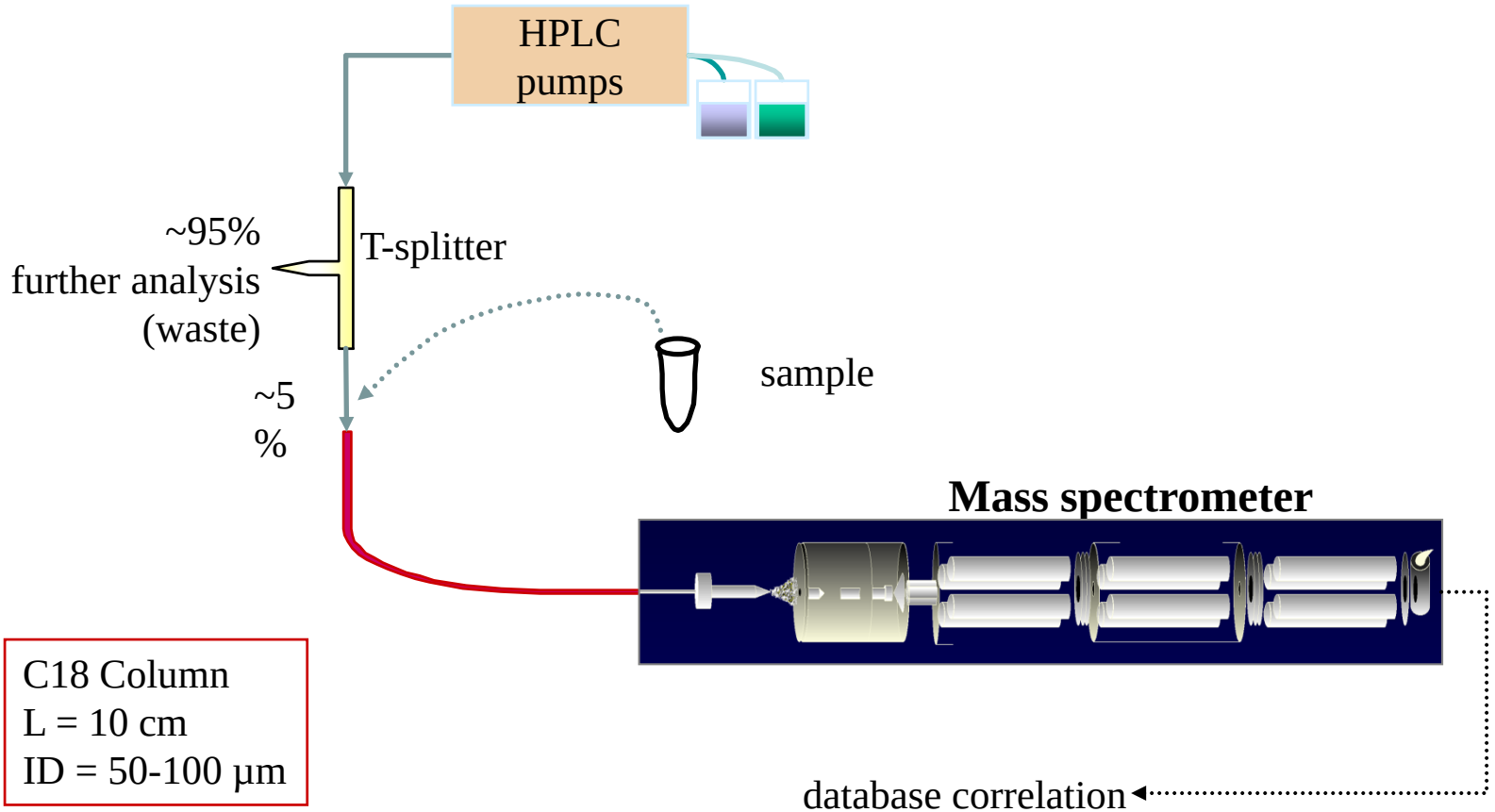
Matched peptides shown in **Bold Red**

```
    1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
   51 FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
  101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
  151 KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQE CCQAEDKGAC
  201 LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
  251 FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
  301 CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
  351 GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
  401 KHLVDEPQNL IKQNCDQFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
  451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
  501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
  551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
  601 STQTALA
```

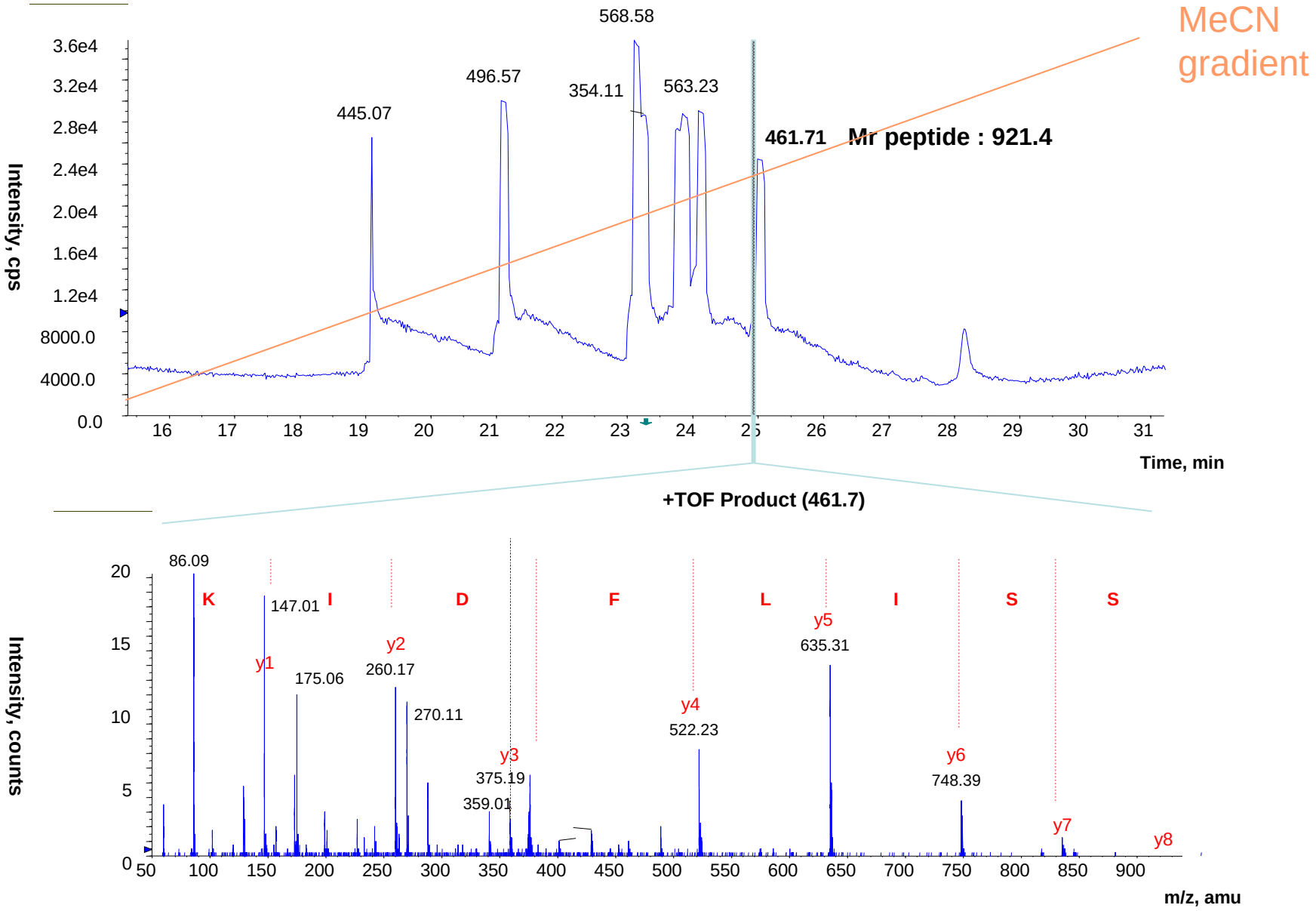
Done

- Protein identification by MS/MS

Experimental set-up : nanoLC-MS/MS

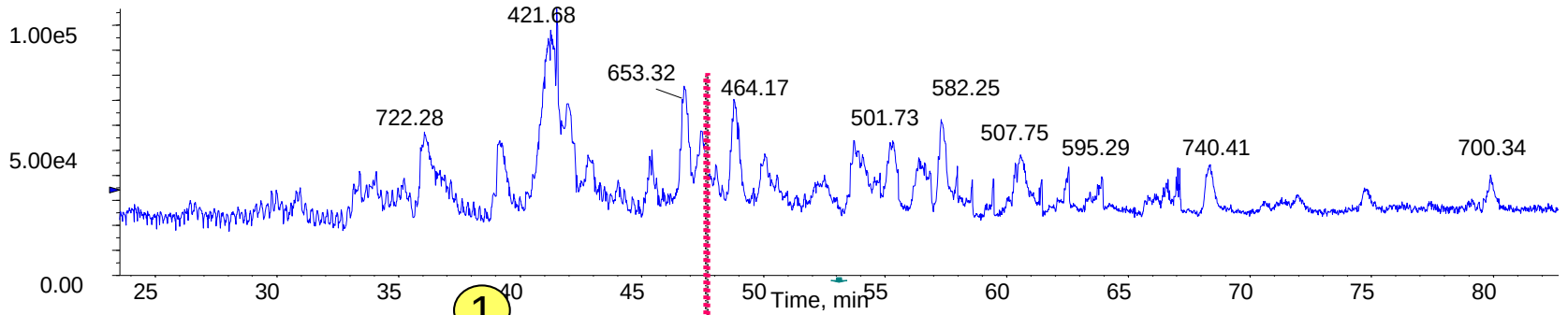


An on-line LC-ESI-MS experiment with automatic data acquisition

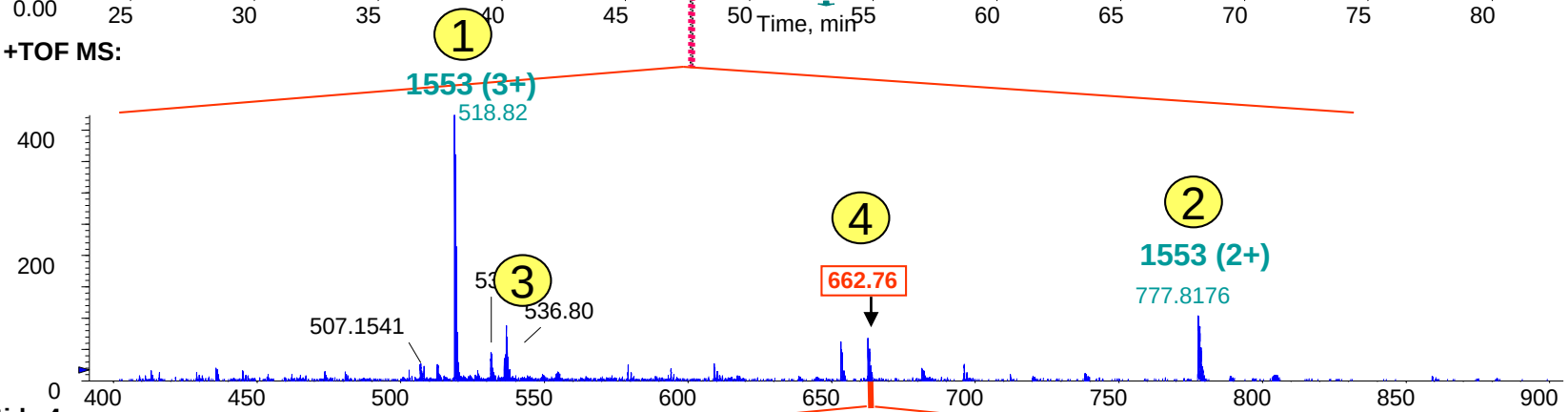


Automated LC- MS/MS run

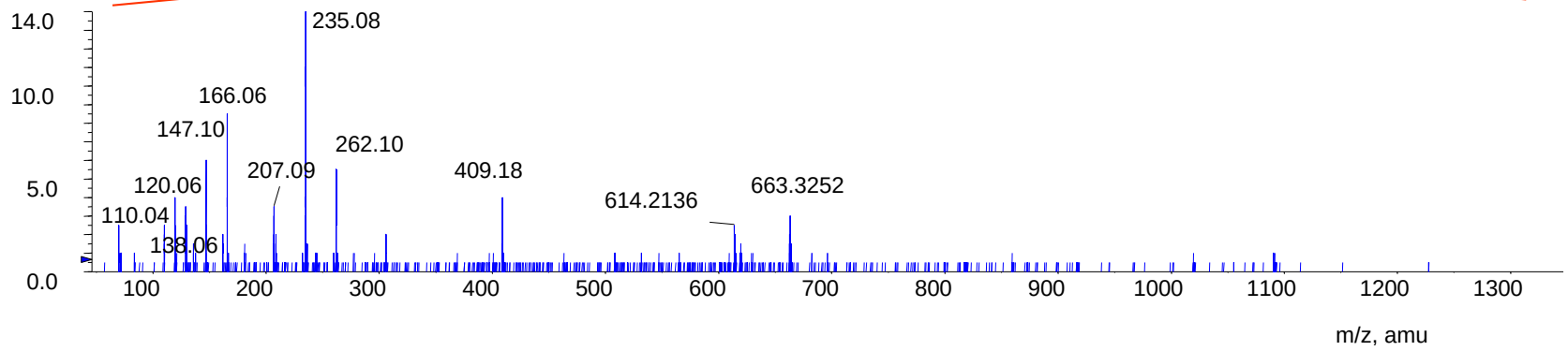
Chromatogram :
Total ion current vs. time



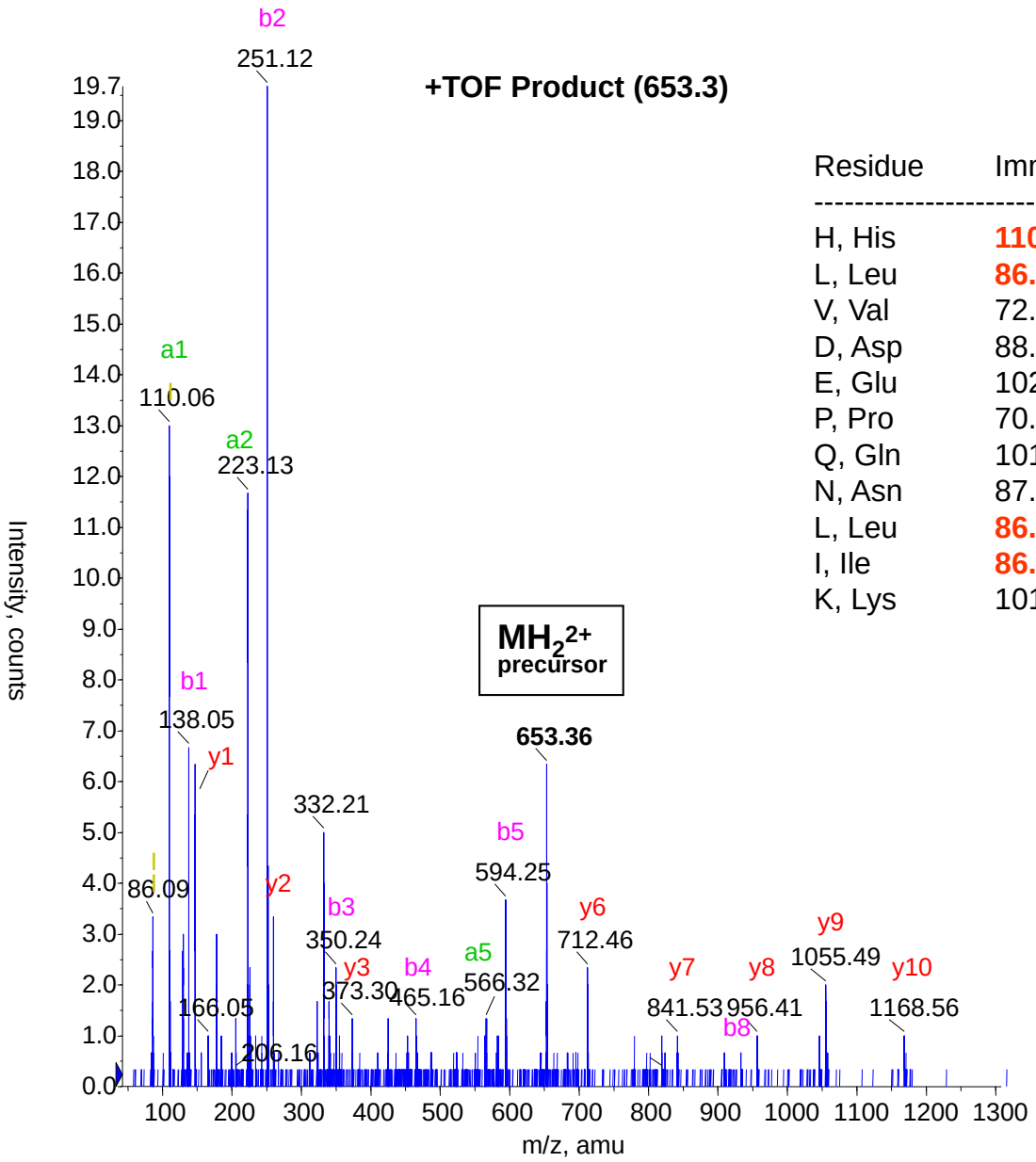
Full scan : +TOF MS:



MS/MS peptide 4 :
+TOF Product (662.8):



Matching of MS / MS data



Residue	Immonium	a	b	y
H, His	**110.07**	**110.07**	**138.06**	1305.71
L, Leu	**86.09**	**223.15**	**251.15**	**1168.65**
V, Val	72.08	322.22	**350.21**	**1055.57**
D, Asp	88.03	437.25	**465.24**	**956.50**
E, Glu	102.05	**566.29**	**594.28**	**841.47**
P, Pro	70.06	663.34	691.34	**712.43**
Q, Gln	101.07	791.40	819.39	615.38
N, Asn	87.05	905.44	**933.44**	**487.32**
L, Leu	**86.09**	1018.53	1046.52	373.28
I, Ile	**86.09**	1131.6157	1159.61	**260.19**
K, Lys	101.10	1259.7106	1287.70	**147.11**

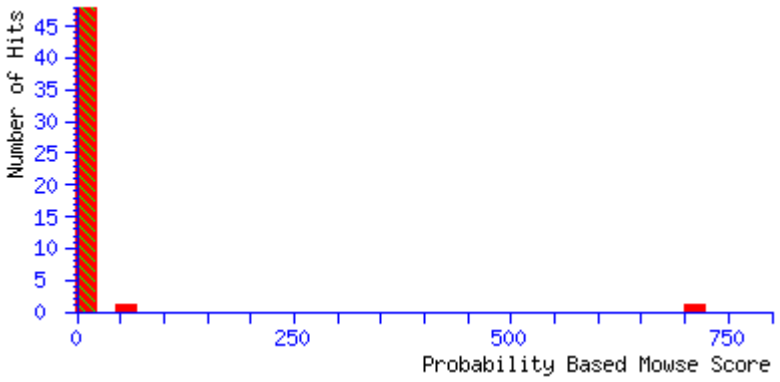
Black : predicted
Red : predicted and detected

Mascot search output

Mascot Search Results

Significant hits:

ALBU_BOVIN (P02769) Serum albumin precursor (Allergen Bos d 6).
ALBU_CANFA (P49822) Serum albumin precursor (Allergen Can f 3).
VWF_PIG (Q28833) Von Willebrand factor precursor (vWF) (Fragment).
CIQ3_BOVIN (P58126) Voltage-gated potassium channel protein KQT-like 3
RYR2_RABIT (P30957) Ryanodine receptor 2 (Cardiac muscle-type ryanodin
K2CA_BOVIN (P04263) Keratin, type II cytoskeletal 68 kDa, component IA
ALFB_RABIT (P79226) Fructose-bisphosphate aldolase B (EC 4.1.2.13) (Li



ALBU_BOVIN Mass: 71244 Total score: 711 Peptides matched: 12
(P02769) Serum albumin precursor (Allergen Bos d 6).

	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
15	461.80	921.58	921.48	0.10	0	55	1	AEFVEVTK
19	501.80	1001.58	1001.58	0.01	0	31	1	LVVSTQTALA
31	569.80	1137.58	1137.49	0.09	0	72	1	CCTESLVNR
33	582.30	1162.58	1162.62	-0.04	0	76	1	LVNELTEFAK
47	653.40	1304.78	1304.71	0.08	0	90	1	HLVDEPQNLIK
58	722.40	1442.78	1442.63	0.15	0	87	1	YICDNQDTISSK
60	739.80	1477.58	1477.52	0.07	0	43	1	ETYGDMADCCEK
61	740.40	1478.78	1478.79	-0.00	0	61	1	LGEYGFQNALIVR
64	751.90	1501.78	1501.61	0.18	0	37	1	EYEATLEECCAK
74	547.30	1638.88	1638.93	-0.05	1	85	1	KVPQVSTPTLVEVSR
97	627.70	1880.08	1879.91	0.16	0	34	1	RPCFSALTPDETYVPK
98	636.70	1907.08	1906.91	0.16	0	42	1	LFTFHADICTLPDTEK

Orthogonal datasets and confidence levels

PMF :

one MS spectrum
→ one dataset (peak list)

MS/MS :

n MS/MS spectra
→ n orthogonal datasets

Database : 100'000 sequences 500 spectra

→Probability of one (any) spectrum "accidentally" matching a sequence (wrong match) : 1/100'000 x 500 = 5.10^{-3} (0.005)

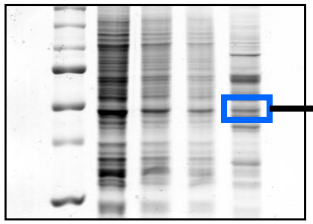
Probability of 2 spectra "accidentally" matching the same sequence (wrong match) : 5.10^{-3} x 5.10^{-3} = 2.5×10^{-5}

Much higher confidence of identification with at least two peptides matching the same protein sequence

Every peptide is unambiguously assigned to its "parent " sequence, therefore many proteins can be identified in one sample during one run

Summary : Typical Analytical Workflow

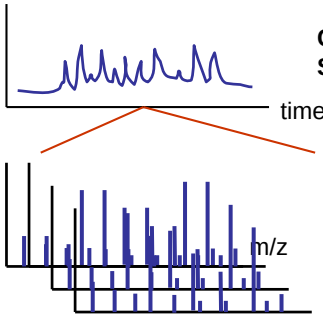
Biological question



Protease digestion Peptide extraction

Nano-HPLC

MS/MS



Chromatographic Separation (reversed-phase)

Tandem mass spectra of 50-2000 peptides

Database searching Software (MASCOT)

Protein sequence database

Output :

- **Protein identification in simple/complex mixtures**
- **Extensive sequence coverage and peptide mapping**
- **Analysis of modified peptides possible**

Database matches

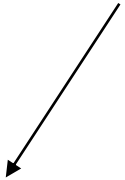
DHX9_HUMAN	ATP-dependent RNA helicase A
NFM_HUMAN	Neurofilament triplet M protein
Q9BQG0	Hypothetical protein
MYO6_HUMAN	Myosin VI.
TP2A_PIG	DNA topoisomerase II, alpha isozyme
Q7Z5Y2	Rho-interacting protein 3.
FLIH_HUMAN	Flightless-I protein homolog.
TP2B_MOUSE	DNA topoisomerase II, beta isozyme
S3B1_HUMAN	Splicing factor 3B subunit
Q8VCW5	Similar to alpha internexin neuronal
Q8CHF9	MKIAA0376 protein (Fragment).

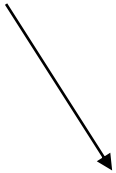
Q7Z5Y2 **Mass:** 118789 **Total score:** 178
Peptides matched: 6
Rho-interacting protein 3.

Mr(calc)	Score	Peptide
930.48	42	EGLTVQER
1032.54	11	NWIQTIMK
1206.63	29	FSLCILTPEK
1369.75	24	LSTHELTSLLEK
1406.77	55	FFILYEHGLLR
1775.88	16	QVPIAPVHLSSEDGGDR

Caveat : Protein identification IS NOT protein characterisation

Two peptides are enough to identify a protein
But
We are still identifying two peptides, not the entire protein



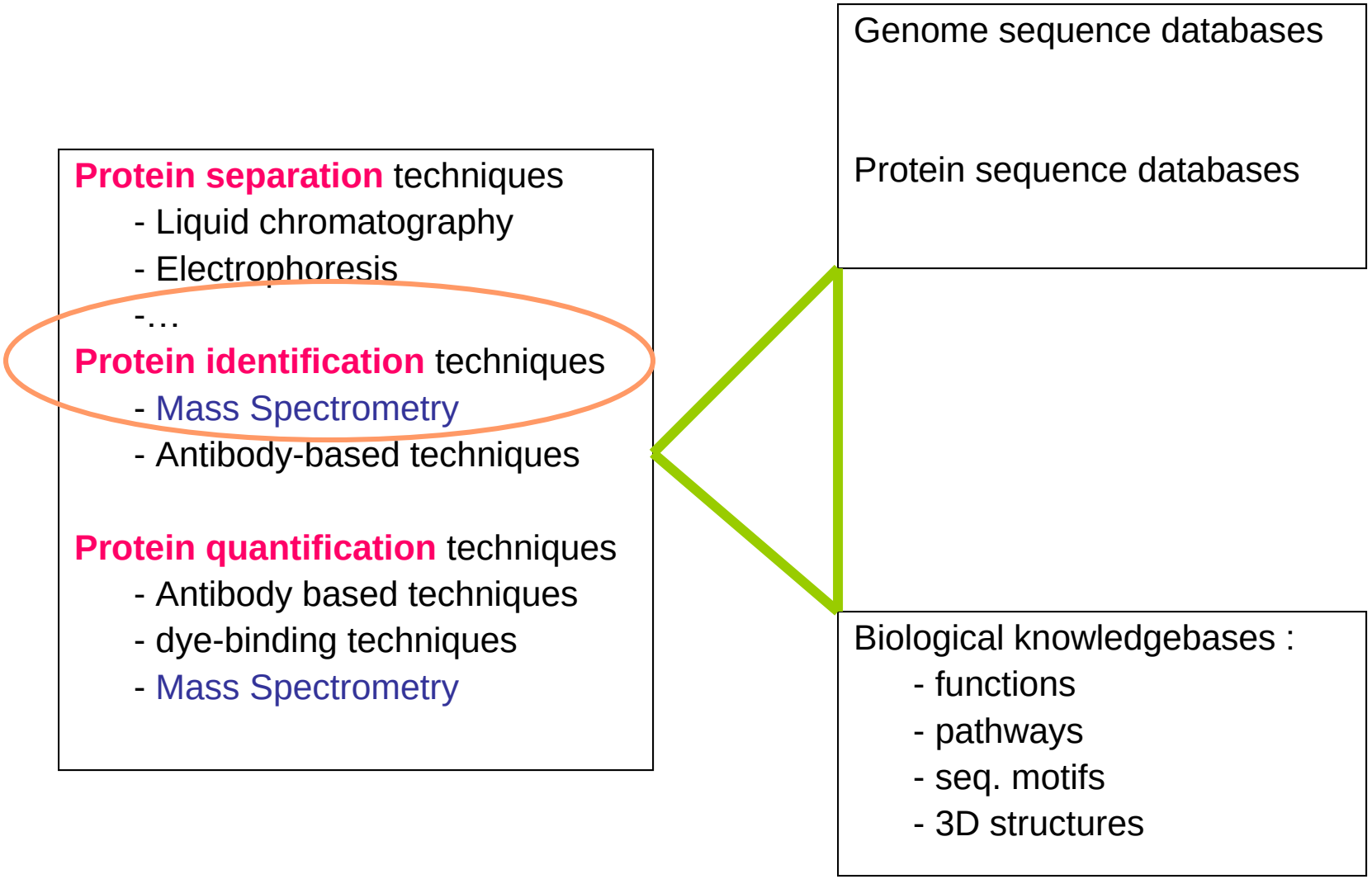


Highly similar sequences cannot be distinguished

For finding PTMs extensive Sequence coverage is essential !!

2
Technology, workflows and applications : what is available

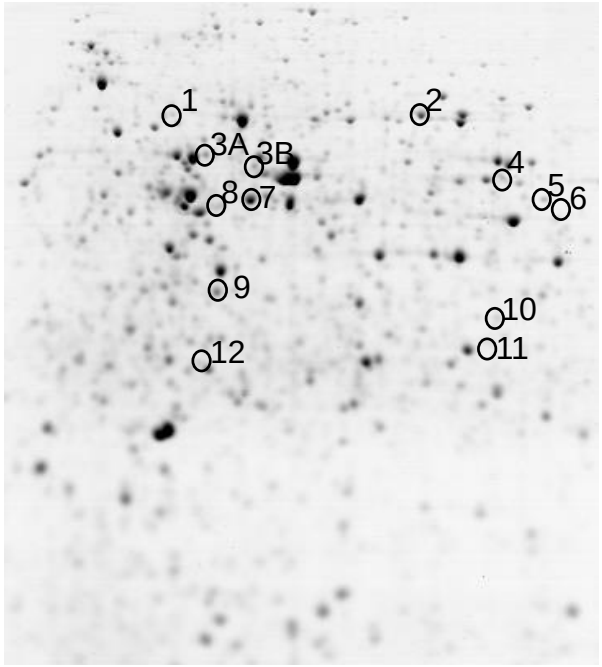
New and old tools



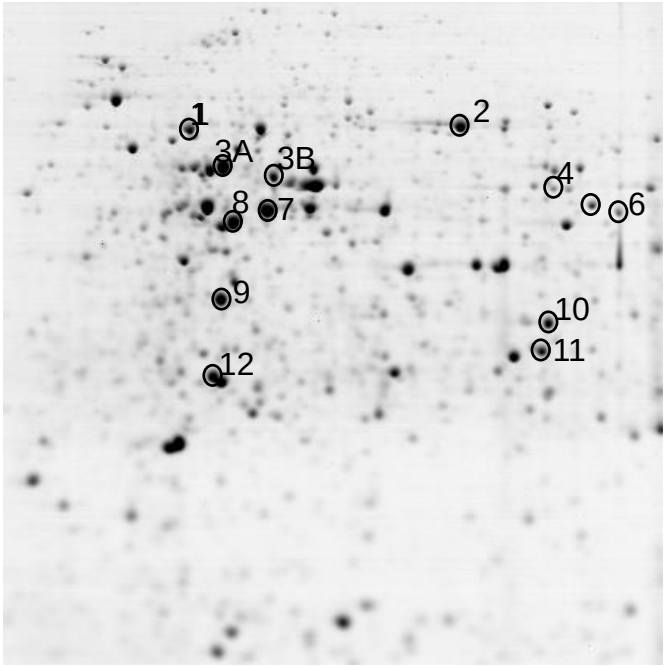
WORKFLOWS 1 :
„classical“
2D-PAGE
+
MALDI TOF

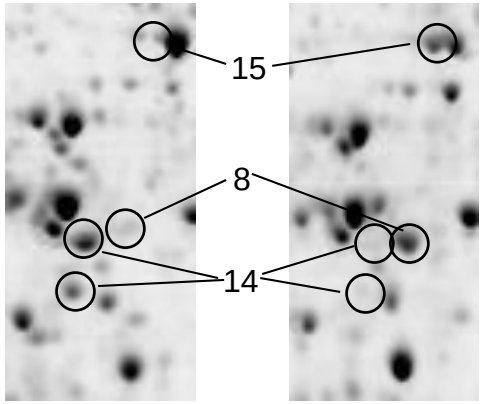
Workflow 1 : adaptation of bacteria to growth conditions

Normal medium



Low Glucose

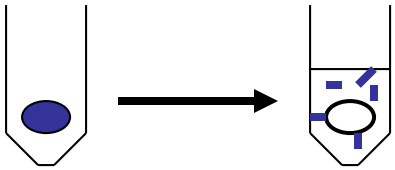




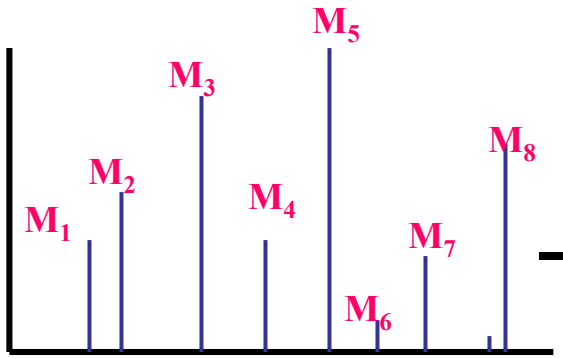
E.Coli adapts to a very low glucose medium by up- and downregulating a set of 15 proteins

Wick LM, et al, Environ Microbiol 3: 588-599, 2001

Workflow 1 : adaptation of bacteria to growth conditions



Tryptic digestion



Peptide Mass Fingerprinting
Search with mass list

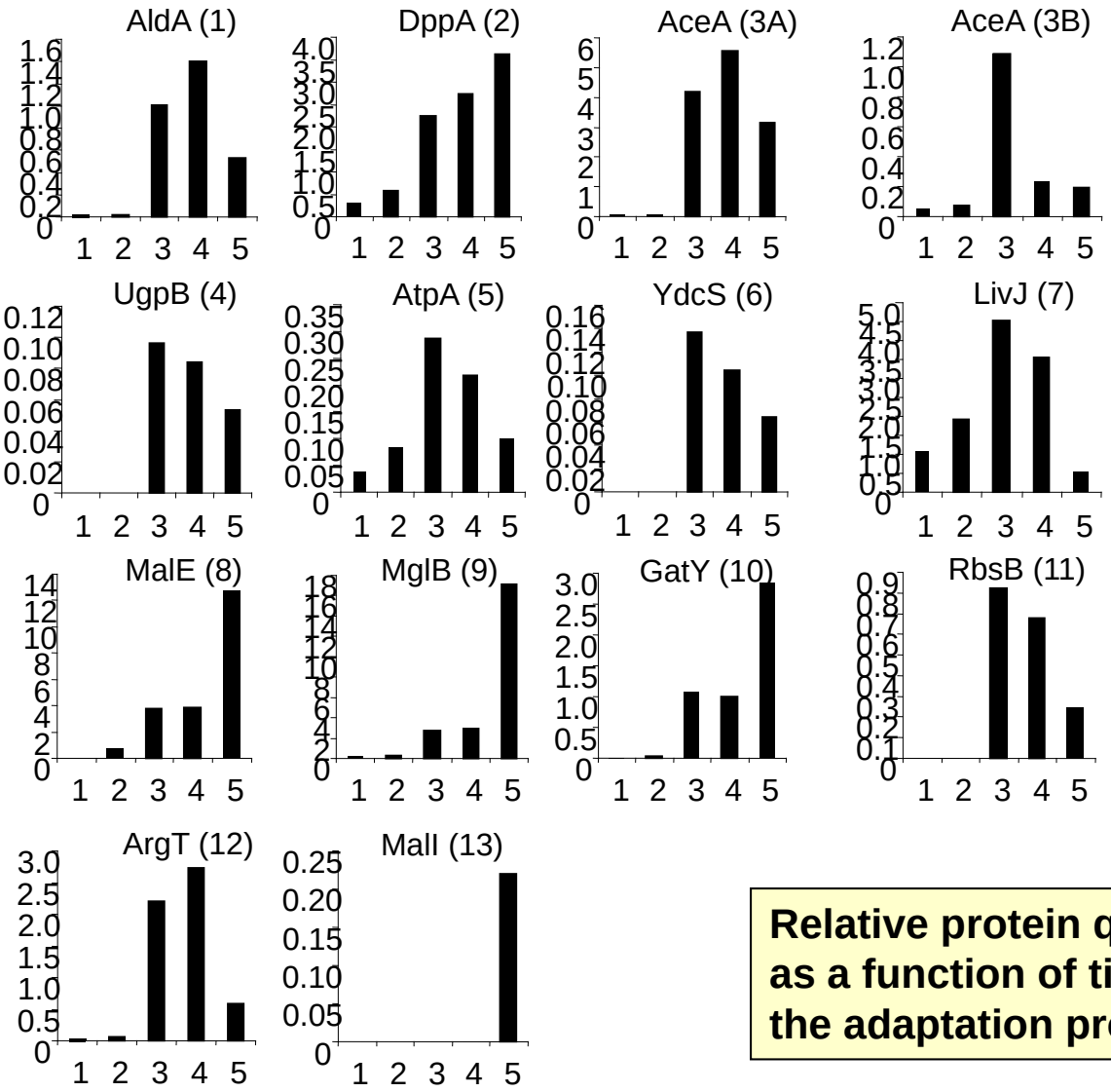
M_1.....M_8

Nr	mw*	pI*	Acc. N.	Name	Function
1	52.2	5.07	P25553	aldA	central metabolism
2	60.3	6.21	P23847	dppA	peptide transport
3	47.5	5.06	P05313	aceA	Central metabolism
4	48.45	6.71	P10904	ugpB	transport
5	55.2	6.02	P00822	atpA	proton transport/energy
6	?	?	P76108	ydcS	transport
7	41.3	7.03	P02917	livJ	amino acid transport
8	40.7	5.22	P02928	malE	sugar transport
9	33.36	5..25	P02927	mglB	sugar transport
10	61.5	5.81	P37192	gatY	catabolism
11	30.9	7.76	P02925	rbsB	sugar transport
12	25.8	5.22	P09551	argT	amino acid transport

Metabolic enzymes and transport proteins affected

Utilisation of alternative sources of energy

WORKFLOW 1 : quantitation and kinetics



Time points :

1) Start – batch culture
2) Adapted bacteria put back into high-glucose batch culture
3) 40 hr adaptation culture
4) 156 hr adaptation culture
5) 500 hr adaptation culture

Relative protein quantities as a function of time during the adaptation process

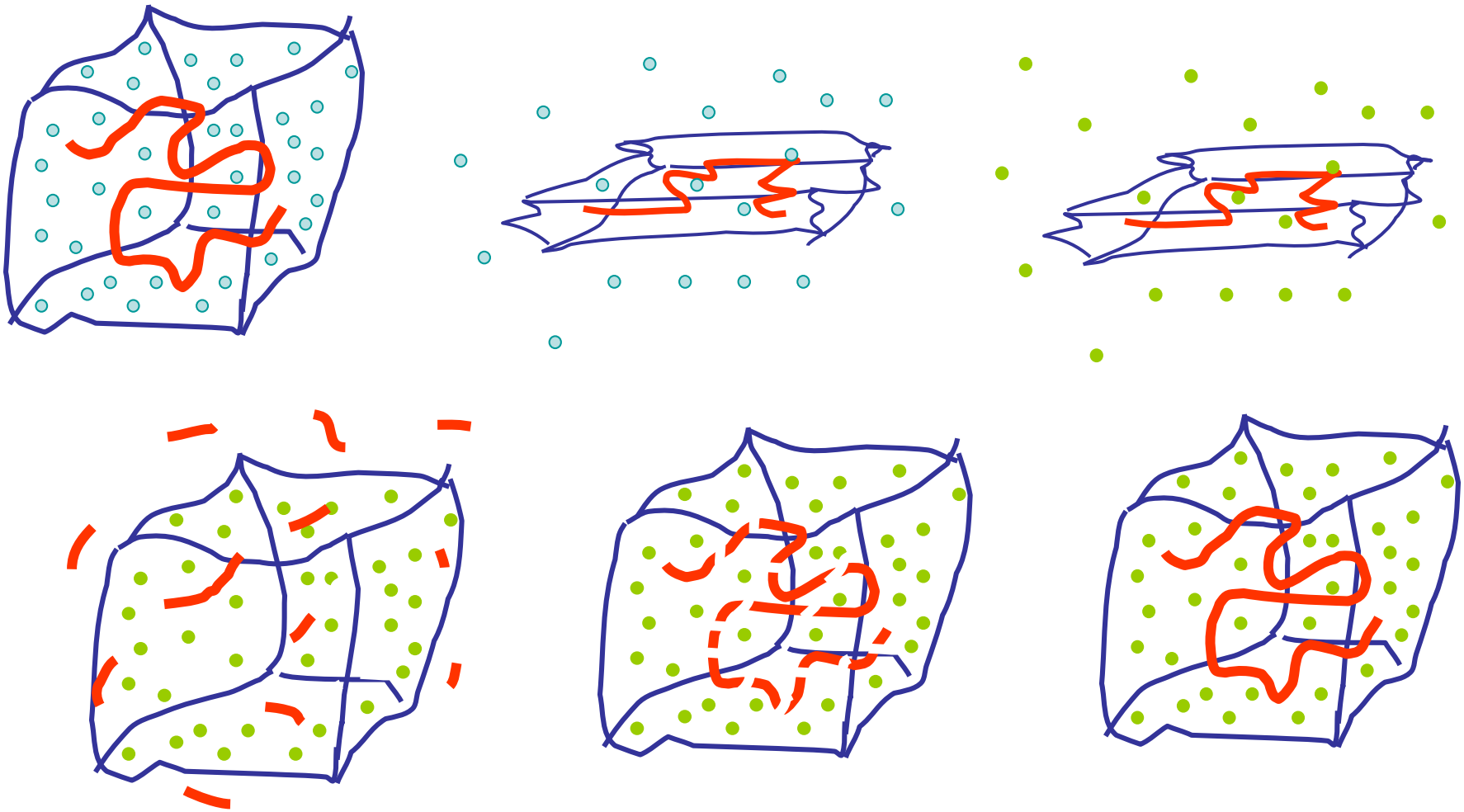
Why is SDS-PAGE such a good preparation method?

- Ideal interface to biology
- Analytical and micropreparative
- Robust
- Solid phase chemistry of proteins
- Easy, low-tech
- Removal of contaminants :
  - At the loading point
  - After migration during fix / staining steps

Disadvantages ►

- protein digestion in gel: non quantitative
- peptide sequence recovery: usually incomplete
- whole protein recovery: poor

In-gel digestion: solid phase chemistry of proteins

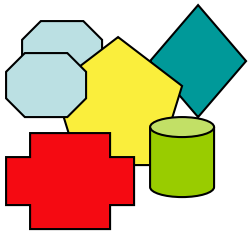


WORKFLOWS 2 :

General shotgun protein identification techniques

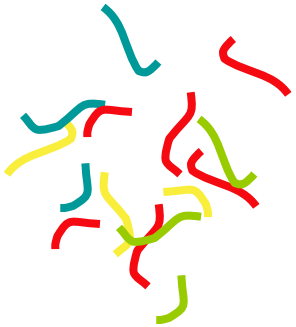
example :
Affinity pull-down
+
1D-PAGE
+
LC-MS/MS

Shotgun sequencing from complex mixtures

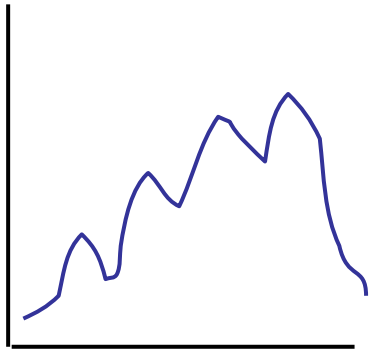


Multiprotein complex

Denaturation,
Proteolytic
digestion



Complex peptide mixture
(1000-20000 species)



Rp-LC-MSMS run

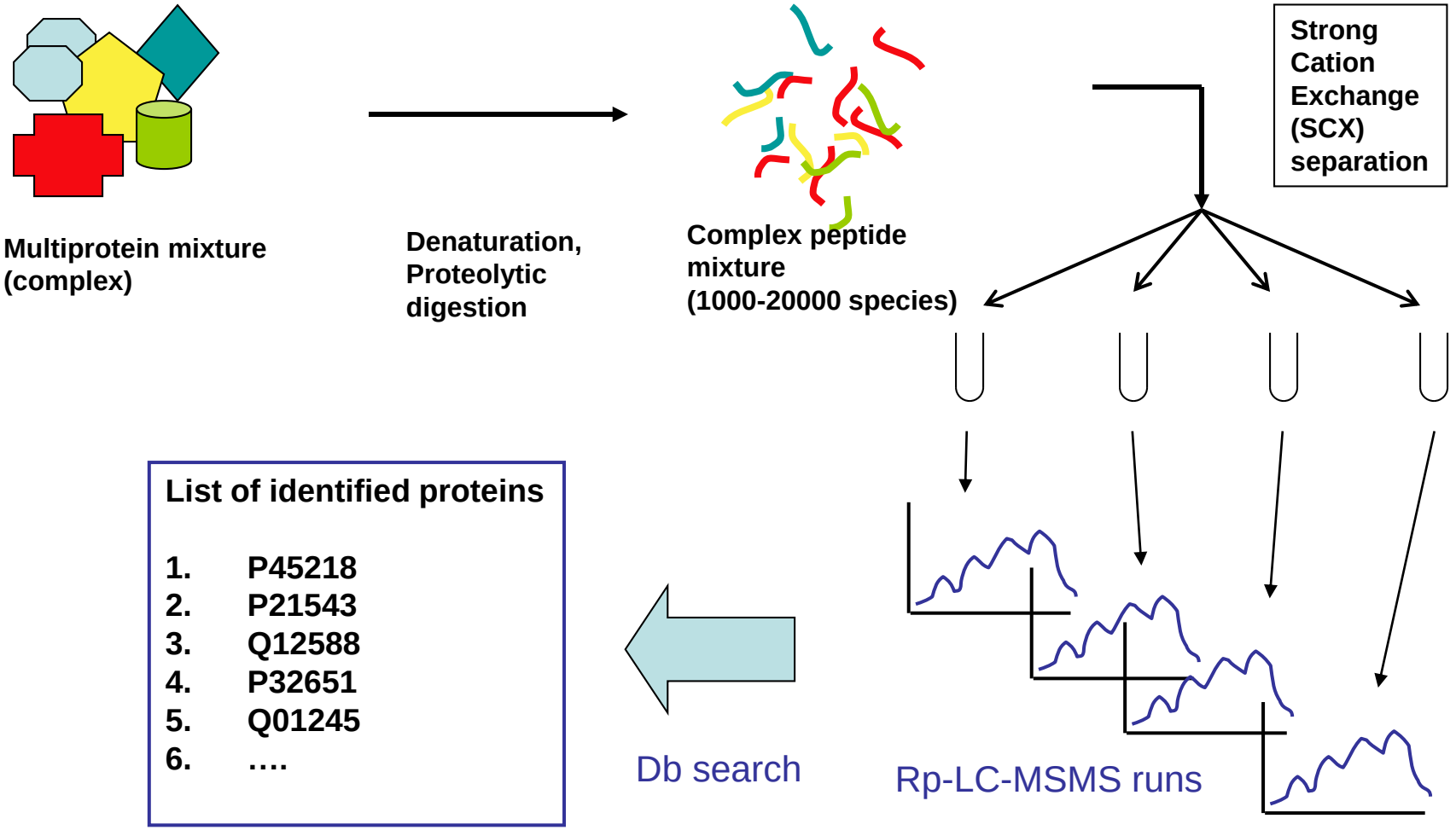
Db search

List of identified proteins

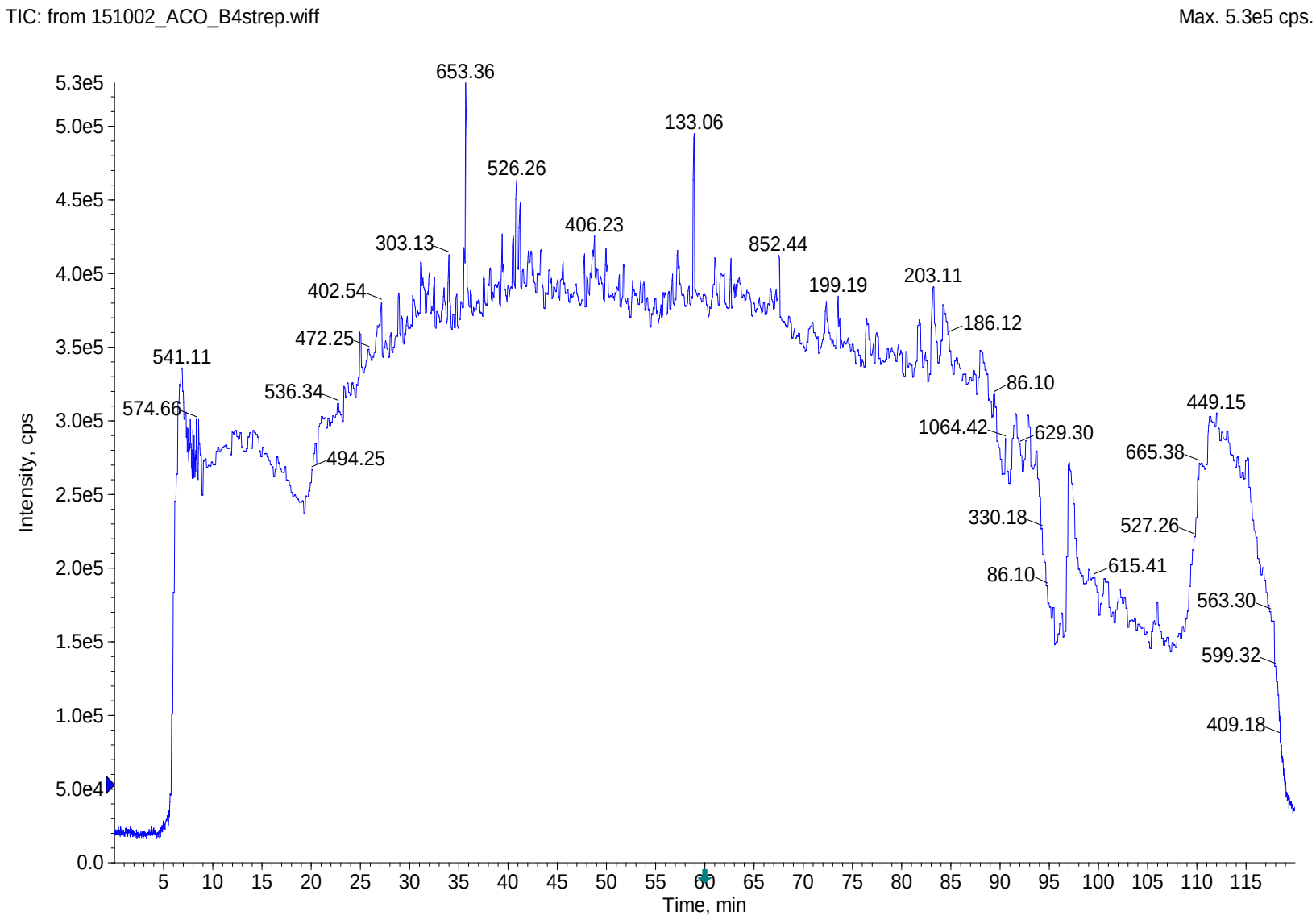
1. **P45218**
2. **P21543**
3. **Q12588**
4. **P32651**
5. **Q01245**
6. **….**

Alternative : MuDPIT (Multi Dimensional Protein Identification Technology)

Post-digestion separation of peptides by two-dimensional liquid chromatography instead of separation of proteins



A VERY complex mixture – direct analysis (no separation)



A VERY complex mixture still gives results, but...

Mascot Search Results

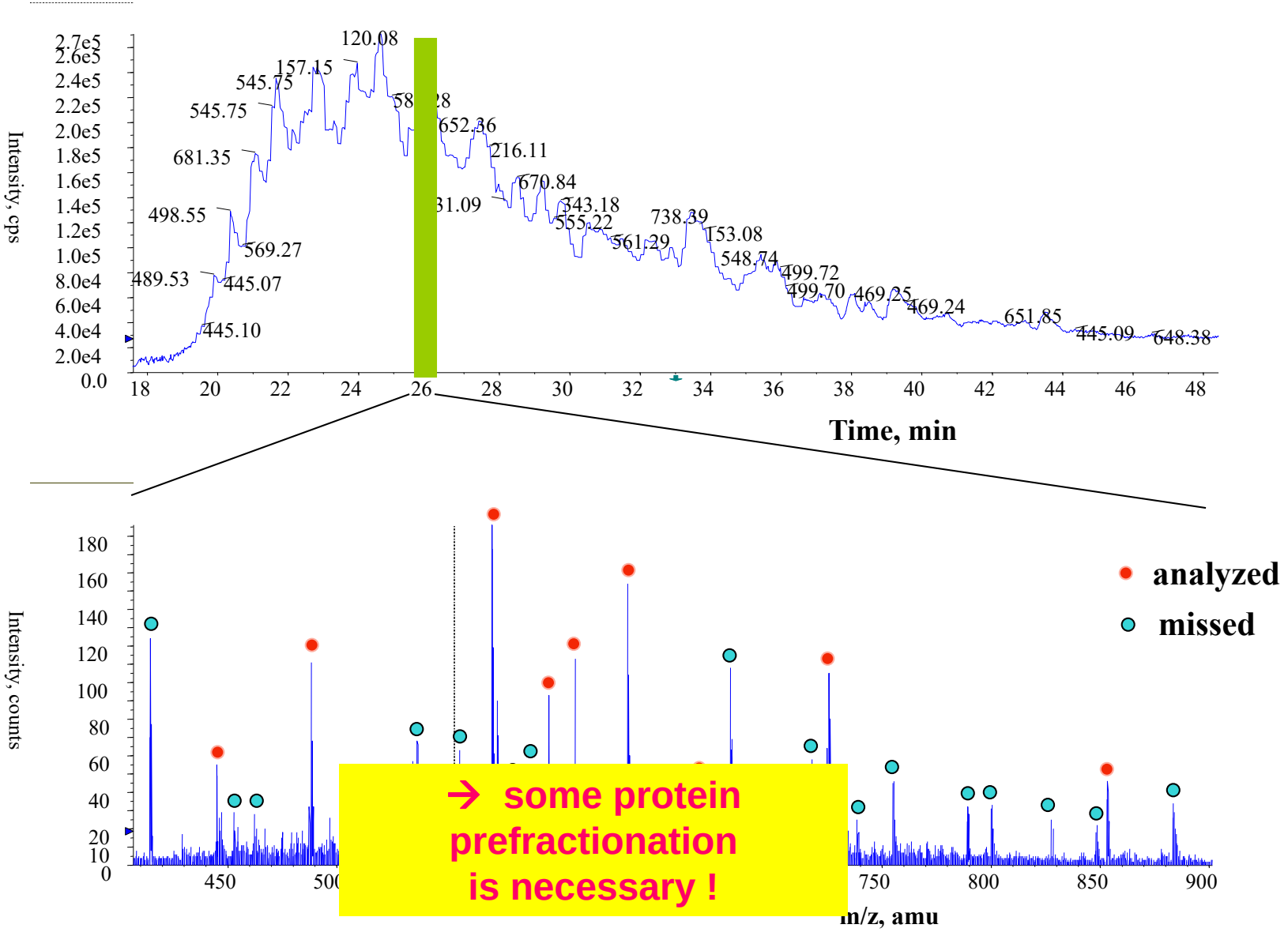
User : MQ
Email :
Search title :151002_ACO_B4strep.wiff : Angelos frac B4 IP strept
MS data file : C:\DOCUME~1\paf\LOCALS~1\Temp\mas5D.tmp
Database : Sprot 4028 (114033 sequences; 41888693 residues)
Taxonomy : Mammalia (mammals) (23838 sequences)
Timestamp : 17 Oct 2002 at 08:09:30 GMT

Significant hits:

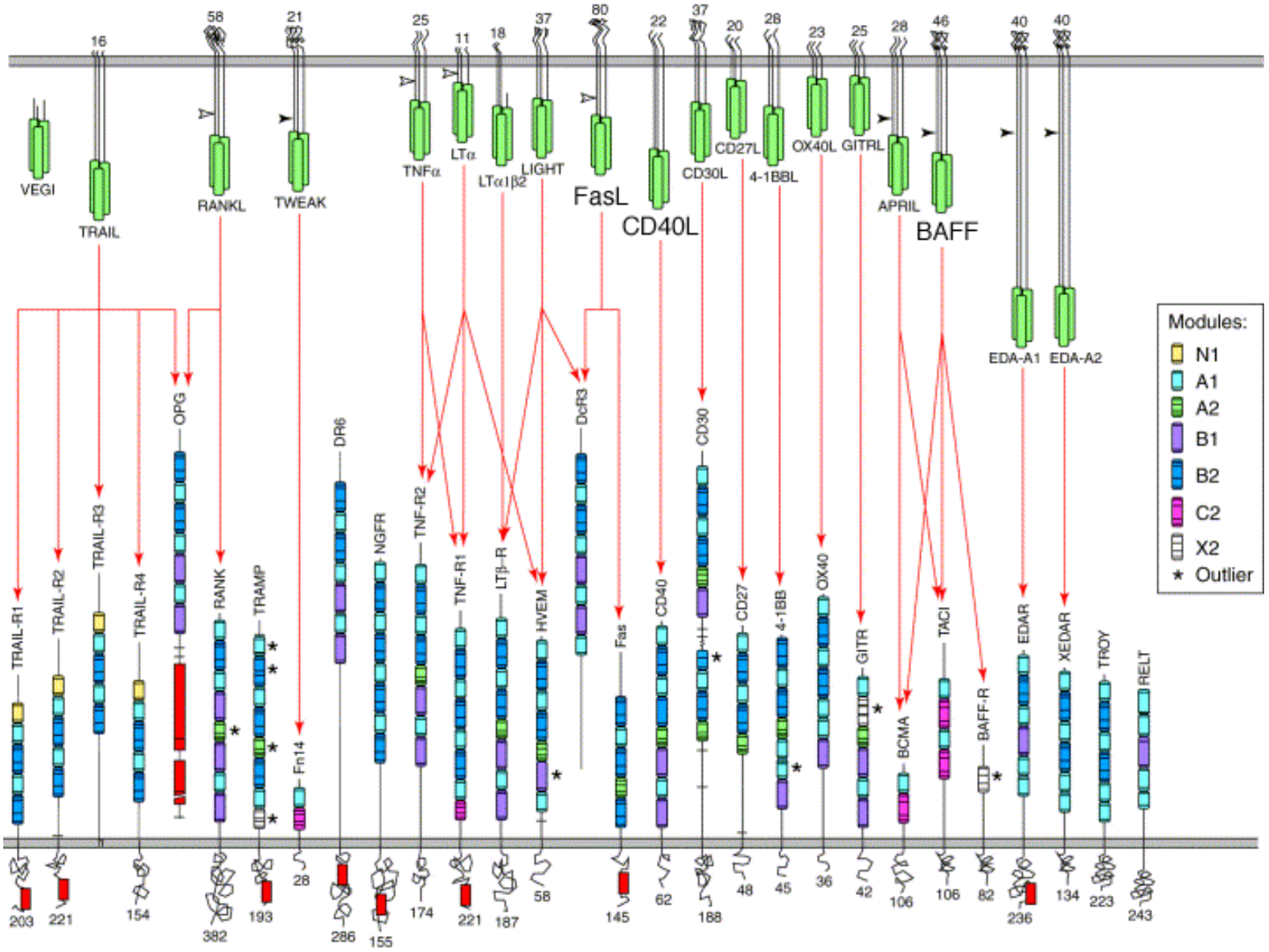
ALBU_BOVIN (P02769) Serum albumin precursor (Allergen Bos d 6).
DNM1_HUMAN (P26358) DNA (cytosine-5)-methyltransferase 1 (EC 2.1.1.37)
AC15_HUMAN (P35251) Activator 1 140 kDa subunit (Replication factor C
IF16_HUMAN (Q16666) Gamma-interferon-inducible protein Ifi-16 (Interfe
K1CJ_HUMAN (P13645) Keratin, type I cytoskeletal 10 (Cytokeratin 10) (
K22E_HUMAN (P35908) Keratin, type II cytoskeletal 2 epidermal (Cytoker
ACF7_HUMAN (Q9UPN3) Actin cross-linking family protein 7 (Macrophin) (
AC14_HUMAN (P35250) Activator 1 40 kDa subunit (Replication factor C 4
ALBU_FELCA (P49064) Serum albumin precursor (Allergen Fel d 2).
AC15_MOUSE (P35601) Activator 1 140 kDa subunit (Replication factor C
DYHC_MOUSE (Q9JHU4) Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic
AC12_HUMAN (P35249) Activator 1 37 kDa subunit (Replication factor C 3
EF11_CRIGR (P20001) Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elonga
RYR3_HUMAN (Q15413) Ryanodine receptor 3 (Brain-type ryanodine recepto
K2C1_HUMAN (P04264) Keratin, type II cytoskeletal 1 (Cytokeratin 1) (K
PLE1_RAT (P30427) Plectin 1 (PLTN) (PCN).
AHNK_HUMAN (Q09666) Neuroblast differentiation associated protein AHNA
TRYP_PIG (P00761) Trypsin precursor (EC 3.4.21.4).
ACF7_MOUSE (Q9QXZ0) Actin cross-linking family protein 7 (Microtubule
CENF_HUMAN (P49454) CENP-F kinetochore protein (Centromere protein F)
ALBU_HUMAN (P02768) Serum albumin precursor.
PLE1_HUMAN (Q15149) Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1)
TRI4_HUMAN (Q15650) Activating signal cointegrator 1 (ASC-1) (Thyroid
NF1_HUMAN (P21359) Neurofibromin (Neurofibromatosis-related protein N
NEBU_HUMAN (P20929) Nebulin

. . .

..... how deep are we going ?



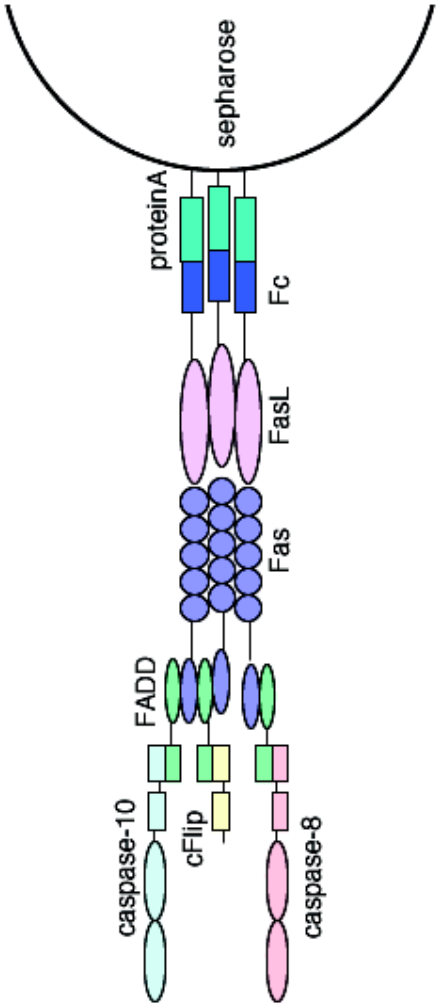
TNF family of ligands and TNF-receptor family



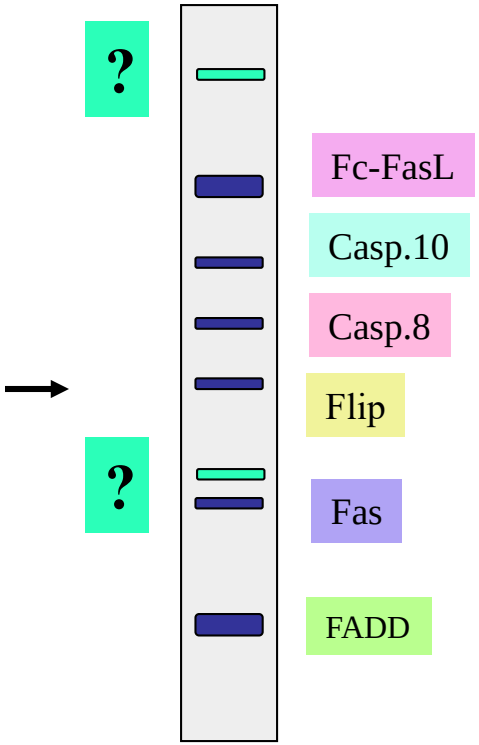
Bodmer JL et al, TIBS. 2002 27(1):19-26

Analysis of apoptotic signalling complexes

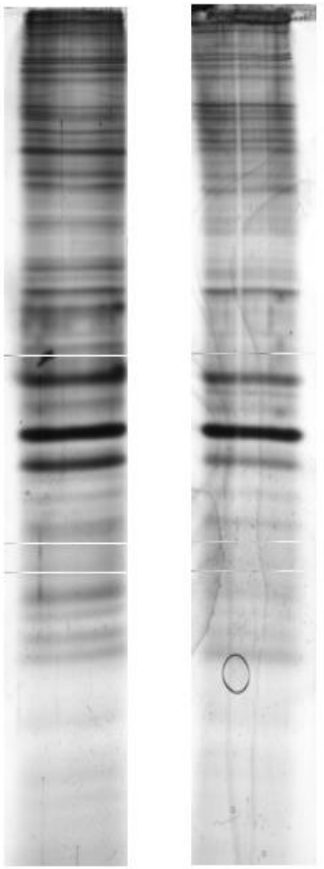
The Fas (CD95) signalling complex (DISC)



Model

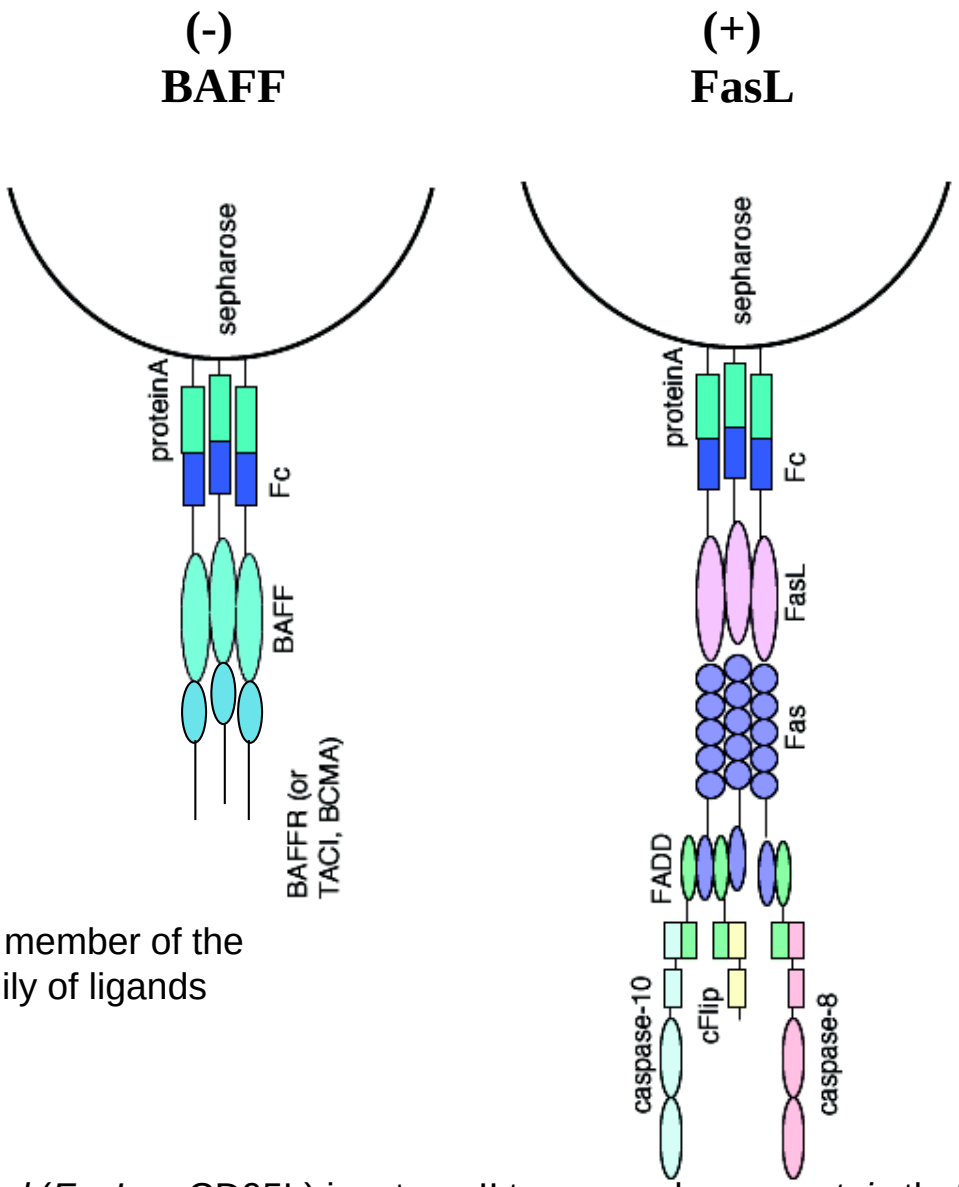


Western blot



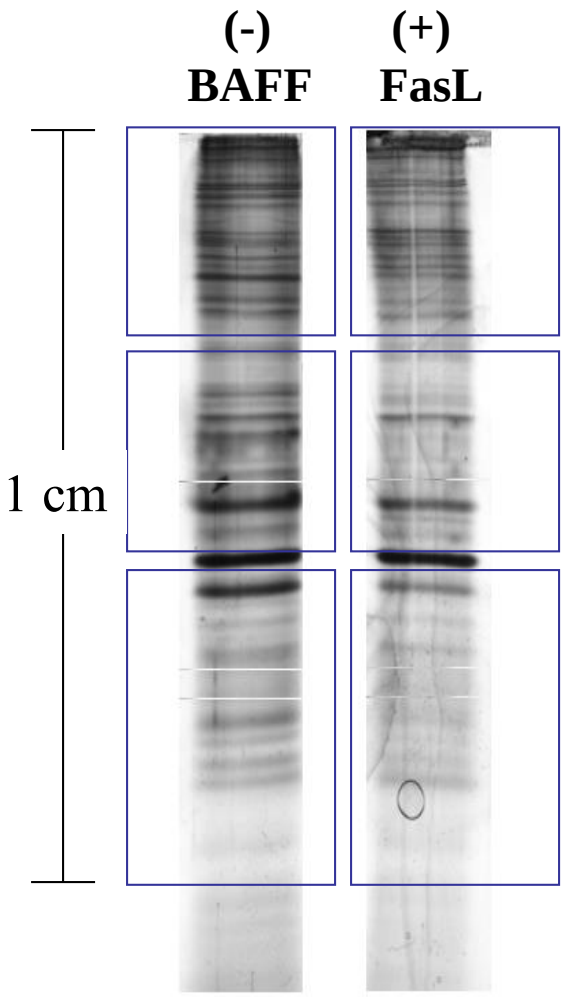
Real life

Analysis of apoptotic signalling complexes : negative control



BAFF, a member of the TNF family of ligands

Fas ligand (*FasL* or CD95L) is a type-II transmembrane *protein* that belongs to the tumor necrosis factor (TNF) family

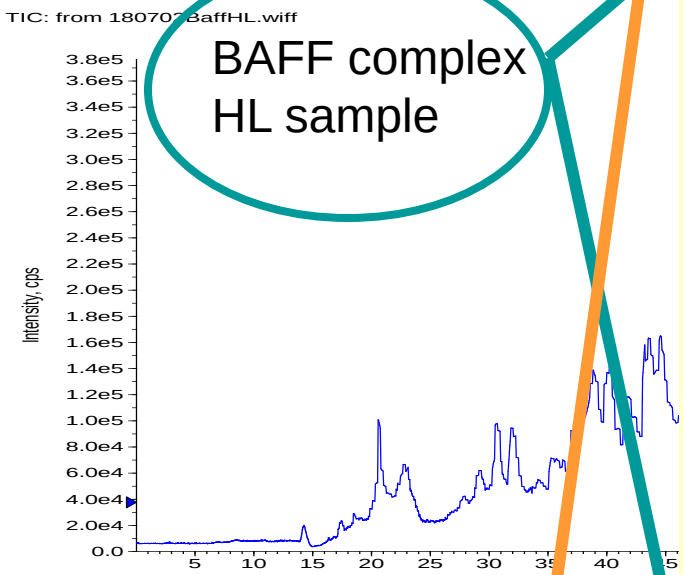


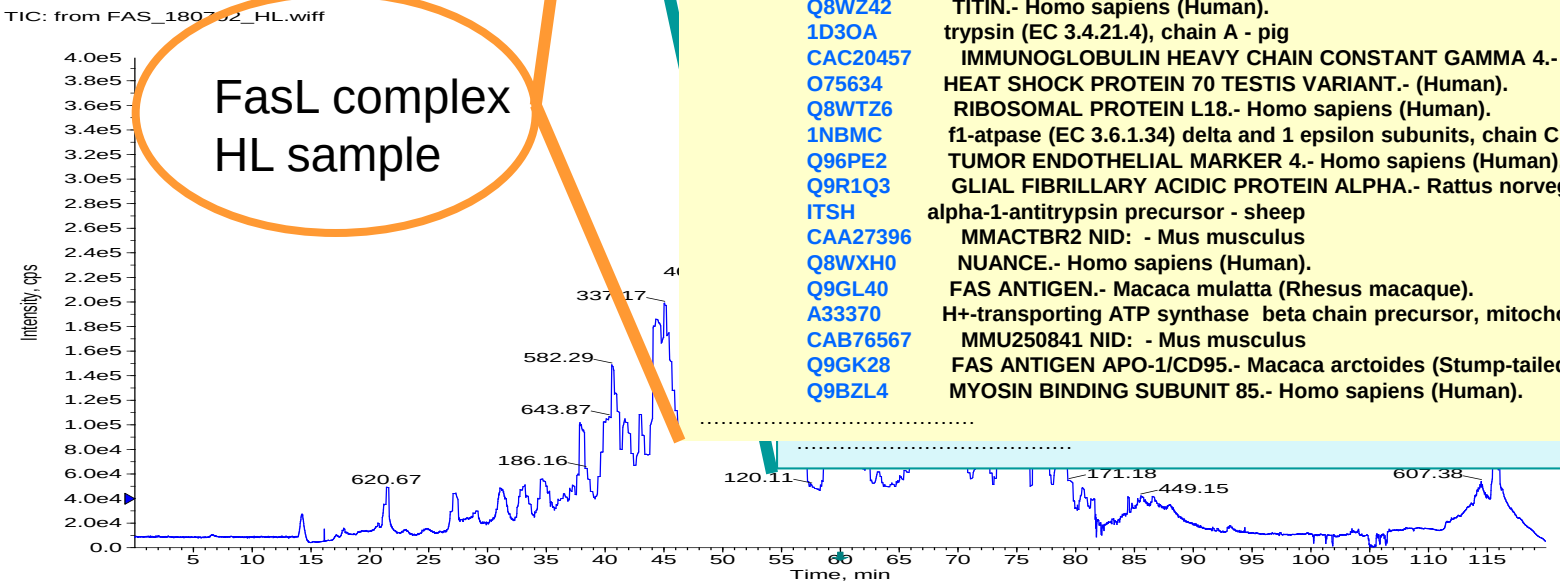
* run 1 cm
* no fix, no stain !
* cut, digest

LC-MS

BAFF complex HL sample

FasL complex HL sample





FasL HL search results

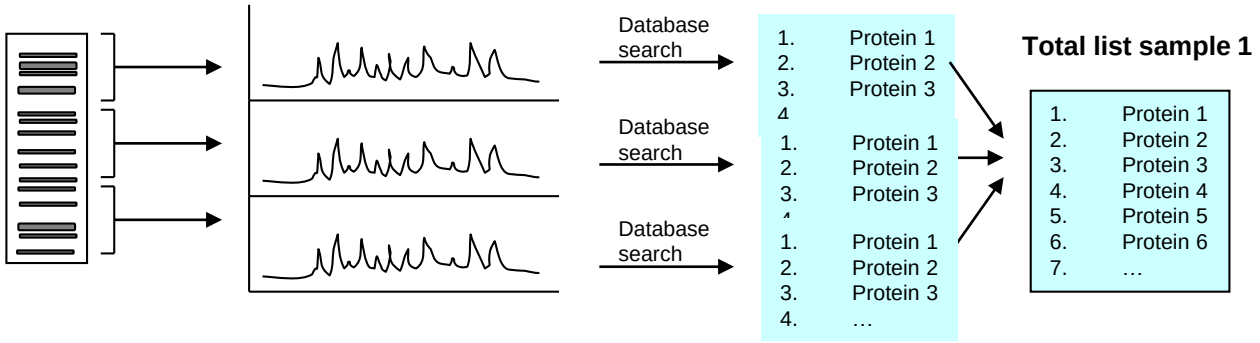
Database : MSDB 200402 (851746 sequences; 265326103 residues)
Taxonomy : Mammalia (mammals) (166849 sequences)

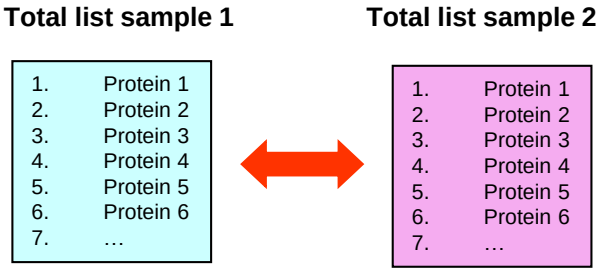
Significant hits:

A37241	52K autoantigen Ro/SS-A - human
Q8WUC1	TUBULIN, BETA 5.- Homo sapiens (Human).
C25437	tubulin beta-3 chain - mouse
CAC39526	SEQUENCE 15 FROM PATENT WO0129232.-HUman
AAH19046	SIMILAR TO IMMUNOGLOBULIN HEAVY CONSTANT GAMMA 3
K1CJ_HUMAN	Keratin, type I cytoskeletal 10 (Cytokeratin 10-(Human).
I37383	FAS soluble protein - human
A24903	tubulin alpha-1 chain - Chinese hamster
A44861	keratin, 67K type II epidermal - human
I38707	Fas ligand - human
AAG41947	AF304164 NID: - Homo sapiens
Q9BDN1	CD95L PROTEIN.- Cercocebus torquatus atys
CAA82315	HSKERAT9 NID: - Homo sapiens
AAB86467	IMMUNOGLOBULIN GAMMA HEAVY CHAIN (Human).
NUCL_HUMAN	Nucleolin (Protein C23).- (Human).
1ATS	heat shock cognate protein 70 kD (44 kD chaperone ATPase
1FC1A	Ig gamma-1 chain C region (Fc fragment), chain A - human
I61769	keratin 6d, type II - human (fragment)
A29904	keratin 5, type II, epidermal - human
A40389	translation elongation factor eEF-1 alpha chain (clone pS1) - rat
AAB46730	HSU86214 NID: - Homo sapiens
CAD23746	IMMUNOGLOBULIN GAMMA HEAVY CHAIN CONSTANT REGION
Q10466	TITIN, HEART ISOFORM N2-B (EC 2.7.1.-) (CONNECTIN).-(Human).
Q8WZ42	TITIN.- Homo sapiens (Human).
1D3OA	trypsin (EC 3.4.21.4), chain A - pig
CAC20457	IMMUNOGLOBULIN HEAVY CHAIN CONSTANT GAMMA 4.- (Human).
O75634	HEAT SHOCK PROTEIN 70 TESTIS VARIANT.- (Human).
Q8WTZ6	RIBOSOMAL PROTEIN L18.- Homo sapiens (Human).
1NBMC	f1-atpase (EC 3.6.1.34) delta and 1 epsilon subunits, chain C - bovine
Q96PE2	TUMOR ENDOTHELIAL MARKER 4.- Homo sapiens (Human).
Q9R1Q3	GLIAL FIBRILLARY ACIDIC PROTEIN ALPHA.- Rattus norvegicus (Rat).
ITSH	alpha-1-antitrypsin precursor - sheep
CAA27396	MMACTBR2 NID: - Mus musculus
Q8WXH0	NUANCE.- Homo sapiens (Human).
Q9GL40	FAS ANTIGEN.- Macaca mulatta (Rhesus macaque).
A33370	H+-transporting ATP synthase beta chain precursor, mitochondrial
CAB76567	MMU250841 NID: - Mus musculus
Q9GK28	FAS ANTIGEN APO-1/CD95.- Macaca arctoides (Stump-tailed macaque).
Q9BZL4	MYOSIN BINDING SUBUNIT 85.- Homo sapiens (Human).

.....

FILTERED RESULTS





Subtract from each list :

1) COMMON CONTAMINANTS (PROTEINS STICKING TO BEADS)

2) LIGAND-COPURIFYING PROTEINS (EX SJOGREN SYNDROME 52 KDA)

3) COMMON HITS (WHAT IS IN BOTH LISTS)

FILTERED RESULTS

MASCOT DATABASE SEARCH Fas (CD95) complex	MASCOT DATABASE SEARCH BAFF receptor complex
ICE8_HUMAN (Q14790) Caspase-8	
RO52_HUMAN (P19474) 52 kDa Ro protein (Sjogren syndrome type A antigen	RO52_HUMAN (P19474) 52 kDa Ro protein (Sjogren syndrome type A antigen
RS3_HUMAN (P23396) 40S ribosomal protein S3	RS3_HUMAN (P23396) 40S ribosomal protein S3
FADD_HUMAN (Q13158) FADD	
K2C1_HUMAN (P04264) Keratin	K2C1_HUMAN (P04264) Keratin
GC1_HUMAN (P01857) Ig gamma-1 chain C region	GC1_HUMAN (P01857) Ig gamma-1 chain C region
CFLA_HUMAN (O15519) CASP8 and FADD-like apoptosis regulator	
RS8_HUMAN (P09058) 40S ribosomal protein S8	
GBLP_HUMAN (P25388) Guanine nucleotide-binding protein	GBLP_HUMAN (P25388) Guanine nucleotide-binding protein
TNR6_HUMAN (P25445) TNF-family receptor CD95	
RL7A_MOUSE (P12970) 60S ribosomal protein L7a	RL7A_MOUSE (P12970) 60S ribosomal protein L7a
RL7_HUMAN (P18124) 60S ribosomal protein L7	RL7_HUMAN (P18124) 60S ribosomal protein L7
ICEA_HUMAN (Q92851) Caspase-10	
PHB_HUMAN (P35232) Prohibitin	
K22E_HUMAN (P35908) Keratin	K22E_HUMAN (P35908) Keratin
CLUS_HUMAN (P10909) Clusterin precursor	CLUS_HUMAN (P10909) Clusterin precursor
RS3A_MOUSE (P97351) 40S ribosomal protein S3A	RS3A_MOUSE (P97351) 40S ribosomal protein S3A
	T13B_HUMAN (Q9Y275) Tumor necrosis factor ligand superfamily member 13 (BAFF)
EF11_HUMAN (P04720) Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elonga	EF11_HUMAN (P04720) Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elonga

- One major evolution of proteomics technologies in the last years has been the introduction of gel-free approaches for large scale protein identification and quantification

- These methods combine isotope labelling, separation techniques and mass spectrometry

WORKFLOWS 3:
isotope labelling strategies

Relative quantification :

Comparison of proteins from samples A vs B

? Which proteins change in amount and how much ?

Applications :

-Healthy vs. diseased tissues

-Healthy vs. diseased body fluids

-Drug treated / untreated cells

-Stimulated / unstimulated cells

-Mutants / wt cells

-……..

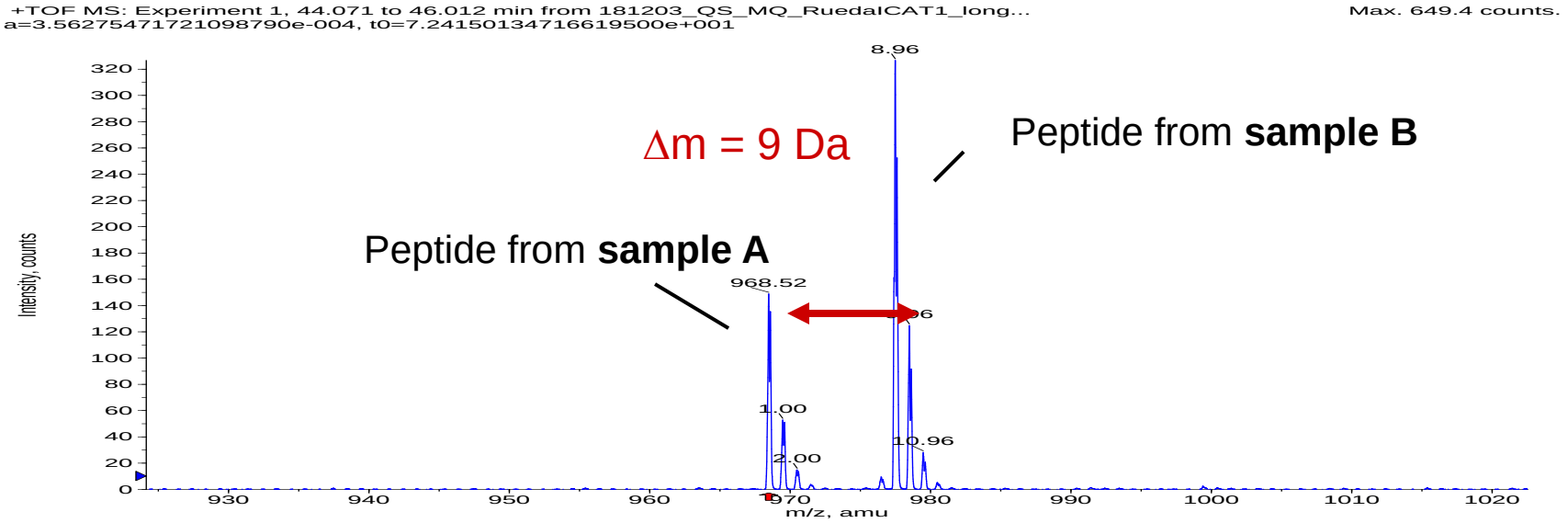
Relative quantitation : stable isotope labelling is very fashionable!

Sample A : light isotope

Sample B : heavy isotope

mix, digest

Quantitate and identify (MS)



How to label ?

-chemically, post protein synthesis

→ "specific" chemical modification of AA side chain

(+) ***any sample can be done***

(-) ***side reactions***

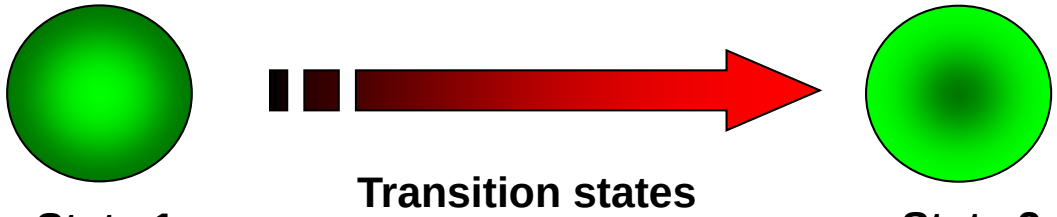
-metabolically, during protein synthesis

→Incorporation of one or more labelled amino acid

(+) "native" proteins

(-) need cultivable organism

Isotope Coded Affinity Tag (ICAT) reagents



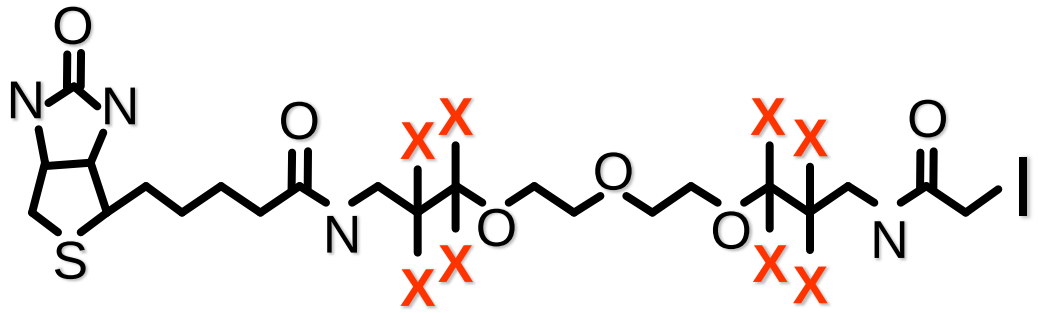
Transition states

State 1

[protein$_1$]
[protein$_2$]
.
.
[protein$_n$]

State 2

[protein$_1$]
[protein$_2$]
.
.
[protein$_n$]



Biotin tag

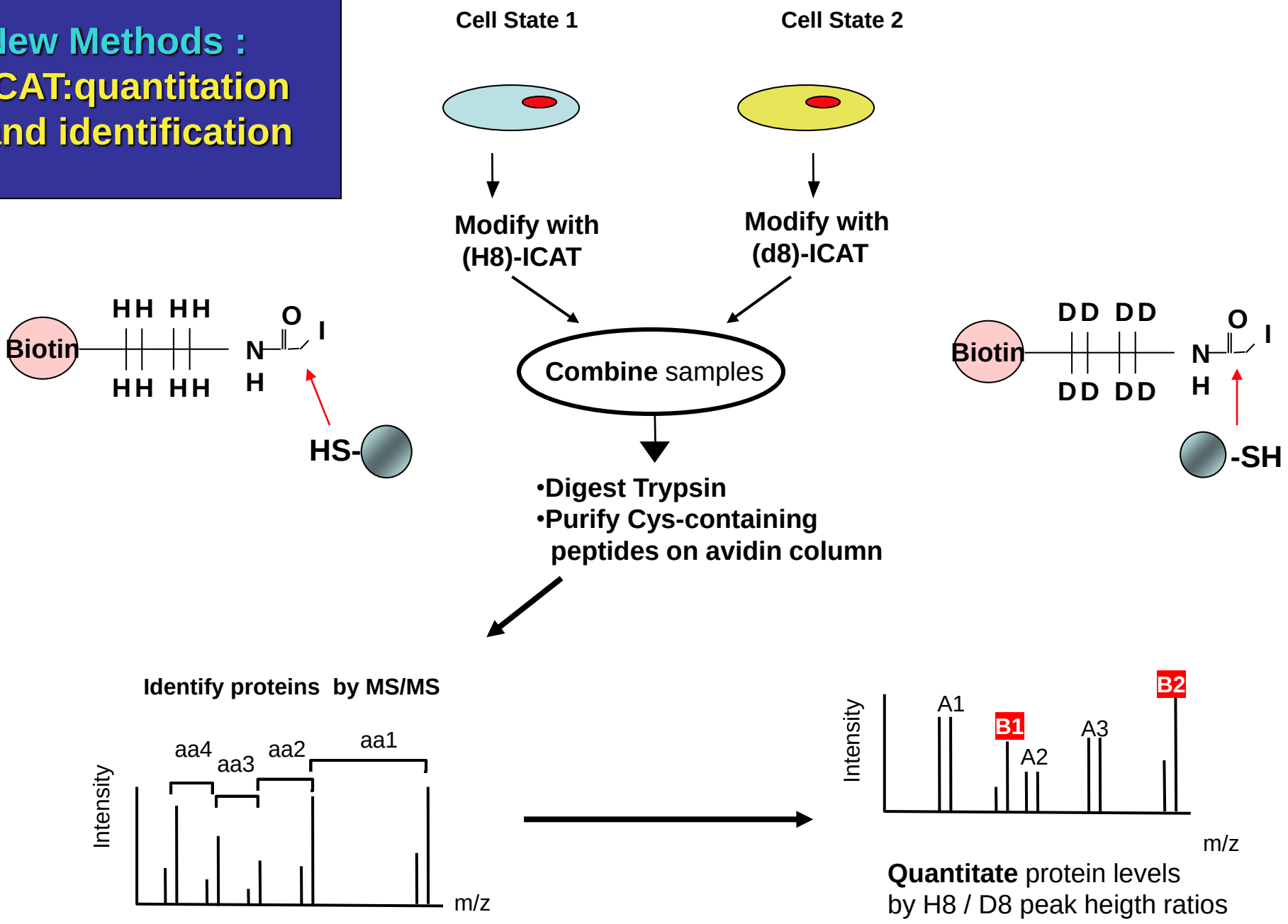
Linker (heavy or light)

Thiol reactive

d0- or d8-ICAT

(X= H or D)

New Methods : ICAT:quantitation and identification



Pair wise ICAT with Multidimensional Chromatography

1)

2)

3)

4)

5)

6)

control

treatment

d0

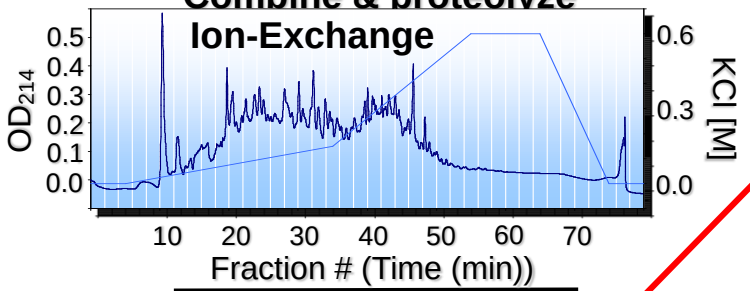
d8

ICAT label

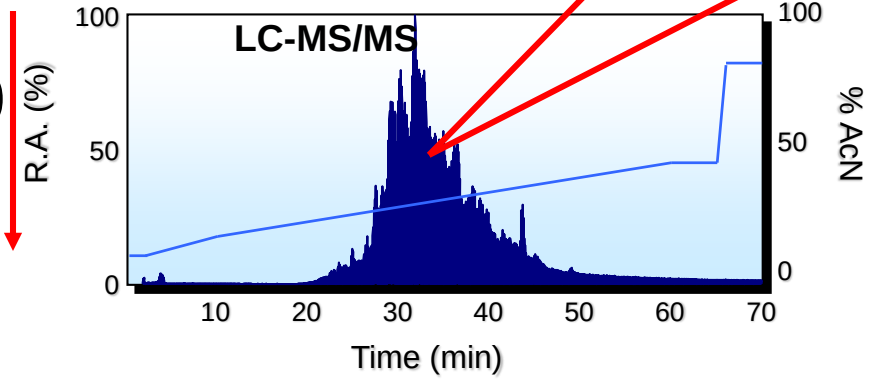
d0/d8

Combine & proteolyze

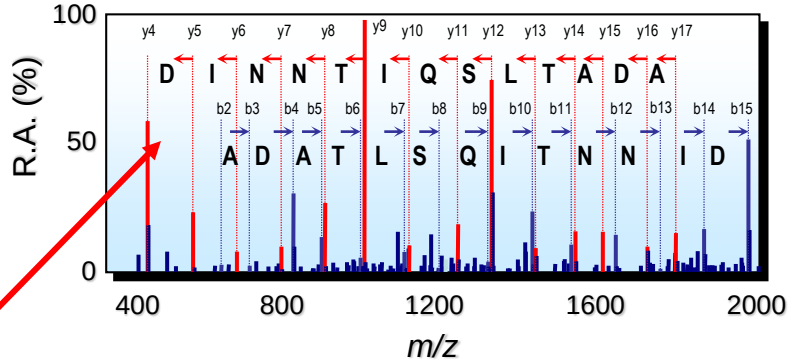
Ion-Exchange



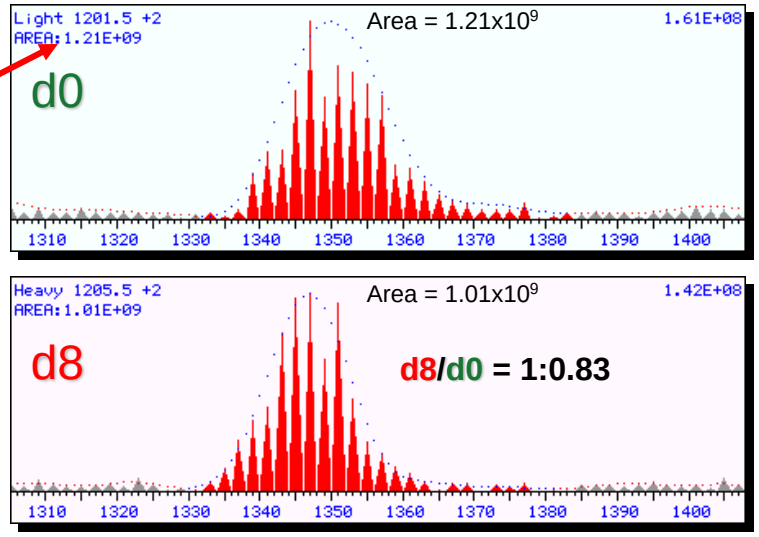
Avidin Affinity Chromatography



A) Identify



B) Quantify



ICAT (+) and (-)

+

- relative protein quantification by MS

- simplification of complex mixtures by selecting a subset of peptides after digestion

- eliminate analytical variability by mixing samples

-

- protein quantification unreliable for weak signals

- affinity purification (avidin) : losses for low amounts

- multiple side reactions possible

~15 different isotope labelling methods developed in the last 5 years !!

Recent ICAT studies (R. Aebersold's group)

Wollscheid B, von Haller PD, Yi E, Donohoe S, Vaughn K, Keller A, Nesvizhskii AI, Eng J, Li XJ, Goodlett DR, Aebersold R, Watts JD.

Lipid raft proteins and their identification in T lymphocytes.

Subcell Biochem. 2004;37:121-52

Yan W, Lee H, Yi EC, Reiss D, Shannon P, Kwieciszewski BK, Coito C, Li XJ, Keller A, Eng J, Galitski T, Goodlett DR, Aebersold R, Katze MG.

System-based proteomic analysis of the interferon response in human liver cells.

Genome Biol. 2004;5(8):R54.

Giglia-Mari G, Coin F, Ranish JA, Hoogstraten D, Theil A, Wijgers N, Jaspers NG, Raams A, Argentini M, van der Spek PJ, Botta E, Stefanini M, Egly JM, Aebersold R, Hoeijmakers JH, Vermeulen W.

A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A.

Nat Genet. 2004 Jul;36(7):714-9.

Ranish JA, Hahn S, Lu Y, Yi EC, Li XJ, Eng J, Aebersold R.

Identification of TFB5, a new component of general transcription and DNA repair factor IIH.

Nat Genet. 2004 Jul;36(7):707-13.

Hardwidge PR, Rodriguez-Escudero I, Goode D, Donohoe S, Eng J, Goodlett DR, Aebersold R, Finlay BB

Proteomic analysis of the intestinal epithelial cell response to enteropathogenic Escherichia coli.

J Biol Chem. 2004 May 7;279(19):20127-36.

Zhang J, Goodlett DR, Peskind ER, Quinn JF, Zhou Y, Wang Q, Pan C, Yi E, Eng J, Aebersold RH, Montine TJ.
Quantitative proteomic analysis of age-related changes in human cerebrospinal fluid.

Neurobiol Aging. 2005 Feb;26(2):207-27.

Marelli M, Smith JJ, Jung S, Yi E, Nesvizhskii AI, Christmas RH, Saleem RA, Tam YY, Fagarasanu A, Goodlett DR, Aebersold R, Rachubinski RA, Aitchison JD.

Quantitative mass spectrometry reveals a role for the GTPase Rho1p in actin organization on the peroxisome membrane.

J Cell Biol. 2004 Dec 20;167(6):1099-112. Epub 2004 Dec 13.

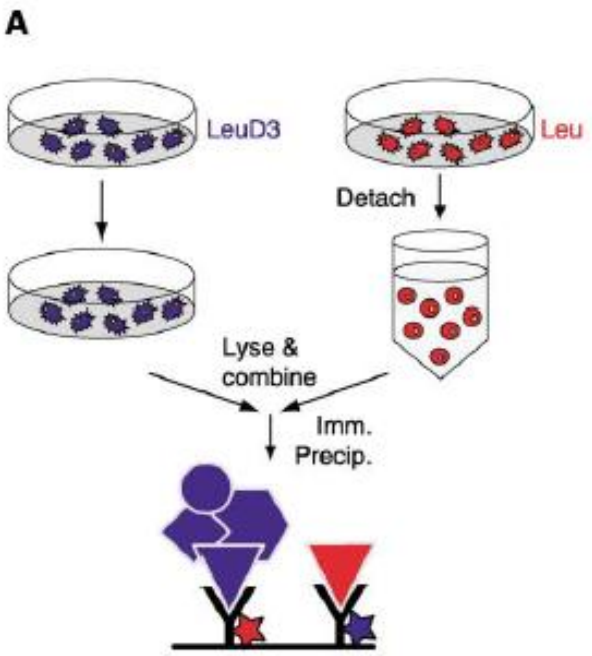
SILAC

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M.

Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.

Mol Cell Proteomics. 2002 May;1(5):376-86.

- Label light / heavy cultures (Leu d0 / d3)
- Stimulate heavy cells
- Mix cells or lysates
- Purify fraction of interest
- Analyse by LC-MS/MS (->ID)
- Quantify signals of ion pairs



SILAC (+) and (-)

+

- relative protein quantification by MS
- eliminate praparative variability by mixing samples immediately after culture
- eliminate analytical variability
- peptides in native state (no side reactions)

-

- protein quantification unreliable for very weak signals
- mass shift variable (dependent on number of residues)
- only feasible with organisms in culture

Recent SILAC articles

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M.
Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.
Mol Cell Proteomics. 2002 May;1(5):376-86.

Blagoev B, Ong SE, Kratchmarova I, Mann M.
Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics.
Nat Biotechnol. 2004 Sep;22(9):1139-45. Epub 2004 Aug 15.

Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON.
Quantitative Phosphoproteomics Applied to the Yeast Pheromone Signaling Pathway.
Mol Cell Proteomics. 2005 Mar;4(3):310-327.

de Hoog CL, Foster LJ, Mann M.
RNA and RNA binding proteins participate in early stages of cell spreading through spreading initiation centers.
Cell. 2004 May 28;117(5):649-62.

Ong SE, Kratchmarova I, Mann M.
Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC).
J Proteome Res. 2003 Mar-Apr;2(2):173-81.

Blagoev B, Kratchmarova I, Ong SE, Nielsen M, Foster LJ, Mann M.
A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling.
Nat Biotechnol. 2003 Mar;21(3):315-8. Epub 2003 Feb 10.

Foster LJ, De Hoog CL, Mann M.
Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors.
Proc Natl Acad Sci U S A. 2003 May 13;100(10):5813-8. Epub 2003 Apr 30.

Other fields

- Proteome subsets
  - Phosphoproteome
  - Ubiquitinated proteins
  - …
- Clinical proteomics (marker discovery)
  - Too vast to summarise

- Proteome imaging
  - MALDI of tissues