Laboratory of Functional Genomics and Proteomics
National Centre for Biomolecular Research
Faculty of Science, Masaryk University

Methods in Genomics and Proteomics

Mass Spectrometry in Proteomics

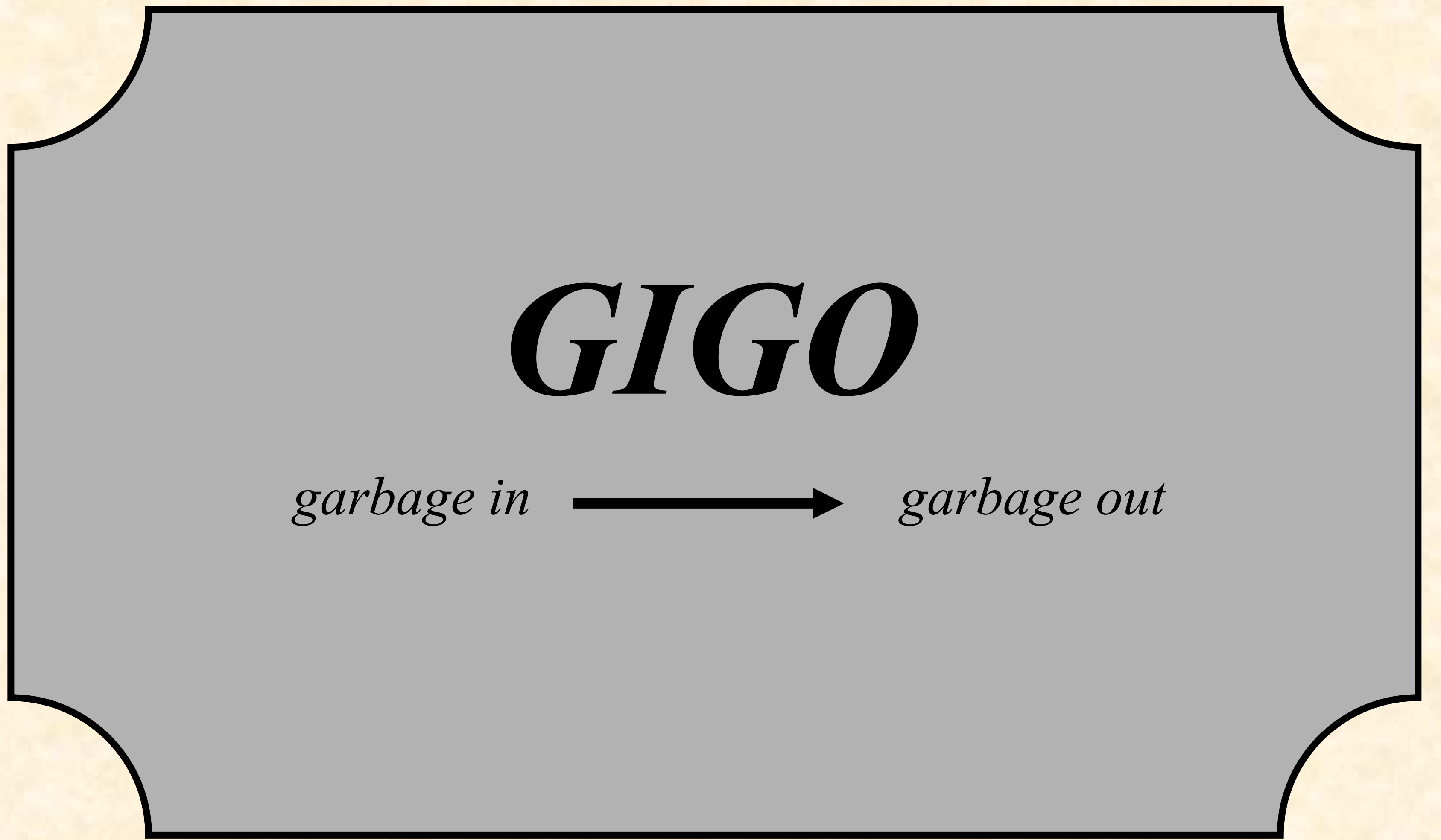
CG980

Zbyněk Zdráhal

RG Proteomics, CEITEC-MU
Proteomics CF , CEITEC-MU
NCBR, FS MU
zdrahal@sci.muni.cz

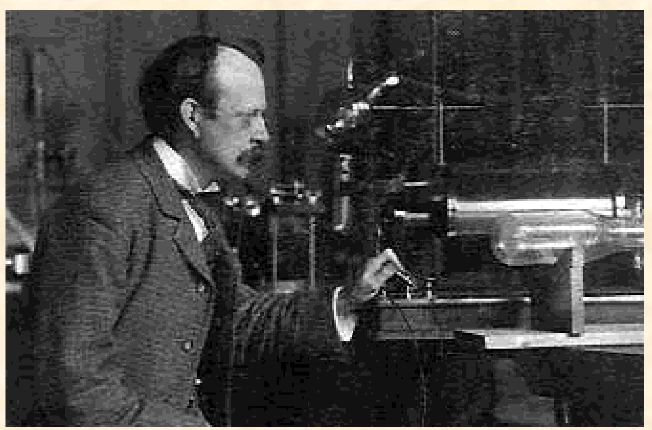
Introduction

Sample preparation is a base of good results



Main applications of mass spectrometry in proteomics

- **Intact mass measurements**
- **Protein identification** (incl. protein complexes, de novo sequencing)
- **Analysis of protein modifications**
- **Protein quantification**
- **MS imaging**
- ***Protein structure elucidation* (complementary to NMR)**



J.J. Thomson - father of mass spectrometry, Nobel prize for physics, 1906

Mass Spectrometry Basics

Mass Spectrometry (MS)

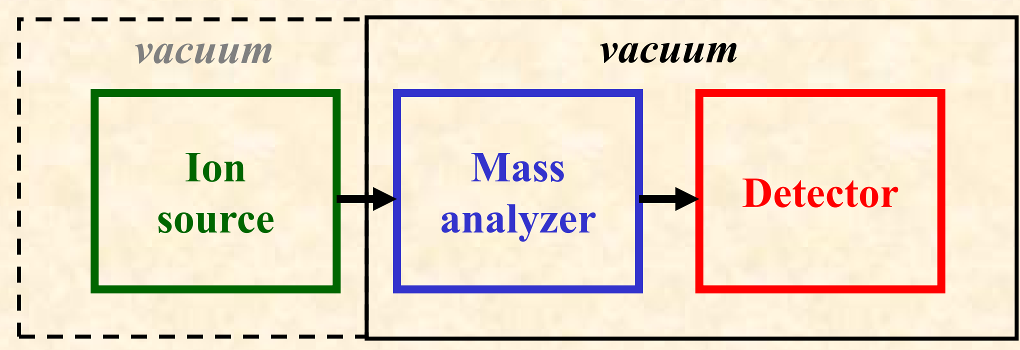
Method principle:

- measurement of *m/z* ion ratio of analyte

m – ion mass
z – charge number

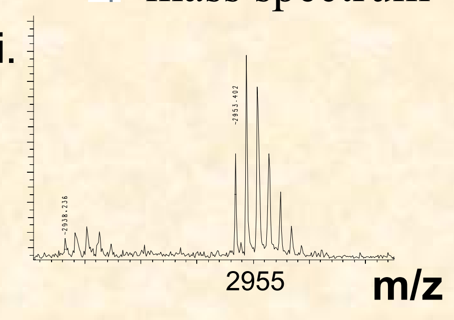
Basic steps of MS analysis:

- ionization of analyte molecules (fragments)
- *ion* separation according to their *m/z*
- ion detection



Analysis outcome:

- mass spectrum – dependence of ion intensity on their *m/z* allowing **determination of ion mass** in case of molecular ion **mass of the whole molecule**



Note: Apart from selected types of ionization, all steps of MS analysis take place in vacuum to prevent ions from unwanted collisions during their way from ion source to detector (mean free path of molecules)

Landmark in MS of biomolecules

New „soft“ ionization techniques (in the middle of 80s in the 20th century)

basic prerequisite for wide use of MS in biomolecule analysis, mainly proteins

(Nobel prize 2002)

Koichi Tanaka
Shimadzu Corp., Kyoto, Japan

MALDI

Matrix Assisted Laser Desorption/Ionization

KARAS M., HILLENKAMP F.
Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons
Anal. Chem., 60 (20): 2299-2301 (1988)

John B. Fenn
Virginia Commonwealth University, Richmond, USA

ESI

ElectroSpray Ionization

Mass spectrometry of proteins

Most widely used technique in proteomics

MALDI

Most often in combination with Time-of- Flight analyzer **TOF**

(matrix-assisted laser desorption/ionization time-of-flight mass spectrometry)

MALDI - TOF MS
MALDI – TOF/TOF MS

ESI

In combination with several different analyzer types

Ion trap – **IT**
triple quadruple and its variants - **QQQ, Q-TOF, Q-LIT,**
Ion Cyclotron Resonance **- ICR**
Orbitrap and its combinations – e.g. **IT-Orbitrap**

MALDI-TOF MS analysis

matrix is low mass compound capable to absorb laser radiation

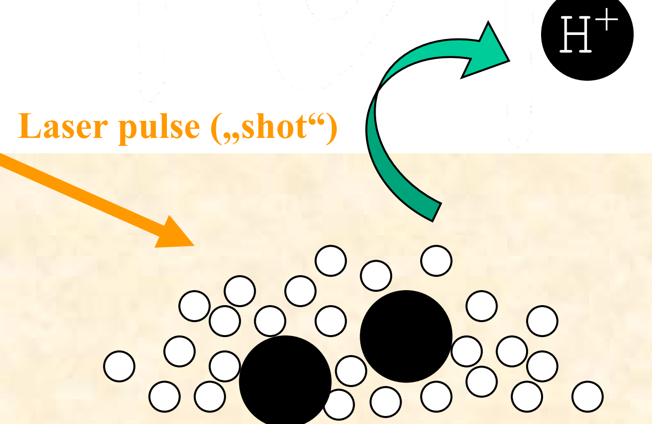
e.g. Dihydroxybenzoic acid (for UV laser)

Příprava vzorku:

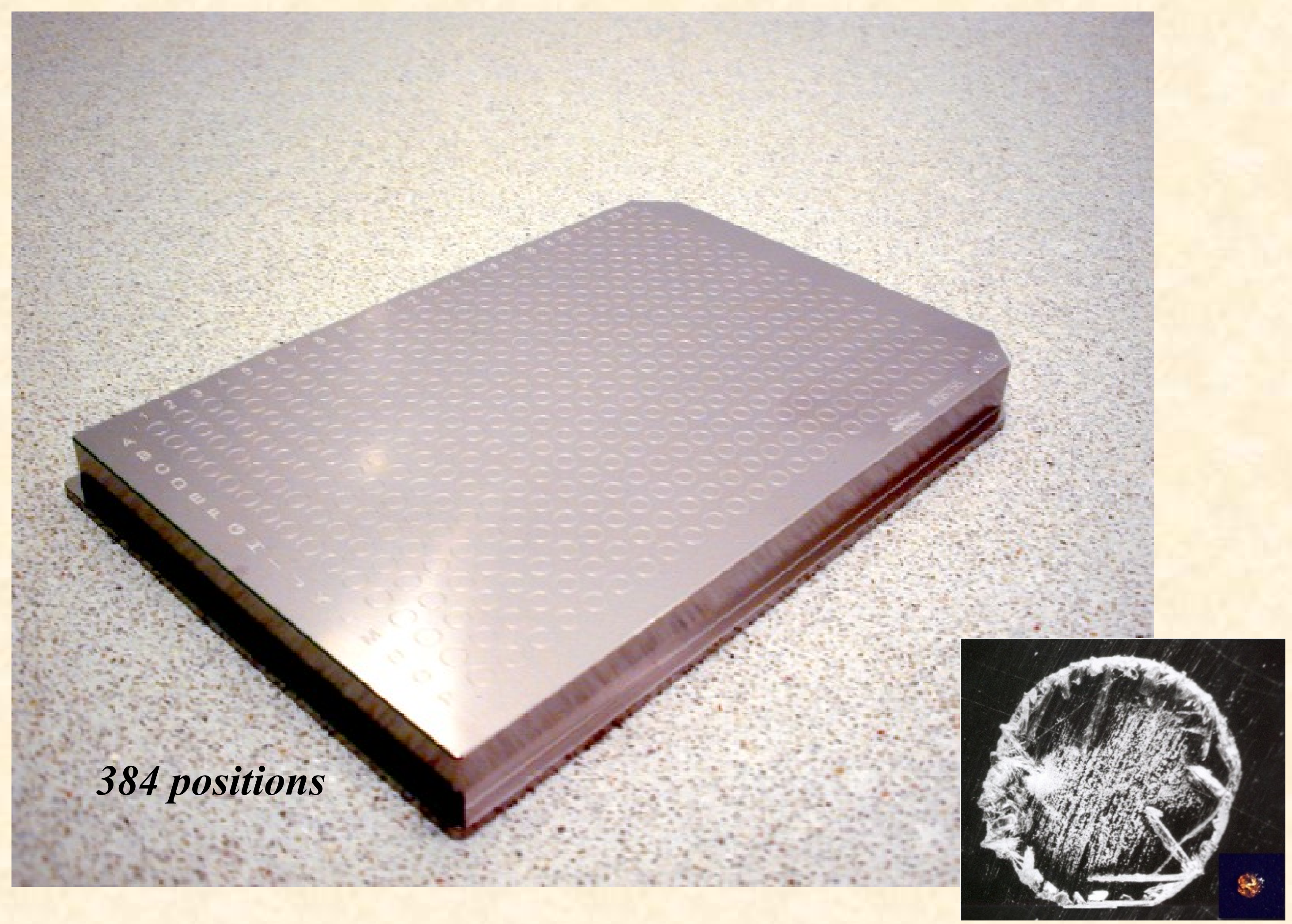
- sample is mixed with excess of matrix (in solution)
- mixture is deposited on a sample target and is allowed to dry
- sample co-crystallize with matrix during drying
- MS analysis

Result:

- Soft ionization without unwanted fragmentation
- Simple spectra
- Saving of sample on the sample target (for additional analysis)



Target for sample deposition prior MALDI-MS



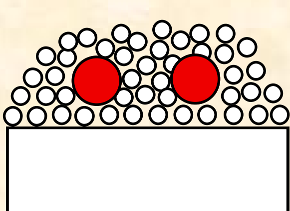
Detail of sample cocrystallized with matrix (DHB) prepared for analysis

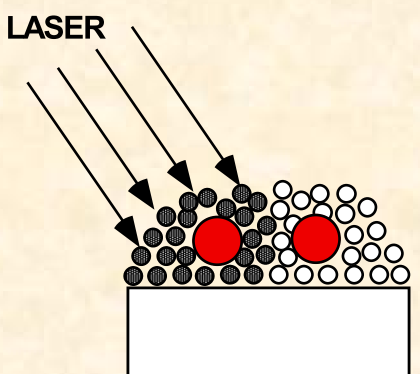
Desorption-ionization process

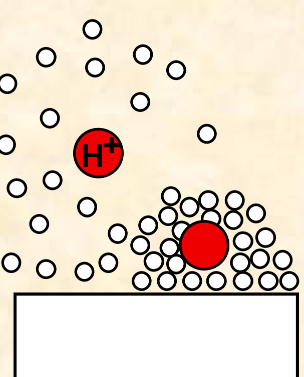
Sample embedded in light-absorbing matrix

LASER-excitation of matrix molecules

Sample desorption and protonation







pictures by courtesy of Dr. Sauerland (Bruker)

Time-of-Flight Analyzer (TOF)

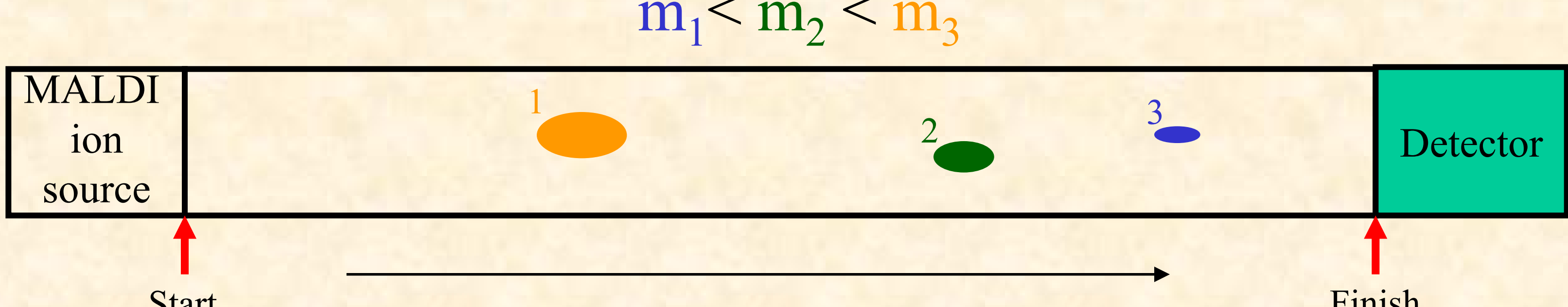
Separation of ions according to time of flight in the analyzer, time is recalculated to mass

$$E = \frac{1}{2} mv^2$$
$$V = s/t$$

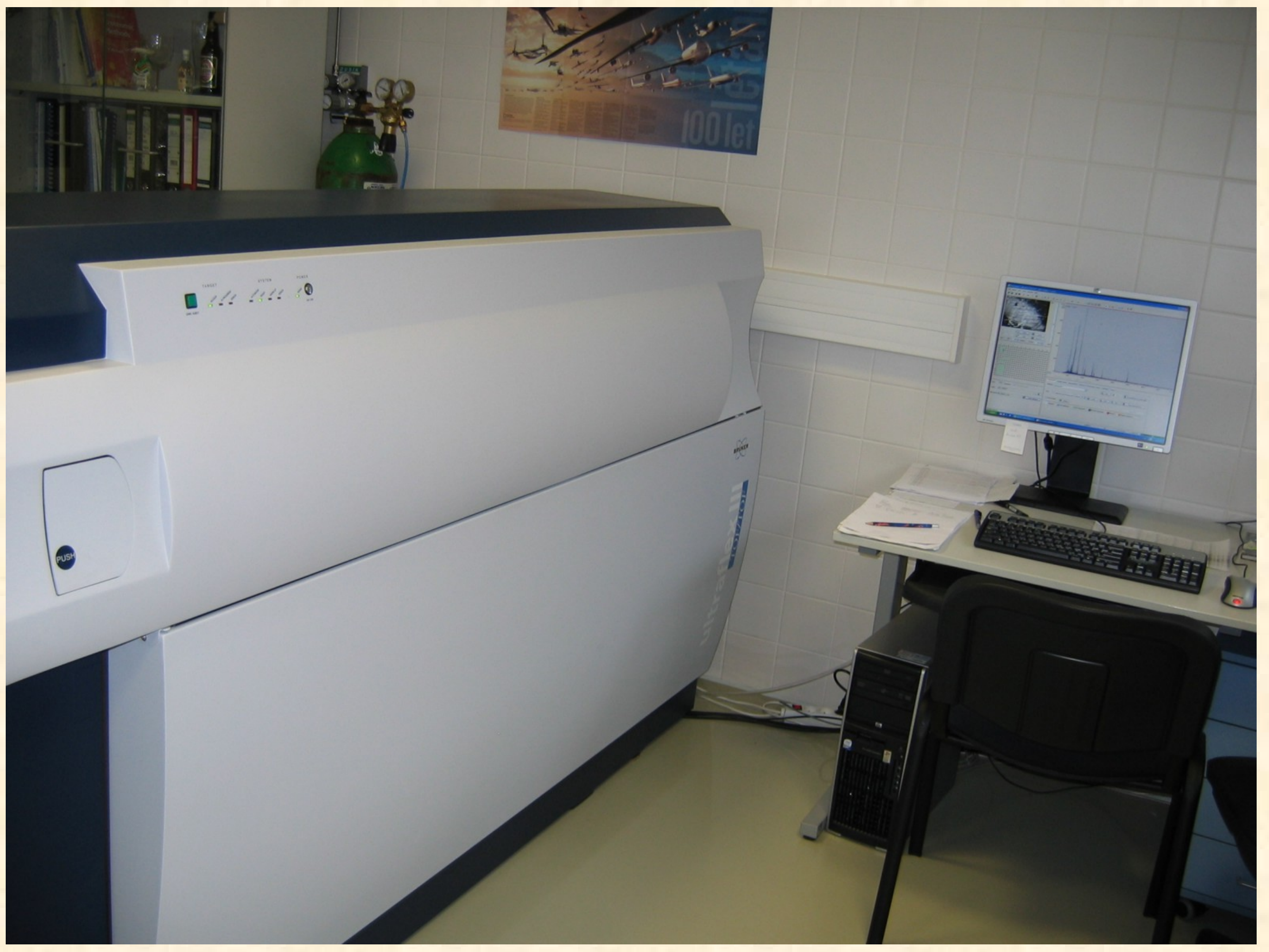
E – ion energy
m – ion mass
v – ion velocity
s – flight path
t – ion flight time

Prerequisite: Ions have to receive **the same kinetic energy** before entering analyzer drift zone which they enter simultaneously and their flight time is measured in a detector

$m_1 < m_2 < m_3$

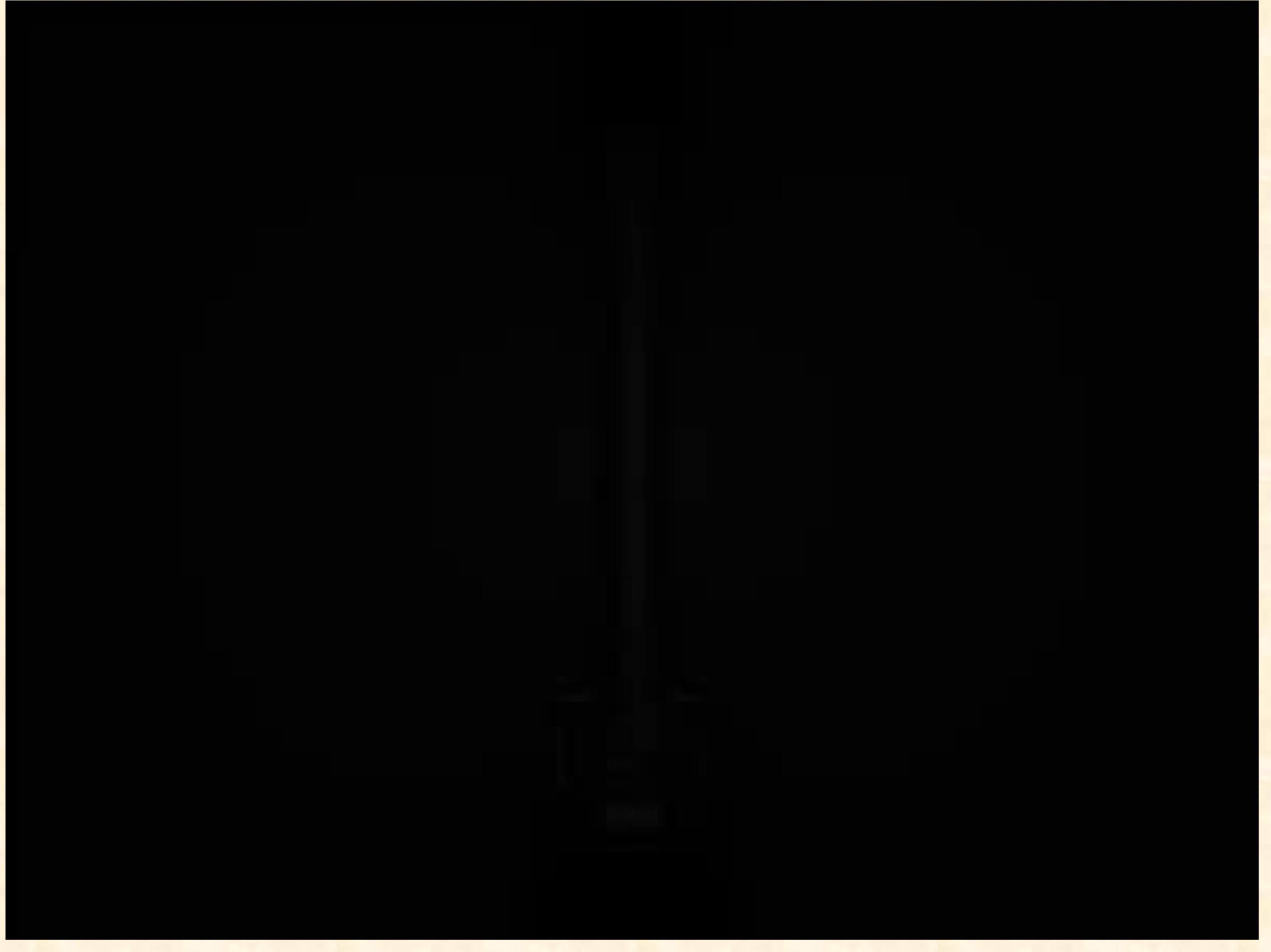


MALDI - MS/MS instrument



MALDI-TOF/TOF mass spectrometer Ultraflex III (Bruker)

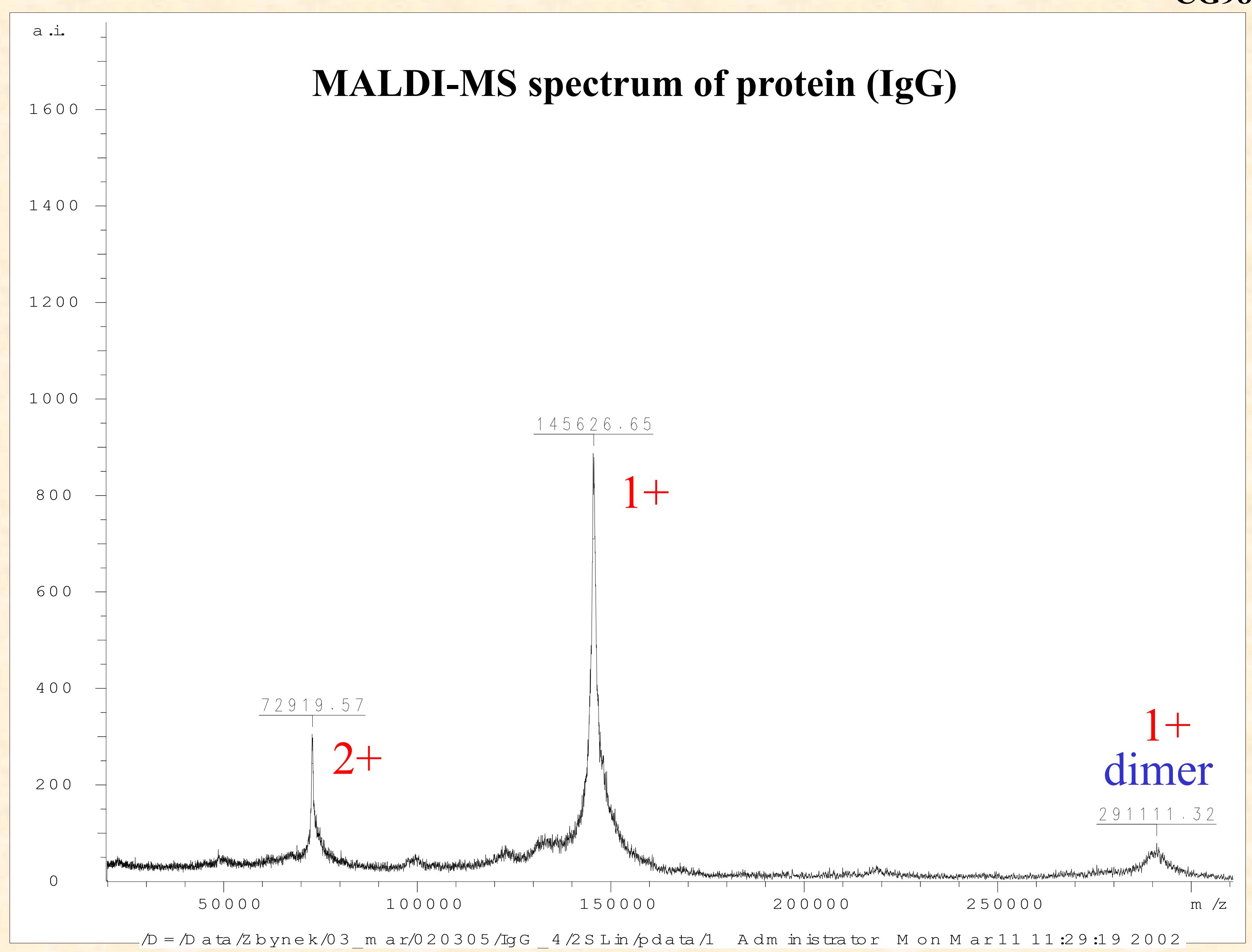
MALDI – TOF MS animation

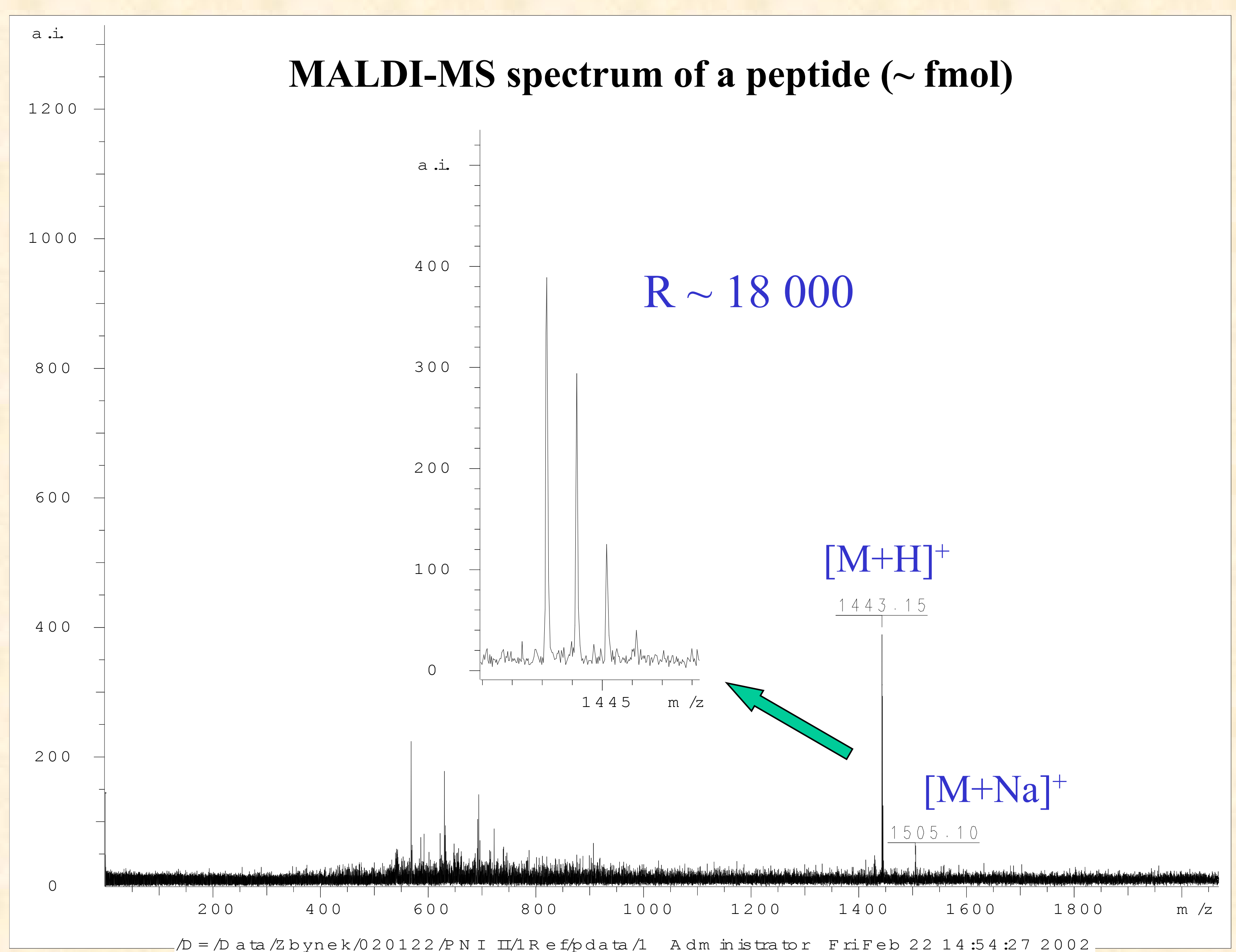


Task: find out more about MALDI-TOF MS basics on web

use expression „***time-of-flight mass spectrometry***“

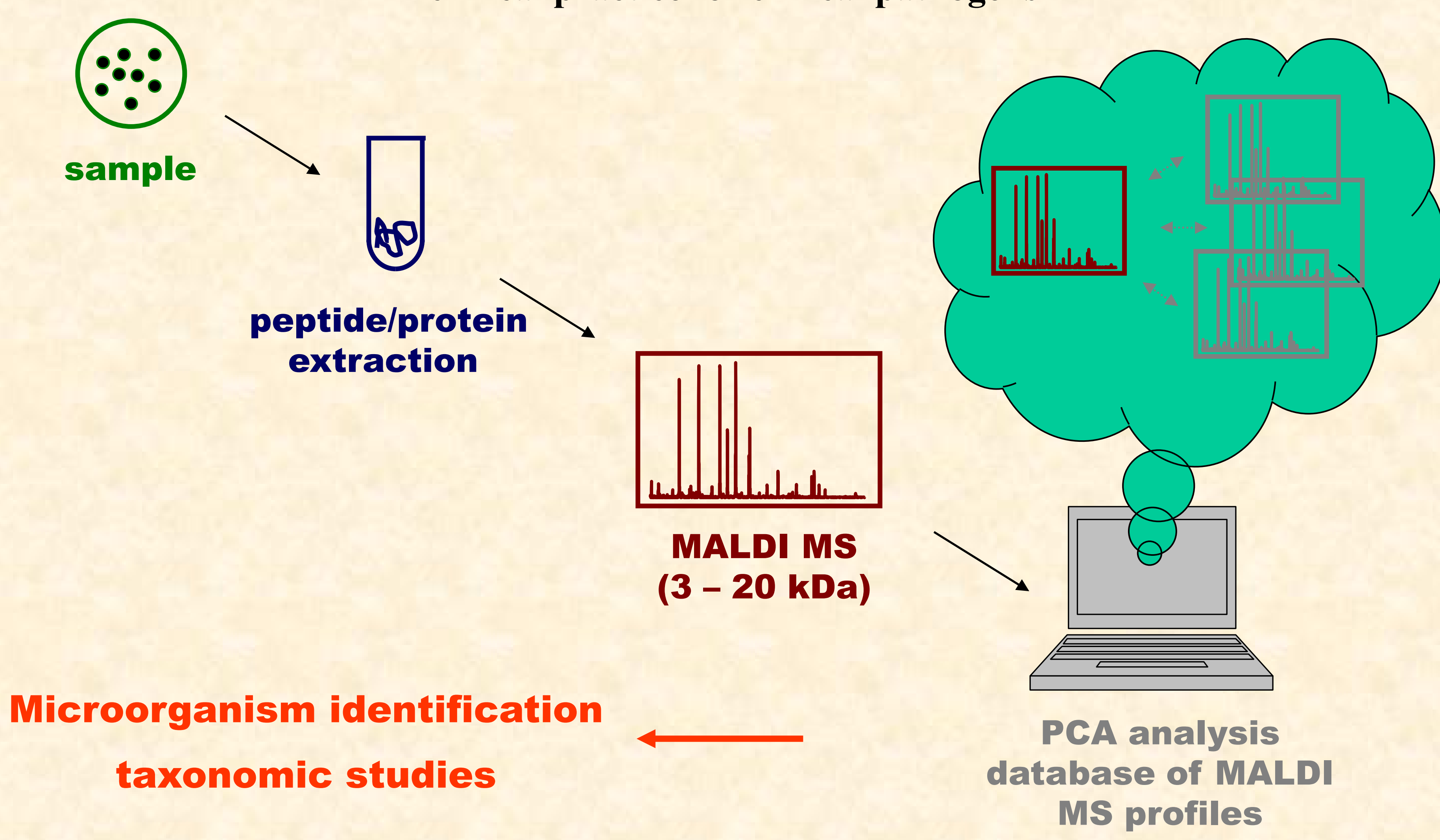
CG980





MALDI-MS spectrum of a peptide (~ fmol)
a.i.
1200
1000
800
600
400
200
0
200
400
600
800
1000
1200
1400
1600
1800
m/z
R ~ 18 000
a.i.
400
300
200
100
0
1445
m/z
[M+H]+
1443.15
[M+Na]+
1505.10
/D=/Data/Zbynek/020122/PNI_II/1Ref/pdata/1 Administrator Fri Feb 22 14:54:27 2002

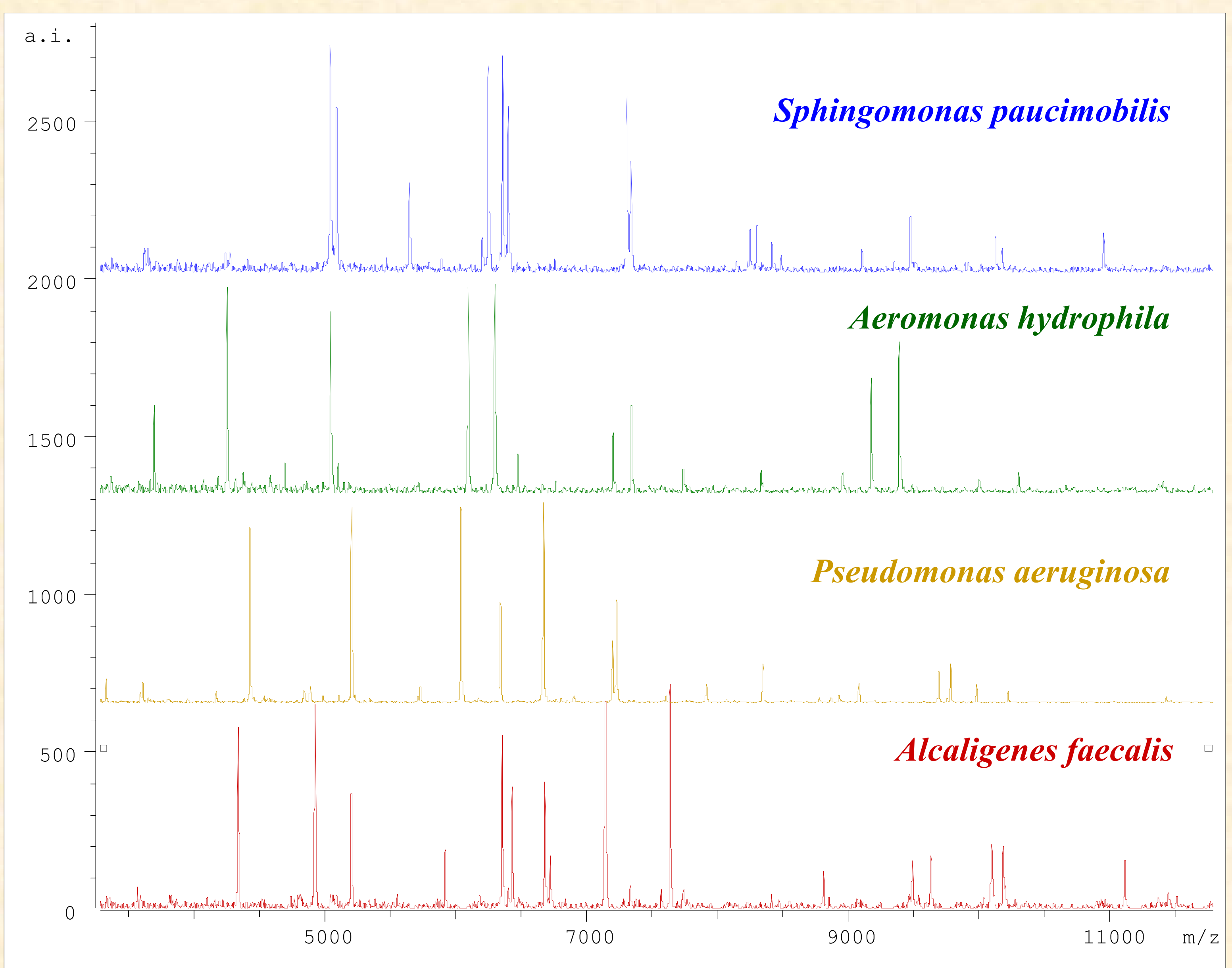
MALDI-MS profiling
microorganism identification
In clinical practice for clinical pathogens



The identification is based on comparison of the measured profile with profiles in database

MALDI-MS profiling

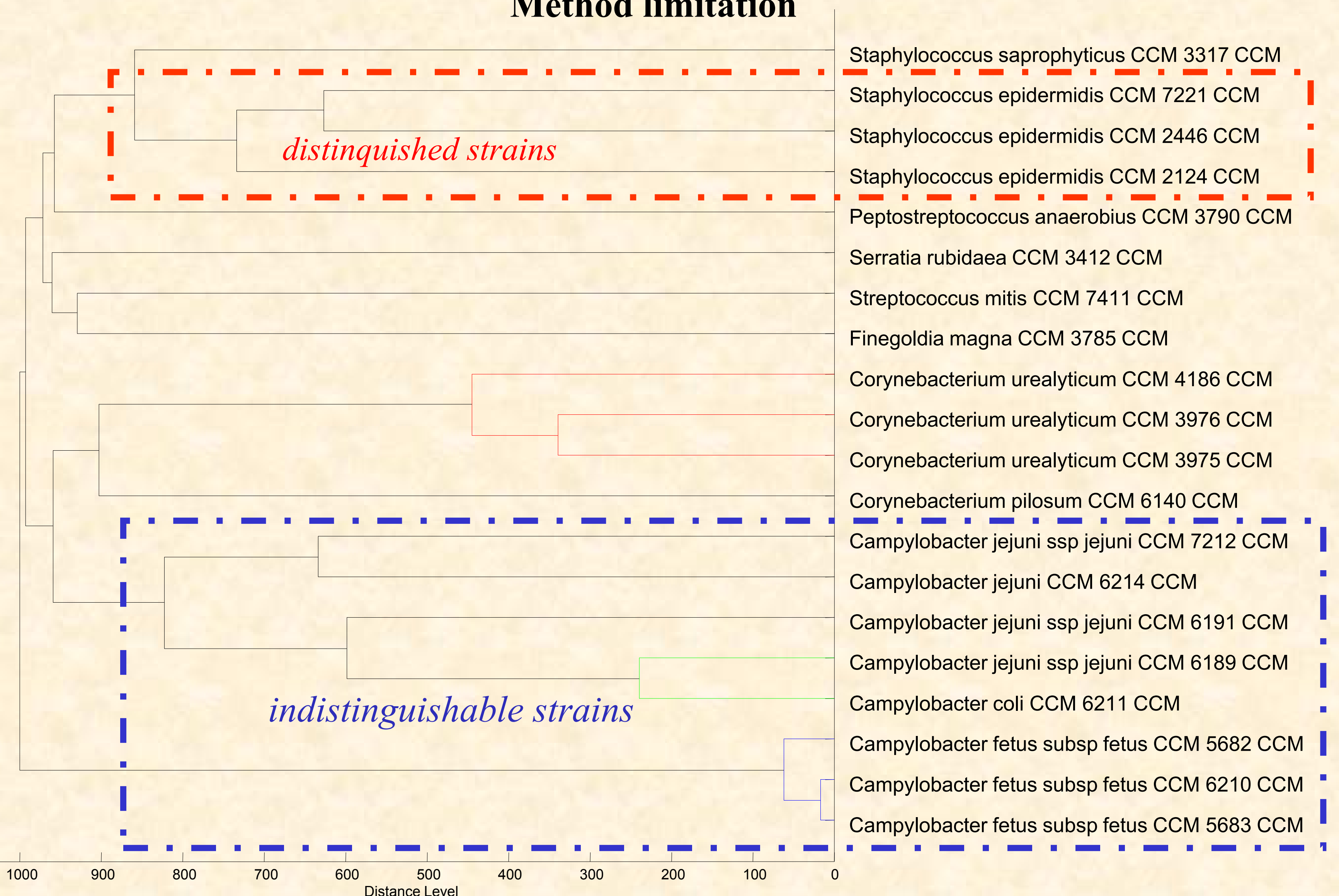
MALDI-MS spectra (profiles) of selected bacteria species



CG980

MALDI-MS profiling

Method limitation



In general, the method can not discriminate bacteria at strain level

ESI ionization

Sample preparation:

- sample has to be in solution
- sample solution is introduced into the ion source by spray needle

Sample ionization:

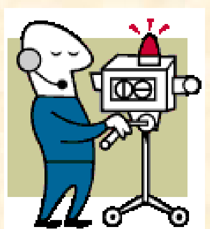
- sample solution is sprayed by spray needle in ion source chamber **under atmospheric pressure**
- ionization proceeds within the spray of liquid droplets by applying strong electric field
- charged liquid droplets are formed, which are transformed to multiply-charged ions during evaporation
- ions are transported into vacuum part of the instrument via transfer line and subjected to MS analysis

Result:

- Soft ionization without unwanted fragmentation
- Multiply charged ions
- Easy on-line connection with separation techniques (LC, CE)

LC - liquid chromatography; CE - capillary electrophoresis

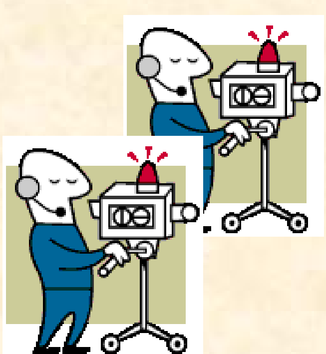
Ion trap operation



MS scan

measurement of *m/z* ratio of analyzed compounds

- ion capture in an ion trap
- sequential ejecting of ions from the trap according to *m/z*
- ion detection



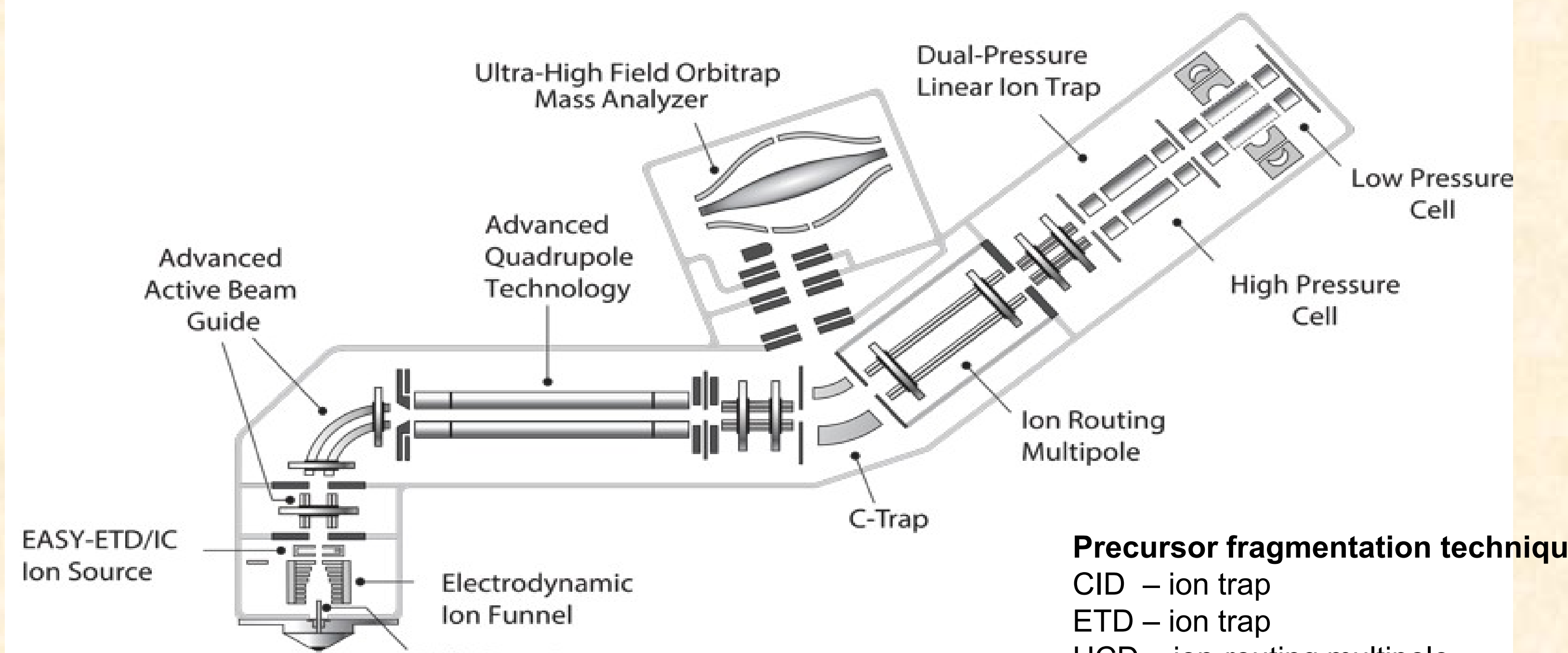
MS/MS scan

targeted fragmentation of selected ions (precursors)

- ion capture in an ion trap
- ejecting of all ions except ions with selected *m/z*
- excitation and fragmentation of selected ions
- detection of formed fragment ions (product ions)

Orbitrap Fusion™ Lumos Tribrid

example of hybrid mass spectrometer



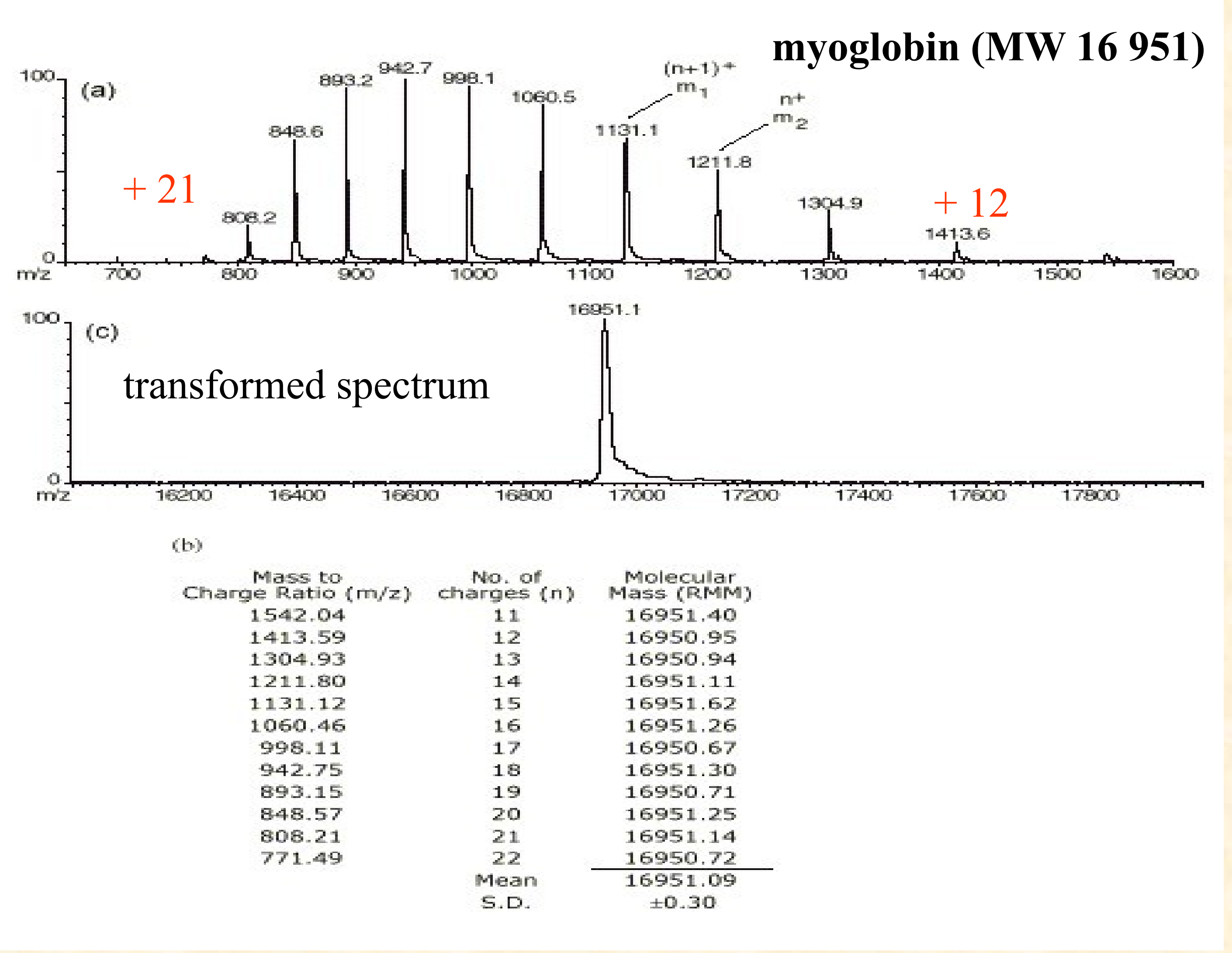
Precursor fragmentation techniques:
CID – ion trap
ETD – ion trap
HCD – ion-routing multipole
EThcD – ion trap/ion-routing multi pole

Ion separation/detection:
Ion trap – low resolution
Orbitrap – high resolution

ETD HD – high dynamic range ETD providing significantly increased fragment ion coverage

Resolution Orbitrap
15,000–500,000 (FWHM) at *m/z* 200

ESI–MS spectrum of protein



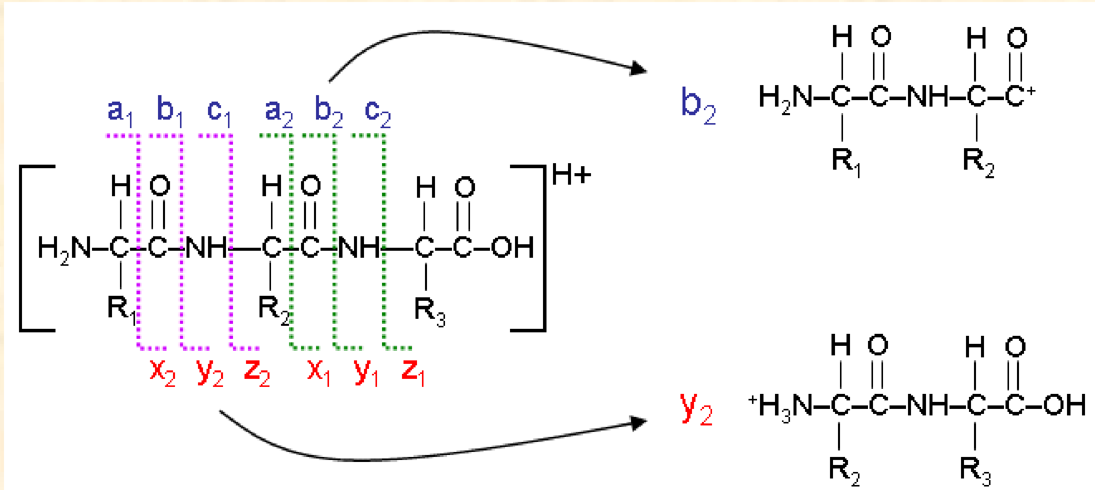
Mass to Charge Ratio (m/z)	No. of charges (n)	Molecular Mass (RMM)
1542.04	11	16951.40
1413.59	12	16950.95
1304.93	13	16950.94
1211.80	14	16951.11
1131.12	15	16951.62
1060.46	16	16951.26
998.11	17	16950.67
942.75	18	16951.30
893.15	19	16950.71
848.57	20	16951.25
808.21	21	16951.14
771.49	22	16950.72
	Mean	16951.09
	S.D.	±0.30

Serie of multiply-charged ions is formed differing in number of charges

MS/MS fragmentation of peptides

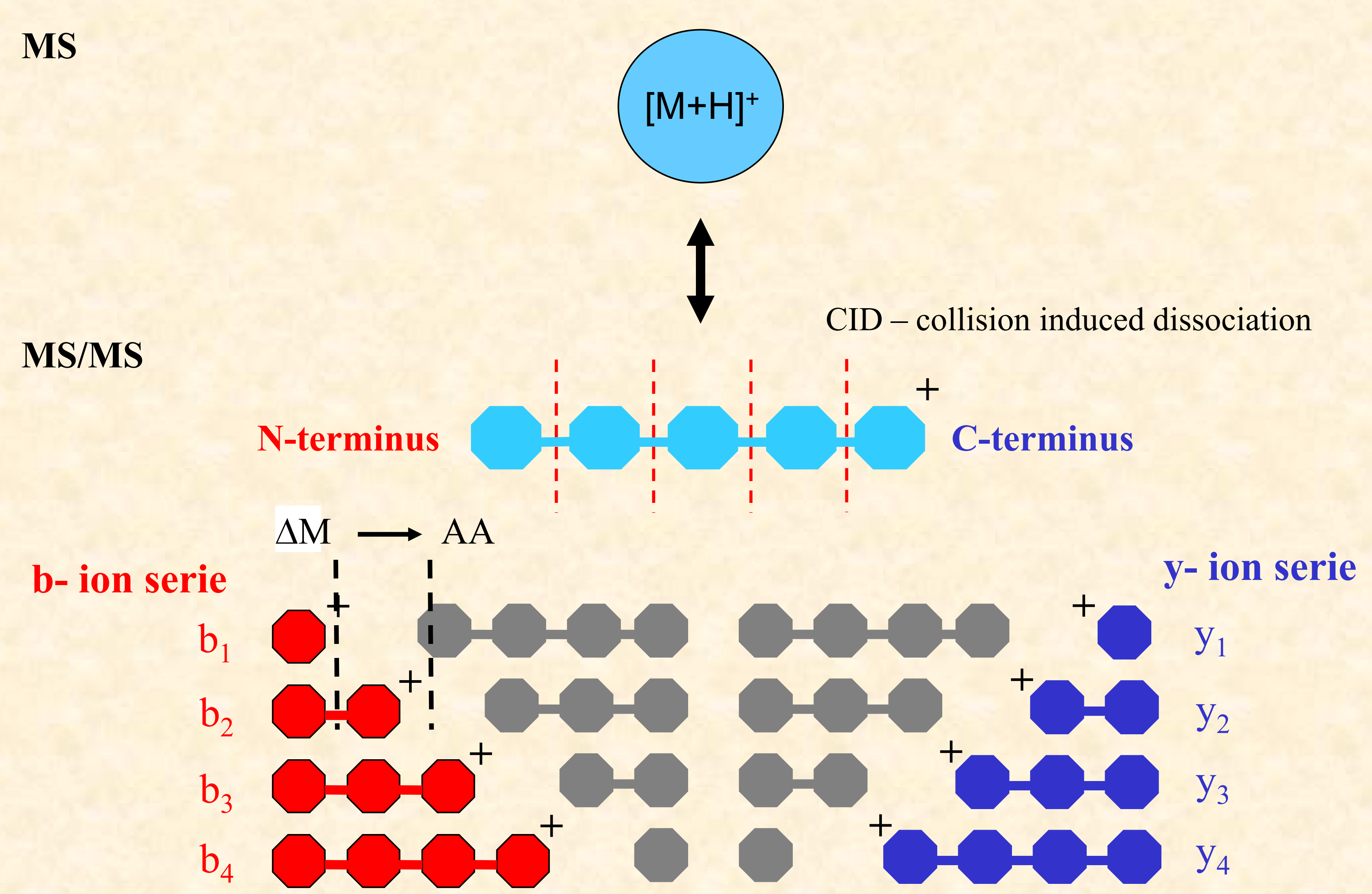
- peptides consist of individual aminoacids which are connected by peptide bond
- during fragmentation (CID), peptide is fragmented preferentially at peptide bond and thus:
  all peptide bonds might be fragmented (in each precursor molecule different ones) forming set of fragments with various number of aminoacids
  differences in *m/z* (or mass) of „neighbouring fragments“ determines type of terminal aminoacid in the longer fragment
- serie of fragment ions are formed (**b – y**, a – x, c – z) which can be used for *de novo* primary structure elucidation; moreover they are predictable and they can be used for database search based protein identification even if they are not complete

Outline of tripeptide fragmentation



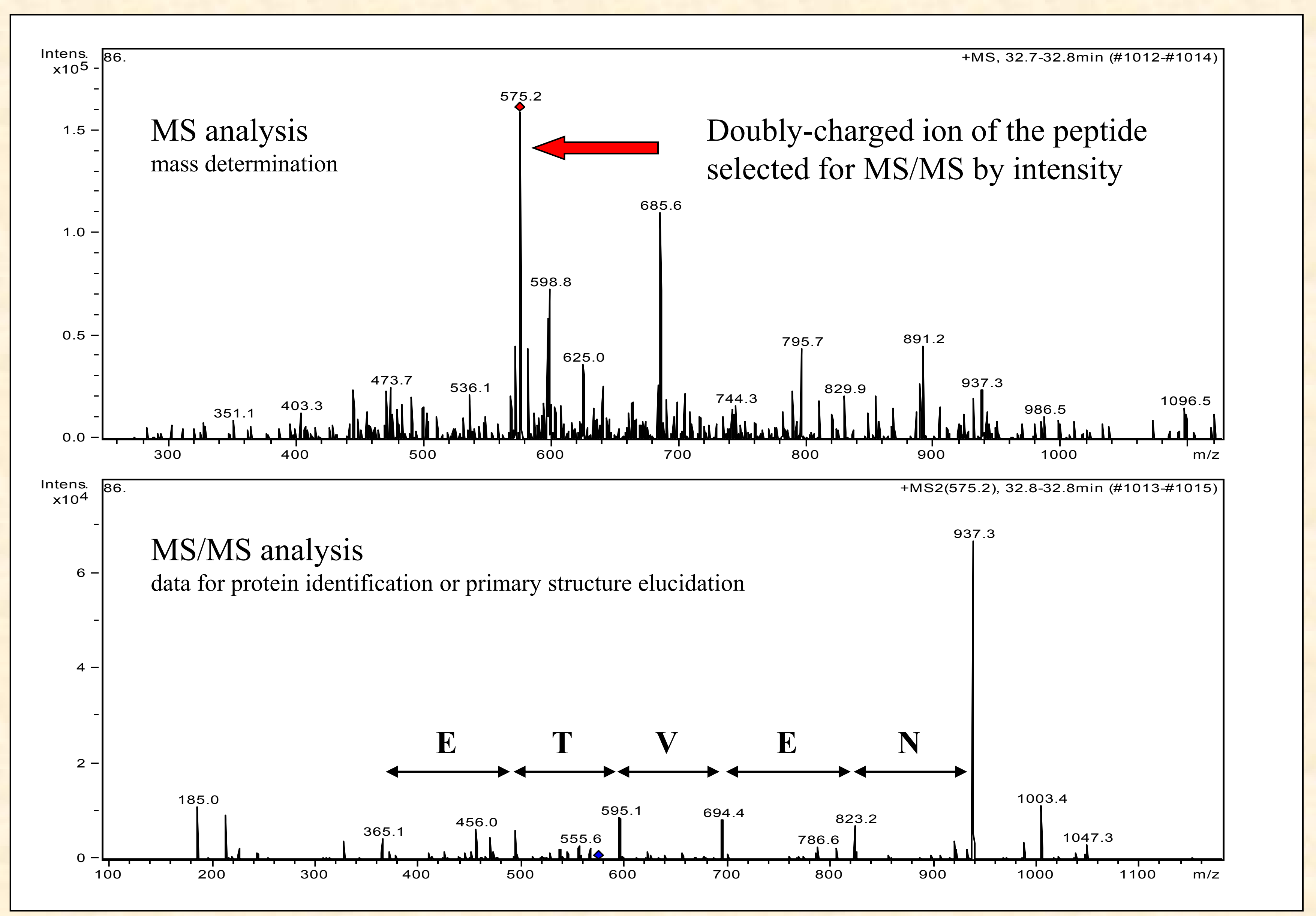
Roepstorff P. and Fohlman, J., *Biomed. Mass Spectrom*., 11 (11), 601 (1984)

MS *vs* MS/MS of peptides (CID)



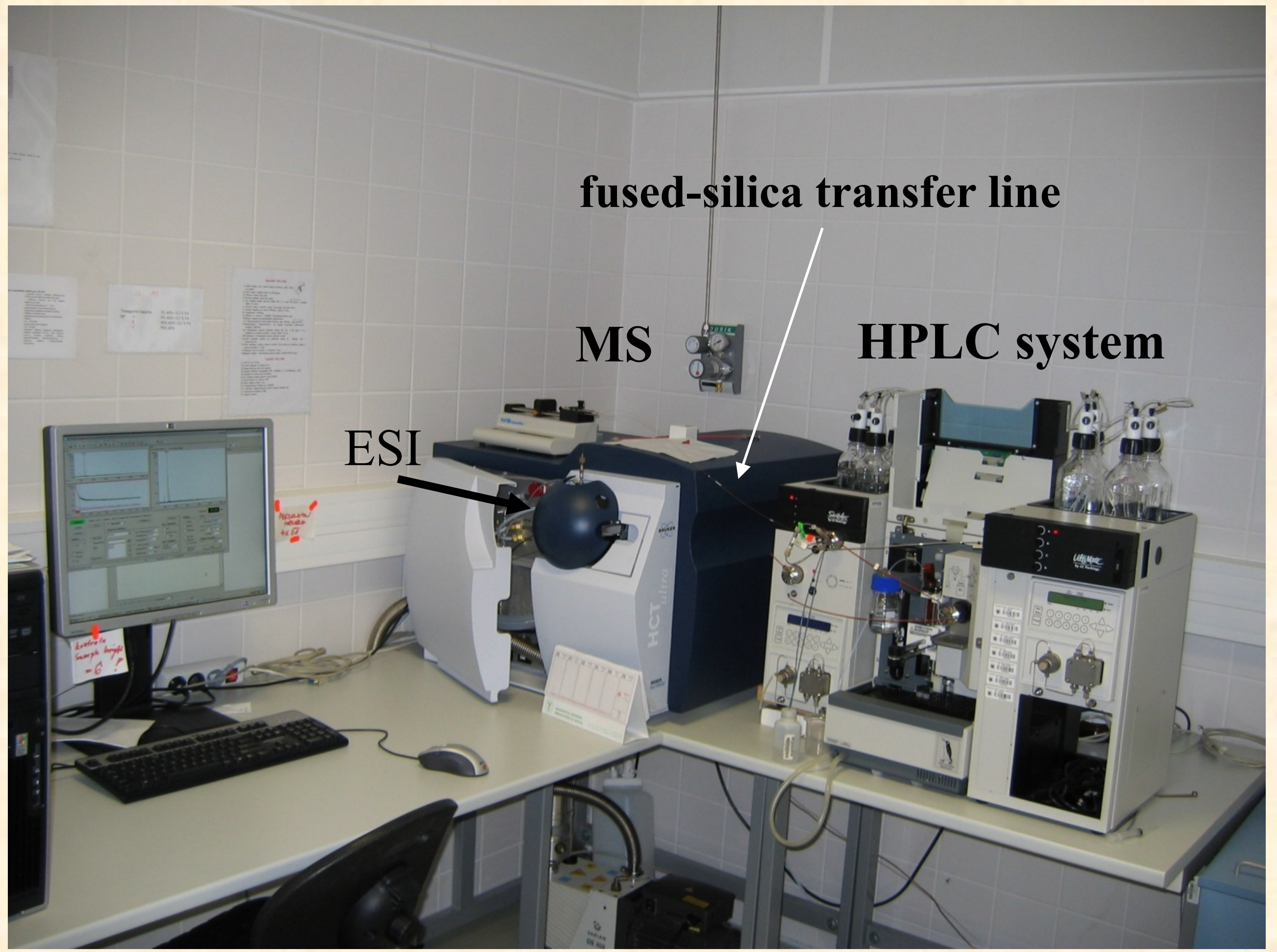
fragmentation maps for individual peptides

ESI-MS a MS/MS spectra of a peptide (MW 1148.5)



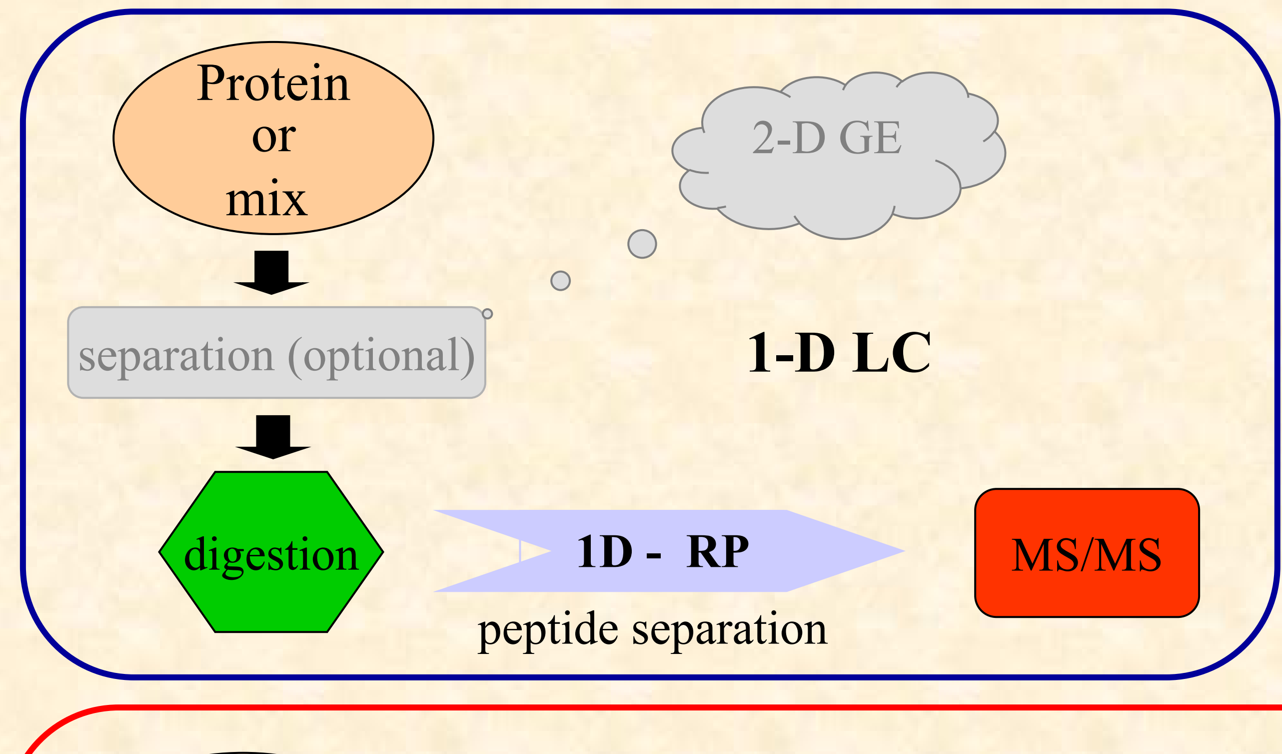
LC-MS systems
On-line

LC-MS system



ESI-IT mass spectrometer HCT Ultra (Bruker) connected on-line to capillary liquid chromatograph Ultimate (LC Packings)

Basic LC-MS/MS separation schemes



Protein mix

2-D LC
MudPit

digestion

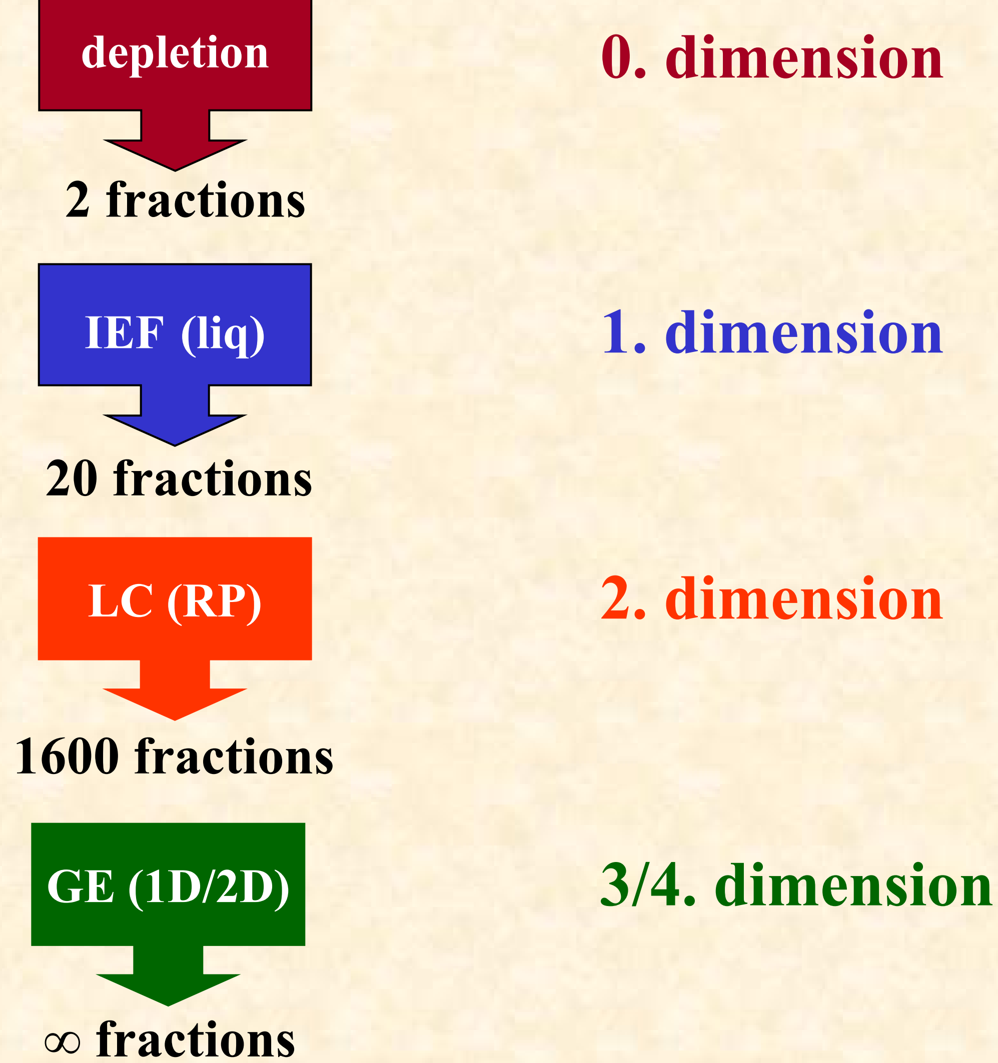
1D - SCX

step by step
fractionation

2D - RP

MS/MS

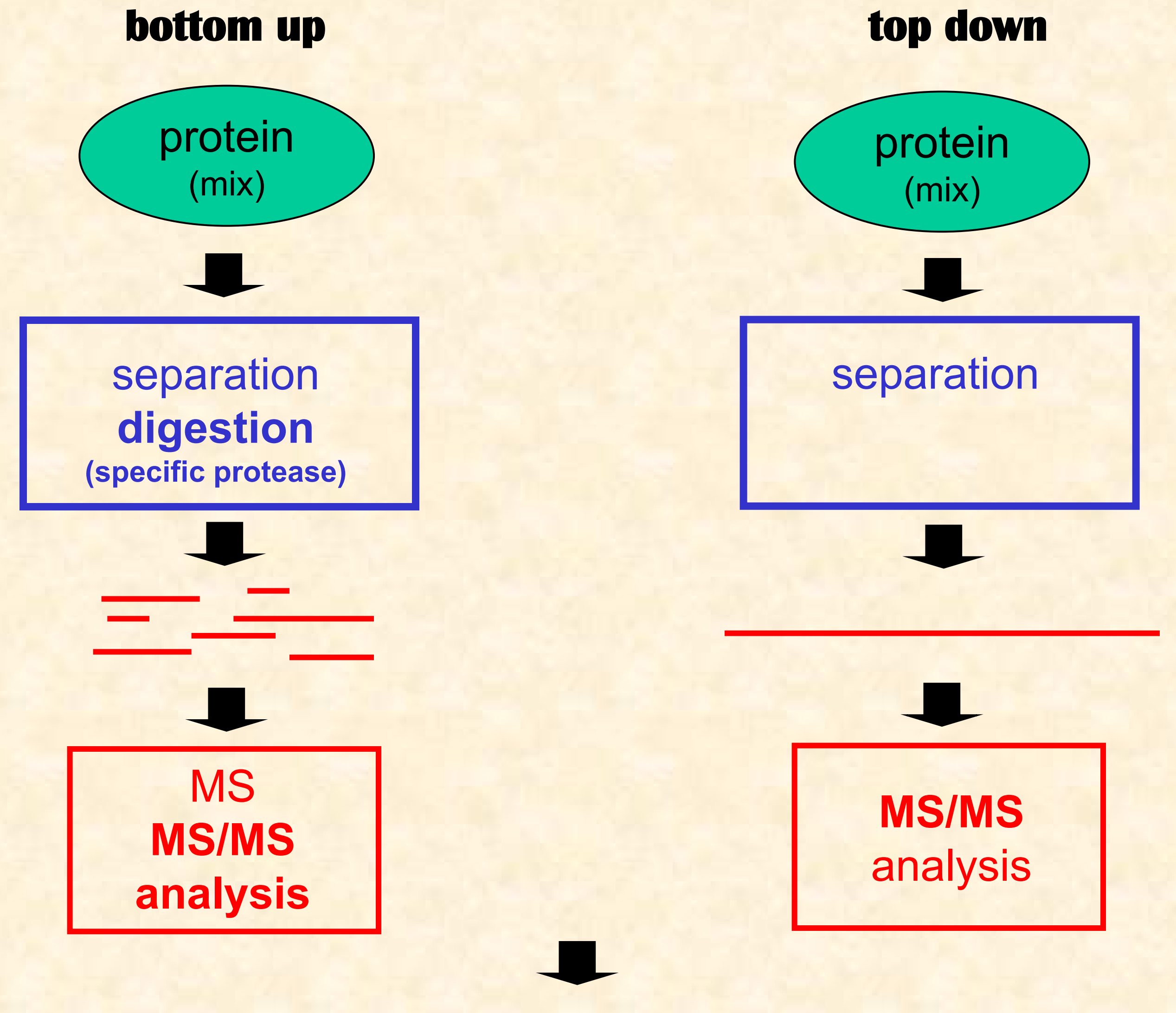
Blood plasma
(3500 – 9000 proteins ??)
example of multidimensional separation



from H. Wang, Molecular & Cellular Proteomics, 2005, ***4****, 618–625.*

Protein Identification using Mass Spectrometry

Protein identification using mass spectrometric data



Identification (DB search, *de novo*)

Protein identification using MS
bottom up
proteins with known primary sequence

Common approaches:

peptide mapping

protein separation → **specific digestion** → **MS peptides** → **comparison of obtained peptide map with sequence database**

gel electrophoresis
liquid chromatography
isoelectric focusation

MALDI-TOF MS

MS/MS Ion Search

specific digestion → **Separation of digested peptides** → **MS/MS peptides** → **comparison of fragmentation maps of individual peptides with sequence database**

liquid chromatography
capillary electrophoresis

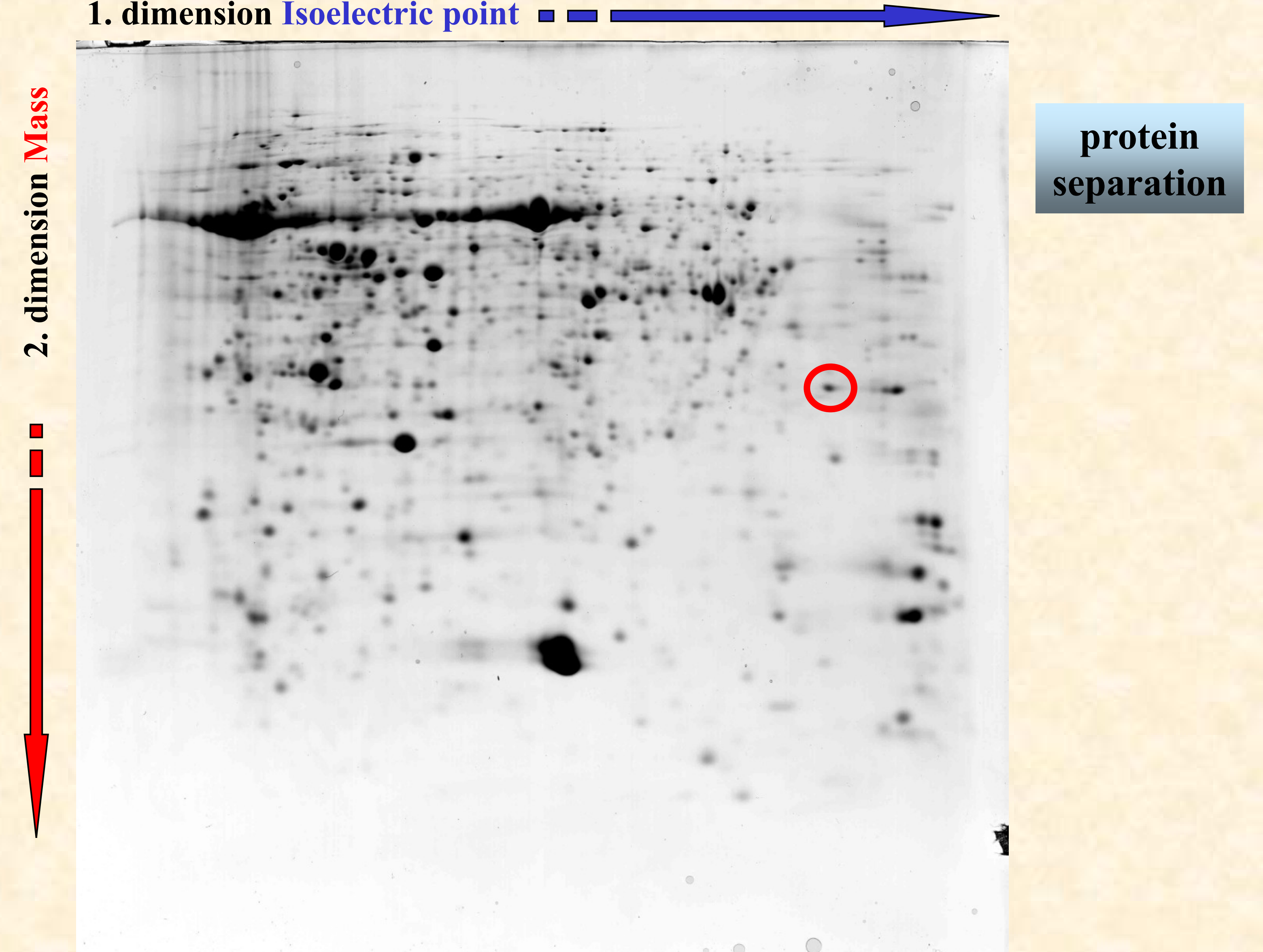
ESI-MS

Protein identification by peptide mapping
(peptide mass fingerprinting)

Blue approach

CG980

Separation of protein mixture by two-dimensional gel electrophoresis



Digestion

enzymatic digestion of protein results in set of peptides
specific protease is preferred

Trypsin
cleaves after **lysine (K)** and **arginine (R)**, if proline does not follow

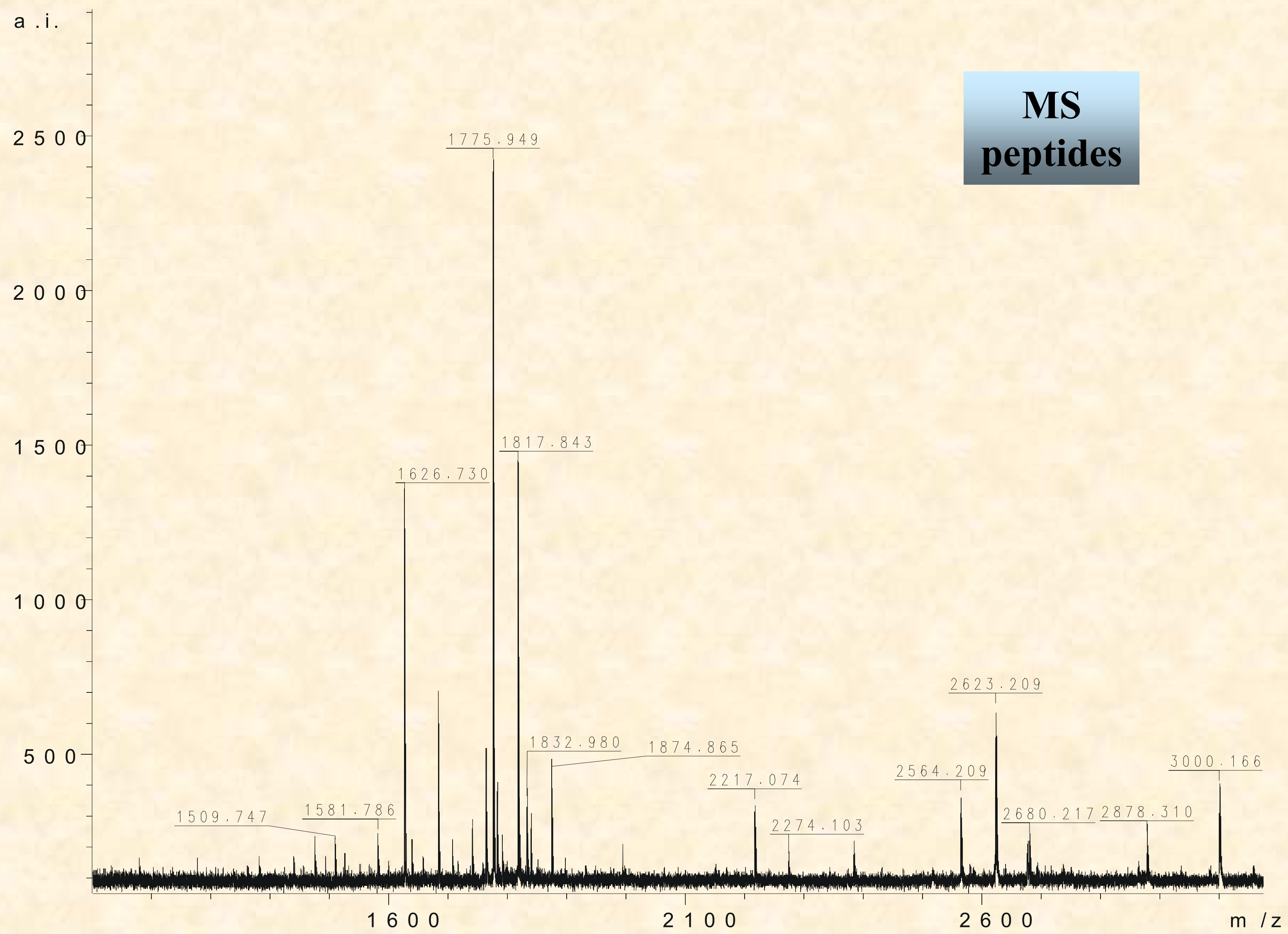
QNGVQMLSPSEIPQRDWFPSDFTFGAATSAYQIEGAWNEDGKGESNWDHFCHNHPERILD
GSNSDIGANSYHMYKTDVRLLKEMGMDAYRFSISWPRILPKGTKEGGINPDGIKYYRNLI
NLLLENGIEP

QNGVQMLSPSEIPQR	1-15	1683.848 Da
DWFPSDFTFGAATSAYQIEGAWNEDGK	16-42	3010.317 Da
GESNWDHFCHNHPER	43-57	1864.757 Da
ILDGSNSDIGANSYHMYK	58-75	1984.907 Da
TDVR	76-79	490.262 Da
...		

Specific digestion

Set of masses of these formed peptides (i.e. peptide map) is characteristic for given protein similarly as fingerprint for human individual.

MALDI - TOF MS spectrum of peptides after protein digestion



MS spectrum contains masses of peptides formed by digestion of selected protein

Protein identification– peptide mapping database searching

Measured peptide map (set of masses (or m/z) of peptides formed by digestion of analysed protein) is searched against database of protein sequences using database search engines.

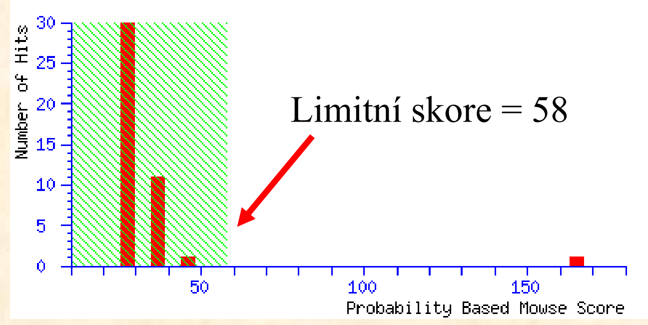
Database search engine calculates theoretical peptide map for each protein sequence in database (applying cleavage rules for selected protease) and stepwise compares experimentally obtained peptide map of our analysed protein with *in-silico* calculated peptide maps.

Searching results in a list of proteins with most similar peptide maps. Similarity extent is given by score, all protein candidates with score value higher than the limit significant value (calculated by software) are considered as identified by search engine.

comparison of obtained peptide map with sequence database

Result of database searching
peptide mapping

Mascot Search Results



Database : MSDB 20021127 (1019653 sequences)
Timestamp : 26 Jan 2003 at 10:36:50 GMT
Top Score : 165 for S18600, glutamate-ammonia ligase ....

Score

1. S18600 **Mass:** 47780 **Total score:** **165** **Peptides matched:** 12
   glutamate-ammonia ligase (EC 6.3.1.2) precursor, chloroplast (clone lambdaAtgsl1) - Arabidopsis thaliana
2. S32228 **Mass:** 47714 **Total score:** **76** **Peptides matched:** 7
   glutamate-ammonia ligase (EC 6.3.1.2) precursor - rape - Brassica napus

Sequence Coverage: **44%**

comparison of obtained peptide map with sequence database

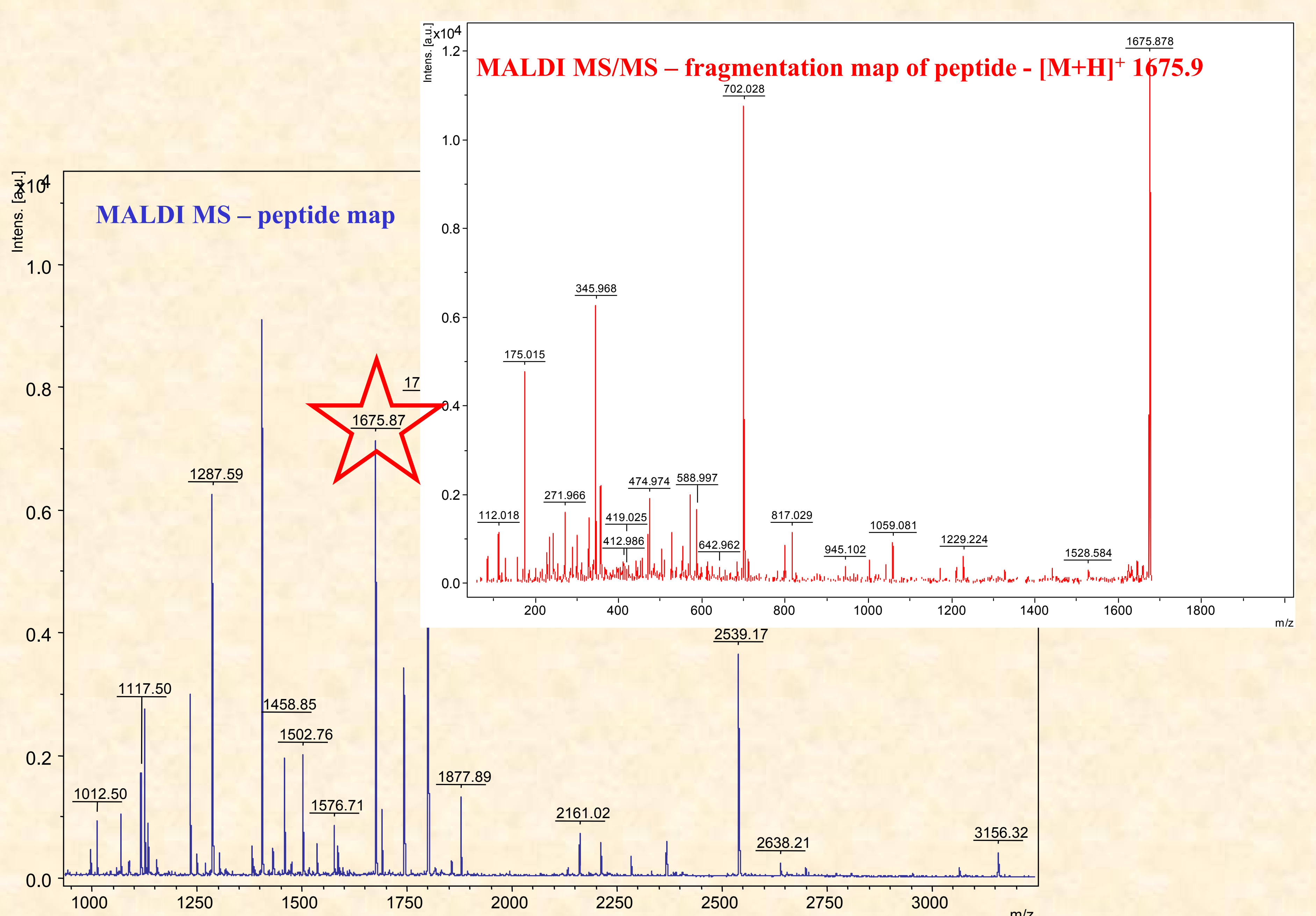
```
1    MAQILAASPT CQMRVPKHSS VIASSSKLWS SVVLKQKKQS NNKVRGF
51   ALQSDNSTVN RVETLLNLDT KPYSDRIIAE YIWIGGSGID LRSKSRT
101  PVEDPSELPK WNYDGSSTGQ APGEDSEVIL YPQAIFRDPF RGGNNILVIC
151  DTWTPAGEPI PTNKRAKAAE IFSNKKVSGE VPWFGIEQEY TLLQQNVKWP
201  LGWPVGAFPG PQGPYYCGVG ADKIWGRDIS DAHYKACLYA GINISGTNGE
251  VMPGQWEFQV GPSVGIDAGD HVWCARYLLE RITEQAGVVL TLDPKPIEGD
301  WNGAGCHTNY STKSMREEGG FEVIKKAILN LSLRHKEHIS AYGEGNERRL
351  TGKHETASID QFSWGVANRG CSIRVGRDTE AKGKGYLEDR RPASNMDPYI
401  VTSLLAETTL LWEPTLEAEA LAAQKLSLNV
```

www.matrixscience.com

Sequence regions in red corresponds to assigned peptides from measured peptide map

Protein identification based on MS data *vs* MS/MS data

CG980



Protein identification by LC-MS/MS

Green approach

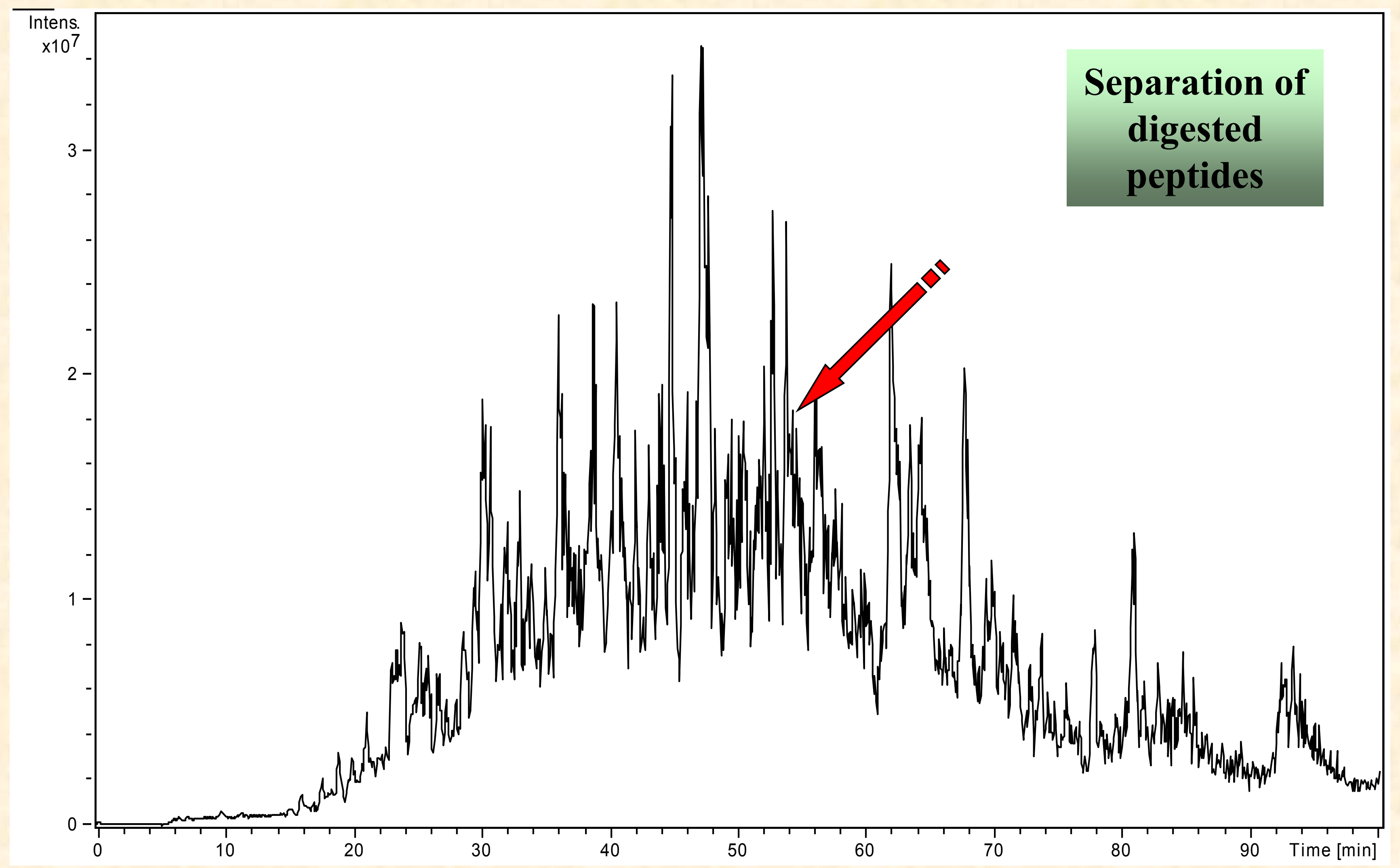
Protein digestion

specific digestion

In difference of „blue approach“ this time whole complex protein mixture is digested altogether, again using specific protease (usually trypsin).

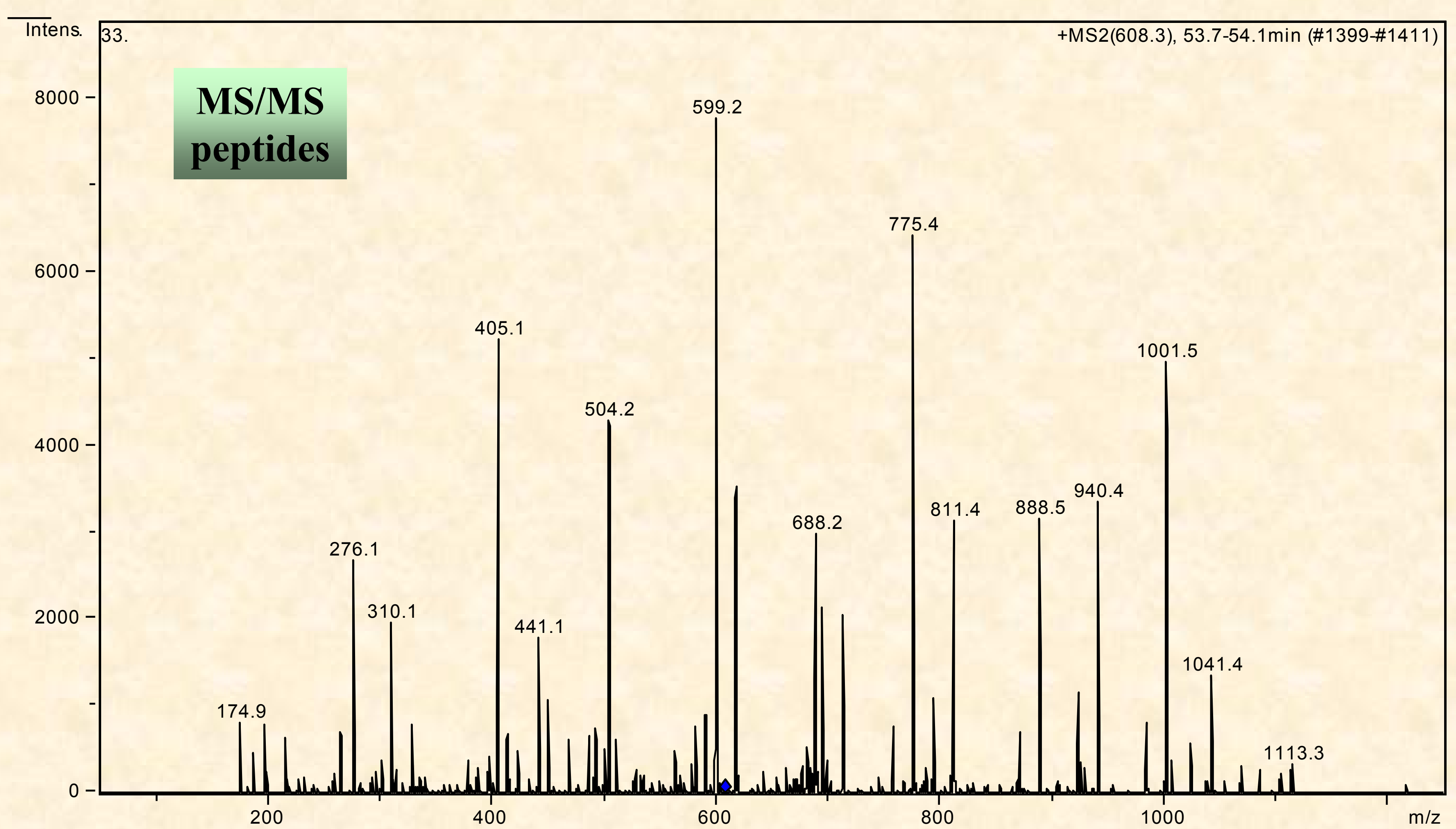
This peptide mixture is separated (frequently multidimensionally depending on sample complexity) and subjected to MS/MS analysis.

Separation of tryptic peptides formed by digestion of protein mixture



Digest of human blood plasma sample separated by liquid chromatography (1D separation) connected to mass spectrometer (LC-MS/MS)

MS/MS spectrum of tryptic peptide (m/z 608.3, 2+)



MS/MS spectrum contains the peptide fragments formed by collision induced dissociation in ion trap. These fragments carry specific information about peptide sequence and allow identification.

Protein identification based on – MS/MS data database searching

Measured fragmentation maps (i.e. sets of masses (or m/z) of fragments formed during MS/MS of individual peptides) are searched against database of protein sequences by search engine.

At first, database search engine prepares theoretical peptide map for a protein sequence in database, subsequently, it calculates theoretical fragmentation map for each peptide of corresponding peptide map (according to given fragmentation rules) and then these *in-silico* prepared fragmentation maps are compared with our experimentally obtained fragmentation maps of analyzed peptides. The engine performs this operation for each protein sequence in database.

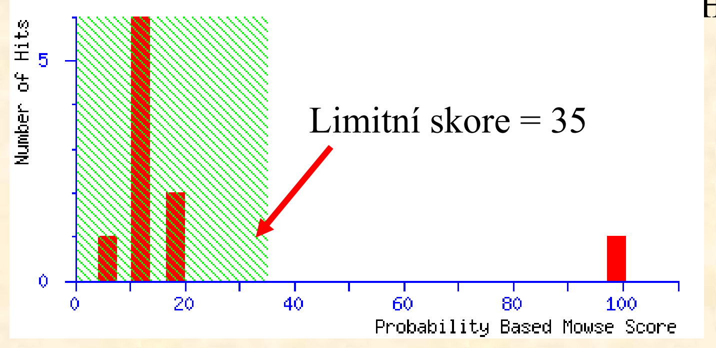
Software calculates individual score for each peptide, score value higher than limit score determines signifikant similarity between theoretical and measured fragmentation map – significant peptide identification.
In final, search engine assort peptides to corresponding protein sequences (the more peptides with significant score per protein – the more reliable protein identification). The software also calculate protein score which is derived from individual peptide score as a tool for setting up results.

comparison of fragmentation maps of individual peptides with sequence database

Result of database searching
MS/MS data

Mascot Search Results

Database : SwissProt 51.2 (243975 sequences; 89639744 residues)
Taxonomy : Homo sapiens (human) (15175 sequences)
Timestamp : 16 Dec 2006 at 16:05:59 GMT
Significant hits: AACT_HUMAN Alpha-1-antichymotrypsin precursor (ACT) – Homo sapiens



comparison of fragmentation maps of individual peptides with sequence database

Protein score

Peptide Summary Report

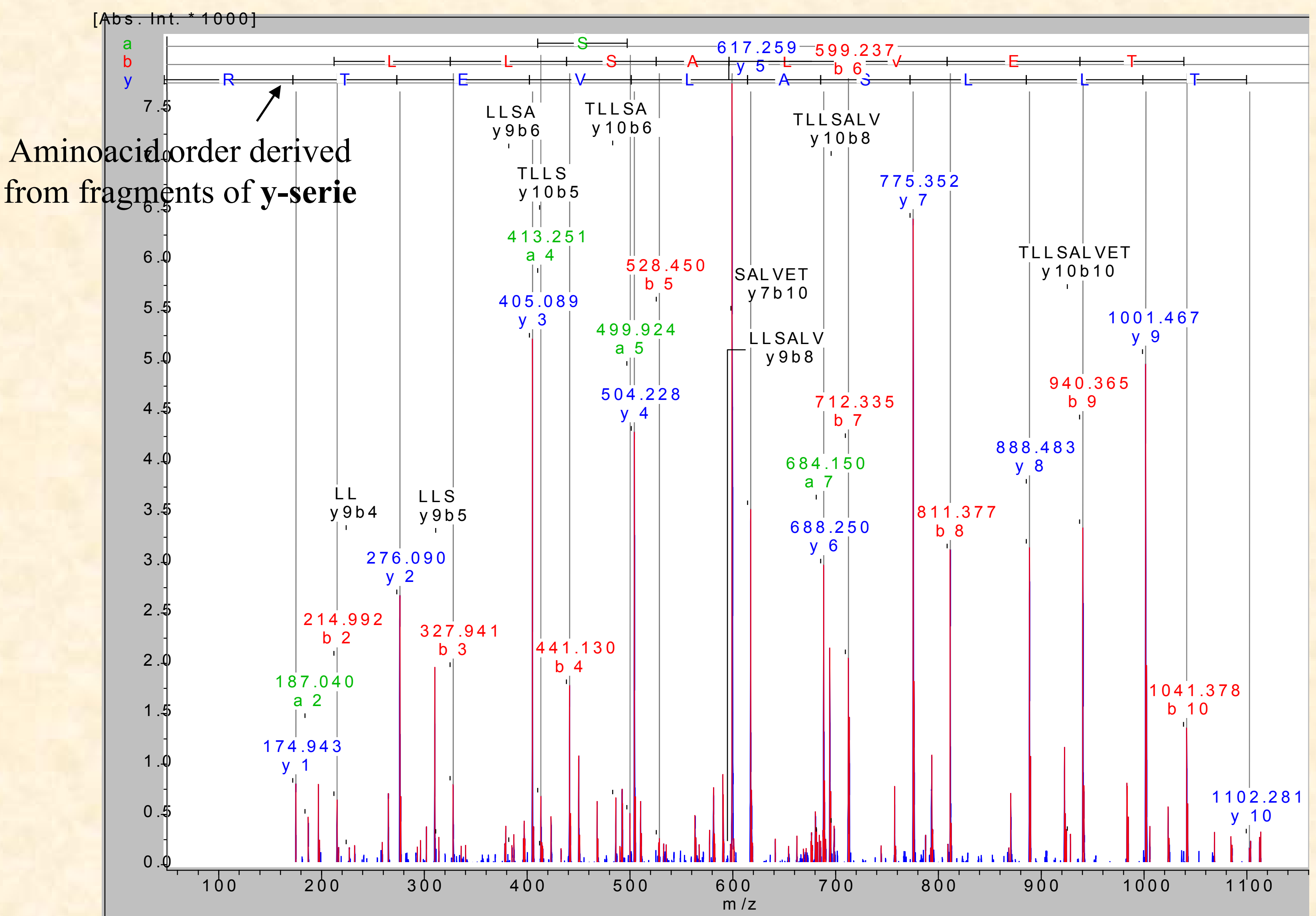
1. AACT_HUMAN Mass: 47621 Score: 99 Queries matched: 1
Alpha-1-antichymotrypsin precursor (ACT) Homo sapiens (Human)

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
1	608.3000	1214.5854	1214.7234	-0.1380	0	99	2.2e-08	1	K.ITLLSALVETR.T

Individual peptide score

Thanks to variability of primary protein structure it is possible to determine identity of protein based on fragments (MS/MS spectrum) of a single peptide.

Processed MS/MS spectrum



Differences in *m/z* (resp. masses) of neighbouring fragment ions of corresponding serie (b, y) enables to determine individual aminoacids and their place in sequence.

Protein identification – database searching

comparison of fragmentation maps of individual peptides with sequence database

MS

m/z

database searching

+

protein is identified
(protein with its sequence in database)

-

unknown protein
(sequence is not available, *de novo sequencing*)

other reasons of unsuccessful identification:
low protein concentration, unspecific digestion, unknown modification, low quality of MS data, ...

Characterization
of Posttranslational Modifications

Characterization of protein modifications

- **mutations** (protein isoforms)
- **chemical** (oxidation, deamidation, etc.)
- **posttranslational** (e.g. phosphorylations, glycosylations)

MS in analysis of protein modifications

- **modification type**
- **localozation**
- **site occupancy**

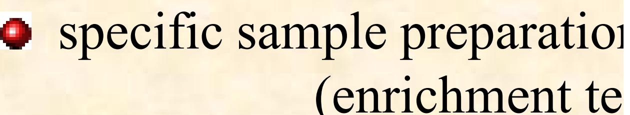
List of modifications and tools:

DeltaMass - https://abrf.org/delta-mass
ExPASy - http://www.expasy.org/proteomics/post-translational_modification

Difficulties in PTMs analysis

- low abundance of modified proteins
- protein occurs frequently in several modification forms
- protein modification status can change during sample preparation
- signal suppresion of modified peptides in MS
  (preferential ionization of unmodified peptides)

To improve success of PTMs analysis



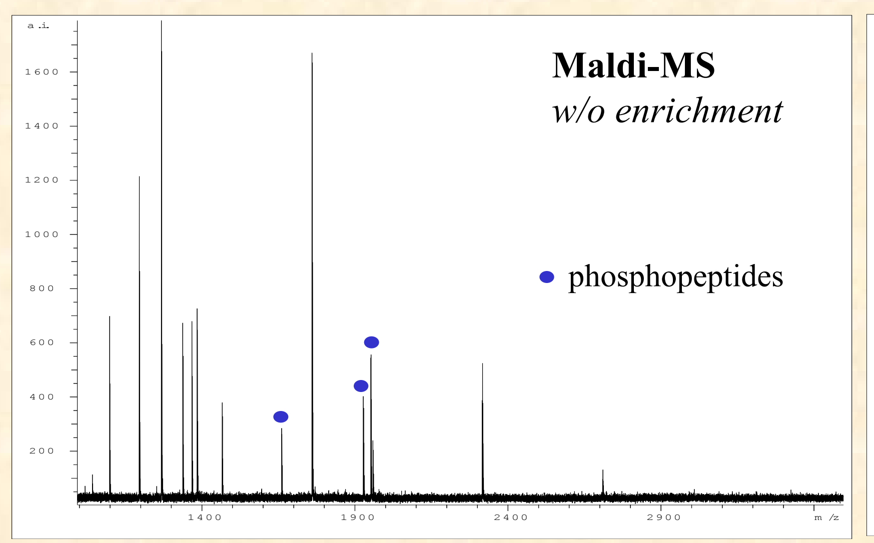
- specific sample preparatio
  (enrichment te treatment etc.)
- specific MS/MS operation modes

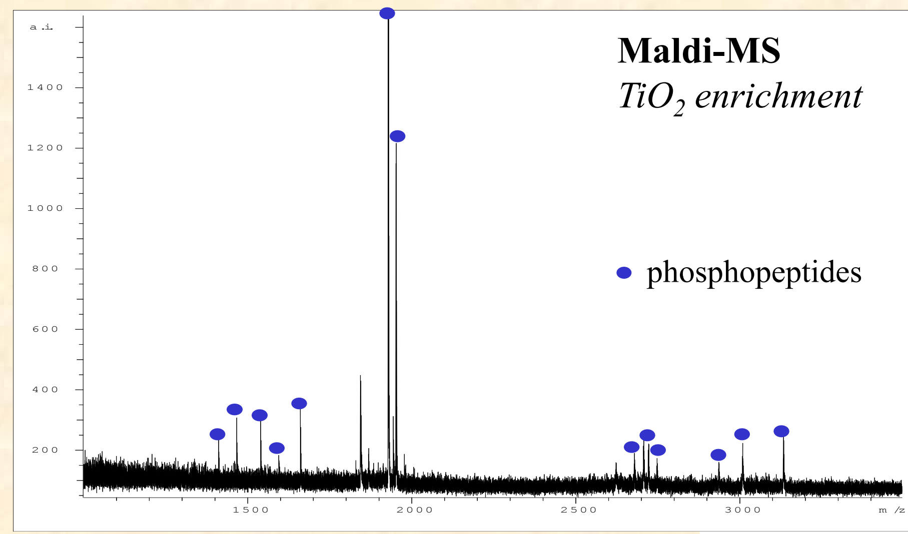
Phosphorylations

sample treatment

- phosphatase inhibitors, denaturation (as soon as possible)
- enrichment of phosphopeptides (proteins)
  - **TiO_2** *(MOAC – „metal oxide affinity chromatography")*
  - **IMAC** („*immobilized metal affinity chromatography*")
  - **SCX** resp. **SAX** or **HILIC** *(„ion exchange or hydrophilic interaction chromatography")*
  - ***immunoprecipitation*** *pomocí specifické protilátky*

I.L. Batalha, Trends in Biotechnology 30 (2), 100-110 (2012)





Phosphorylations

MS analysis

dedicated MS/MS fragmentation techniques preserving phosphogroup at aminoacid residue

CID (limited)

ETD (ECD)

electron transfer (capture) dissociation

HCD

higher-energy collision dissociation

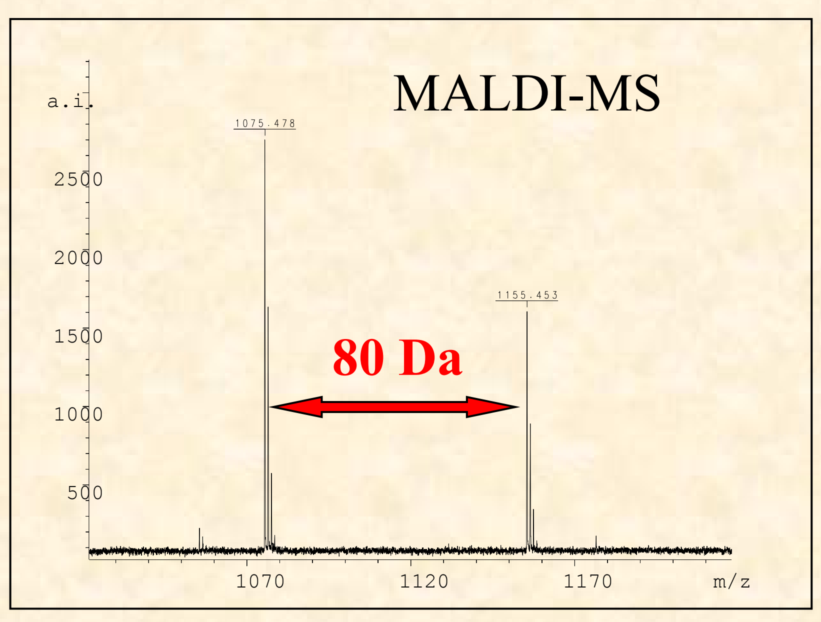
EThcD

electron-transfer/higher-energy collision dissociation

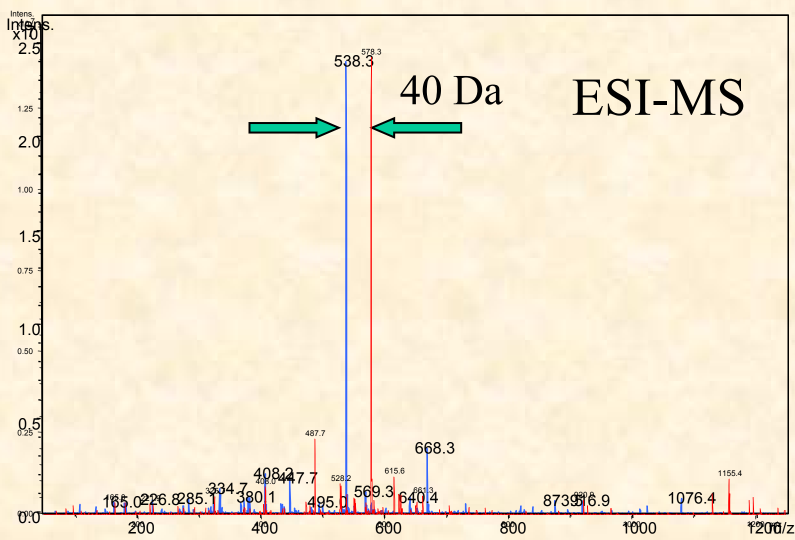
Frese at al., J. Proteome Res., 12, 1520−1525 (2013)

Phosphorylations

how MS see modifications



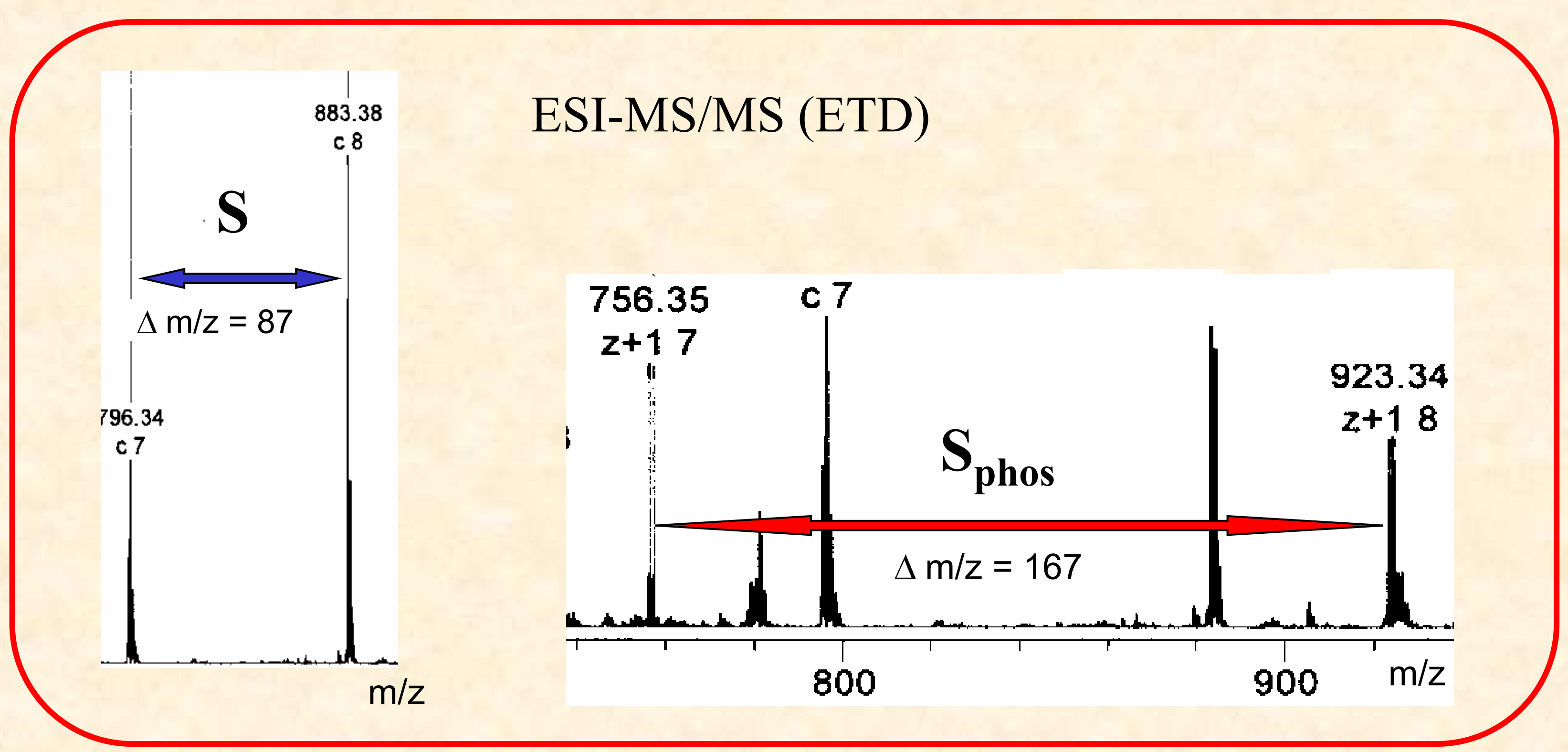
shift in peptide mass in **MS spectrum** corresponding to modification mass **indicates presence** of given type of PTM



Phosphorylations

how MS see modifications

shift in fragment mass in **MS/MS spectrum** corresponding to modification mass **indentifies and localizes** given type of PTM



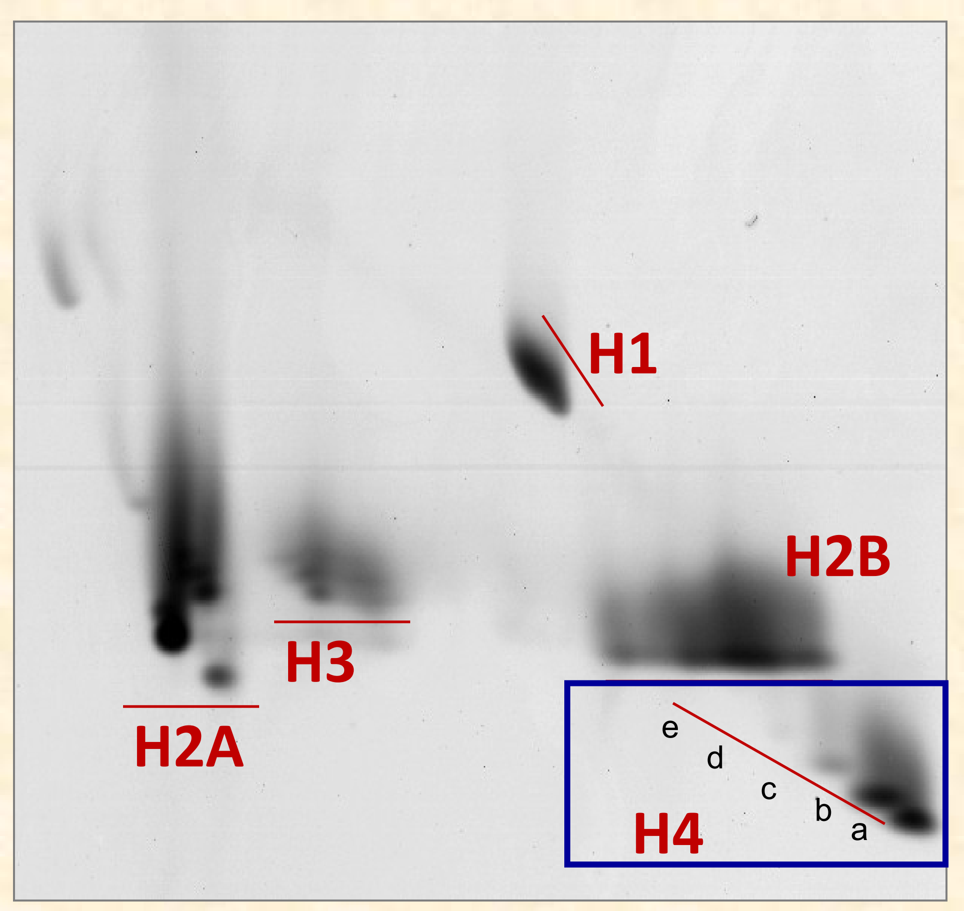
CG980

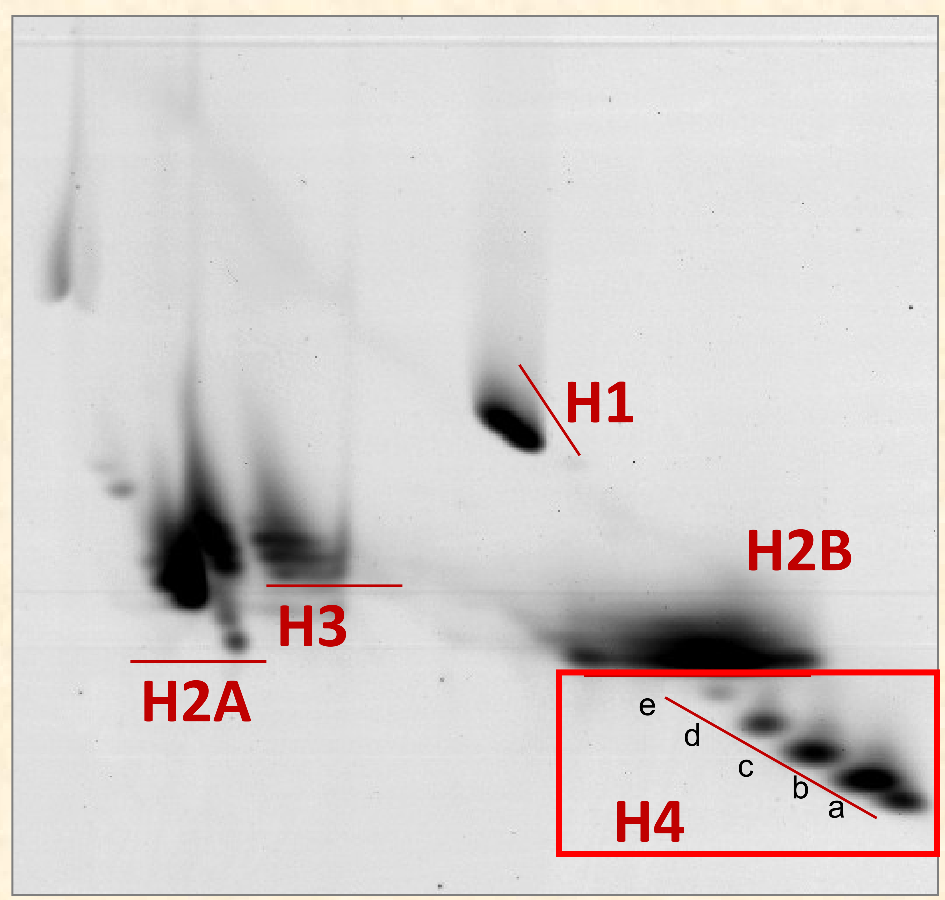
Histone acetylations

2-D gel electrophoresis (AUT-AU) histone extracts

w/o deacetyase inhibitors

with deacetylase inhibitor





Histone acetylations

LC-MS/MS analysis results

w/o deacetyase inhibitors

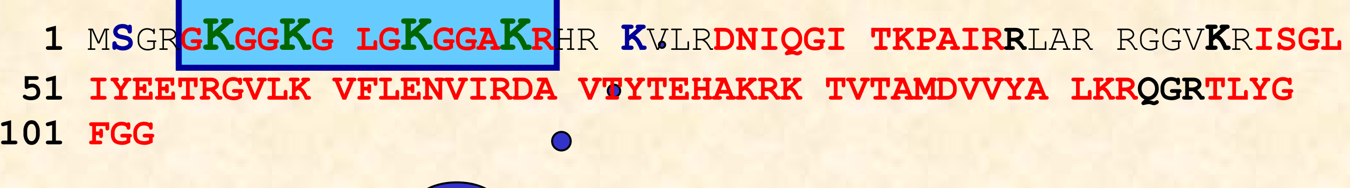
with deacetylase inhibitor

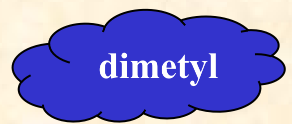
1 x ac GKGGKGLGKGGAKR
2 x ac GKGGKGLGKGGAKR
3 x ac GKGGKGLGKGGAKR
3 x ac GKGGKGLGKGGAKR
4 x ac GKGGKGLGKGGAKR

D C B A

E D C B A

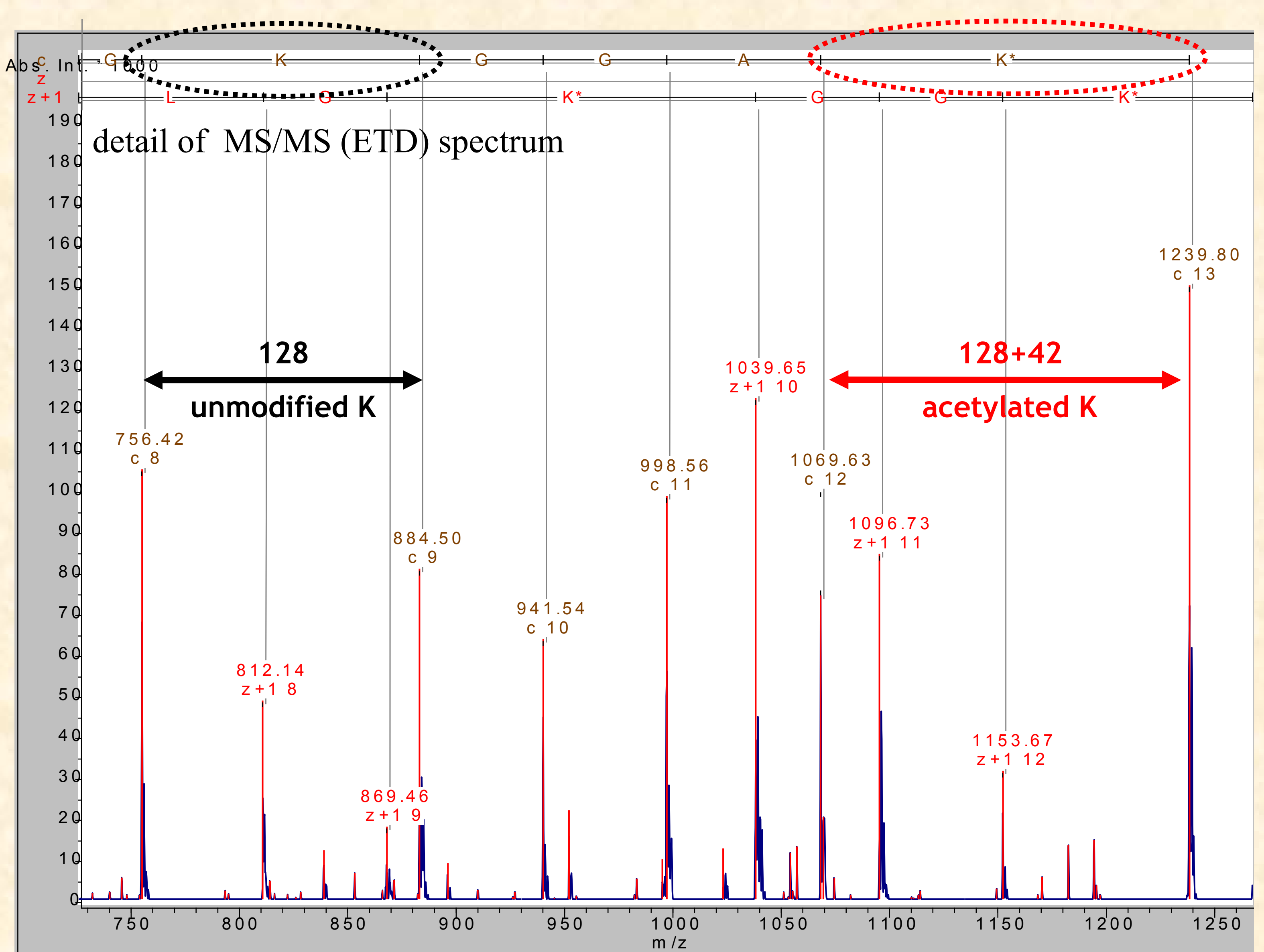
4x Ac
3x Ac
2x Ac
1x Ac
0





Histone acetylations

distinquishing modification site in peptide GKGGKG LGKGGAKR (3x Ac)



Protein quantification

Protein quantification by MS

general approaches

methods based on application of isotopic labels

- absolute quantification

  (determination of amount/concentration of given protein using additon of internal standard with known concentration)

- relative quantification

  (comparison of changes in protein levels between two or more samples)

Isotopic labels are introduced to proteins at different stages of experiment: during cell cultivation or by chemical reaction after protein isolation or after digestion (peptides).

ICAT reagents
Heavy reagent: d8-ICAT (X = deuterium)
Light reagent: d0-ICAT (X = hydrogen)
Biotin
Linker (heavy or light)
Thiol-specific reactive group

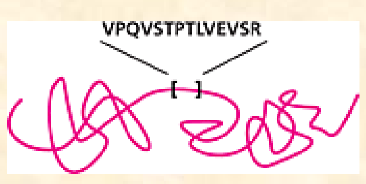
label free methods based on advanced processing of MS (MS/MS) data

Protein quantification by MS

absolute quantification

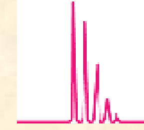
- **AQUA Peptide Selection**

Select an optimal tryptic peptide and stable isotope amino acid from the sequence of your protein of interest



- **Order selected peptide labeled (^{15}N, ^{13}C)**

Optimize LC-MS/MS separation protocol for quantitation

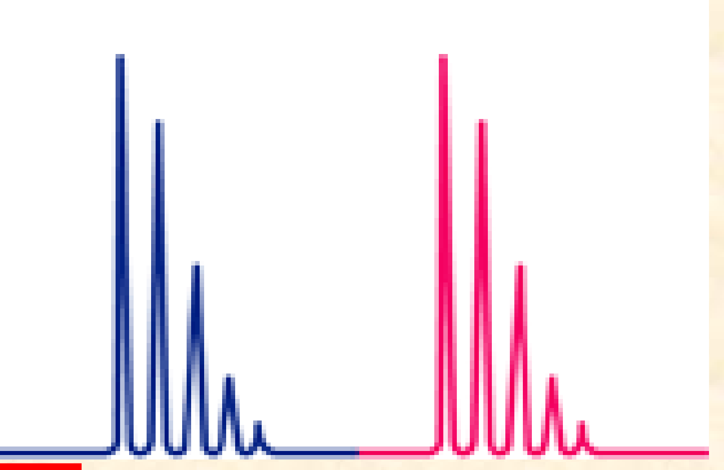


AQUA

- **Adding labeled peptide to protein mix**

- **Digest**

- **Analyze by LC-MS/MS to quantitate protein of interest**

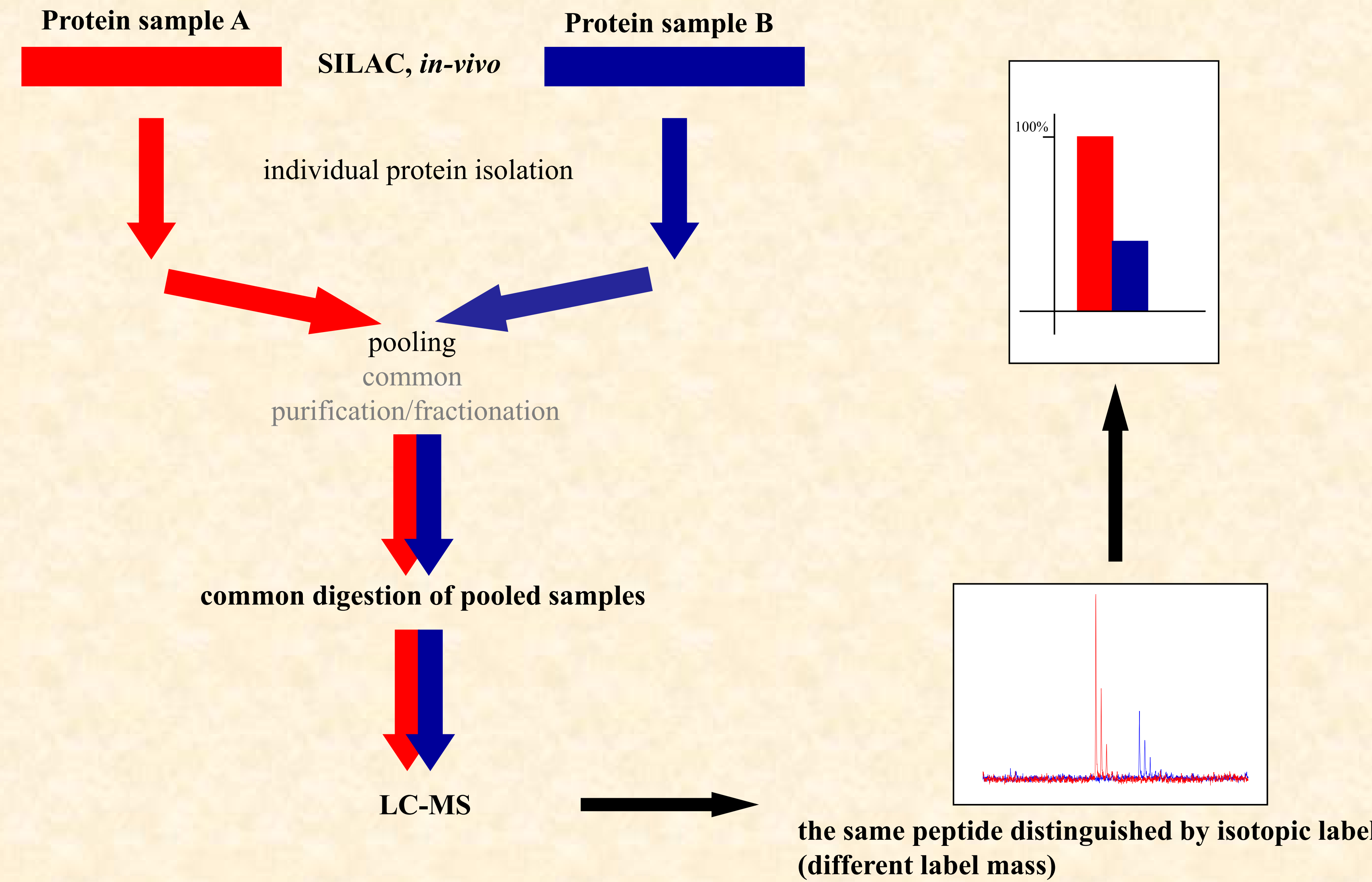


only for protein(s) selected in advance

Protein quantification by MS

relative quantification

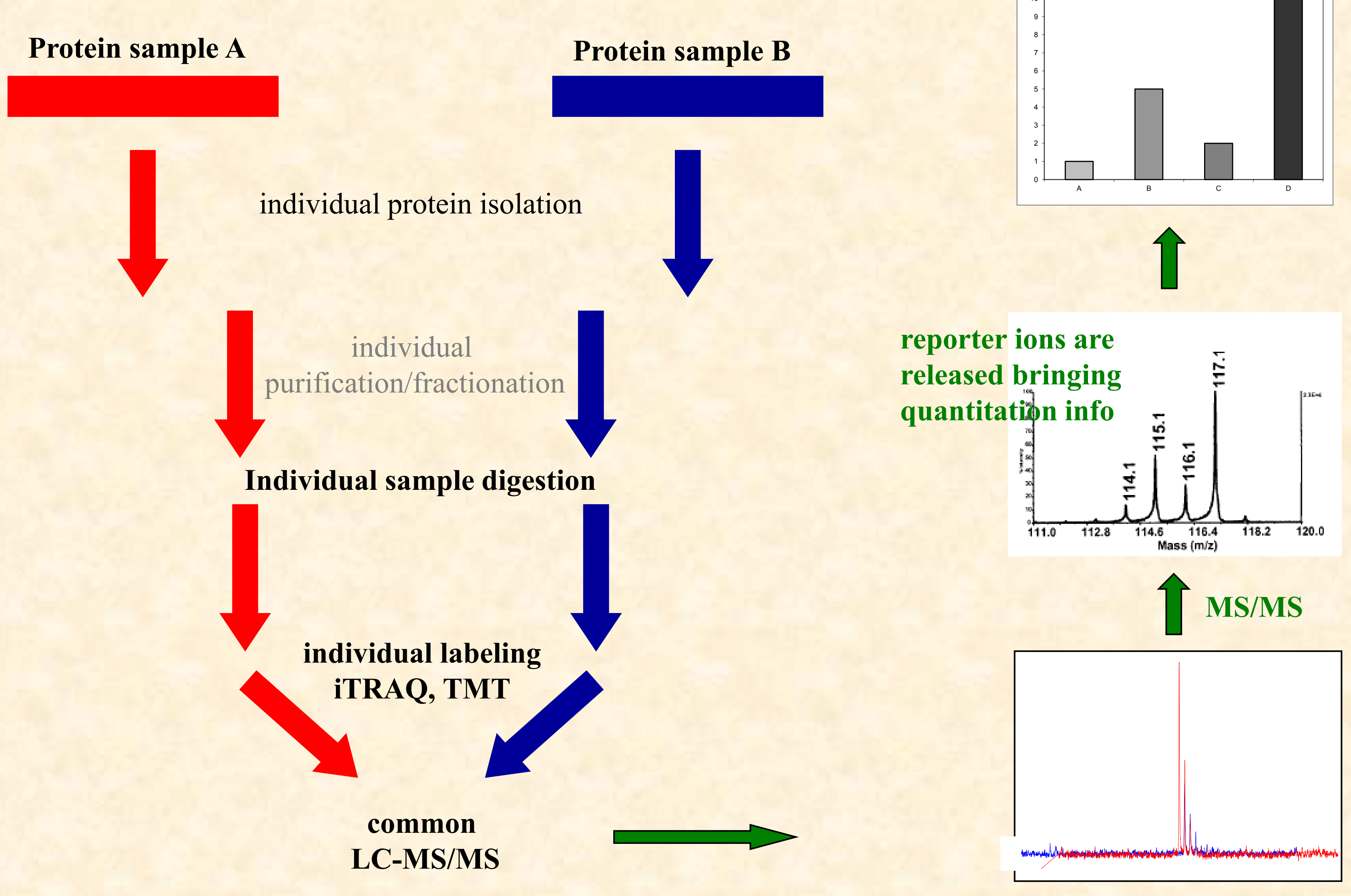
CG980



Protein quantification by MS

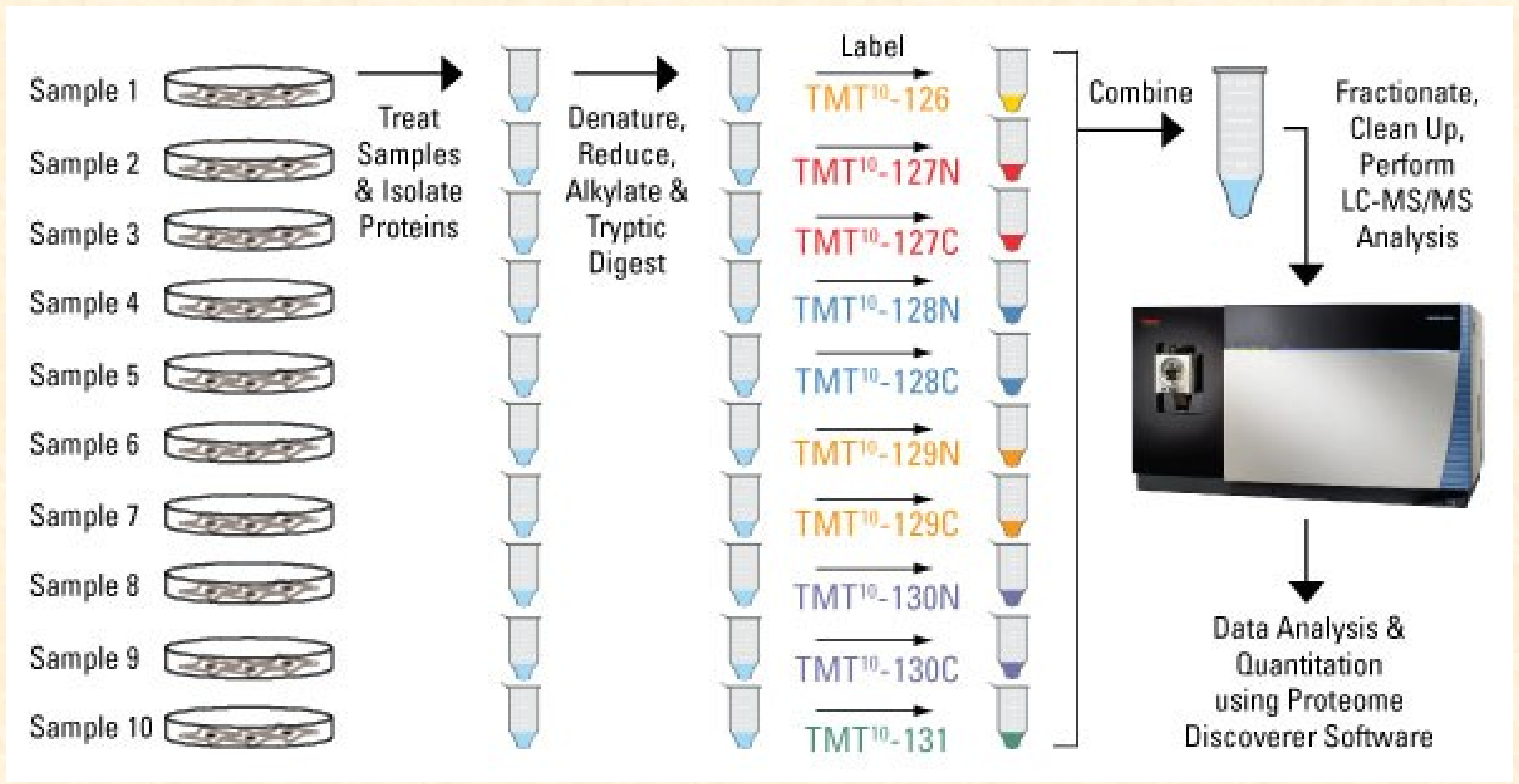
CG980

relative quantification



Protein quantification by MS

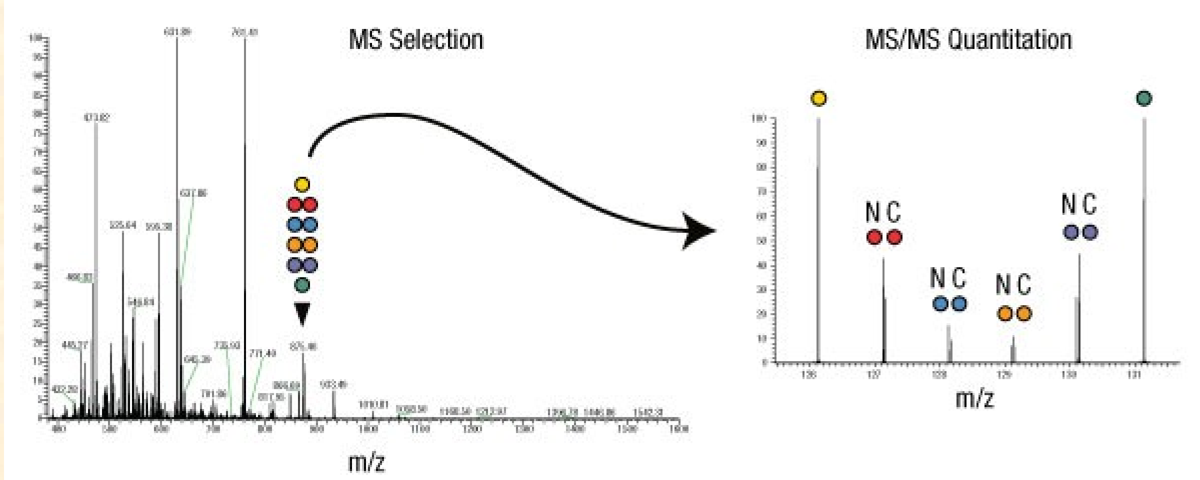
relative quantification – TMT labeling

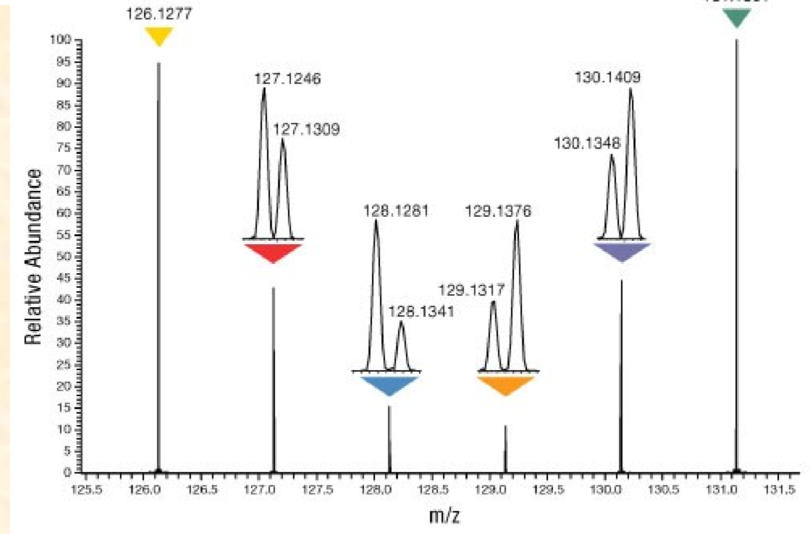


Thermo Fischer Scientific

Protein quantification by MS

relative quantification – TMT labeling





Reagents contain different numbers and combinations of 13C and 15N isotopes in the mass reporter. The different isotopes result in a 10-plex set of tags that have mass differences in the reporter that can be detected using **high resolution** Orbitrap MS instruments.

Thermo Fischer Scientific

… and this is the end