



Laboratoř funkční genomiky a proteomiky

Národní centrum pro výzkum biomolekul

Přírodovědecká fakulta MU

CEITEC-MU



CEITEC

HMOTNOSTNÍ SPEKTROMETRIE V PROTEOMICE

Zbyněk Zdráhal

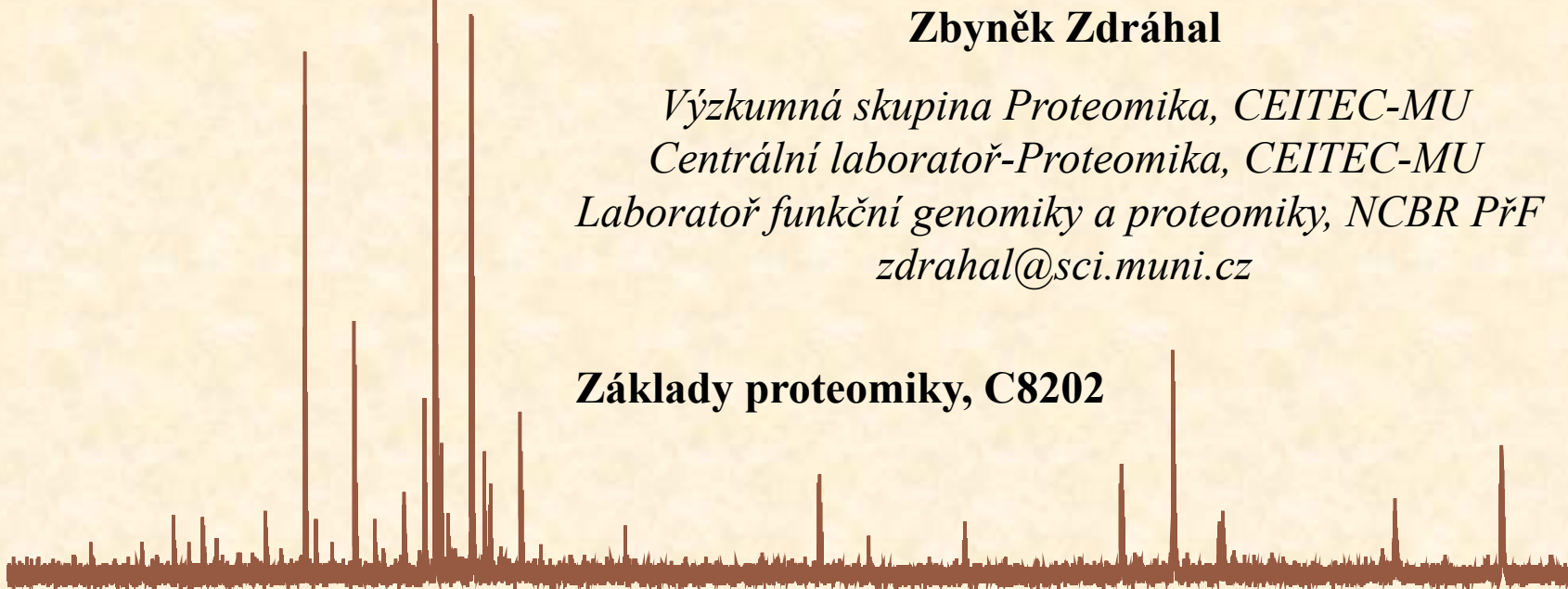
Výzkumná skupina Proteomika, CEITEC-MU

Centrální laboratoř-Proteomika, CEITEC-MU

Laboratoř funkční genomiky a proteomiky, NCBR PŘF

zdrahal@sci.muni.cz

Základy proteomiky, C8202



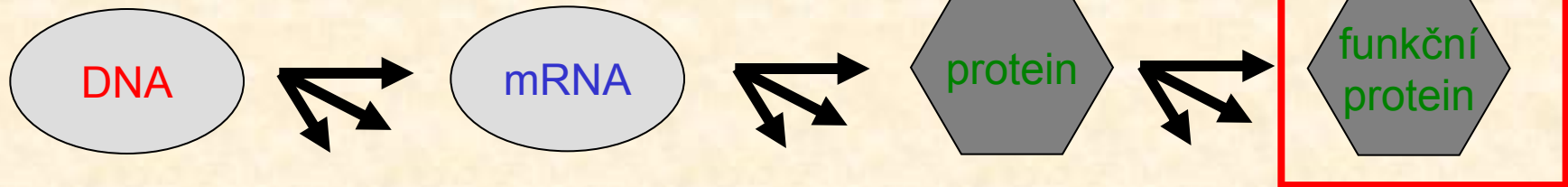
Úvod



Proteomika – disciplína zabývající se analýzou proteomu

co se může stát

co se opravdu děje



transkripce
 posttranskripční úpravy
 (alternativní sestřih aj.)

translace

Posttranslační úpravy
 posttranslační modifikace
 (fosforylace, glykosylace aj.)

Proteinové komplexy

- z každého genu může vzniknout několik proteinů, resp. jejich forem, které nelze indikovat na základě analýzy DNA resp. mRNA
- neexistuje přímá korelace mezi obsahem mRNA a výsledným obsahem proteinů



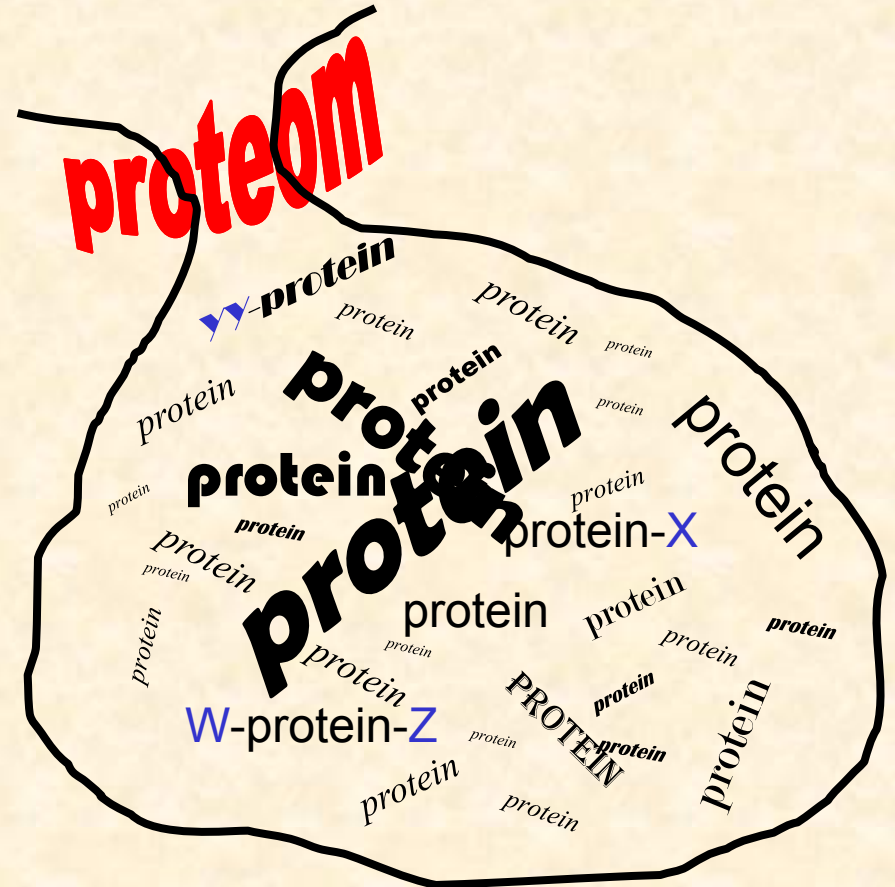
The Desperate Man, Gustave Courbet, 1844-45



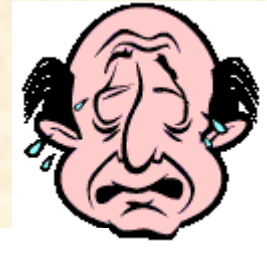
X

Analýza proteomu

charakterizace všech proteinů
včetně všech jejich forem v buňce
(tkáni, organismu) v daném čase
za daných podmínek



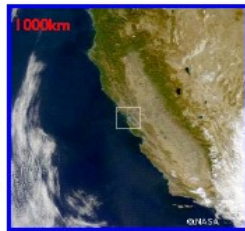
Úskalí analýzy proteomu



10¹⁰ Really Is Wide Dynamic Range (Here on a linear scale)



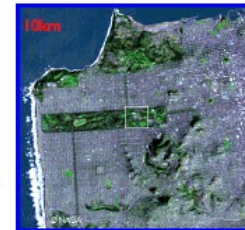
10



9



8



7



6



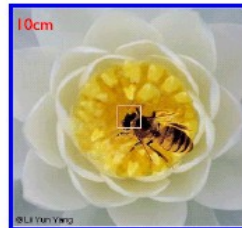
5



4



3



2



1



[esc.html](#))

žiku

otlivých

Komplexita proteomu

Estimates from deep proteomics and transcript profiling suggest that [redacted]

[redacted] (i.e., about 10,000 of the 20,000

human genes)²⁸. Assuming [redacted]

and allowing for detection [redacted]

of a single proteoform, one then multiplies the [redacted]

[redacted]

estimate based on trends in data [redacted]

to be [redacted]

diversity are needed, and an [redacted]

estimate of the number of human genes using [redacted]

in the year 2000 (ref. 30).

Box 1 | Calculating the number of theoretical proteoforms

$$\# \text{ theoretical proteoforms} = \prod_{i=1}^n (\text{potential PTMs at AA}_i + 1)$$

Recent analyses of human histone H4 (P62805) [redacted]

[redacted] (see main text). Considering the [redacted]

[redacted] (acetylation, methylation and phosphorylation,

shown below) and a single SNP, [redacted]

are possible.

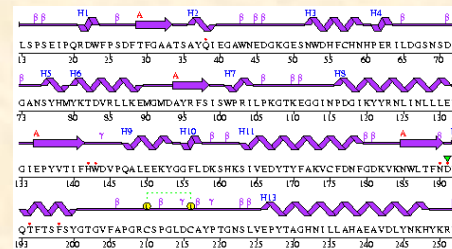
$$2^6 (\text{K5/8/12/16/31/91 ac}) \times 3^1 (\text{R3 me1/2}) \times 4^1 (\text{K20 me1/me2/me3}) \times 2^5 (\text{S1/S46/Y51/T79/Y87 ph}) \times 2 (\text{N-term ac}) \times 2 (\text{E63Q cSNP}) = 98,304 \text{ proteoforms}$$



Struktura proteinů

Primární struktura (pořadí aminokyselin):

...ALEEKYGGFLDKSHKSI VEDYTYFAKVCDFNFGDKVKNWLT FNEPQTFTSFSYGTGVFAP
GRCSPLDCAYP TGN SLVEPYTAGHNI LLAHAEAVDLYN...

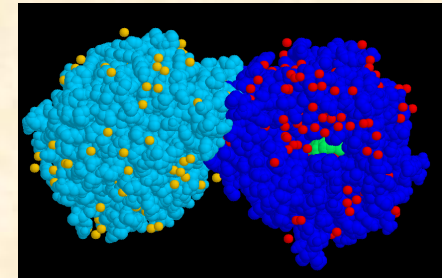


Sekundární struktura (vodíkové můstky)

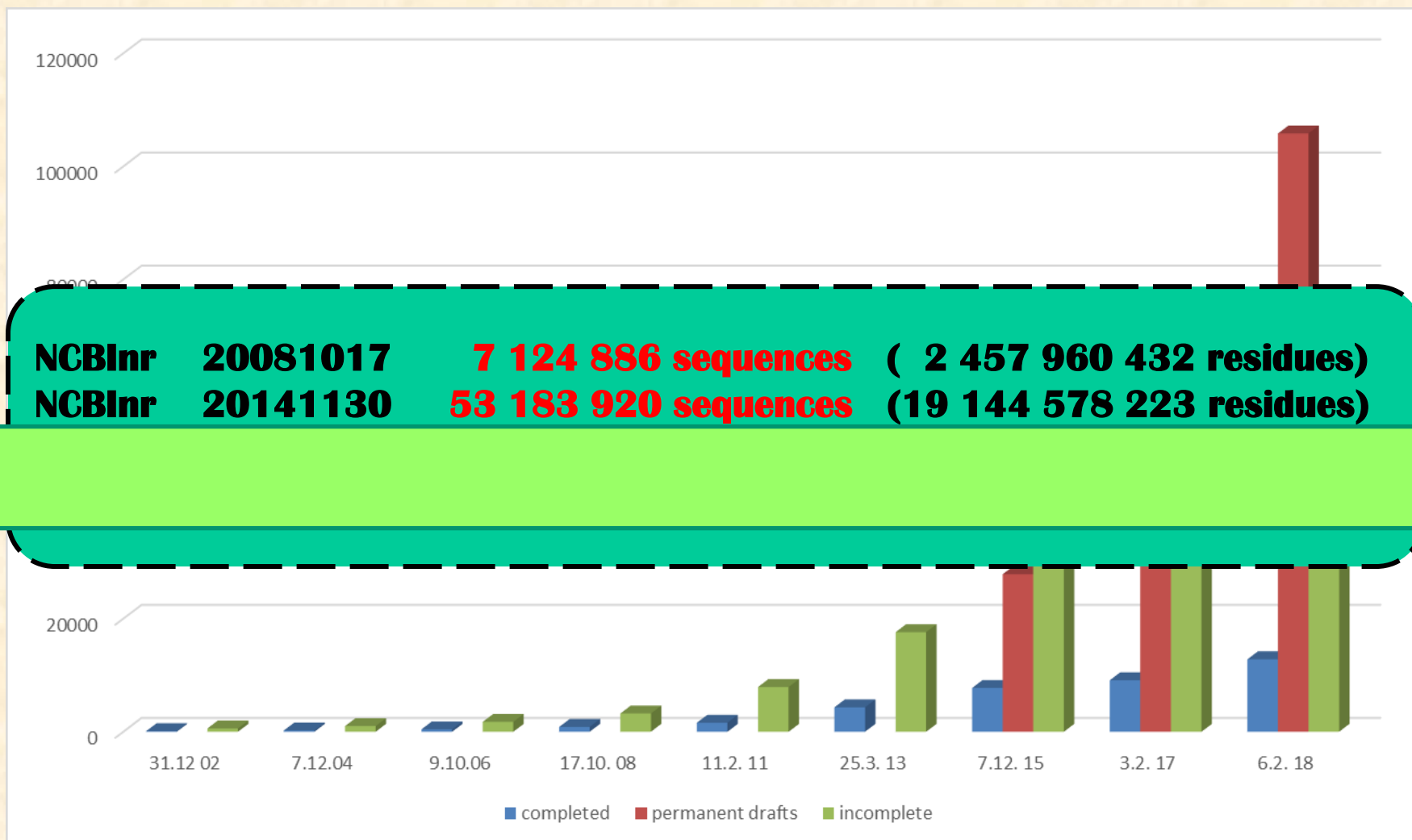
Terciární struktura (skládání)



Kvarterní struktura (asociace podjednotek)



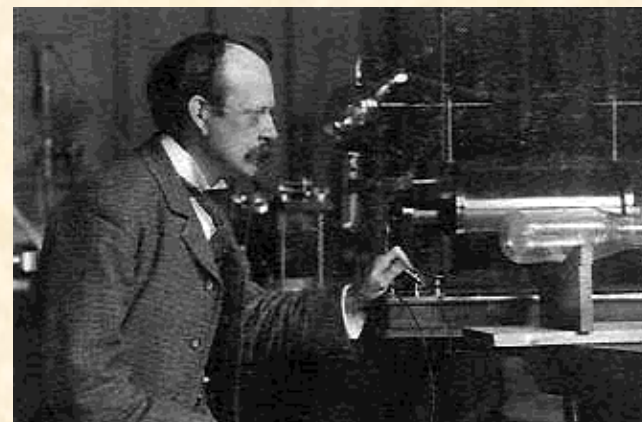
Nárůst znalosti genomů



Možnosti hmotnostní spektrometrie

Základní úkoly MS v analýze proteinů:

- **Určení hmotnosti intaktních proteinů**
- **Identifikace proteinů** (proteinové komplexy)
- **Určení druhu a místa posttranslační modifikace**
- **Kvantifikace proteinů**
- **MS imaging**
- **Studium prostorové struktury** (komplementární k NMR)



*Otec hmotnostní spektrometrie J.J. Thomson,
Nobelova cena za fyziku, 1906*

GIGO

garbage in → *garbage out*

Základy hmotnostní spektrometrie



Princip metody:

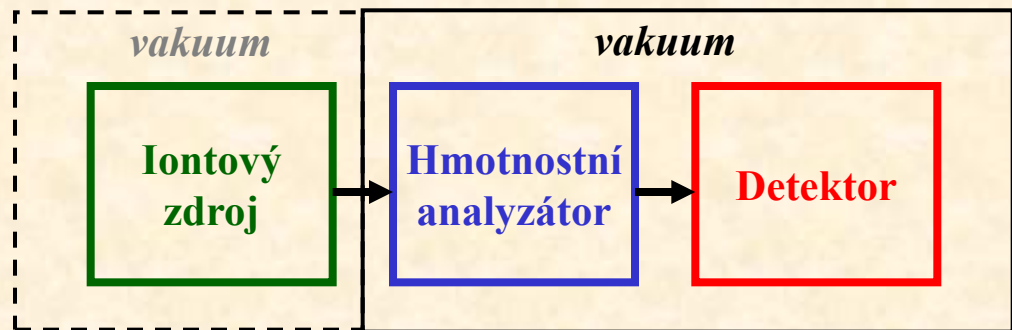
- měření poměru m/z iontů analyzované látky

m - hmotnost iontu

z - počet nábojů

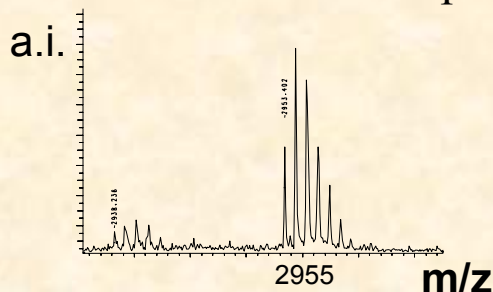
Základní operace:

- ionizace molekul analyzované látky
- separace iontů podle jejich m/z
- detekce iontů



Výsledek:

- hmotnostní spektrum - závislost intenzity iontů na jejich m/z



takto **určení hmotnosti iontů**,
v případě molekulárního iontu

hmotnost celé molekuly

Pozn. Kromě vybraných typů ionizace, všechny kroky MS analýzy probíhají ve vakuu, aby bylo zamezeno nechtěným srážkám analyzovaných iontů po cestě ze zdroje do detektoru (střední volná dráha molekul)

Průlom v MS proteinů

Nové „šetrné“ ionizační techniky (polovina 80. let 20. století)

základní předpoklad pro široké využití MS pro analýzu biomolekul
(Nobelova cena 2002)



Koichi Tanaka
Shimadzu Corp., Kyoto, Japan

MALDI

ionizace laserovou desorpcí za účasti matrice

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

ionizace elektrosprejem

Hmotnostní spektrometrie proteinů

v současnosti nejrozšířenější technika pro charakterizaci proteomu

MALDI

nejčastěji v kombinaci s průletovým hmotnostním analyzátozem – **TOF**
(*matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*)

MALDI - TOF MS

MALDI – TOF/TOF MS

ESI

standardně v kombinaci s iontovou pastí – **IT/Orbitrap**
(*electrospray ion trap mass spectrometry*)

ESI - IT MS

QQQ, Q-TOF, Q-LIT, ICR....

MALDI-TOF MS analýza

matrice je nízkomolekulární látka schopná absorpce laserového záření

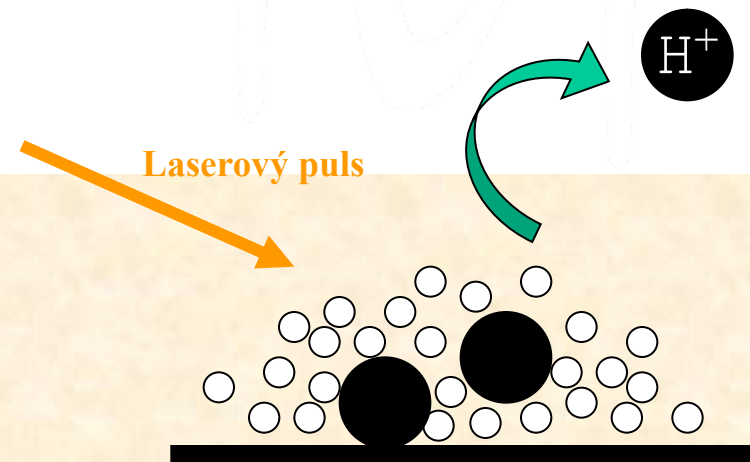
např. kyselina dihydroxybenzoová (pro UV laser)

Příprava vzorku:

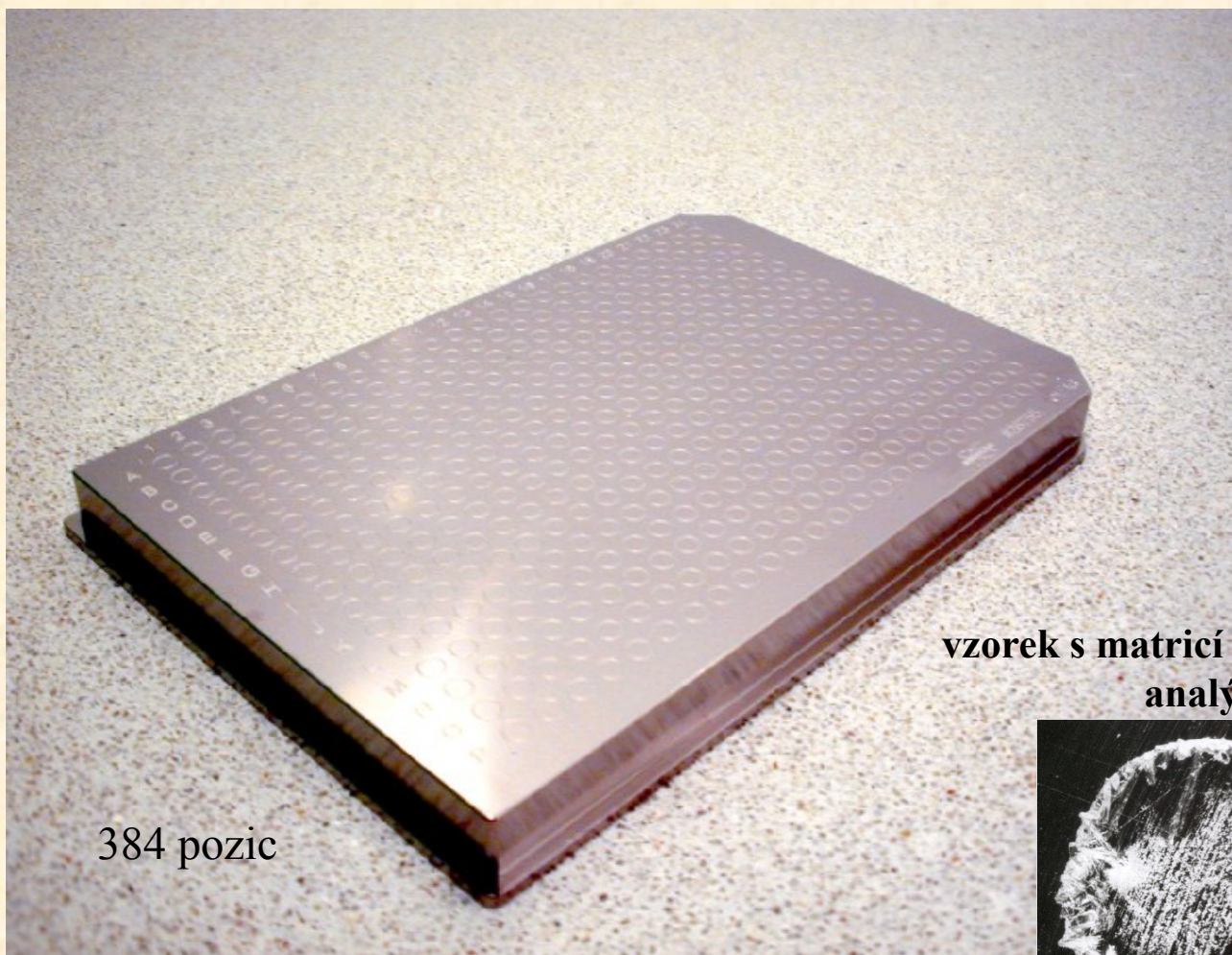
- ▶ vzorek proteinu je smíchán s nadbytkem matrice
- ▶ směs je nanesena na vzorkovací destičku a nechá se zaschnout
- ▶ vzniknou směsné krystaly matrice se vzorkem
- ▶ vlastní MS analýza

Výsledek:

- Šetrná ionizace bez fragmentace
- Jednoduchá spektra
- Uchování vzorku na vzorkovací desce

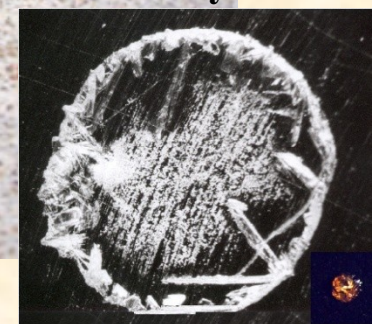


Vzorkovací destička pro MALDI-MS



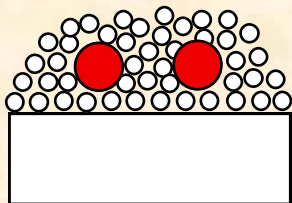
384 pozic

vzorek s matricí připravený na
analýzu

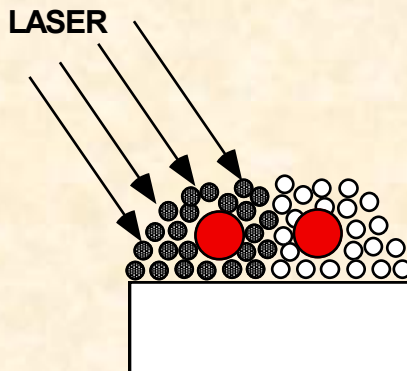


Desorption-ionization process

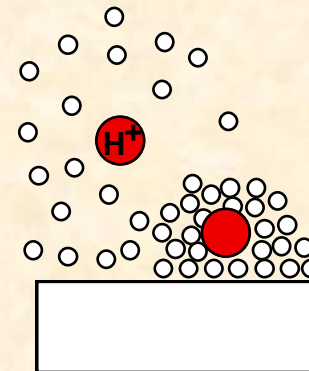
Sample embedded in
light-absorbing matrix



LASER-excitation of
matrix molecules



Sample desorption and
protonation



Průletový analyzátor (Time-of-Flight, TOF)

Separuje ionty dle doby letu analyzátořem, čas je pak přepočten na hmotnost

$$E = \frac{1}{2} m v^2$$

$$V = s/t$$

E – energie iontu

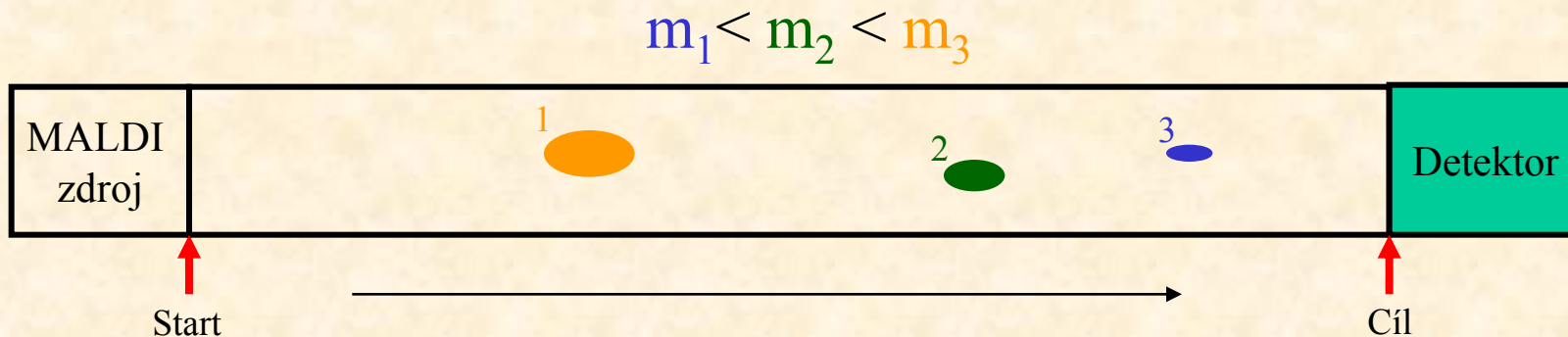
m – hmotnost iontu

v – rychlost iontu

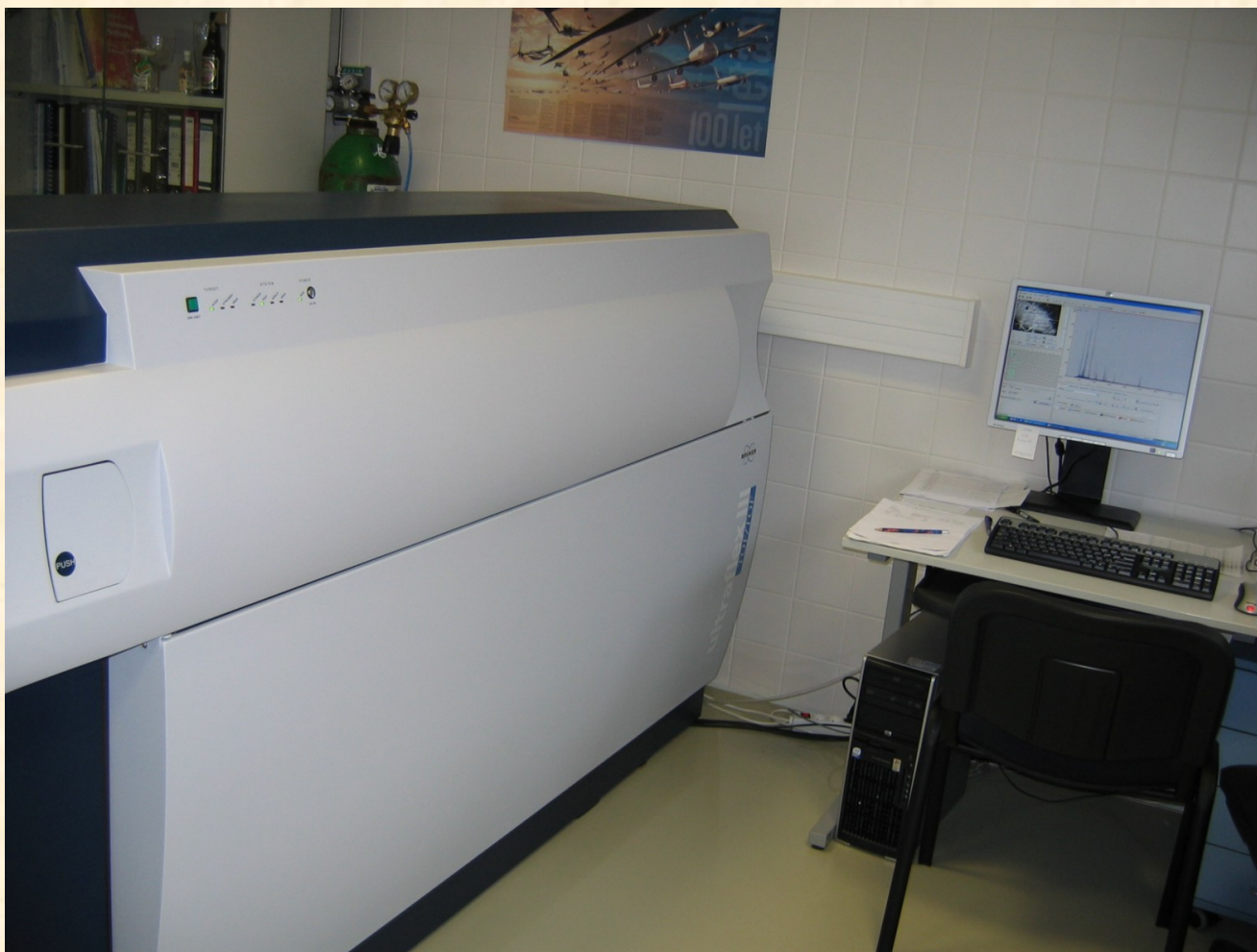
s – dráha letu

t – doba letu iontu

Ionťům je po ionizaci udělena stejná kinetická energie; současně vstupují do trubice analyzátořu a měří se jejich čas dopadu na detektor



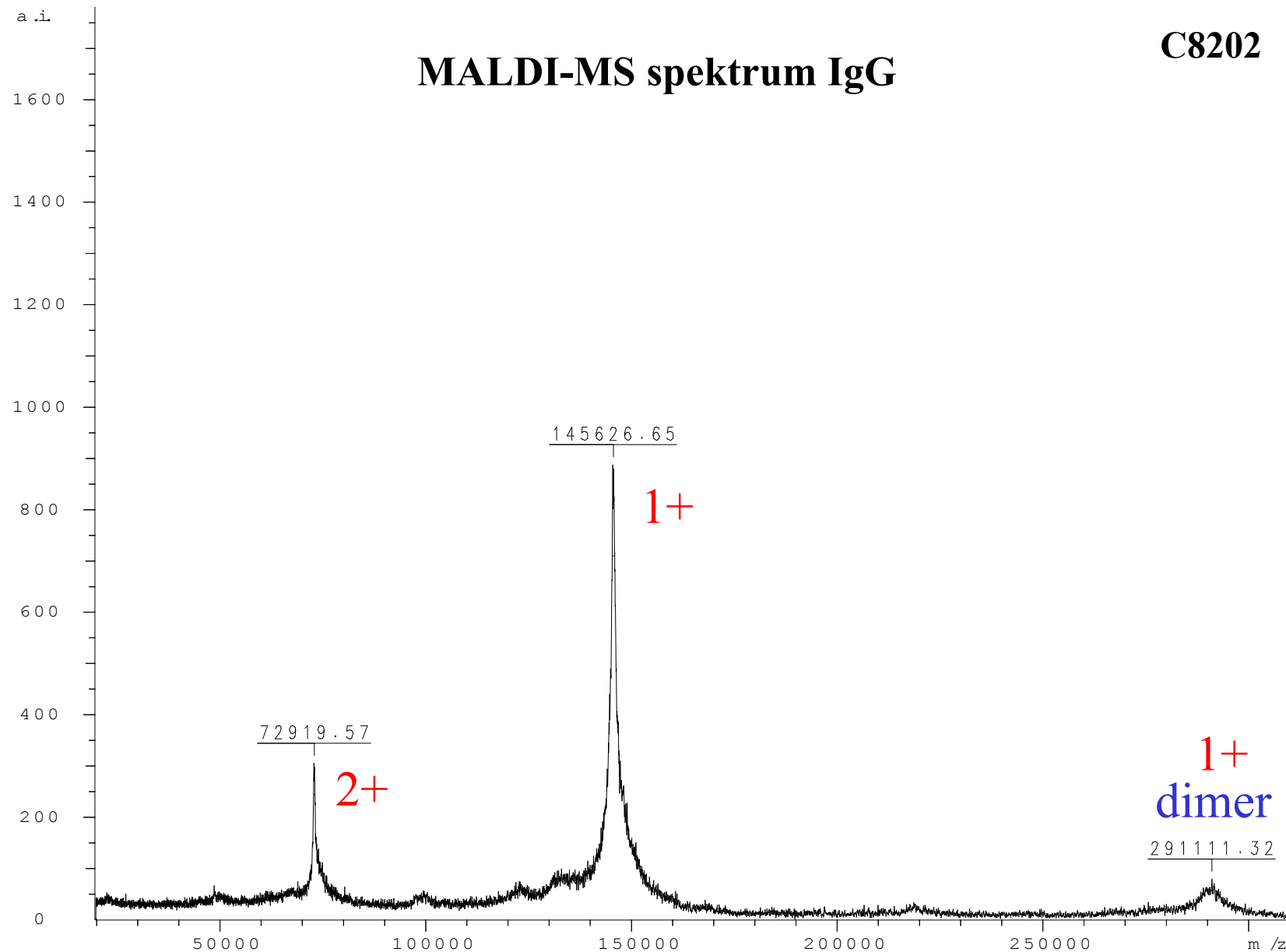
MALDI - MS/MS



MALDI-TOF/TOF hmotnostní spektrometr Ultraflex III (Bruker)

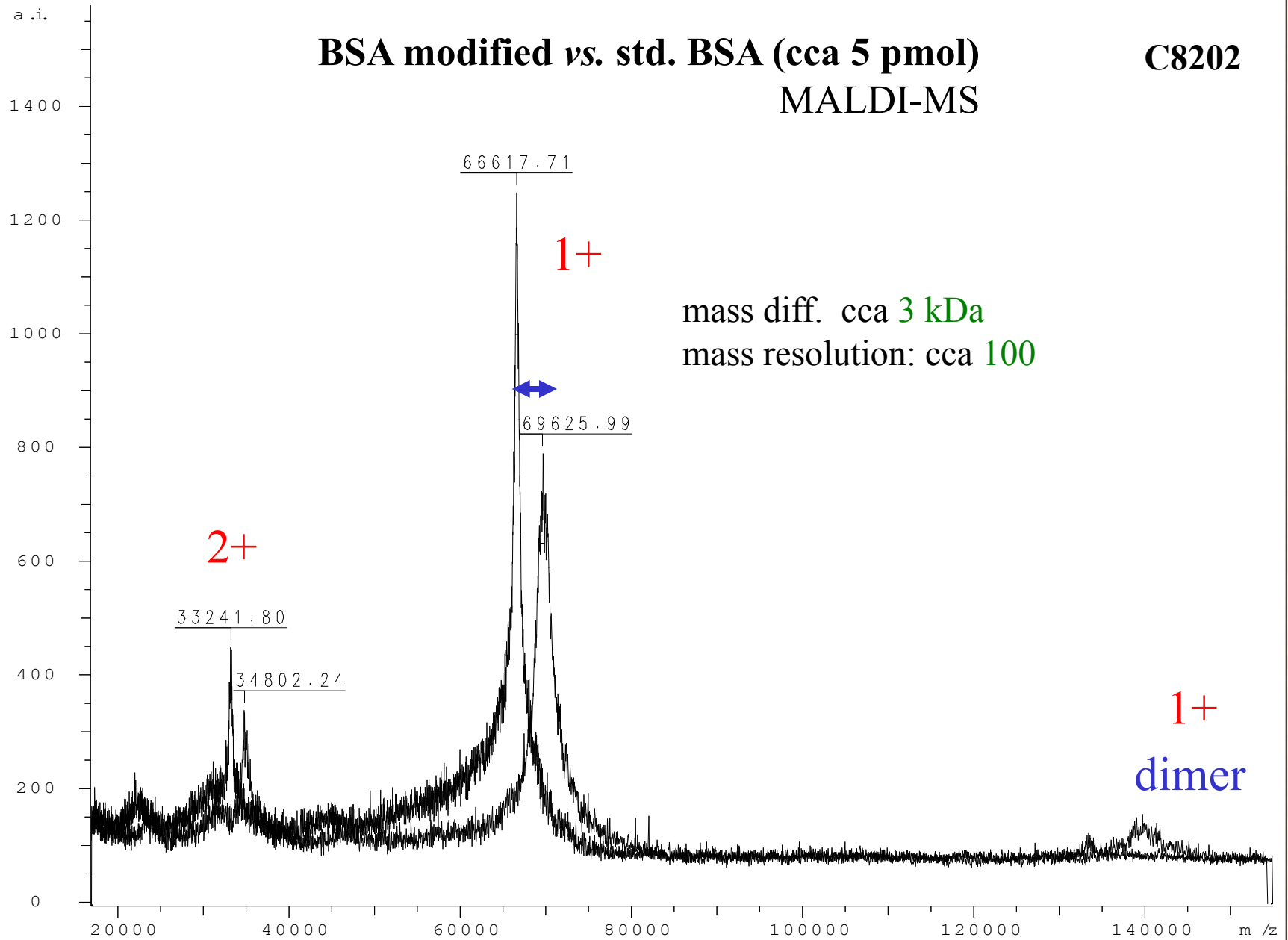
MALDI-MS spektrum IgG

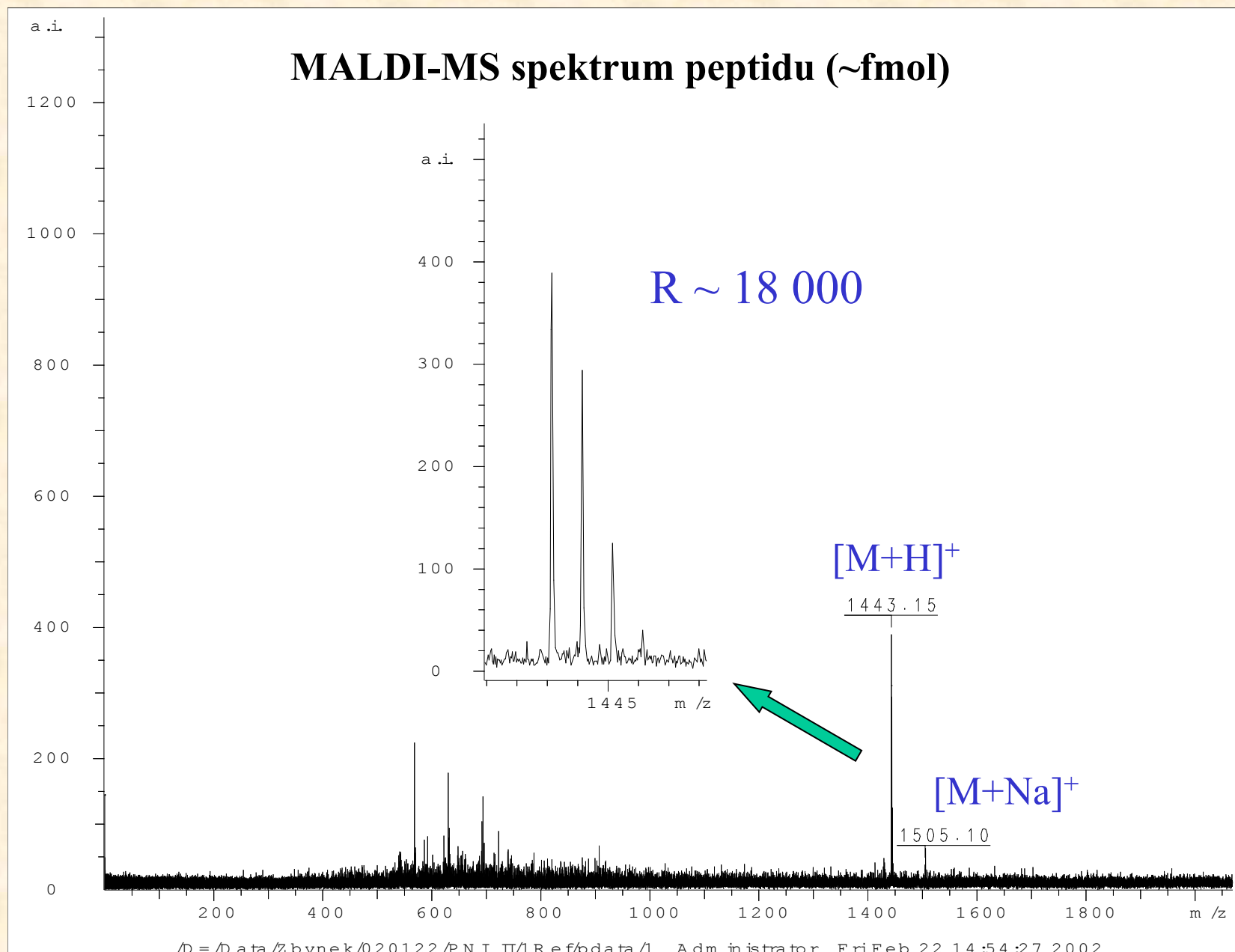
C8202



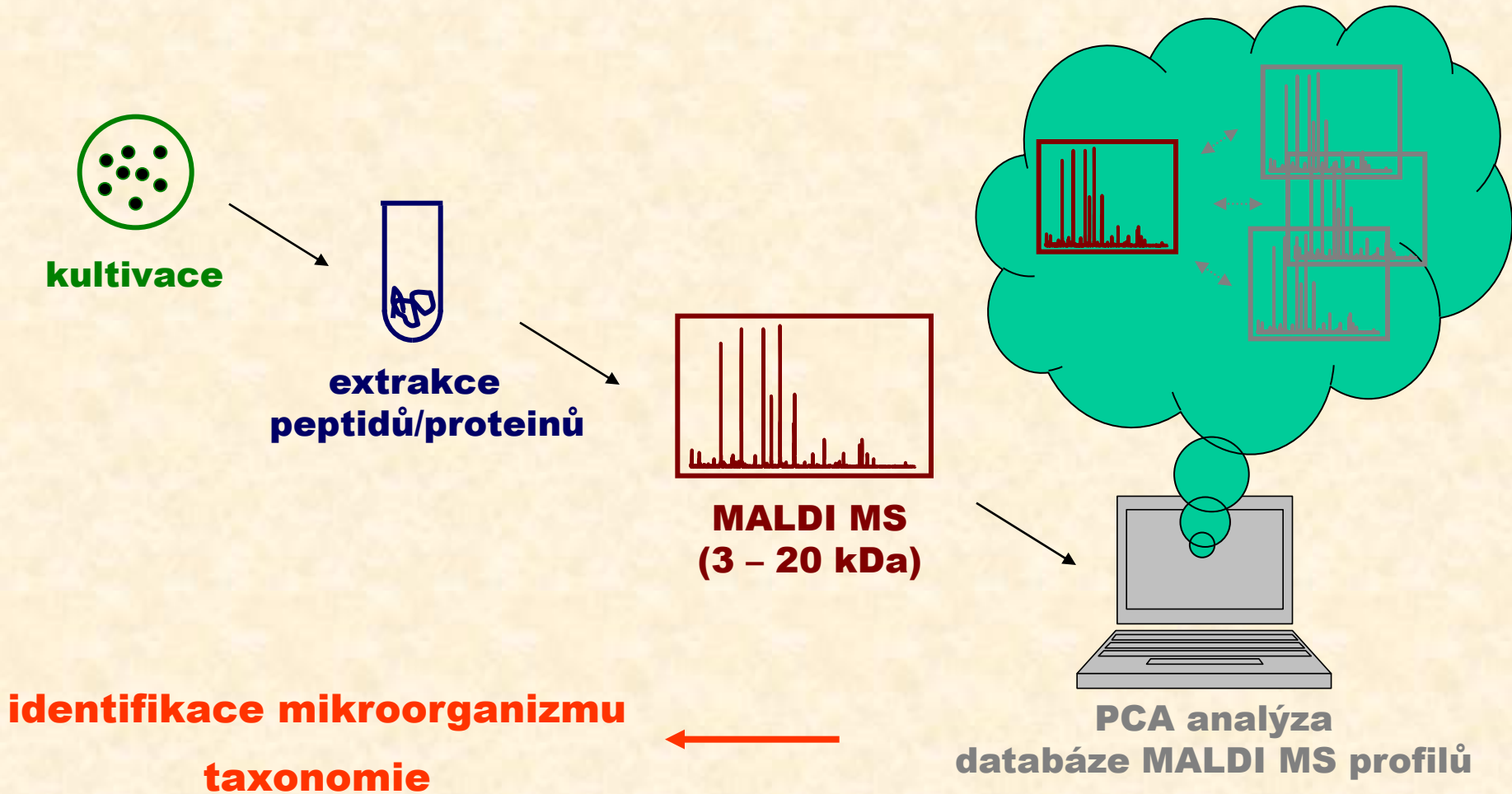
BSA modified vs. std. BSA (cca 5 pmol) MALDI-MS

C8202

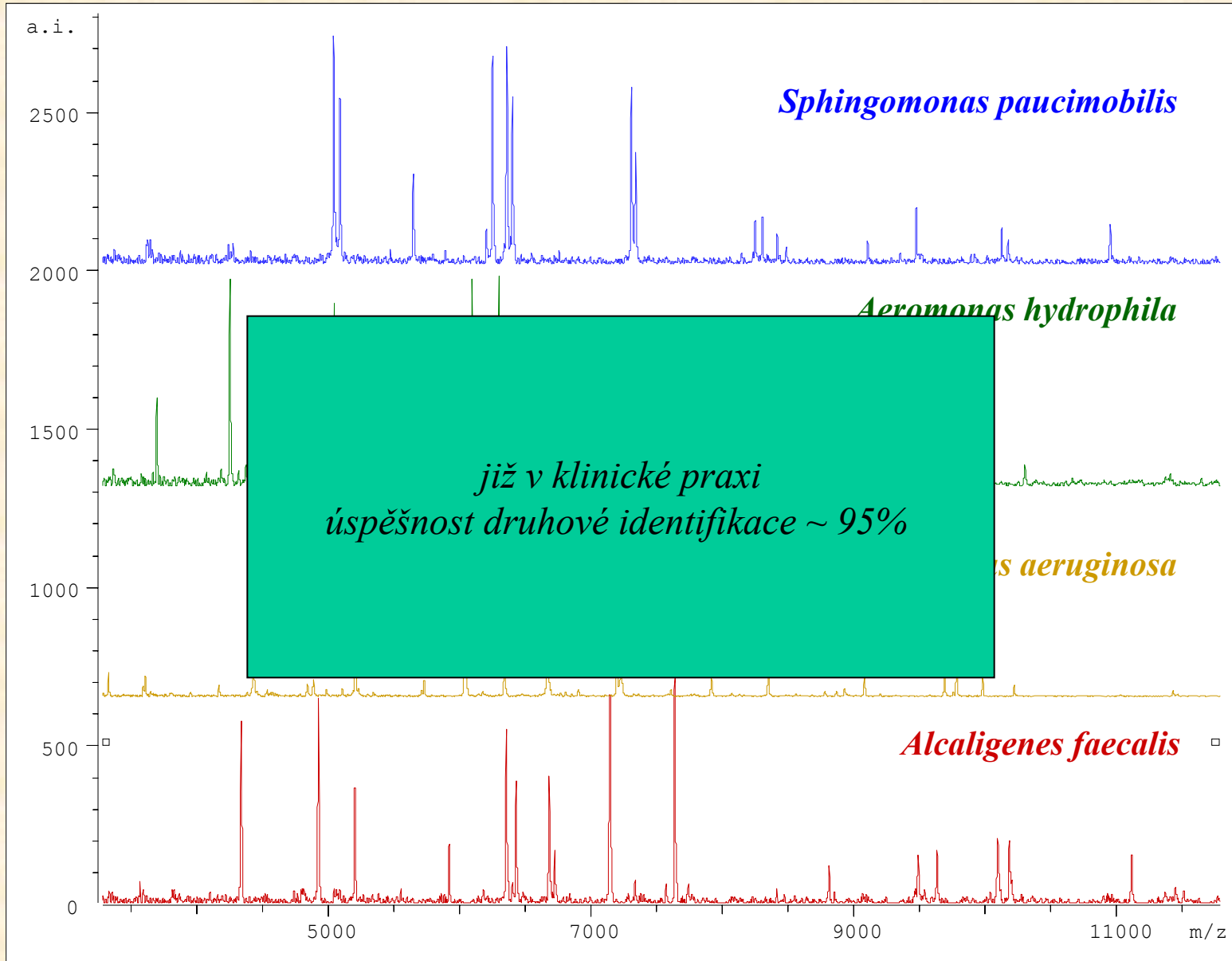




Identifikace mikroorganismů pomocí MALDI-MS

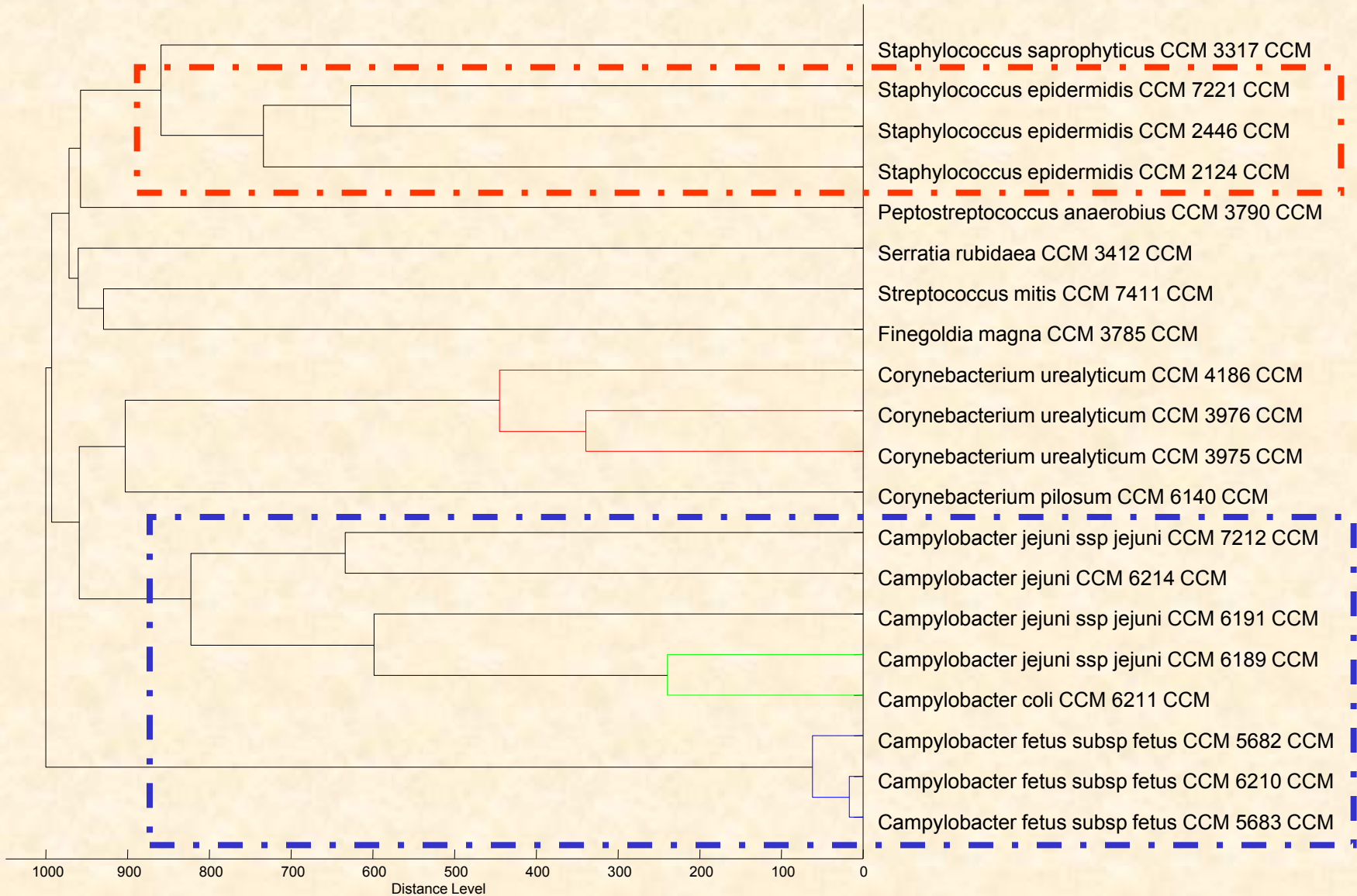


MALDI-MS spektra (profily) vybraných bakterií



MALDI-MS profilování

Grafické vyjádření podobnosti MALDI-MS profilů bakterií



ESI ionizace

Příprava vzorku:

- ▶ vzorek je v roztoku
- ▶ roztok je pak dávkován přes vstupní kapiláru do iontového zdroje

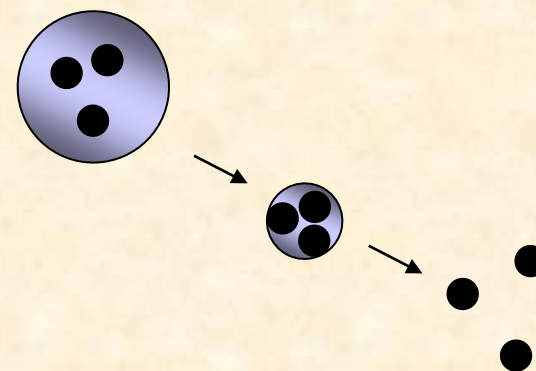
Ionizace vzorku:

- ✚ roztok vzorku je vstupní kapilárou zmlžován v komoře iontového zdroje **za atmosferického tlaku**
- ✚ ionizace probíhá ve spreji působením vloženého elektrického napětí
- ✚ vznikají nabité kapičky kapaliny, které během odpařování přechází na vícenásobně nabité ionty
- ✚ ionty jsou pak vtaženy do vakuové části spektrometru přechodovou kapilárou a dále analyzovány

Výsledek:

- Šetrná ionizace bez fragmentace
- vícenásobně nabité ionty
- Snadné spojení se separačními technikami (LC, CE)

LC - kapalinová chromatografie; CE - kapilární elektroforéza

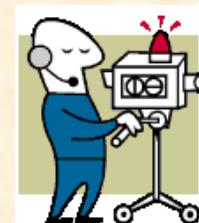


Iontová past

MS scan

Měření hmotností resp. m/z analyzovaných iontů

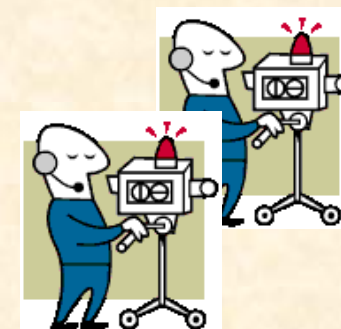
- zachycení iontů
- postupné vypuzování iontů z pasti podle m/z
- detekce iontů



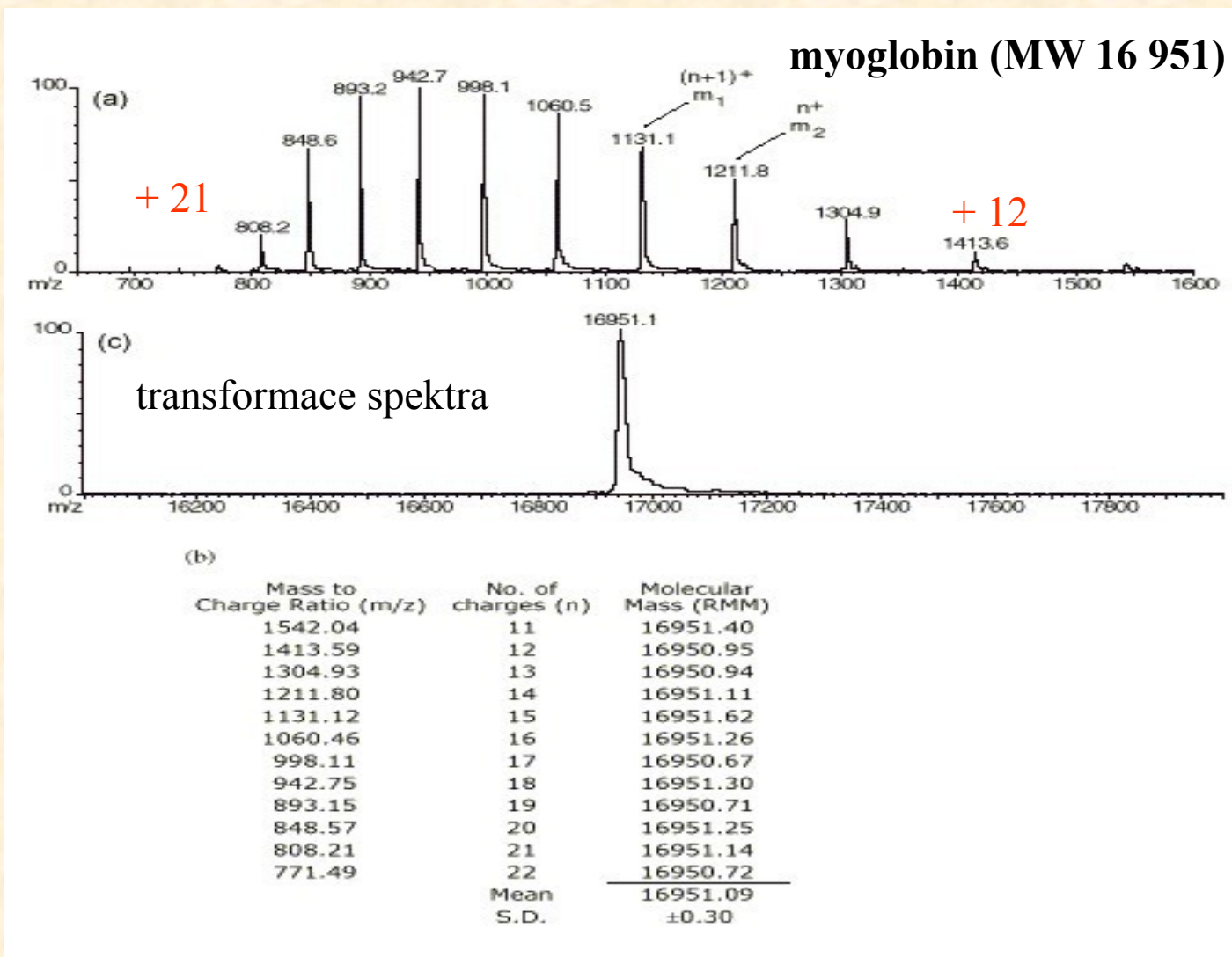
MS/MS scan

Analýza fragmentů vybraných iontů

- zachycení iontů
- vypuzení všech iontů z pasti mimo iontů s požadovaným m/z
- excitace a fragmentace vybraných iontů
- detekce fragmentů (dceřinných iontů)

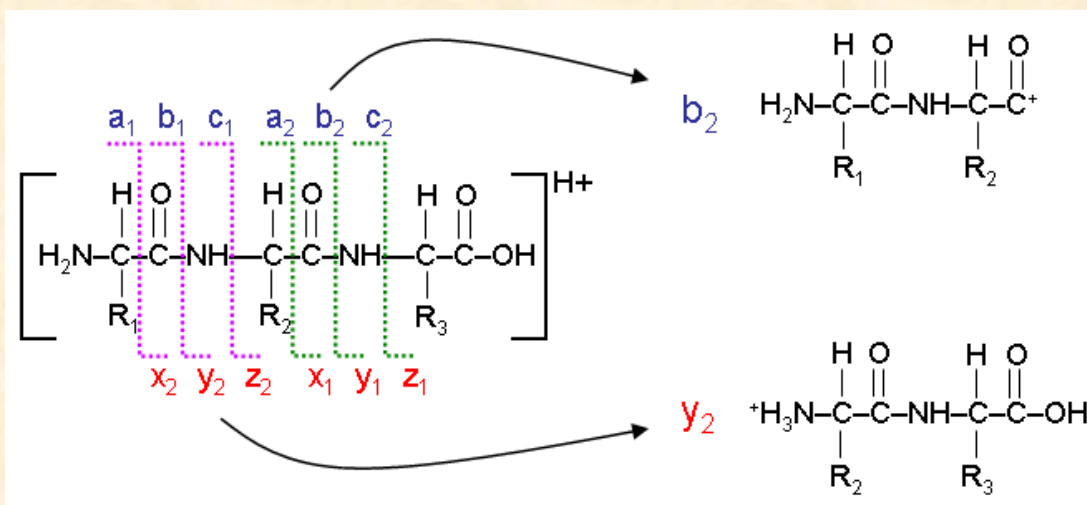


ESI-MS spektrum proteinu



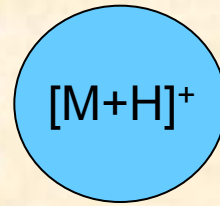
MS/MS peptidů

- ❖ peptidy jako polymerní látky jsou spojeny peptidovou vazbou
- ❖ při fragmentaci je peptid štěpen přednostně na peptidové vazbě a to tak, že v ideálním případě dojde k rozštěpení všech peptidových vazeb, takže vznikne soubor fragmentů s jednou, dvěma, třemi ... aminokyselinami, z rozdílů v m/z (resp. hmotnosti) „sousedních fragmentů“ lze odvodit druh aminokyseliny
- ❖ vznikají predikovatelné serie iontů ($b - y$, $a - x$, $c - z$), které lze snadno ze známé sekvence peptidu odvodit

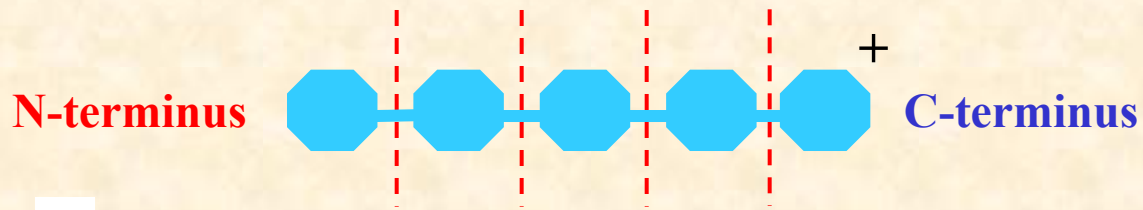


Schema fragmentace tripeptidu

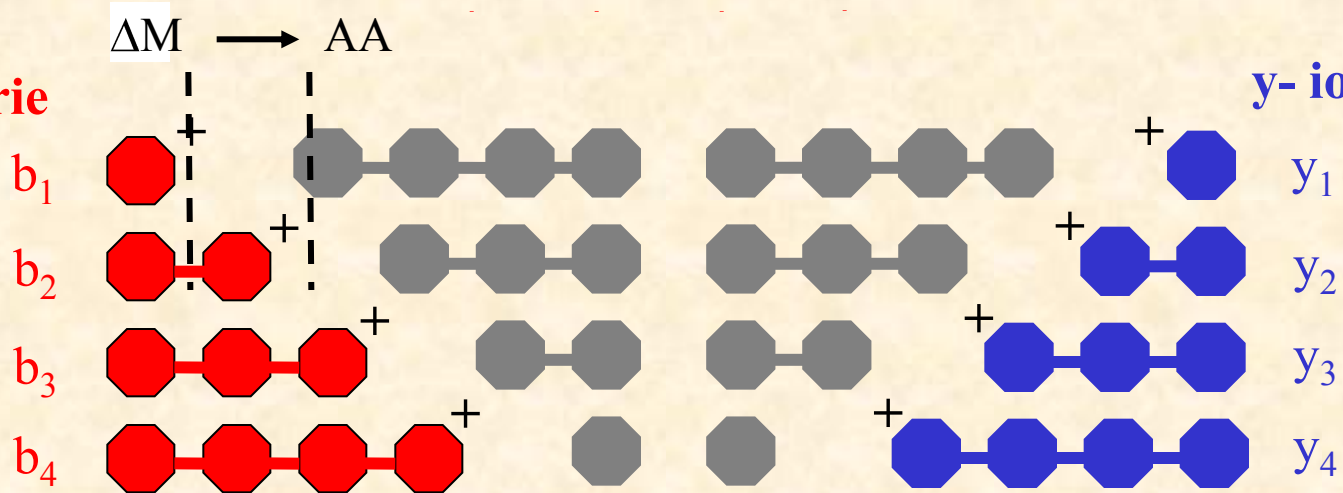
MS peptidů



MS/MS peptidů

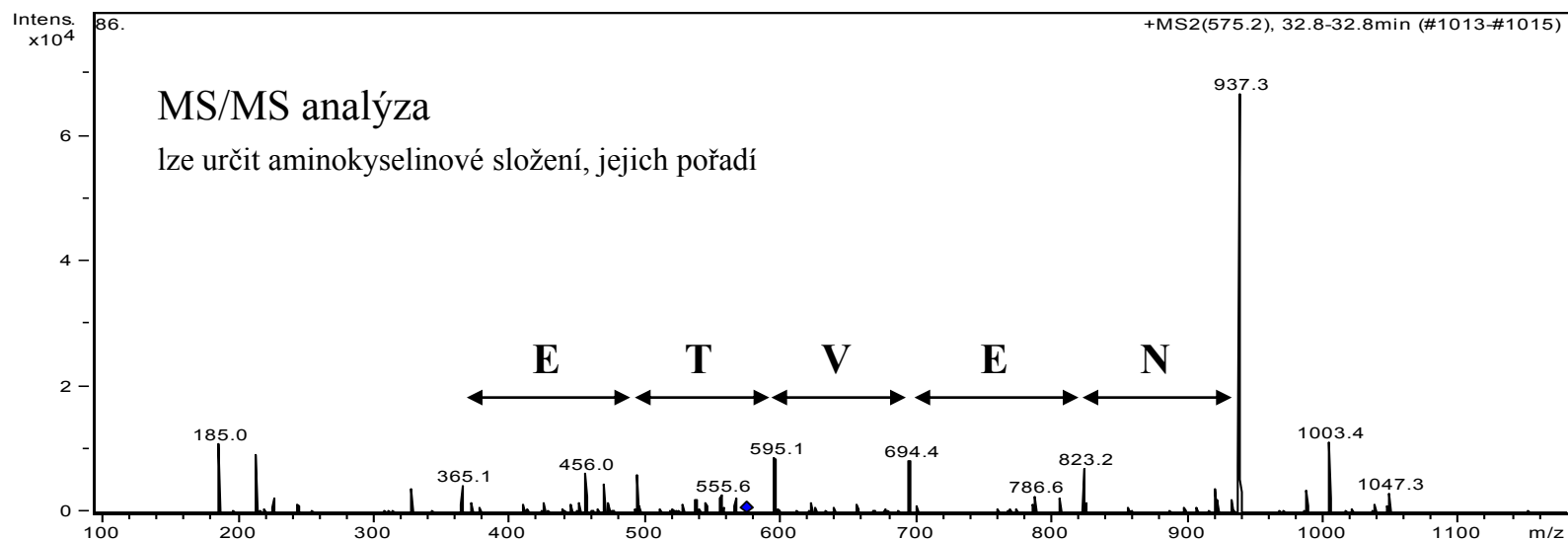
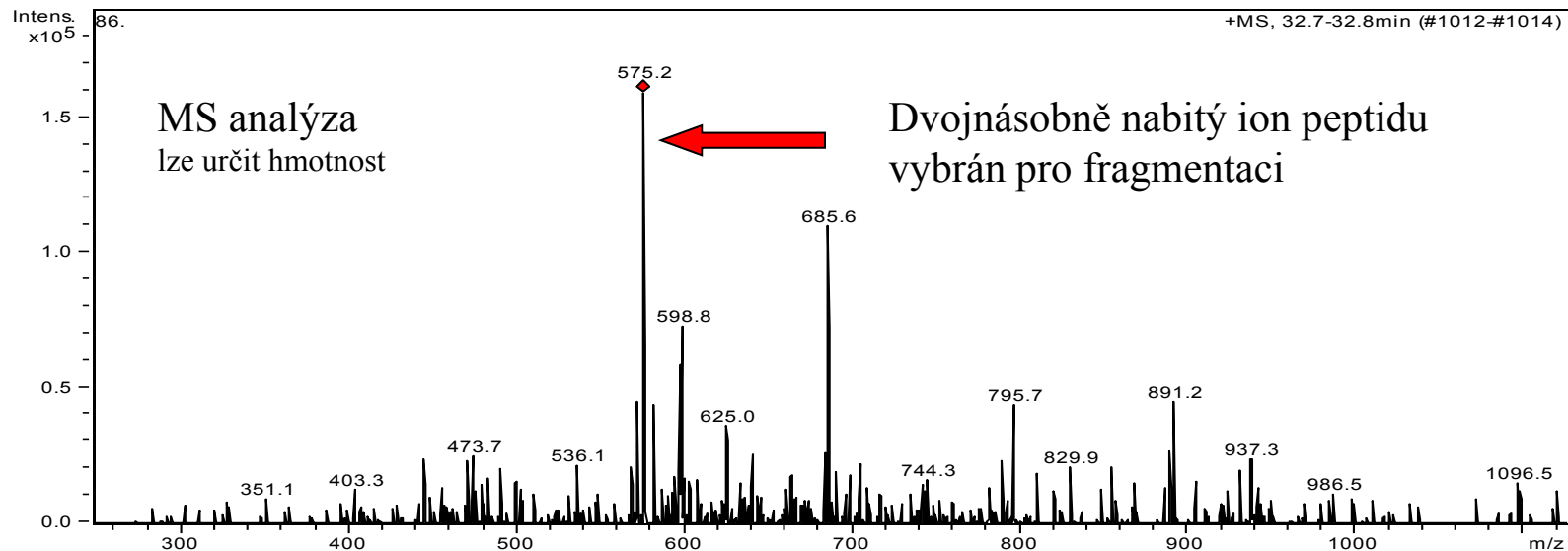


b- ion serie



fragmentační mapy pro jednotlivé peptidy

ESI-MS a MS/MS peptidu (MW 1148.5)



LC

on-line

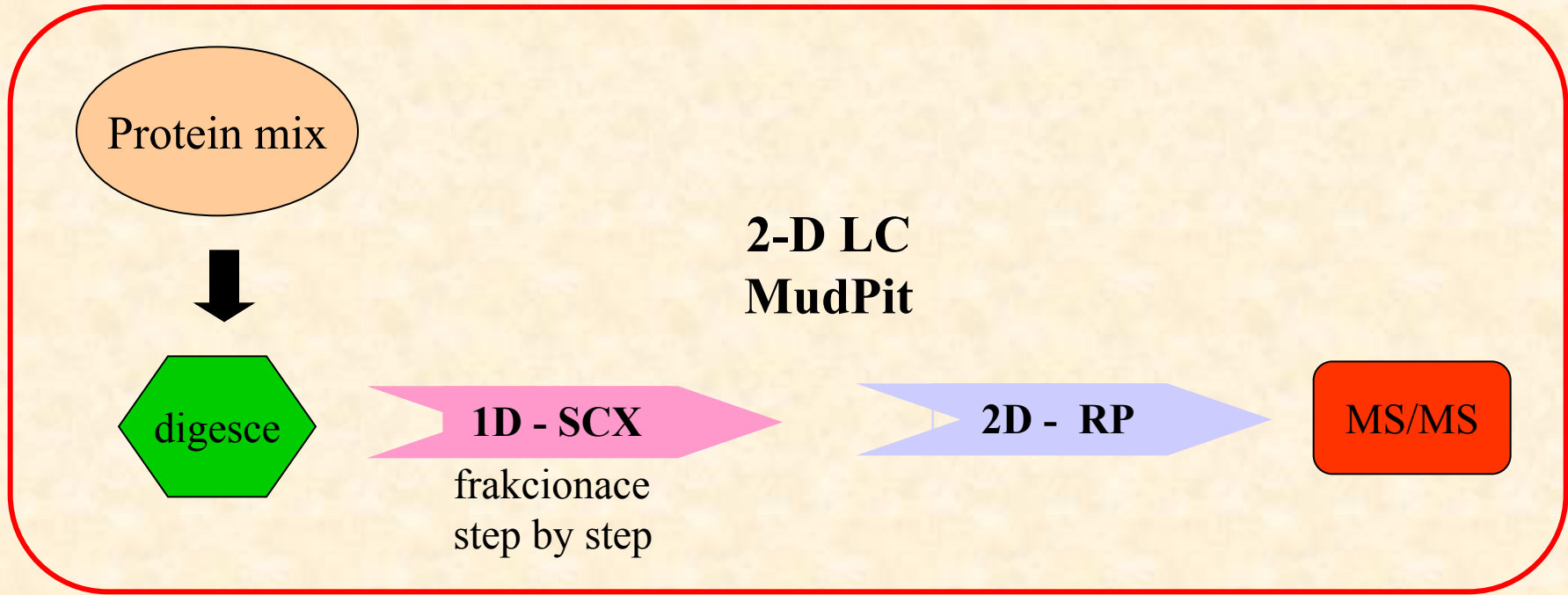
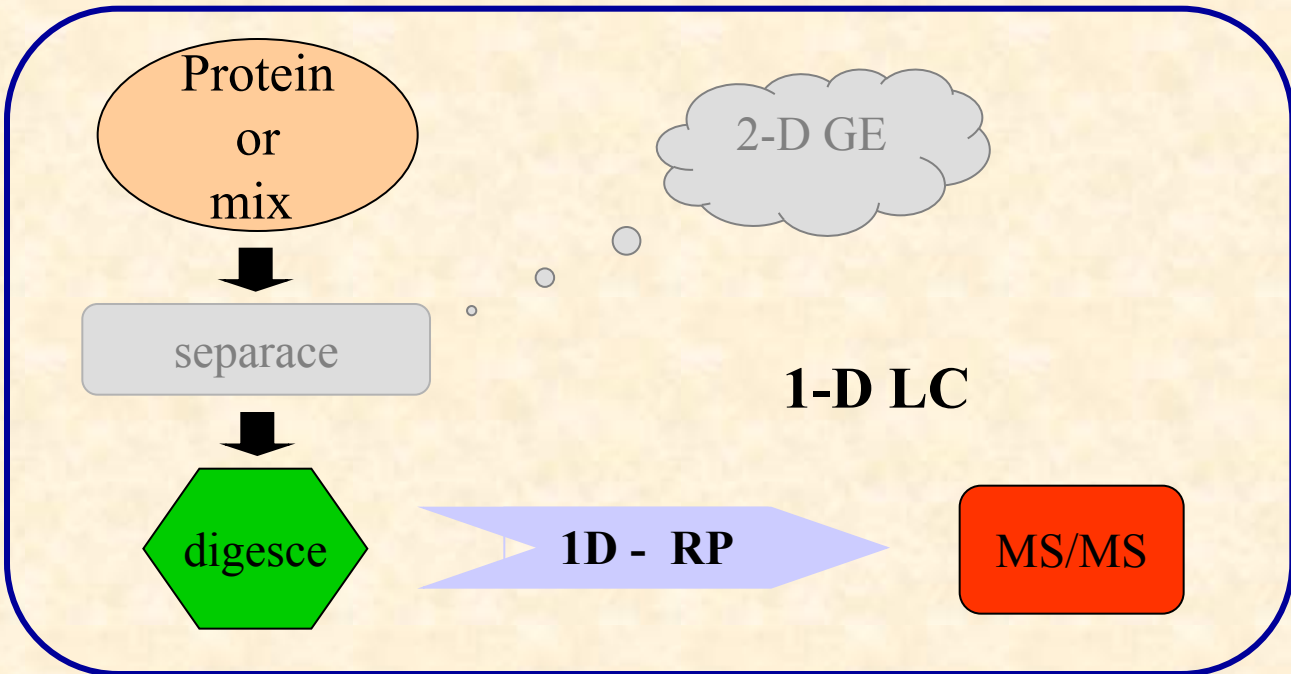
MS



LC-MS/MS



ESI-QTOF hmotnostní spektrometr Impact II (Bruker) spojený s kapalinovým chromatografem Ultimate 3000 RSLCnano (Dionex)



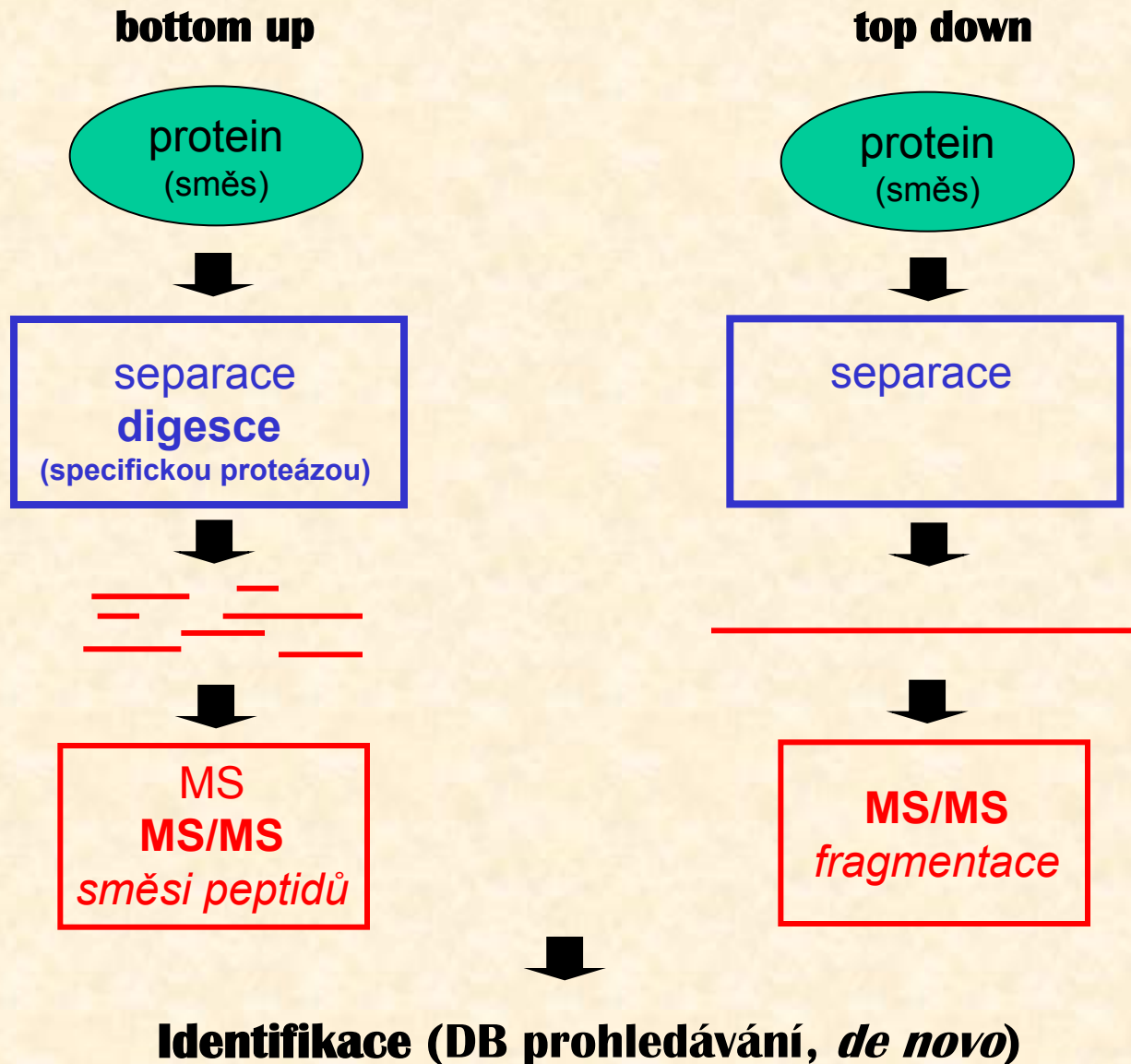
Krevní plazma
(3500 – 9000 proteinů ??)



Metody identifikace proteinů pomocí MS



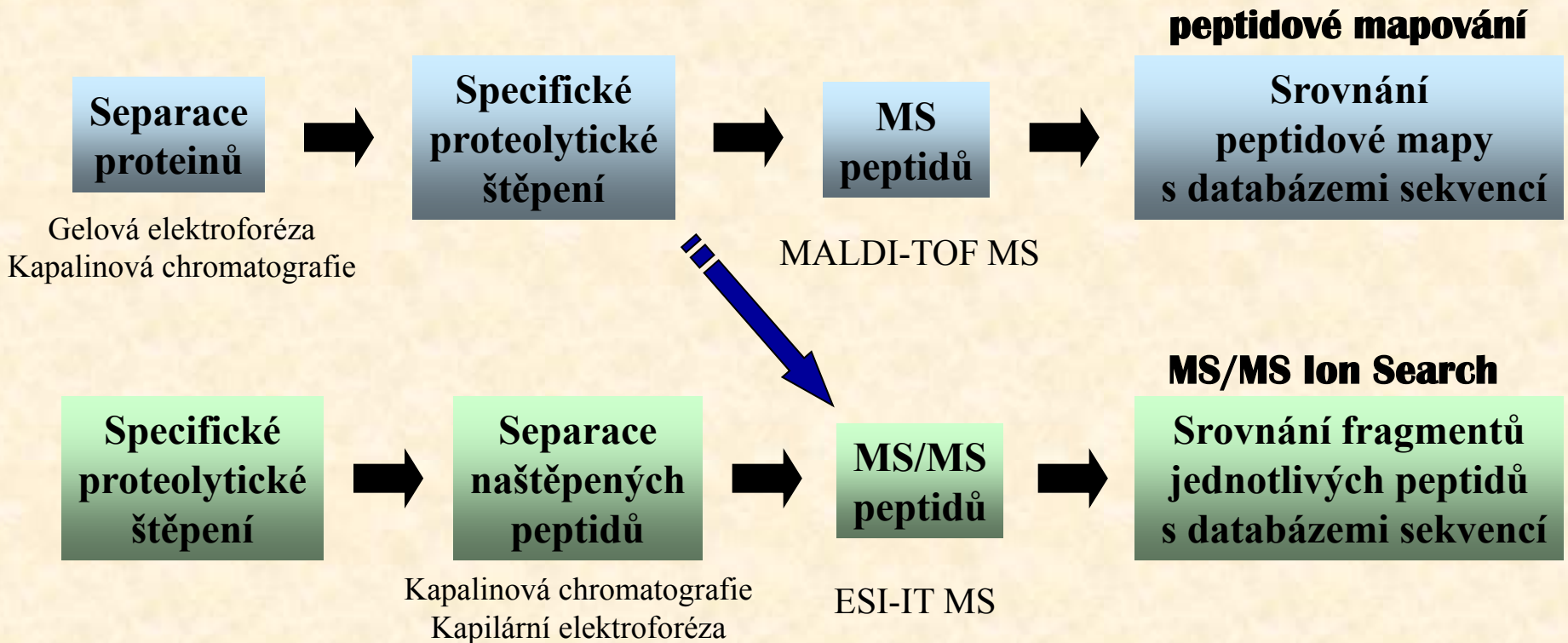
Identifikace proteinů hmotnostní spektrometrií



Identifikace proteinů pomocí MS

bottom up

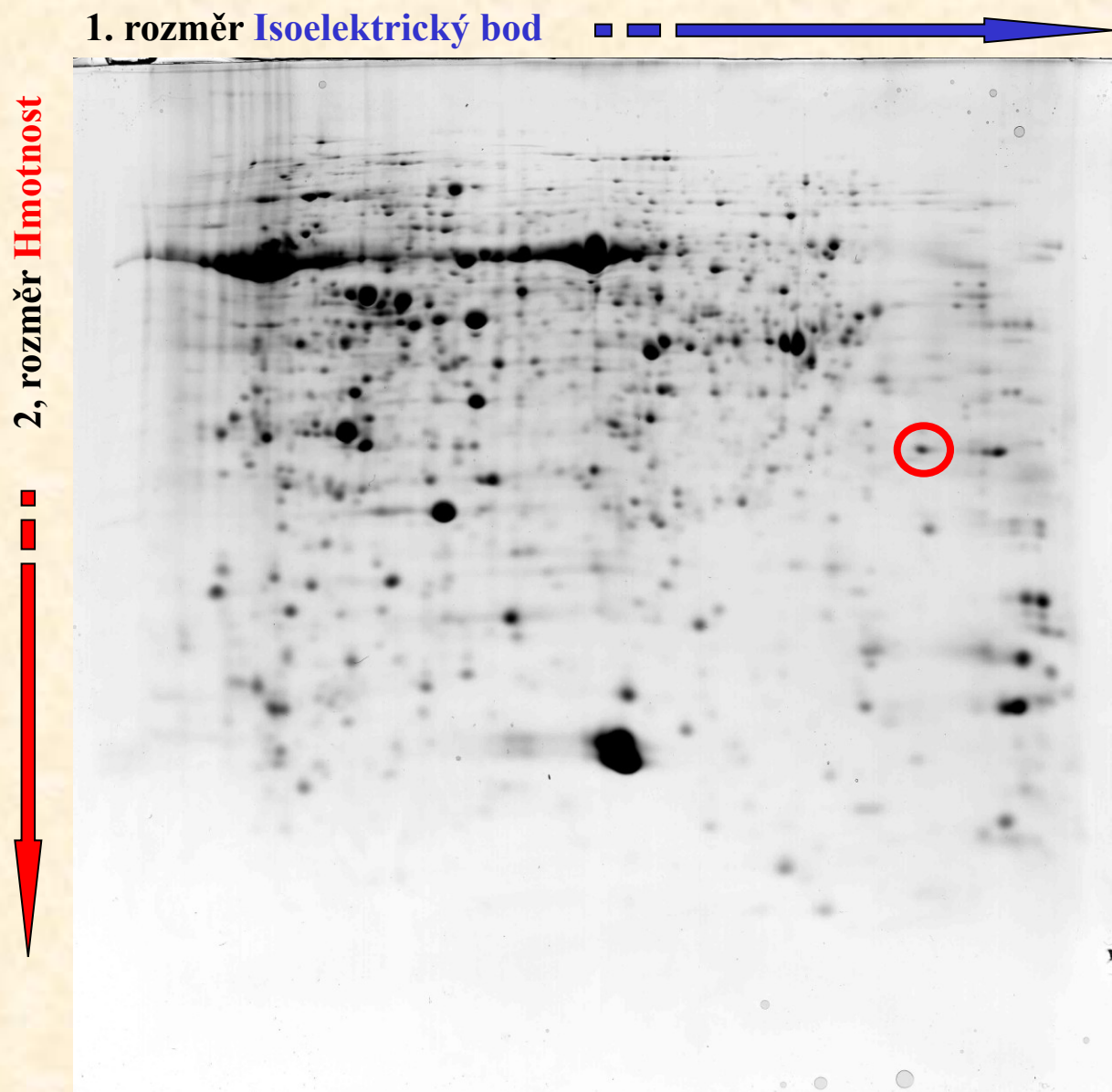
Standardní postupy:



Příklad identifikace proteinu pomocí peptidového mapování
(peptide mass fingerprinting, peptide mapping)

Modrá varianta

Separace proteinové směsi pomocí dvojrozměrné gelové elektroforézy



Separace
proteinů

Proteolytické štěpení (digestce)

enzymatické štěpení proteinu na soubor peptidů specifickou proteázou

Trypsin

štěpí vždy za **lysinem (K)** a **argininem (R)**, pokud není následující aminokyselina prolin

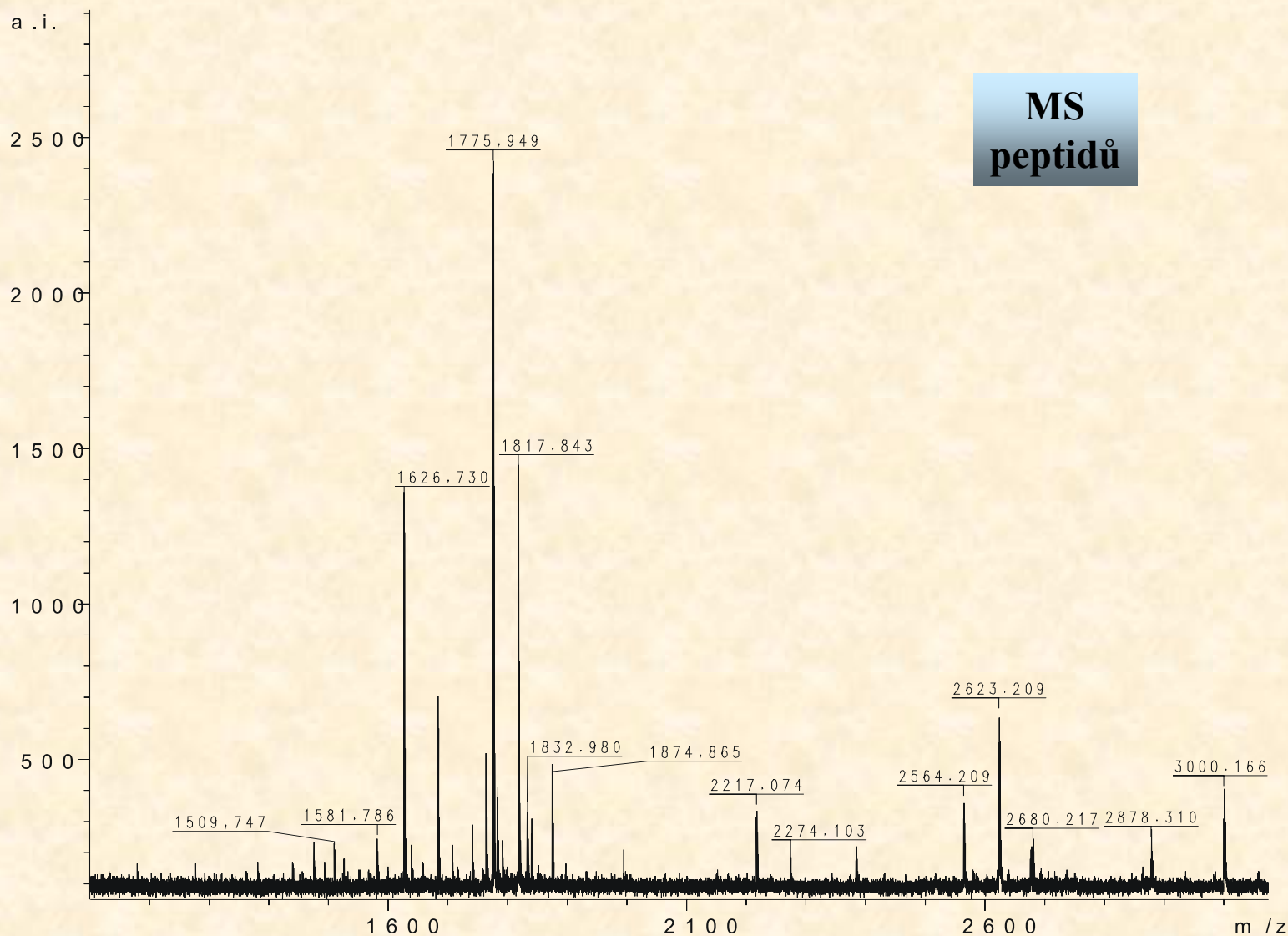
QNGVQMLSPSEI**PQR**DWFPSDFTFGAATSAYQIEGAWNED**GK**GESNWDHFCHNH**PER**ILD
 GSNSDIGANSYHMY**K**TDV**R**LL**K**EMGMDAY**R**FSISW**R**IL**P****K**GT**K**EGGINPDGI**K**YY**R**NLI
 NLLLENGIEP

QNGVQMLSPSEI PQR	1-15	1683.848 Da
DWFPSDFTFGAATSAYQIEGAWNED GK	16-42	3010.317 Da
GESNWDHFCHNH PER	43-57	1864.757 Da
ILDGSNSDIGANSYHMY K	58-75	1984.907 Da
TDV R	76-79	490.262 Da
...		

**Specifické
proteolytické
štěpení**

Soubor hmotností takto vzniklých peptidů (peptidová mapa) je charakteristický pro daný protein podobně jako otisk palce pro člověka

MALDI - TOF MS peptidů po digesti proteinu



MS spektrum obsahuje hmotnosti peptidů vzniklých digesti vybraného proteinu

Identifikace proteinu – peptidové mapování databázové prohledávání

Naměřená peptidová mapa (soubor hmotností peptidů vzniklých digescí analyzovaného proteinu) je prohledávána proti databázi proteinových sekvencí.

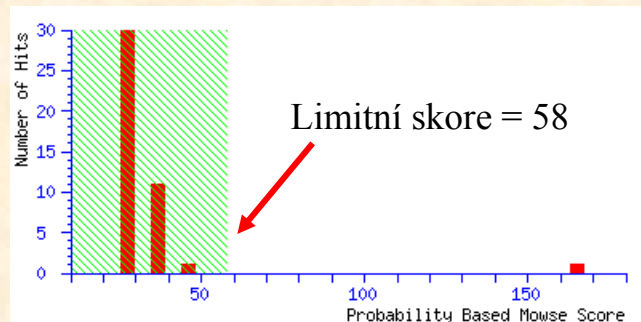
Prohledávací program si vytváří teoretickou peptidovou mapu od každé proteinové sekvence v databázi (dle použité proteázy) a postupně je srovnává s experimentálně naměřenou peptidovou mapou našeho analyzovaného proteinu.

Výsledkem prohledávání je žebříček proteinů s nejpodobnějšími peptidovými mapami. Míra shodnosti je vyjádřena tvz. skore, všechny proteiny s hodnotou skore vyšší než limitní jsou programem považovány za identifikované.

**Srovnání
peptidové mapy
s databázemi sekvencí**

Výsledek databázového prohledávání peptidové mapování

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 ...

1. **S18600** Mass: 47780

glutamate-ammonia ligase (EC 6.3.1.2) precursor, chloroplast (clone lambdaAtgsl1) - Arabidopsis thaliana

2. **S32228** Mass: 47714

glutamate-ammonia ligase (EC 6.3.1.2) precursor - rape - Brassica napus

Total score: **165**

Peptides matched: 12

Total score: **76**

Peptides matched: 7

Sequence Coverage: 44%

Srovnání
peptidové mapy
s databázemi sekvencí

```

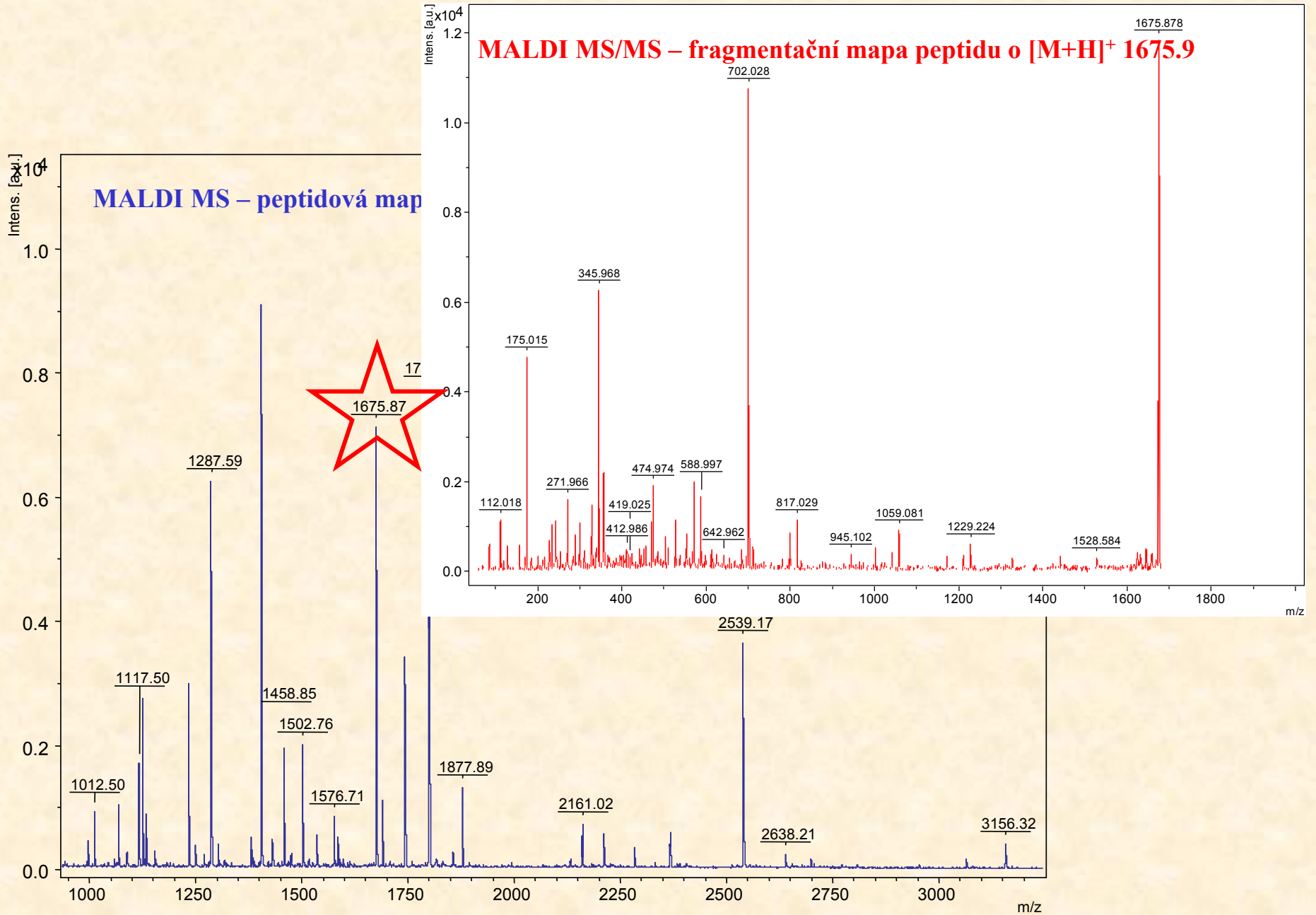
1   MAQILAASPT CQMRVPKHSS VIASSSKLWS SVVLKQKKQS NNKVGRFVRV
51  ALQSDNSTVN RVETLLNLDT KPYSDRIIAE YIWIGGSGID LRSKSRTIEK
101 PVEDPSELPK WNYDGSSTGQ APGEDSEVIL YPQAI FRDPF RGGNNILVIC
151 DTWTPAGEPI PTNKRAKAAE IFSNKKVSGE VPWFGIEQEY TLLQONVKWP
201 LGWPVGAFPG PQGPYYCGVG ADKIWGRDIS DAHYKACLYA GINISGTNGE
251 VMPGQWEFQV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

www.matrixscience.com

Červeně jsou vyznačeny úseky sekvence odpovídající přiřazeným peptidům z naměřené peptidové mapy

Identifikace na základě MS dat vs MS/MS

C8202



Příklad identifikace proteinu pomocí LC-MS/MS

Zelená varianta

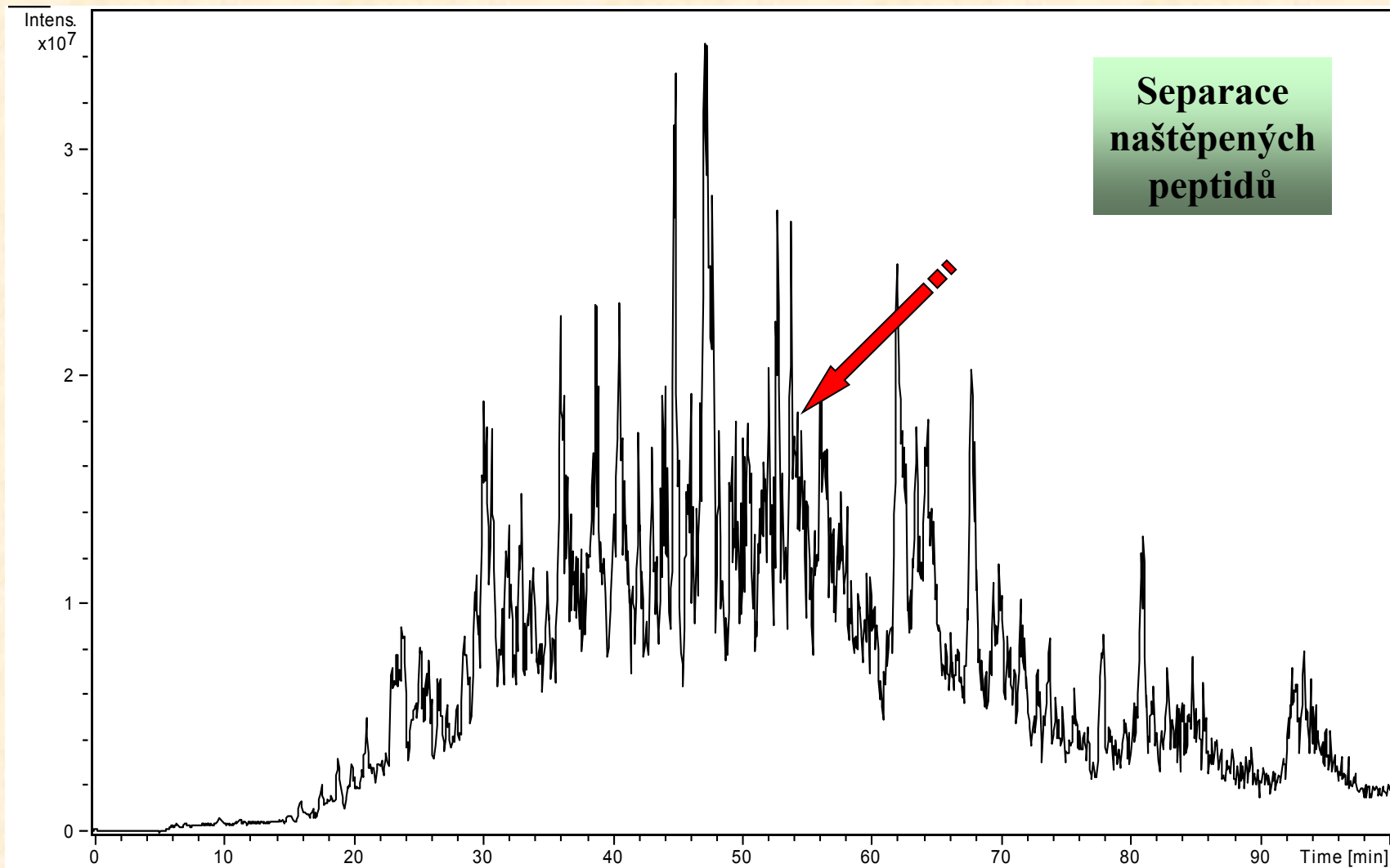
Proteolytické štěpení

**Specifické
proteolytické
štěpení**

Na rozdíl od „modré varianty“ tentokrát je enzymaticky štěpena celá směs proteinů najednou, opět za použití specifické proteázy (obvykle trypsin).

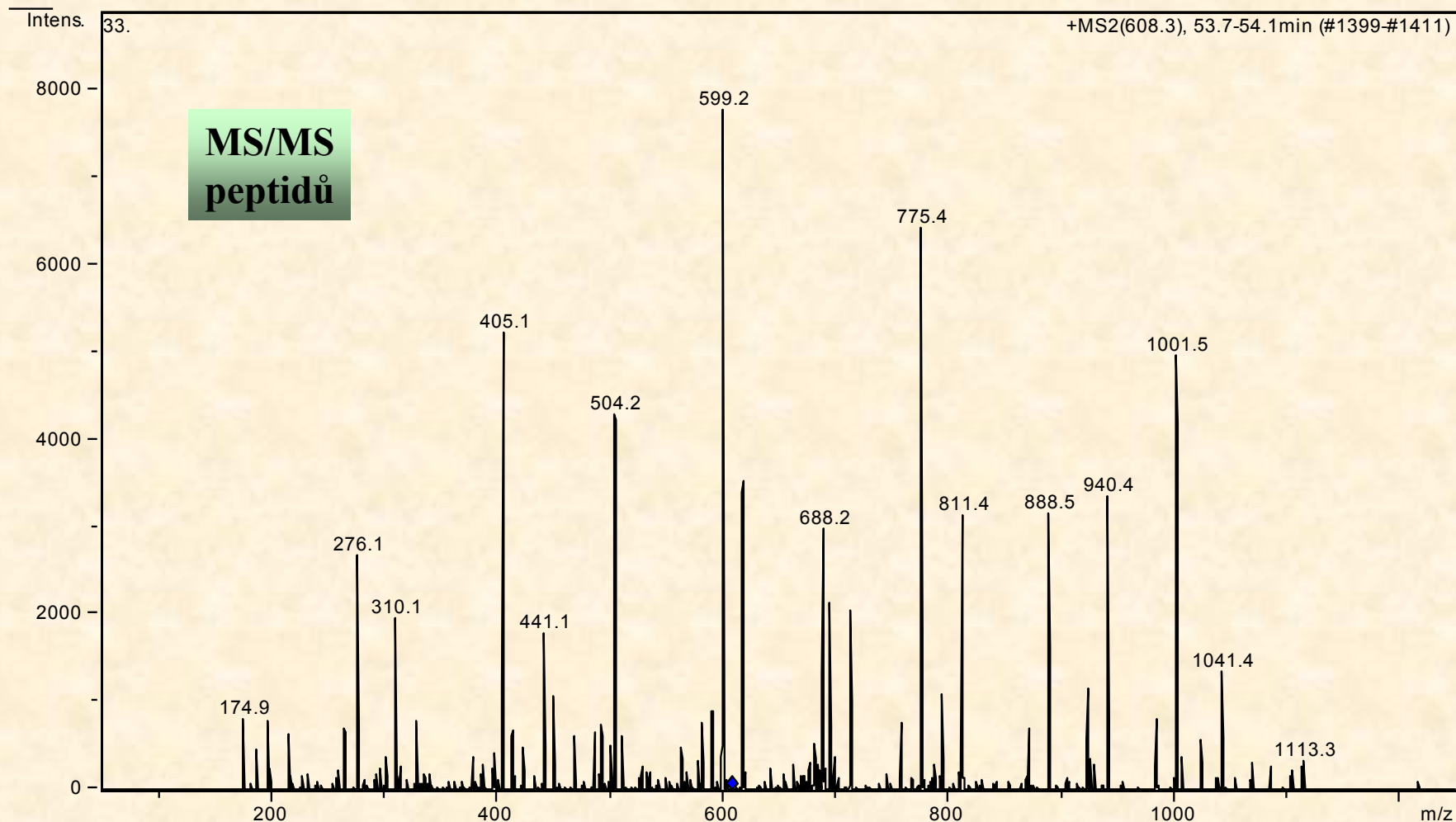
Takto vzniklá směs peptidů je pak separována a podrobena MS/MS analýze.

Separace peptidů tryptického digestu směsi proteinů



Digest vzorku lidské plazmy separovaný kapalinovou chromatografií spojenou s hmotnostní spektrometrií (LC-MS/MS)

MS/MS spektrum tryptického peptidu (m/z 608.3, 2+)



MS/MS spektrum obsahuje fragmenty peptidu vzniklé kolizní disociací v iontové pasti. Fragmenty nesou specifickou informaci o sekvenci peptidu.

Identifikace proteinu – MS/MS databázové prohledávání

Naměřené fragmentační mapy (soubory hmotností fragmentů vzniklých při kolizní disociaci jednotlivých peptidů) jsou prohledávány proti databázi proteinových sekvencí.

Prohledávací program si vytváří teoretickou peptidovou mapu proteinové sekvence v databázi, po té si od každého peptidu příslušné peptidové mapy připraví teoretickou fragmentační mapu a postupně je srovnává s experimentálně naměřenými fragmentačními mapami analyzovaných peptidů. Takto to provede pro každou proteinovou sekvenci v databázi.

Pro každý přiřazený peptid (MS/MS spektrum) je spočítáno individuální skóre, hodnota skóre vyšší než limitní určuje signifikantní shodu mezi teoretickou a naměřenou fragmentační mapou. Míra shodnosti proteinu je vyjádřena celkovým skóre, které je odvozeno z individuálních skóre peptidů přiřazených danému proteinu.

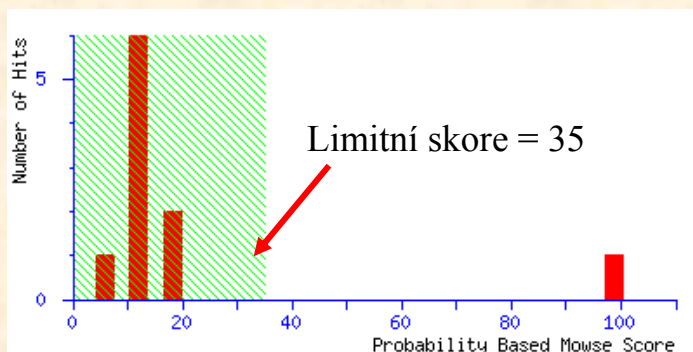
**Srovnání fragmentů
jednotlivých peptidů
s databázemi sekvencí**

Mascot Search Results

Výsledek databázového prohledávání

MS/MS

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



Srovnání fragmentů jednotlivých peptidů s databázemi sekvencí

Peptide Summary Report

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

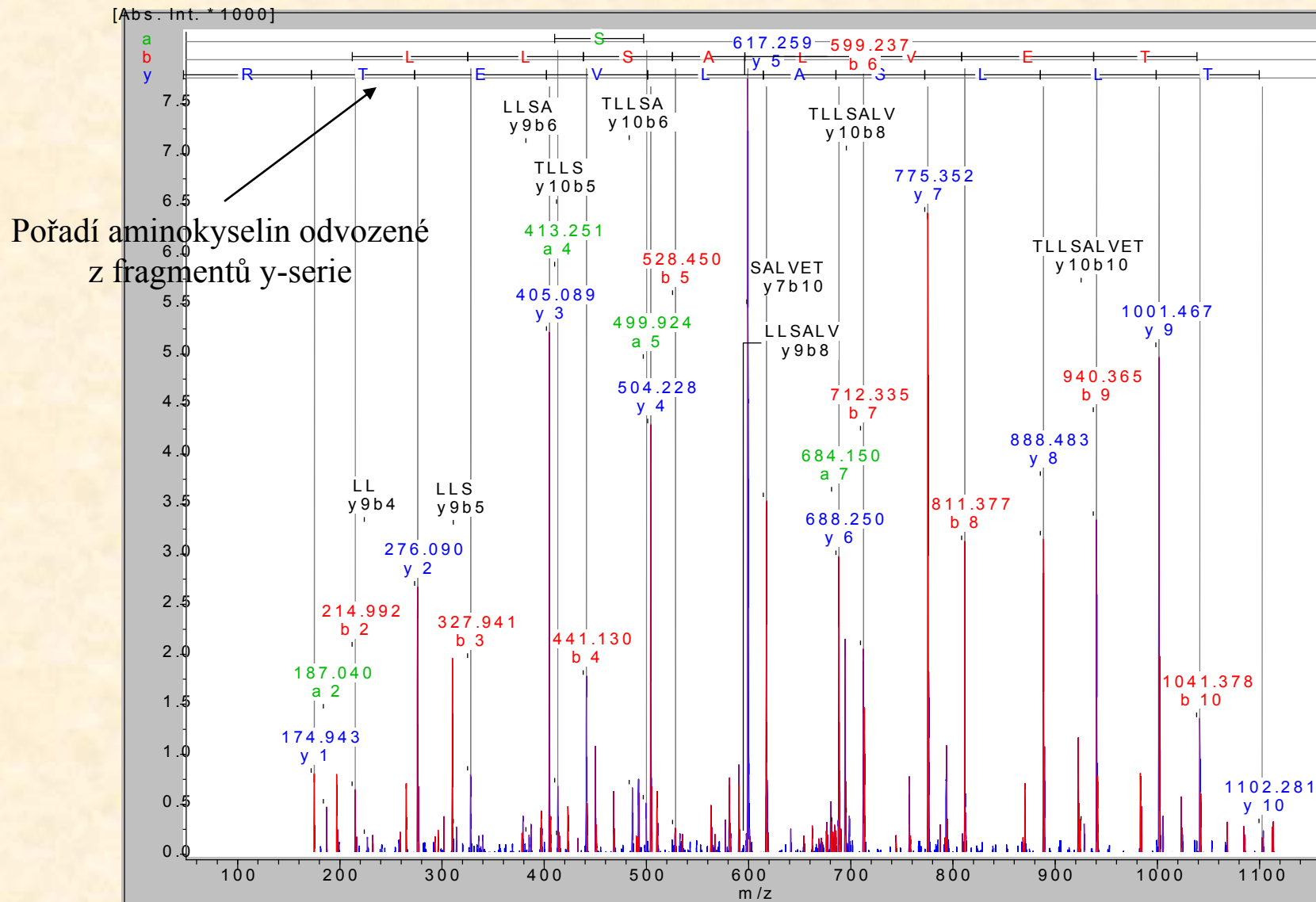
Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
1	608.3000	1214.5854	1214.7234	-0.1380	0	99	2.2e-08	1	K.ITLLSALVETR.T

Celkové skóre

Individuální skóre peptidu

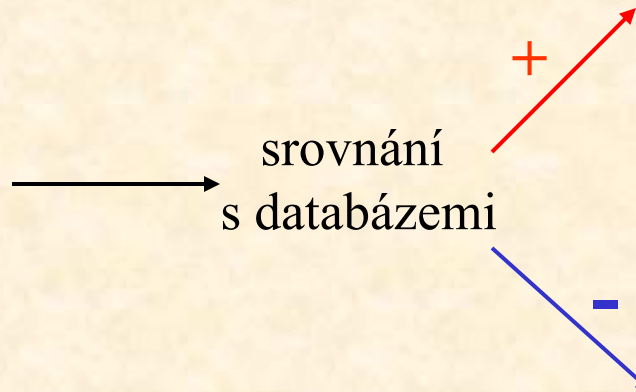
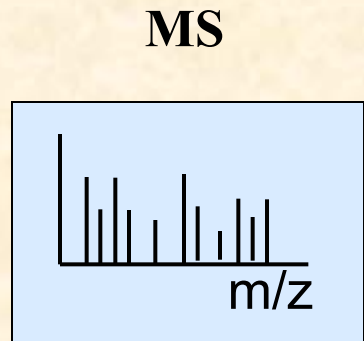
Díky variabilitě primární struktury proteinů lze určit jejich "identitu" na základě fragmentace (MS/MS spektra) jediného peptidu.

Vyhodnocené MS/MS spektrum

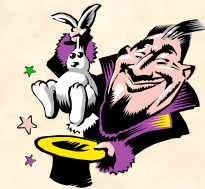


Z rozdílů m/z (resp. hmotnosti) mezi sousedními ionty příslušné serie (b, y) lze zjistit identitu jednotlivých aminokyselin i jejich pořadí.

Identifikace proteinu – databázové prohledávání



protein je identifikován
(protein se sekvencí v databázi)



neznámý protein
(sekvence není k dispozici,
sekvenace *de novo*)






Charakterizace posttranslačních modifikací







Charakterizace modifikací proteinů

Druhy modifikací:

-  **mutace** (izoformy proteinů)
-  **chemické** (oxidace, adukty farmak aj.)
-  **posttranslační** (fosforylace, glykosylace aj.)

Možnosti MS při analýze modifikací proteinů

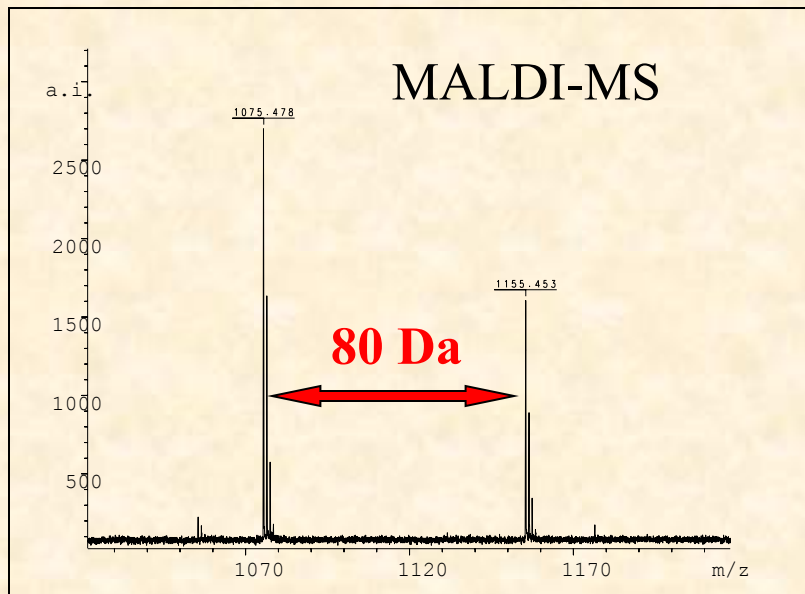
-  **druh modifikace**
-  **vazebné místo**
-  **úroveň** (zastoupení modifikace na daném místě)

-  úprava vzorků (frakcionace, enzym. odbourávání aj.)
-  speciální MS/MS techniky

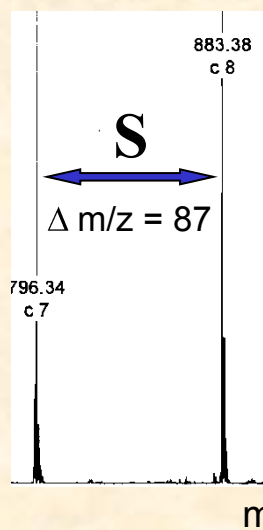
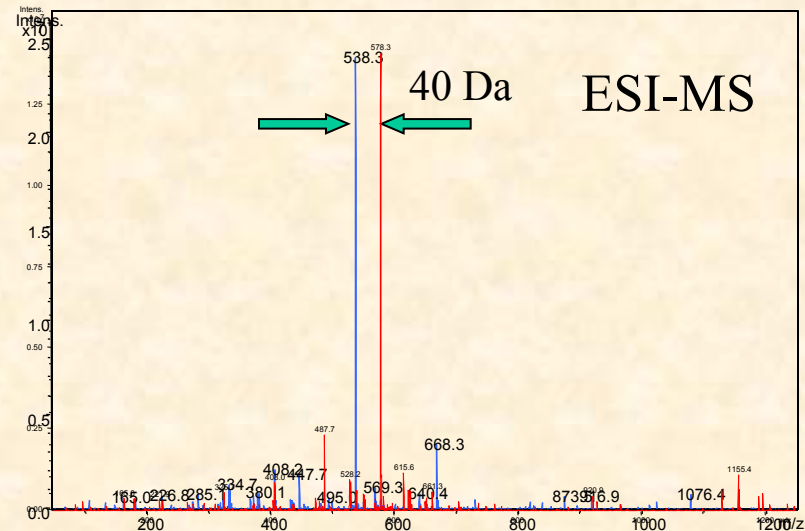
Užitečný přehled modifikací:

DeltaMass <http://www.abrf.org/index.cfm/dm.home>

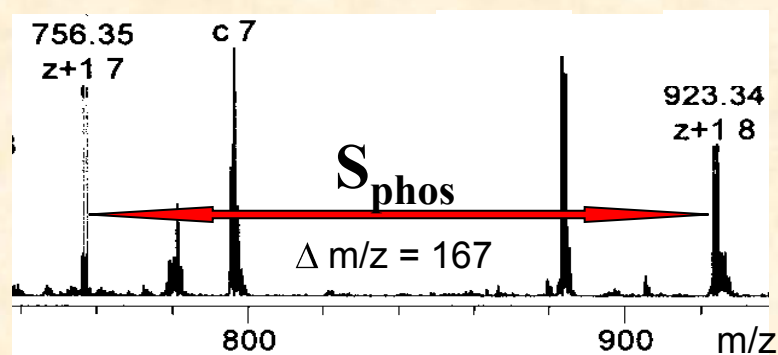
ExpASy - <http://www.expasy.ch/>



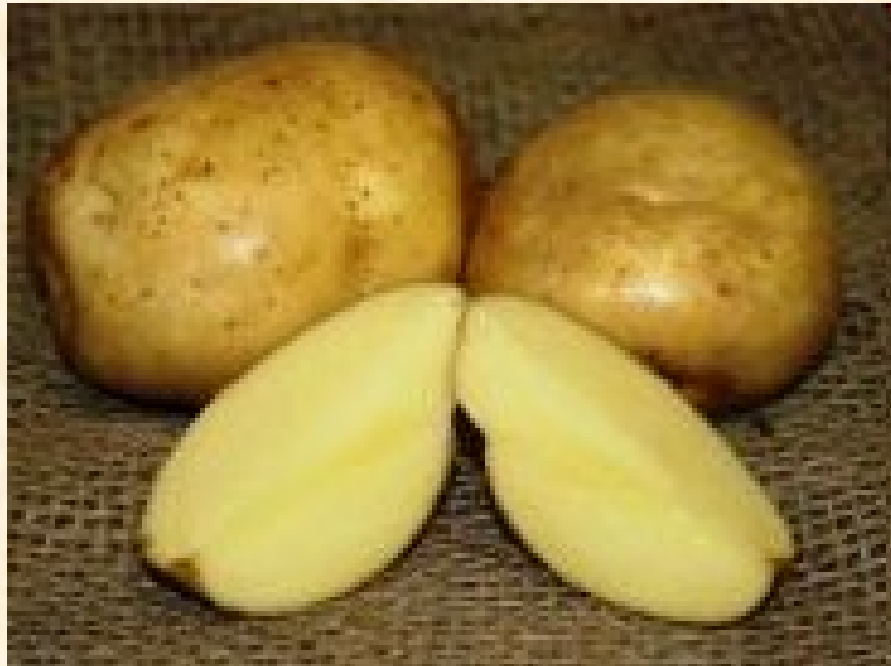
Fosforylace

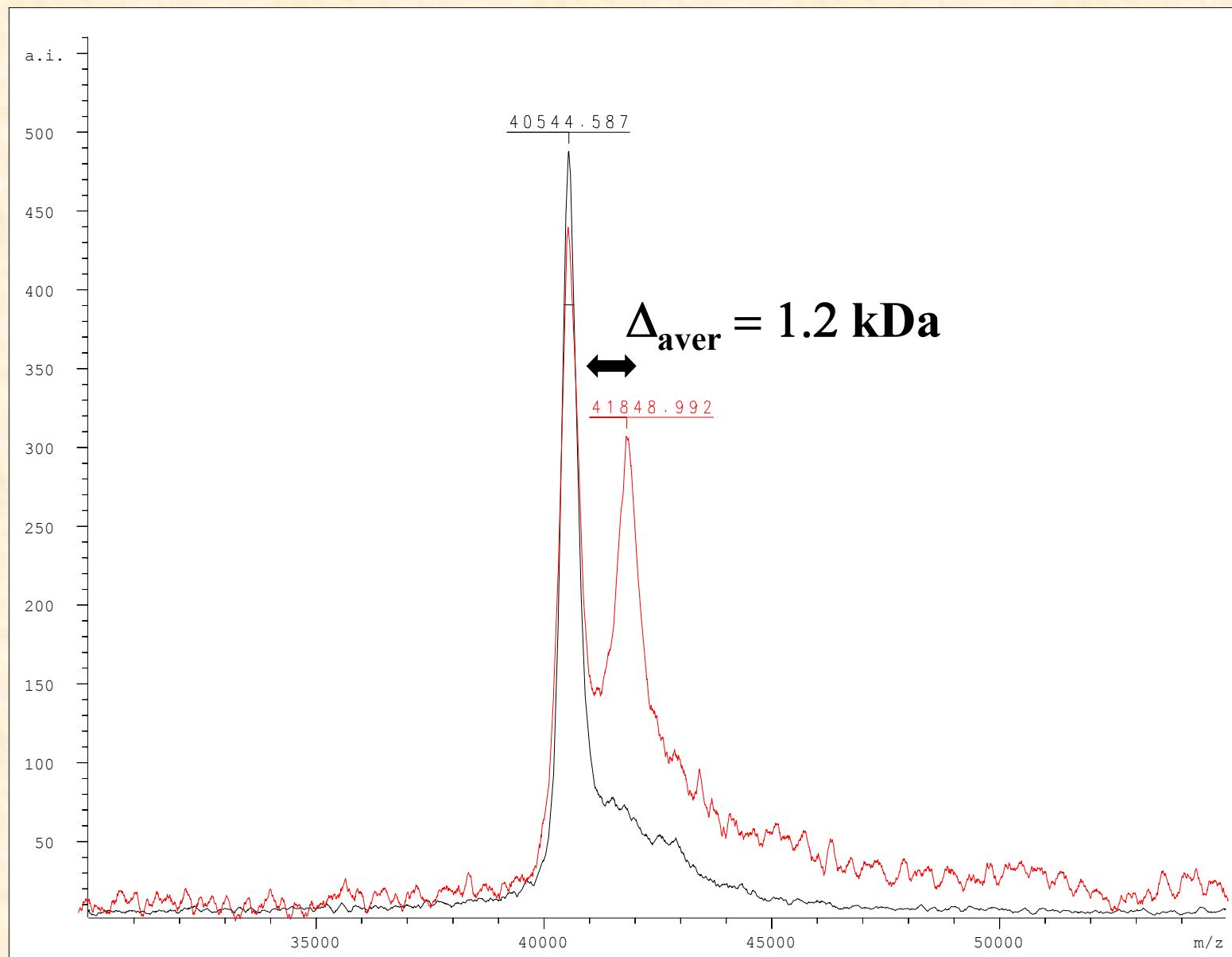


ESI-MS/MS



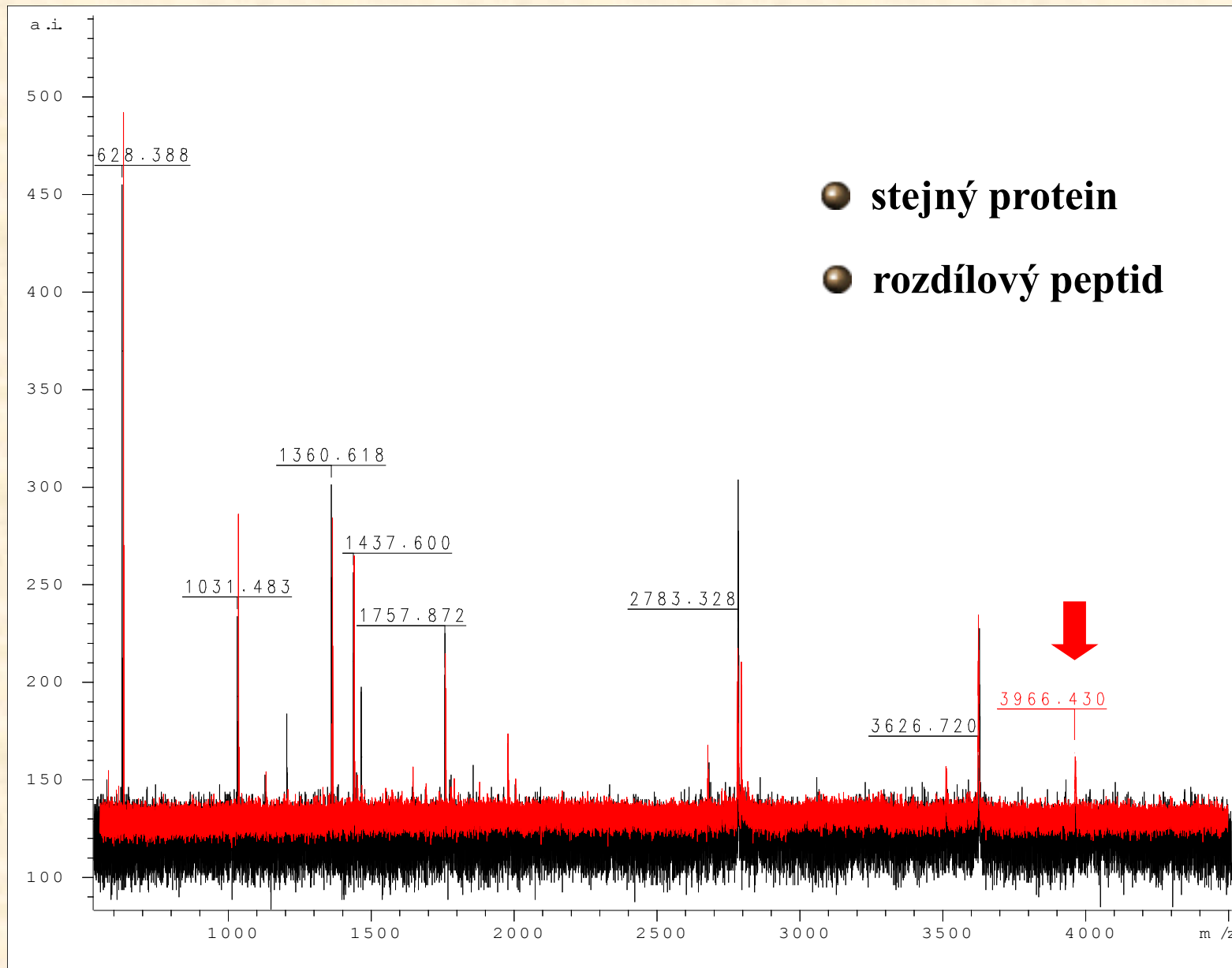
Glykosylace patatinu





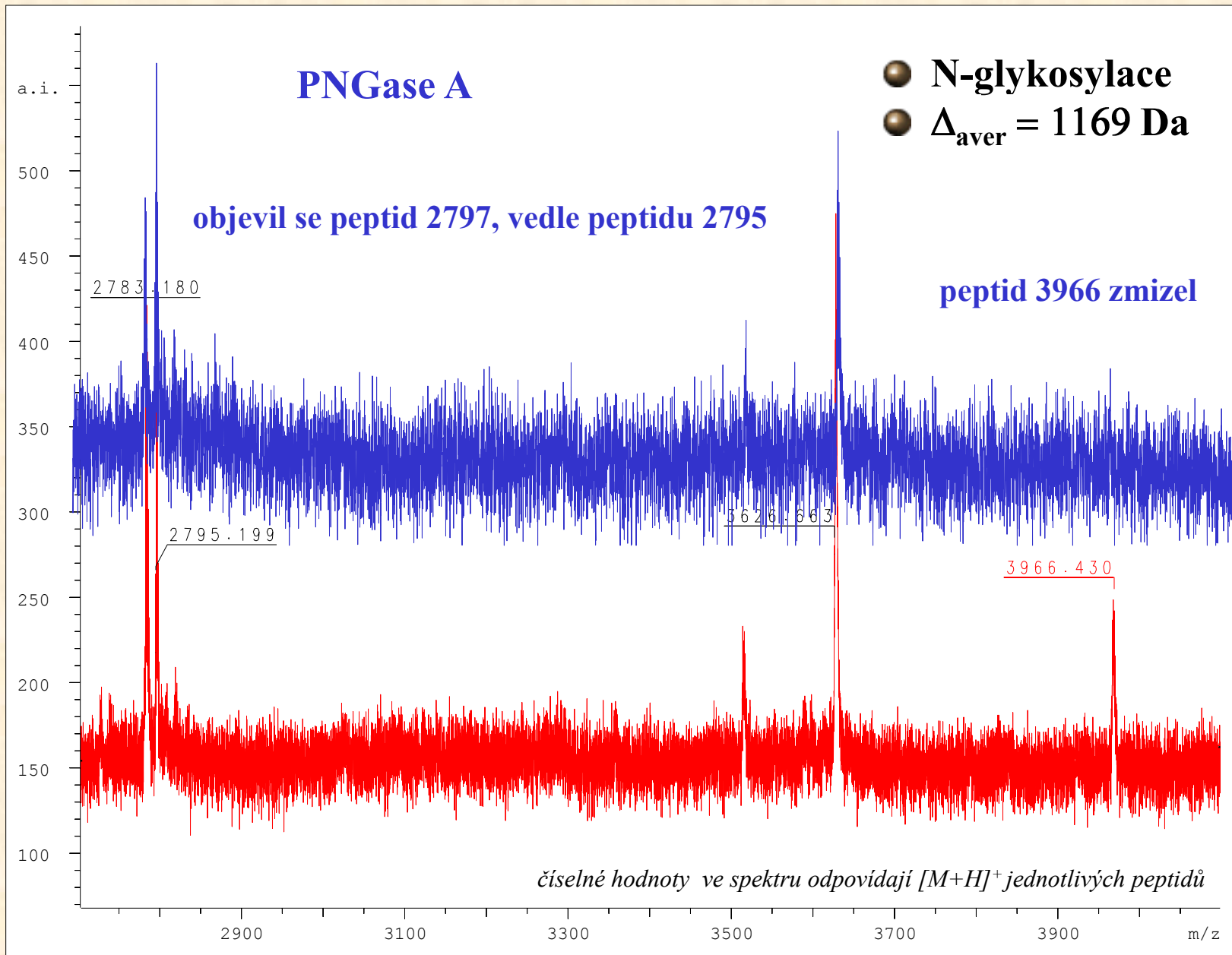
MALDI-MS tryptických digestů

C8202



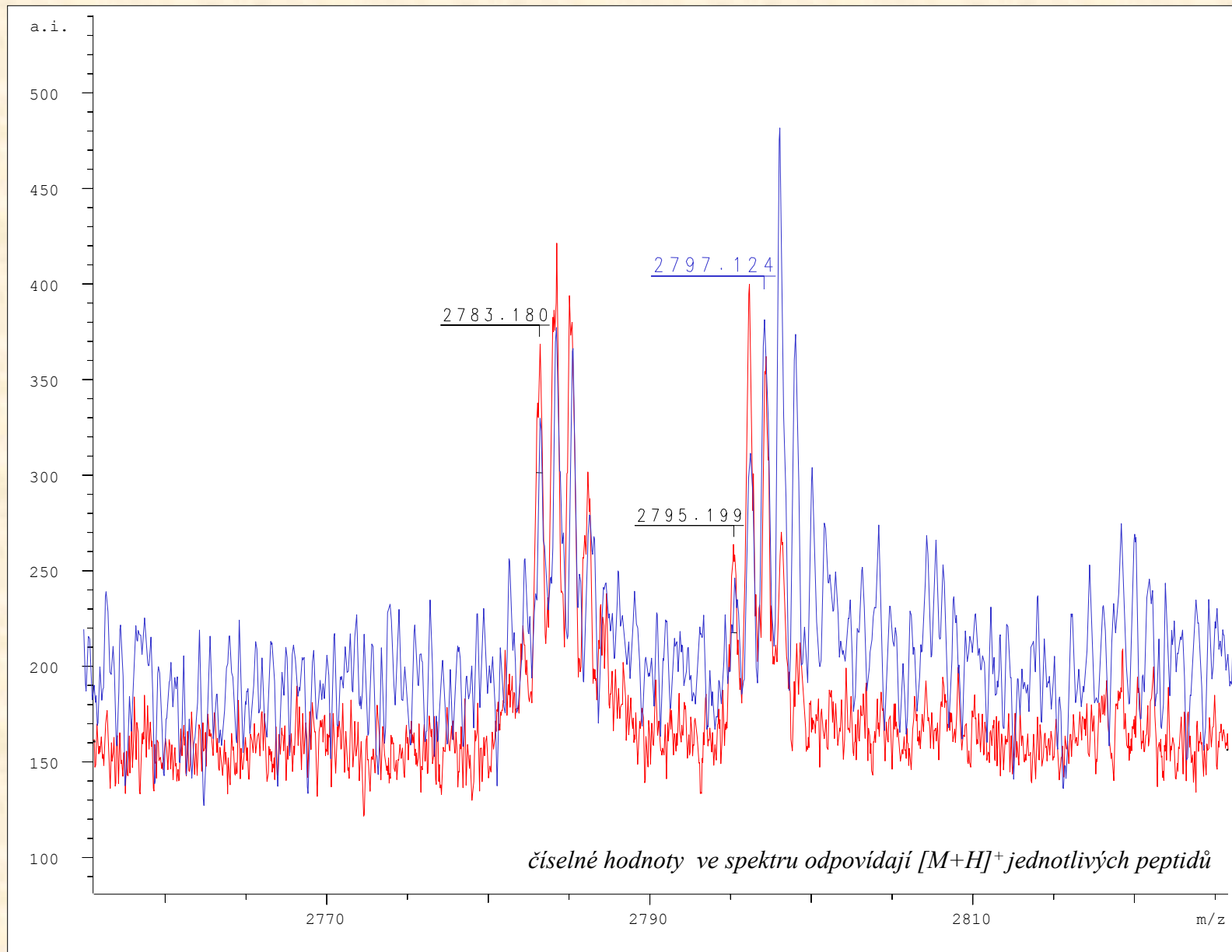
Detail spekter digestů proteinu před a po deglykosylaci

C8202



Detail spekter digestů proteinu **před** a **po** deglykosylaci

C8202



Shrnutí výsledků

tryptický peptid - 2796 Da ...PHIFDYSGS... ,
kde D vzniká z N po deglykosylaci PNGasou A

původní sekvence je tedy ...PHIFNYSGS...
(hmotnost nemodifikovaného peptidu - 2795 Da)

Peptid potvrzen také LCMS/MS analýzou (v glykosylovaném vzorku digestu nebyl nalezen)

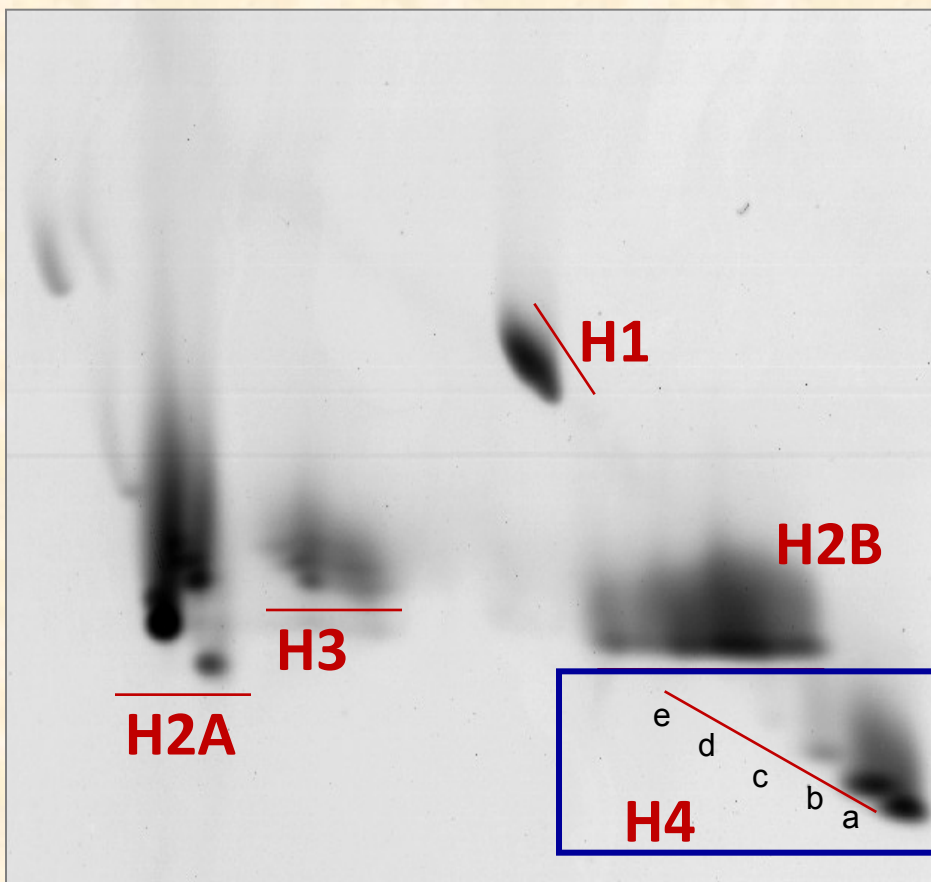
Hmotnost glykanu 1170 Da odpovídá
xylose+fucose+3*mannose+2*N-acetylglukosamin

(U. Sonnewald et al., Planta 179 (1989) 171-180)

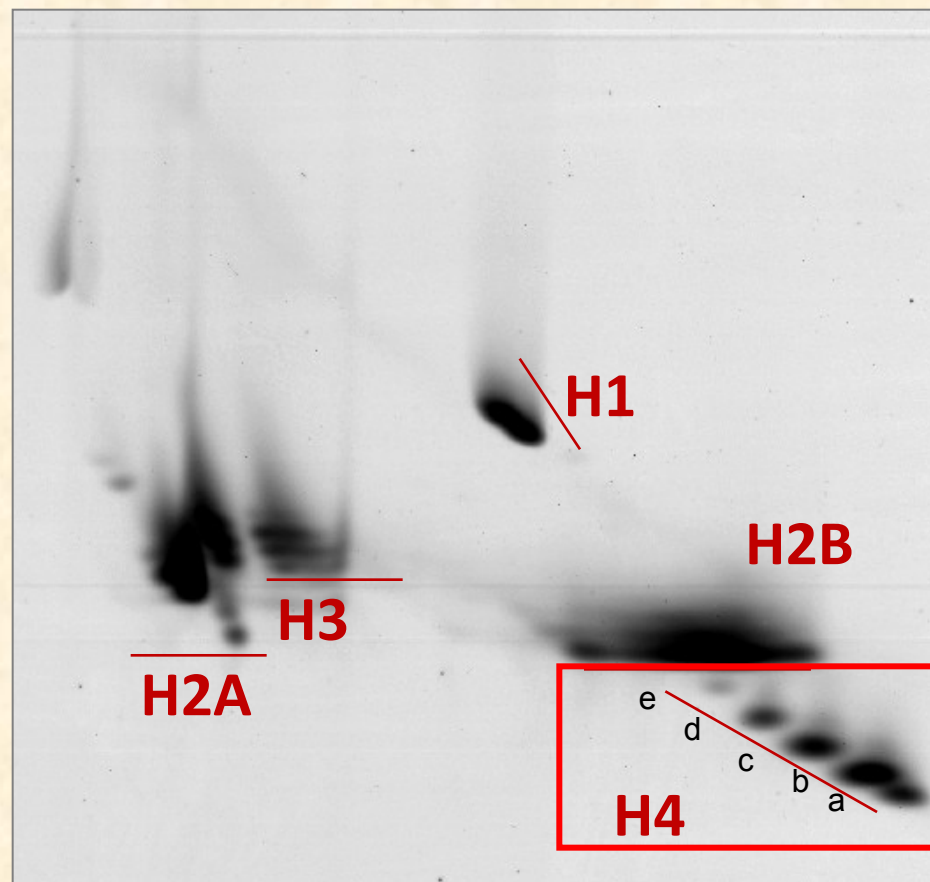
Nebyl dále potvrzen MSMS technikami

2D gelová elektroforéza (AUT-AU) histonových extraktů

bez inhibitoru deacetyláz



s inhibitorem deacetyláz



Histon H4 – detail gelu

C8202

bez inhibitoru deacetyláz

s inhibitorem deacetyláz

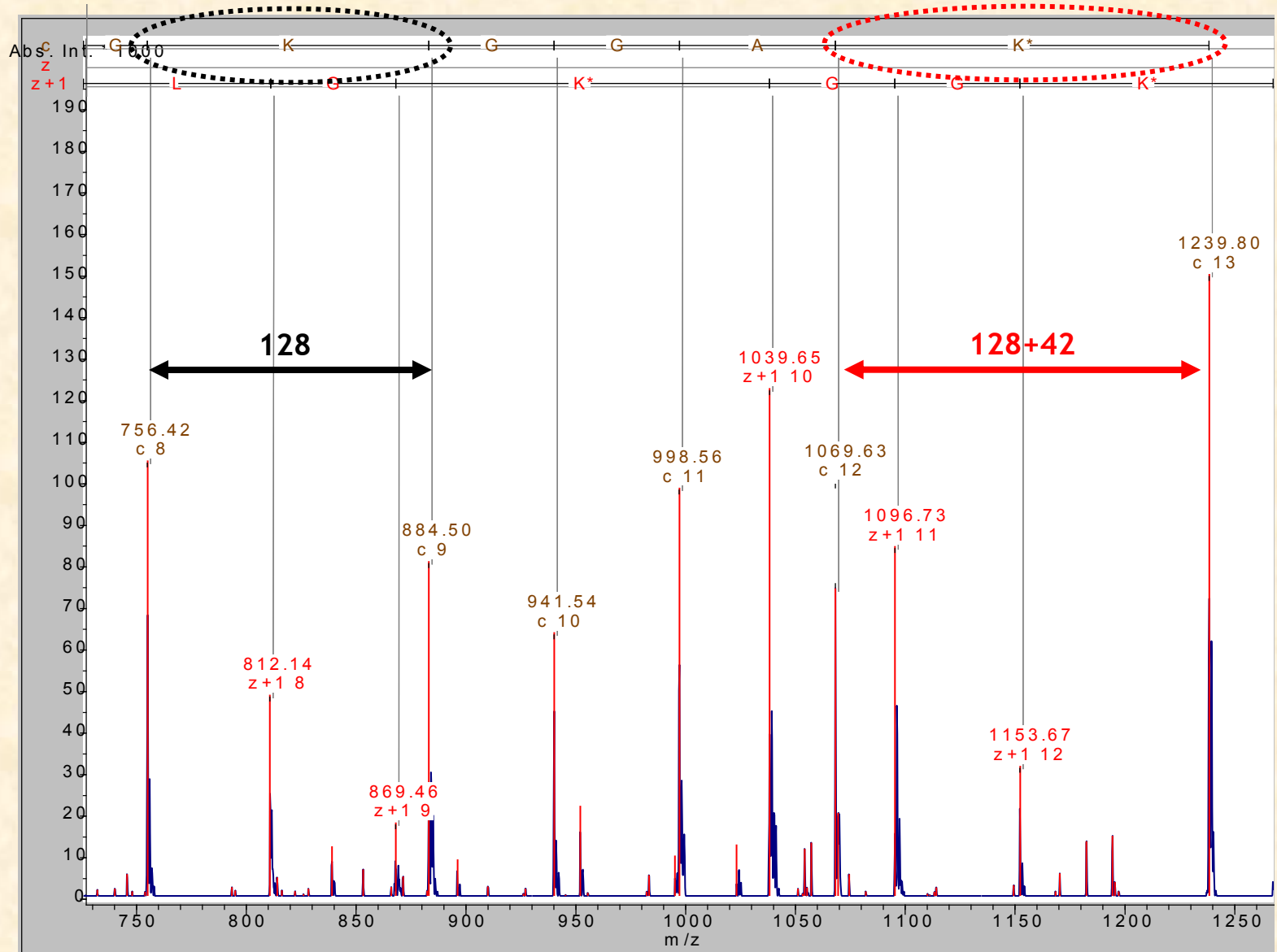


LC-MS/MS analýza

1 MSGRGKGGKGLGKGGAKRHRKVLRDNIQGI TKPAIRRLAR RGGVKRISGL
51 IYEETRGVLK VELENVIRDA YTYTEHAKRK TVTAMDVVYA LKRQGRTLYG
101 FGG

dimetyl

Výřez MS/MS (ETD) spektra peptidu GKGGKG LGKGGAKR (3x Ac)



Kvantifikace proteinů



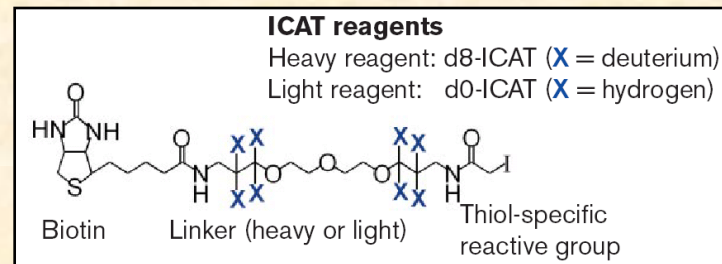
Kvantifikace proteinů pomocí MS

metody s využitím izotopicky značených sloučenin

- absolutní kvantifikace
(přídavek interního standardu o známé koncentraci)

- relativní kvantifikace
(porovnání změny exprese proteinu ve dvou či více vzorcích pomocí značek o různé hmotnosti)

Značky jsou většinou izotopicky značené sloučeniny, které jsou navázány na analyzované proteiny, resp. digestované peptidy nebo jsou do proteinů zabudovány během buněčného růstu.

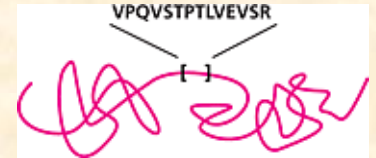


metody založené na statistickém zpracování dat (label free)

Schéma kvantifikačního experimentu *absolutní kvantifikace*

➤ AQUA Peptide Selection

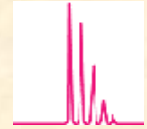
Select an optimal tryptic peptide and stable isotope amino acid from the sequence of your protein of interest



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

Price!!!

Optimize LC-MS/MS separation protocol for quantitation

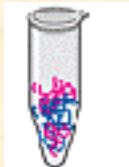


➤ Adding labeled peptide to protein mix

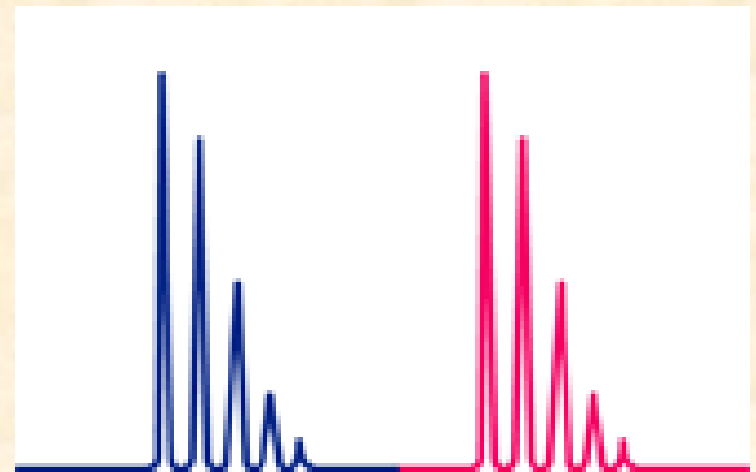


AQUA

➤ Digest



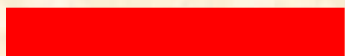
➤ Analyze by LC-MS/MS to quantitate protein of interest



jen pro vybraný protein

Schéma kvantifikačního experimentu *metody relativní kvantifikace*

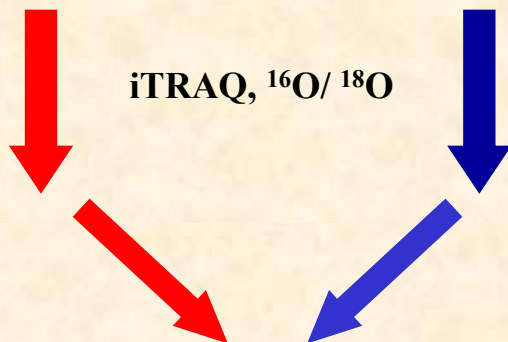
Proteinový vzorek A



Proteinový vzorek B



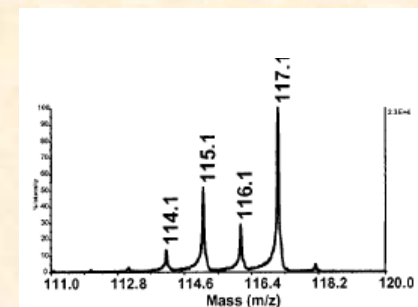
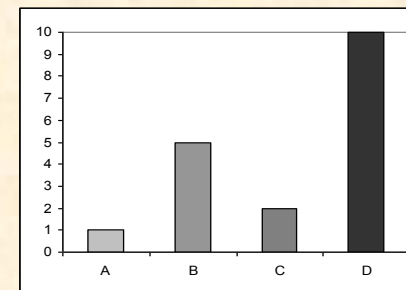
Oddělená digesce vzorků



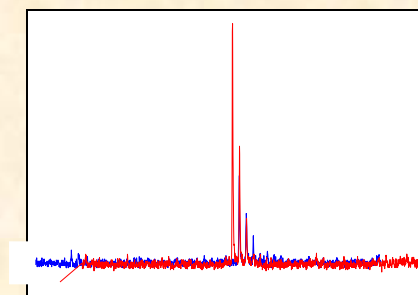
iTRAQ, $^{16}\text{O}/^{18}\text{O}$

LC-MS

LC-MS/MS (iTRAQ)



MS/MS



Stejné peptidy v MS modu neodlišený

A TO JE KONEC

