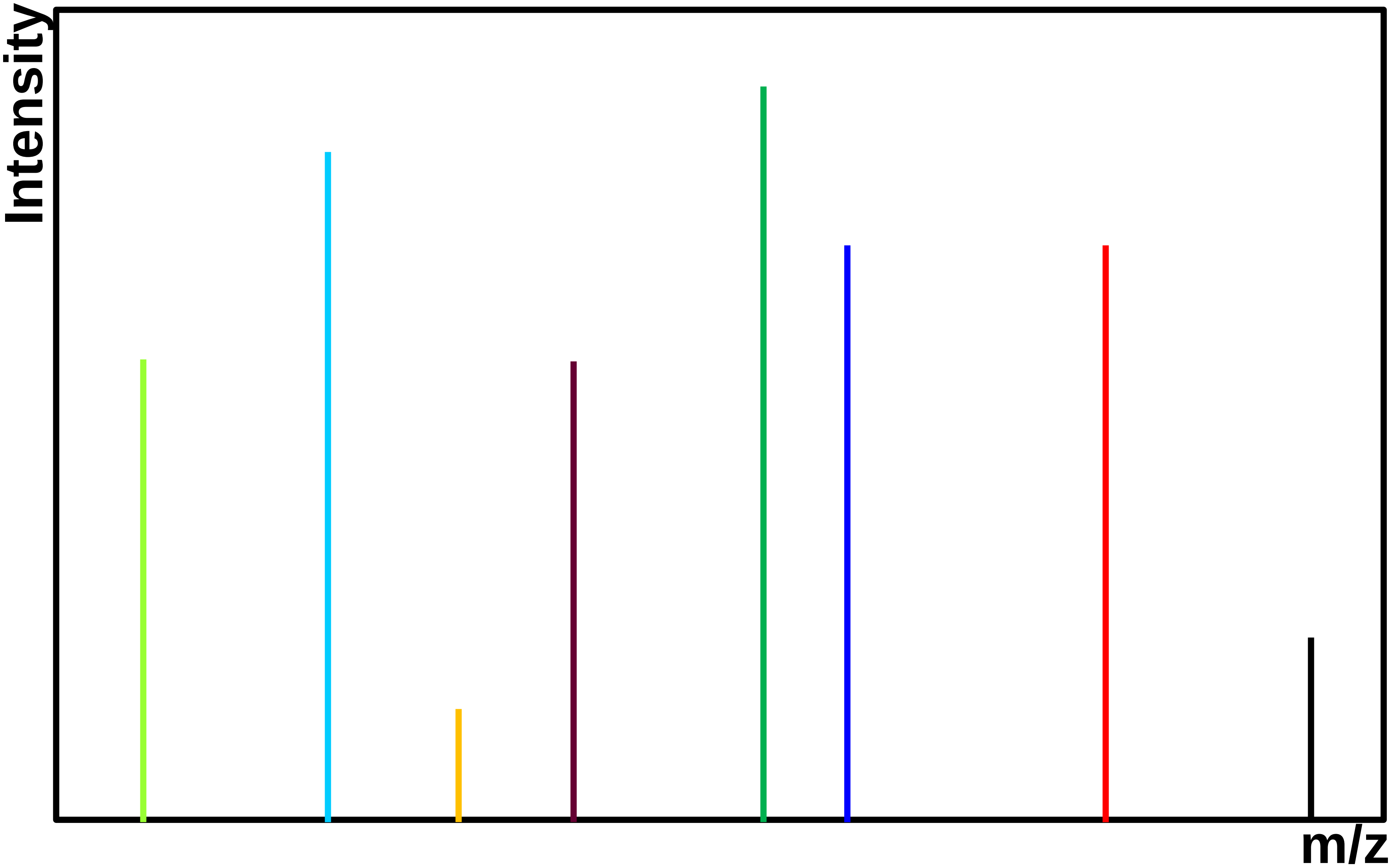
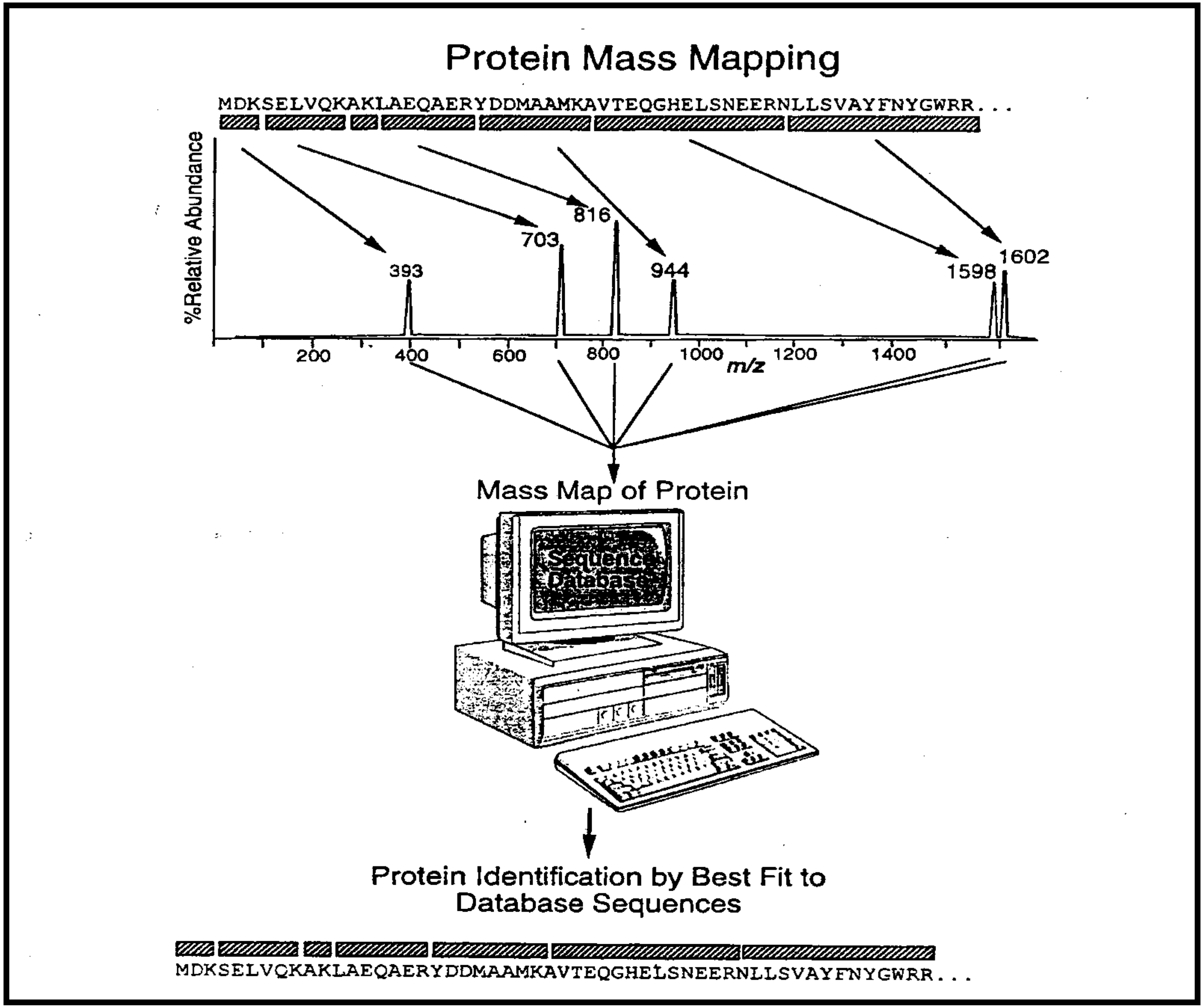
Identifikace jednoho proteinu

ADSMK EETSLINVCPWK VNMGGHDSIYTLR EEIYTLK
MLIILLR TYSHEEDK EWQIDSLAEIR IQPLPMNVSA



Identifikace proteinu - princip



Identifikace proteinů

Protein na gelu nebo roztoku

↓

Štěpení proteinu

↓

Peptidové mapování/
Peptidový fingerprinting

Proteinové štěpení

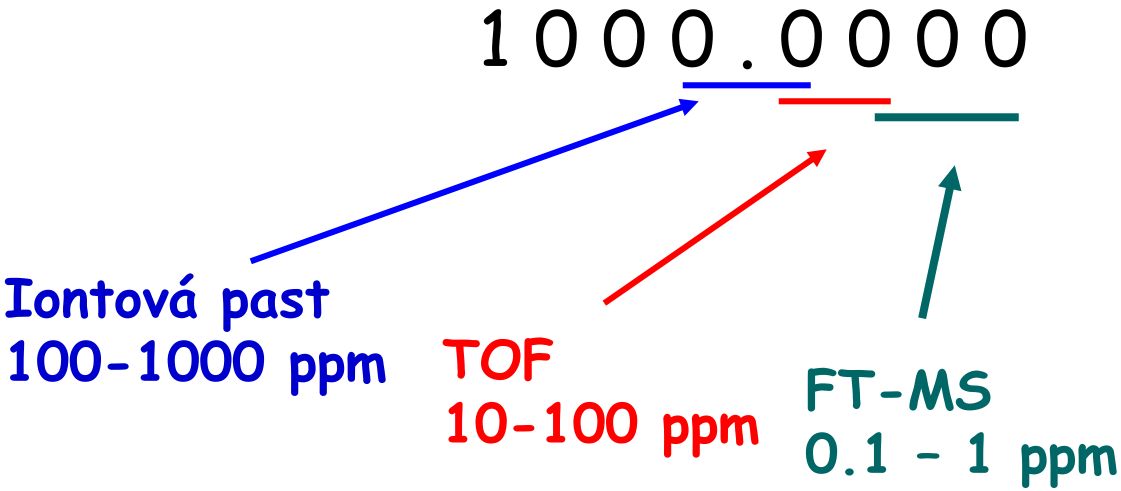
Trypsin	K/R-	\-P
AspN	-D	
GluC	E-, E/D-	\-P
ArgC	R-	\-P
LysC	K-	\-P
CNBr	M-	
Chymotrypsin	*Y/W/K/F-*	*\-P*

Přesnost – správnost měření

- Da – Dalton - 0.1 – 0.0001
- Th – Thompson - 0.1 – 0.0001

- ppm – parts per million – 100 – 0.1

  ppm chyba = 1e6x(naměřená M-teoretická M)/teoretická M
- %



1 0 0 0 . 0 0 0 0

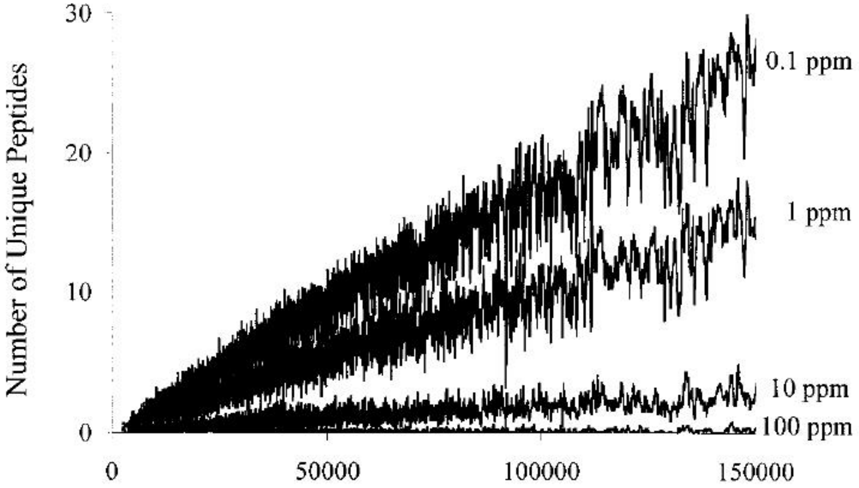
Iontová past
100-1000 ppm

TOF
10-100 ppm

FT-MS
0.1 – 1 ppm

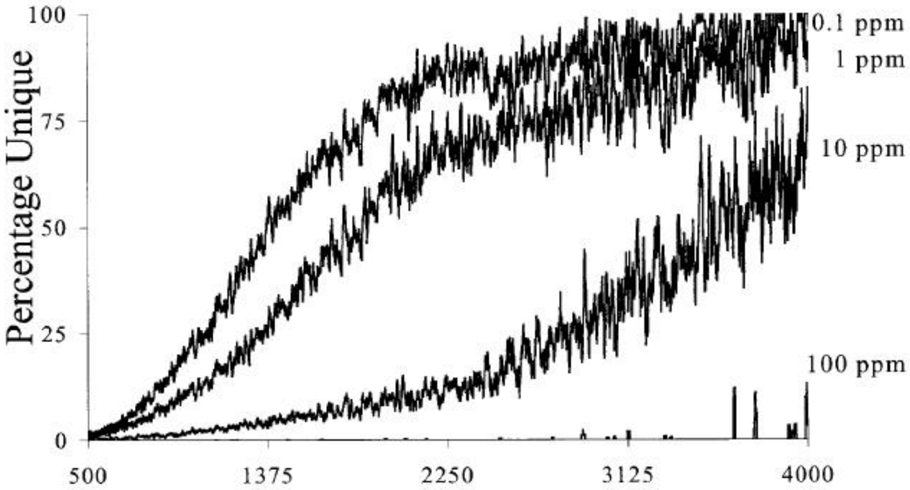
Důležitá je přesnost…

Čím vyšší přesnost měření, tím méně peptidů potřebujeme pro bezchybnou identifikaci



Počet unikátních peptidů jako funkce velikosti proteinu

Unikátnost peptidu jako funkce jeho MW



Conrads TP et al . Anal Chem 2002, 72: 3349.

Peptidové mapování

MASCOT – *Matrix Science - http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF*

MS-Fit – Protein Prospector, *UCSF* **–** http://prospector.ucsf.edu/cgi-bin/msform.cgi?form=msfitstandard

Pro-Found – *Rockefeller University* *http://prowl.rockefeller.edu/prowl-cgi/profound.exe*
http://hs2.proteome.ca/prowl/profound/control.html

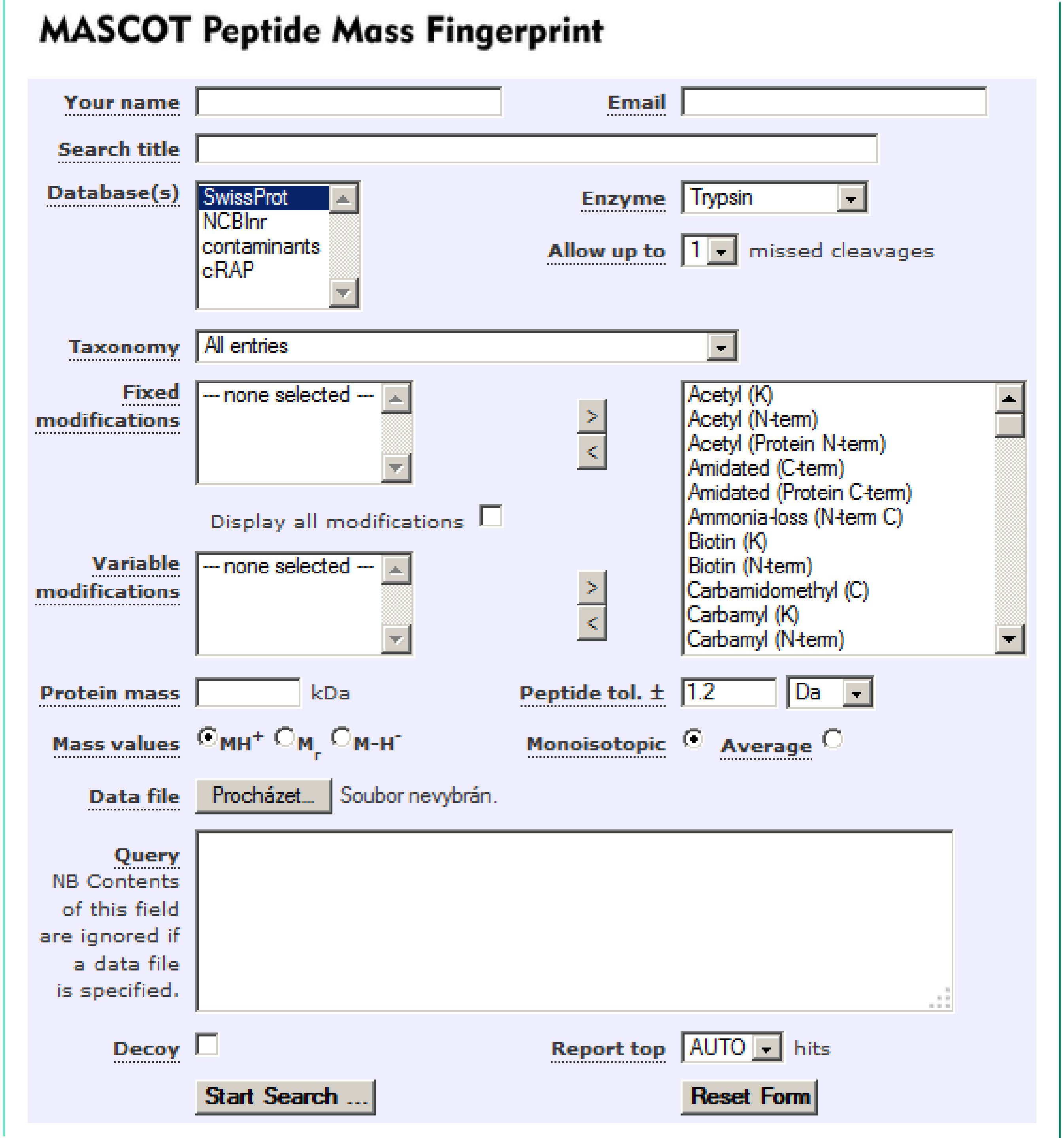
Older engines, not available or unstable today

PepIdent – *SWISS-PROT*

PepSEARCH – *EMBL* ~ PepSea – *Protana*

pepMAPPER – *UMIST*

MOWSE – *MRC/HGMP*



MASCOT Peptide Mass Fingerprint

Your name

Email

Search title

Database(s): SwissProt, NCBInr, contaminants, cRAP

Enzyme: Trypsin

Allow up to 1 missed cleavages

Taxonomy: All entries

Fixed modifications: --- none selected ---

Display all modifications ☐

Variable modifications: --- none selected ---

Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (C-term)
Amidated (Protein C-term)
Ammonia-loss (N-term C)
Biotin (K)
Biotin (N-term)
Carbamidomethyl (C)
Carbamyl (K)
Carbamyl (N-term)

Protein mass kDa

Peptide tol. ± 1.2 Da

Mass values ● MH^+ ○ M_r ○ $M-H^-$

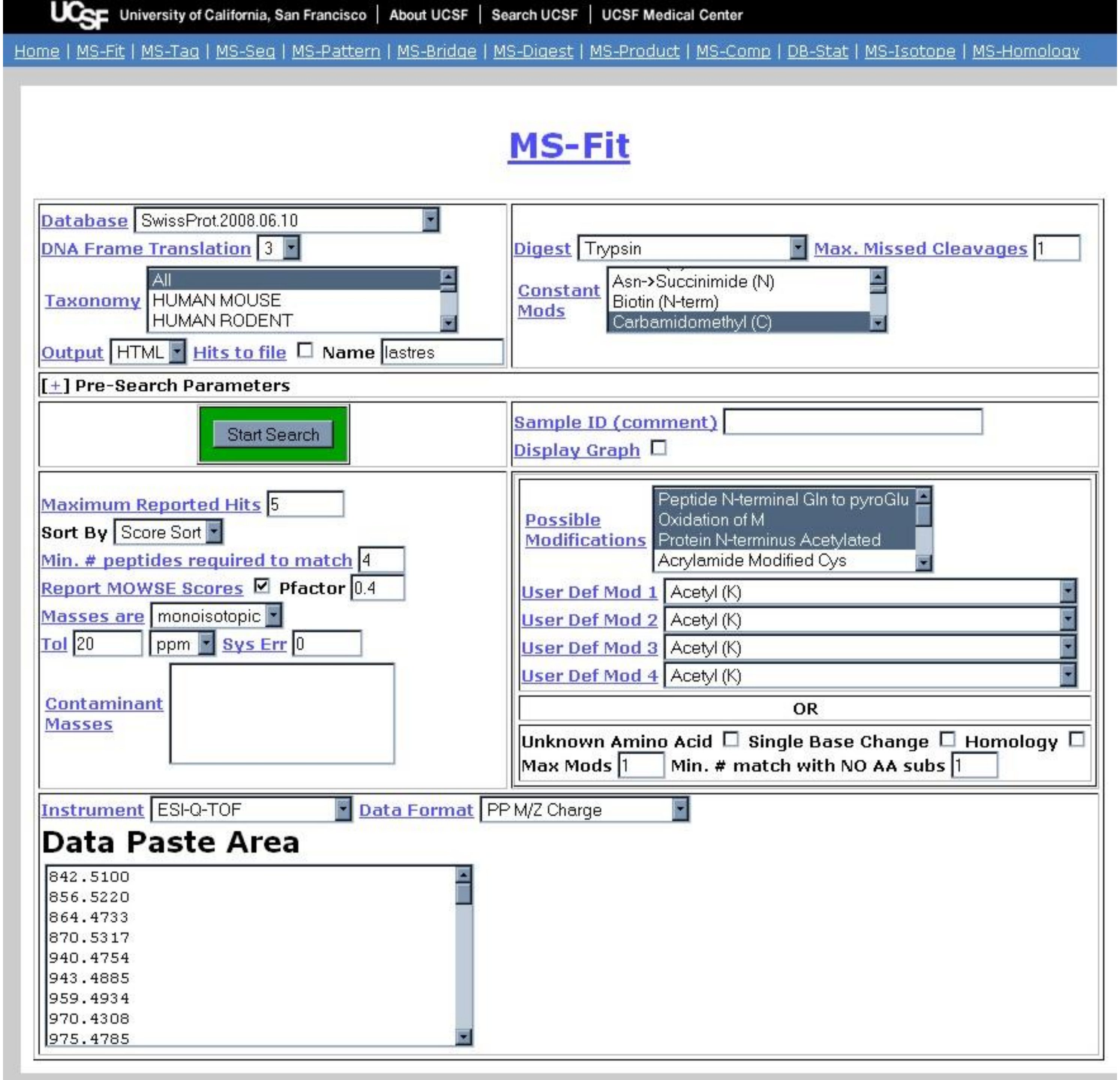
Monoisotopic ● Average ○

Data file Procházet... Soubor nevybrán.

Query
NB Contents of this field are ignored if a data file is specified.

Decoy ☐

Report top AUTO hits



MS-Fit

Database SwissProt.2008.06.10
DNA Frame Translation 3
Taxonomy All
HUMAN MOUSE
HUMAN RODENT
Output HTML Hits to file ☐ Name lastres

Digest Trypsin Max. Missed Cleavages 1
Constant Mods Asn->Succinimide (N)
Biotin (N-term)
Carbamidomethyl (C)

[+] Pre-Search Parameters

Sample ID (comment)
Display Graph ☐

Maximum Reported Hits 5
Sort By Score Sort
Min. # peptides required to match 4
Report MOWSE Scores ☑ Pfactor 0.4
Masses are monoisotopic
Tol 20 ppm Sys Err 0
Contaminant Masses

Possible Modifications Peptide N-terminal Gln to pyroGlu
Oxidation of M
Protein N-terminus Acetylated
Acrylamide Modified Cys
User Def Mod 1 Acetyl (K)
User Def Mod 2 Acetyl (K)
User Def Mod 3 Acetyl (K)
User Def Mod 4 Acetyl (K)

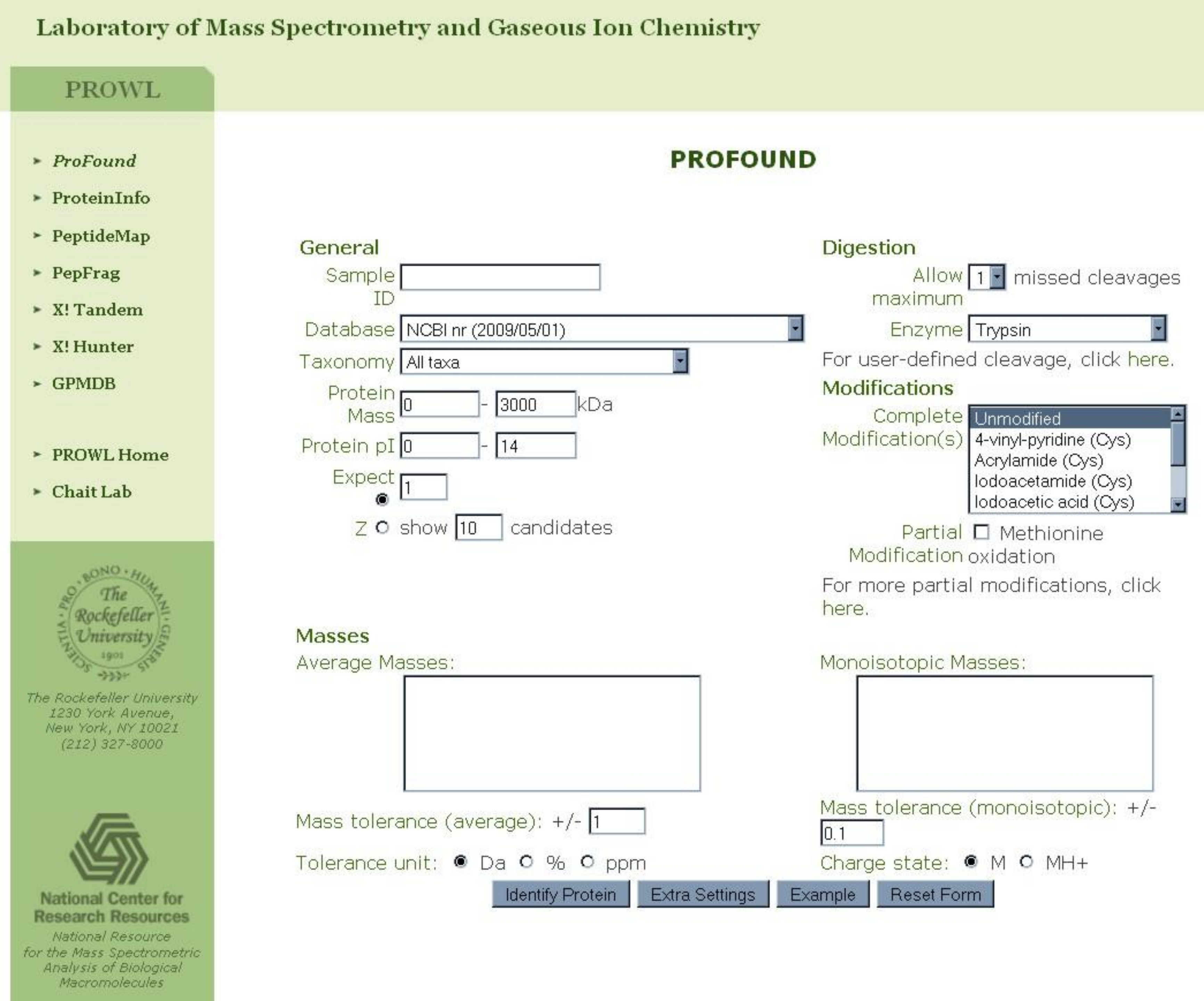
OR

Unknown Amino Acid ☐ Single Base Change ☐ Homology ☐
Max Mods 1 Min. # match with NO AA subs 1

Instrument ESI-Q-TOF Data Format PP M/Z Charge

Data Paste Area

```
842.5100
856.5220
864.4733
870.5317
940.4754
943.4885
959.4934
970.4308
975.4785
```



Laboratory of Mass Spectrometry and Gaseous Ion Chemistry

PROFOUND

General

Sample ID:

Database: NCBI nr (2009/05/01)

Taxonomy: All taxa

Protein Mass: 0 - 3000 kDa

Protein pI: 0 - 14

● Expect: 1

○ Z show 10 candidates

Digestion

Allow maximum: 1 missed cleavages

Enzyme: Trypsin

For user-defined cleavage, click here.

Modifications

Complete Modification(s): Unmodified

Partial Modification: ☐ Methionine oxidation

For more partial modifications, click here.

Masses

Average Masses:

Monoisotopic Masses:

Mass tolerance (average): +/- 1

Mass tolerance (monoisotopic): +/- 0.1

Tolerance unit: ● Da ○ % ○ ppm

Charge state: ● M ○ MH+

Identifikace proteinů peptidovým mapováním

- Než začneme posbíráme maximum informací
  - zdroj vzorku (u tkáňových kultur zjistíme zda nebyly modifikovány virem)
  - postup zpracování, složení pufrů – vliv na modifikace proteinu
  - fyzikálně chemické informace – MW, pI
  - potenciální modifikace
  - typ štěpení, kvalita a čistota enzymu, čistota a stáří chemikálií….

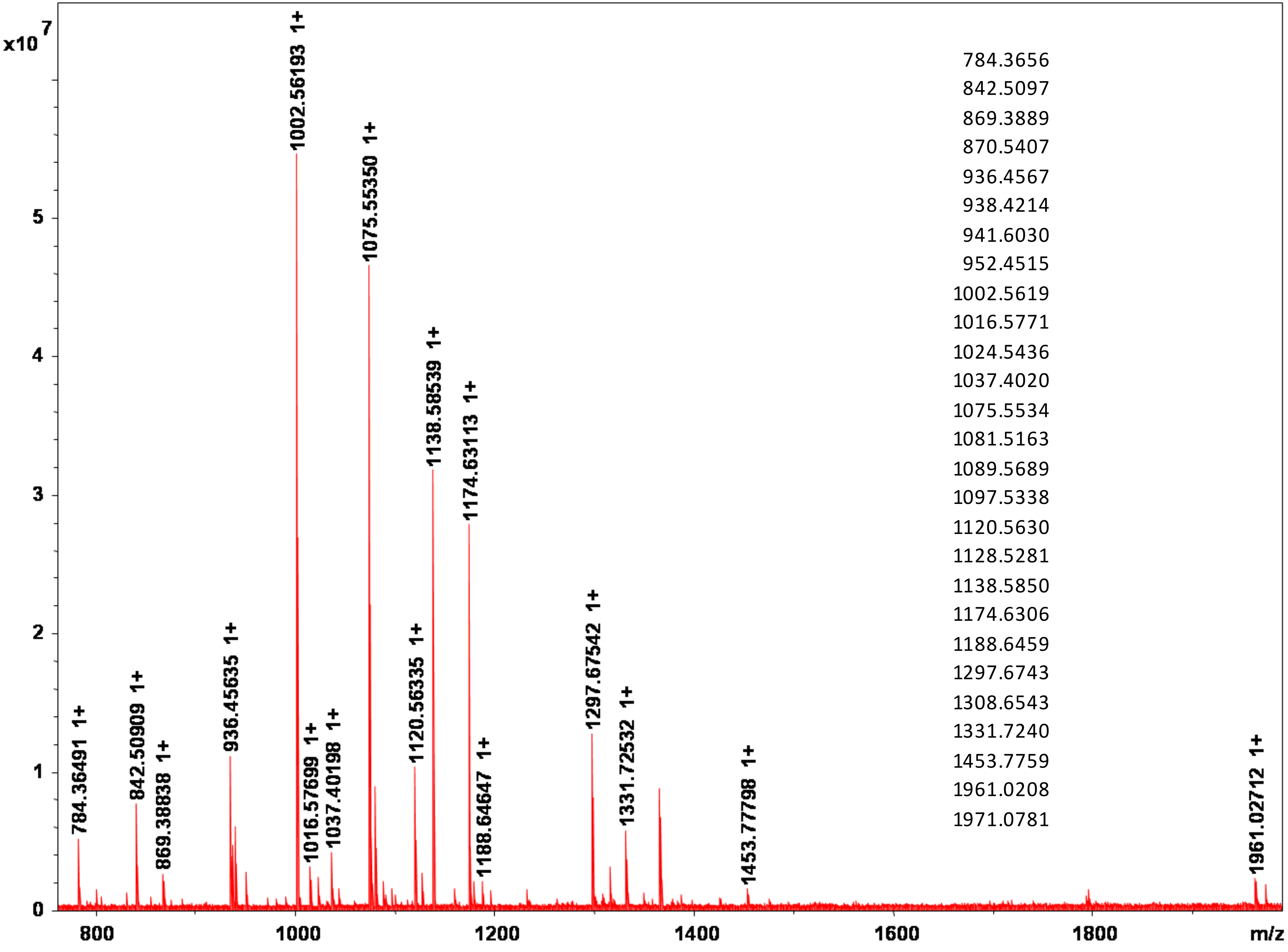
- Důkladně prohlédneme spektrum, odstraníme známé kontaminace a adukty (pozor na odstranění píků ze vzorku – důležitou roli hraje přesnost a rozlišení)
  Keller BO. Anal Chim Acta 2008, 627:71.

- První nástřel provádíme s volnějšími parametry (zkrácené nebo modifikované proteiny mají jiné MW a pI než ty v databázi, organismus není zcela osekvenovaný, atd.)
  pak upřesňujeme (užší taxonomie, MW a pI interval, přesnost, modifikace, atd.) a sledujeme jak se vyvíjí skóre a počet přiřazených peptidů

- Zvolíme jeden z prohledávacích programů
  ProFound – http://prowl.rockefeller.edu/prowl-cgi/profound.exe
  MASCOT – http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF
  MsFit – http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msfitstandard

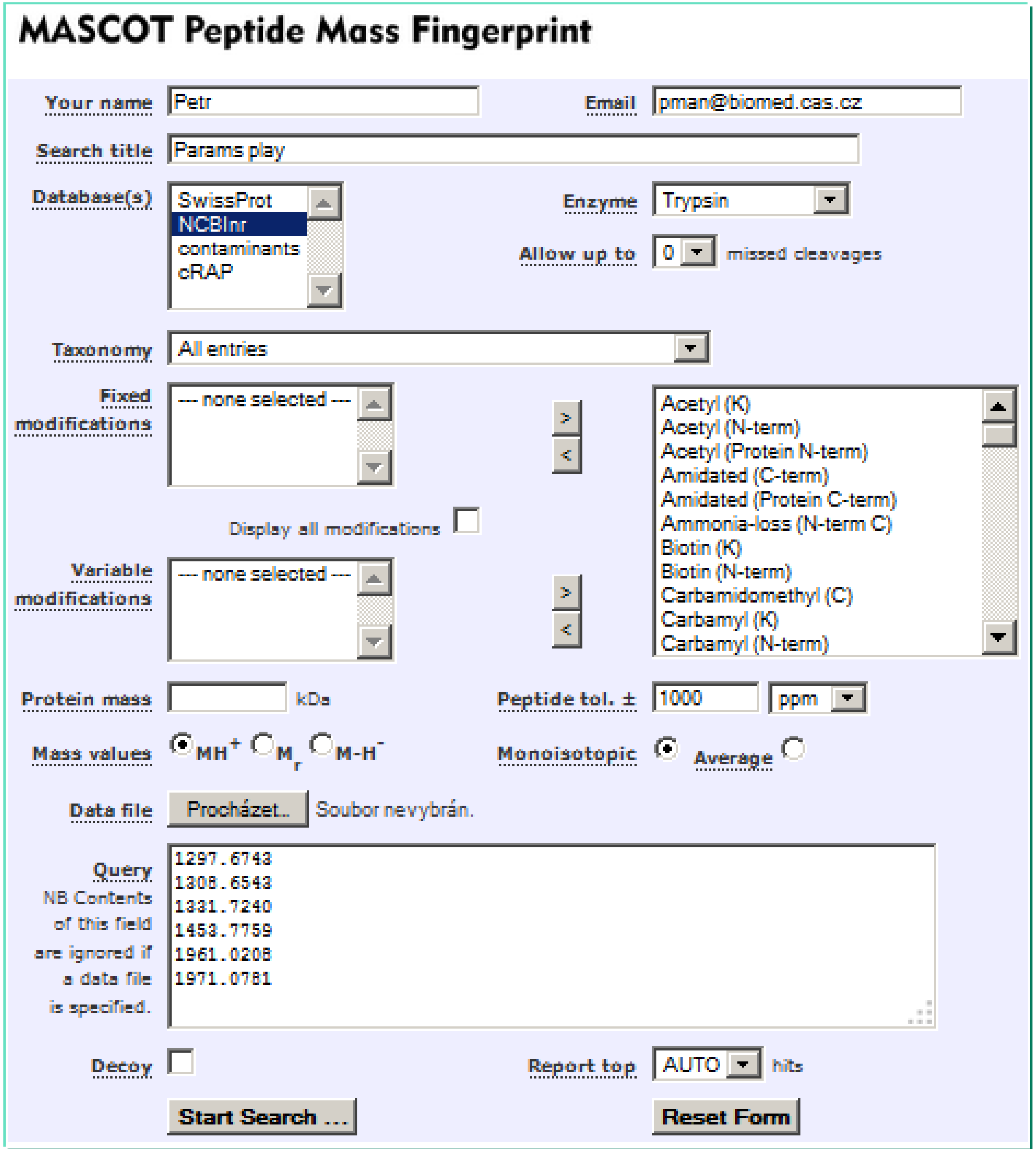
- Nepřiřazené píky buď dále prohledáváme (směs proteinů) nebo se je snažíme přiřadit (nespecifické štěpení, procesing proteinu). Sledujeme i spektrum – je nepřiřazený signál intenzivní nebo ne?

Peptidové mapování – vliv parametrů



784.3656
842.5097
869.3889
870.5407
936.4567
938.4214
941.6030
952.4515
1002.5619
1016.5771
1024.5436
1037.4020
1075.5534
1081.5163
1089.5689
1097.5338
1120.5630
1128.5281
1138.5850
1174.6306
1188.6459
1297.6743
1308.6543
1331.7240
1453.7759
1961.0208
1971.0781

Peptidové mapování – vliv parametrů



MASCOT Peptide Mass Fingerprint

Your name: Petr

Email: pman@biomed.cas.cz

Search title: Params play

Database(s): SwissProt, NCBInr (selected), contaminants, cRAP

Enzyme: Trypsin

Allow up to: 0 missed cleavages

Taxonomy: All entries

Fixed modifications: --- none selected ---

Display all modifications ☐

Variable modifications: --- none selected ---

Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (C-term)
Amidated (Protein C-term)
Ammonia-loss (N-term C)
Biotin (K)
Biotin (N-term)
Carbamidomethyl (C)
Carbamyl (K)
Carbamyl (N-term)

Protein mass: kDa

Peptide tol. ±: 1000 ppm

Mass values: ● MH^+ ○ M_r ○ $M-H^-$

Monoisotopic ● Average ○

Data file: Procházet... Soubor nevybrán.

Query (NB Contents of this field are ignored if a data file is specified.):

```
1297.6743
1308.6543
1331.7240
1453.7759
1961.0208
1971.0781
```

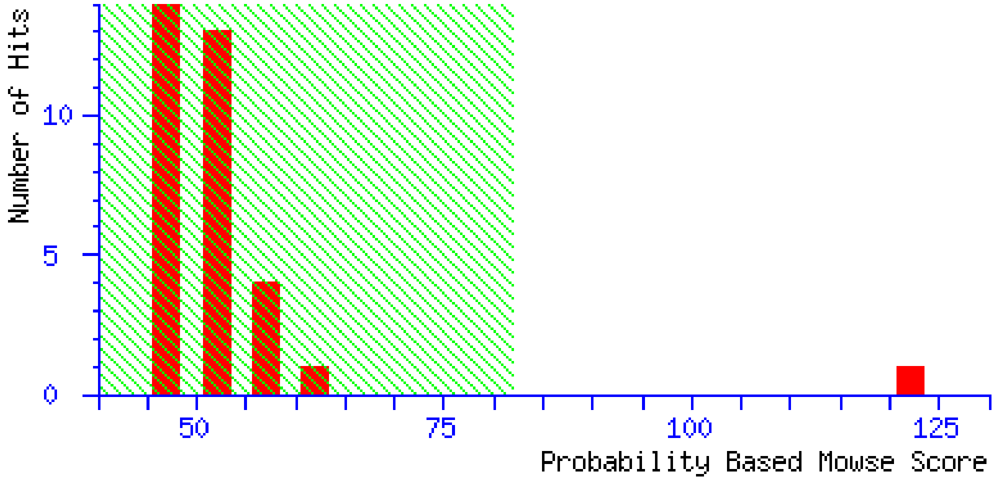
Decoy ☐

Report top: AUTO hits

Peptidové mapování – vliv parametrů

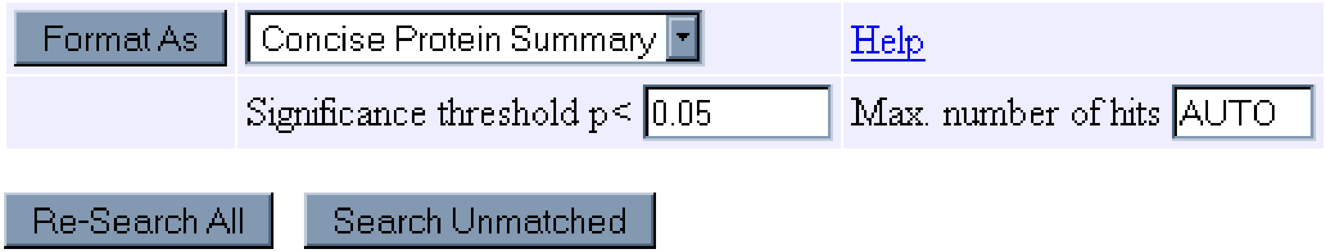
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant ($p<0.05$).



Database	**NCBI**
Taxonomy	**All taxa**
Missed Cleavages	**0**
Fixed Mods	**0**
Variable Mods	**0**
Peptide tol	**1000**

Concise Protein Summary Report



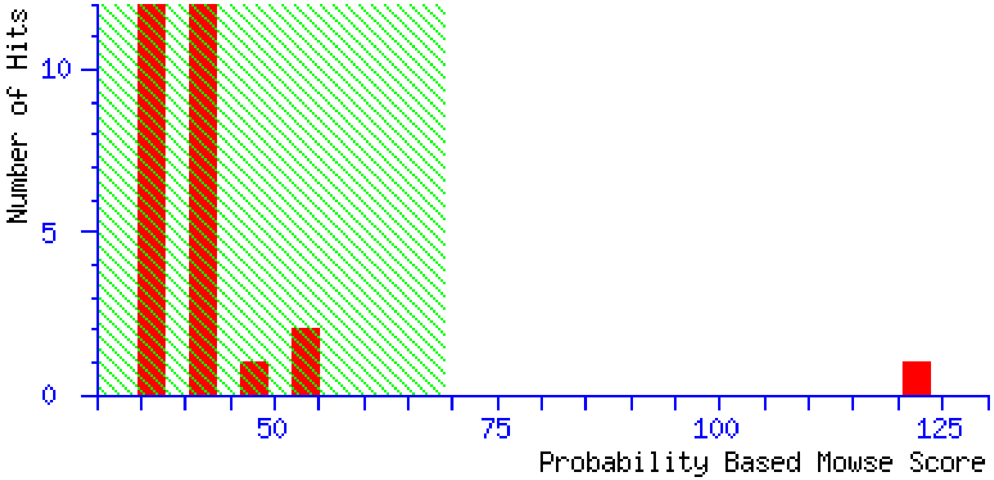
Significance threshold p< 0.05
Max. number of hits AUTO

```
1.     gi|15900172     Mass: 34186     Score: 122     Expect: 6e-06   Queries matched: 12
       DNA-directed RNA polymerase subunit alpha [Streptococcus pneumoniae TIGR4]
```

Peptidové mapování – vliv parametrů

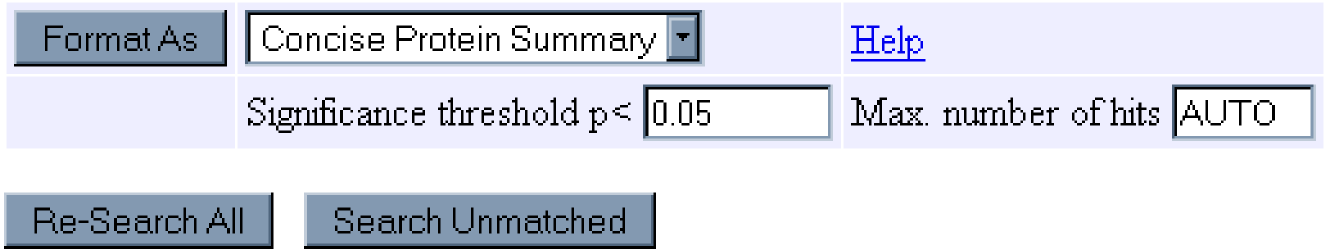
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 69 are significant (p<0.05).



Database	**SwissProt**
Taxonomy	**All taxa**
Missed Cleavages	**0**
Fixed Mods	**0**
Variable Mods	**0**
Peptide tol	**1000**

Concise Protein Summary Report



Significance threshold p< 0.05

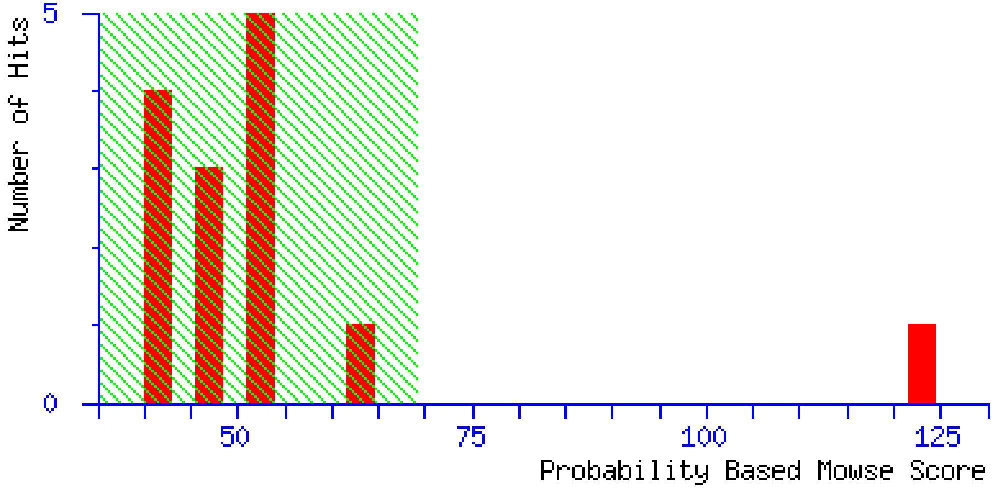
Max. number of hits AUTO

```
1.      RPOA_STRP2      Mass: 34186     Score: 122      Expect: 3.1e-07  Queries matched: 12
        DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2
```

Peptidové mapování – vliv parametrů

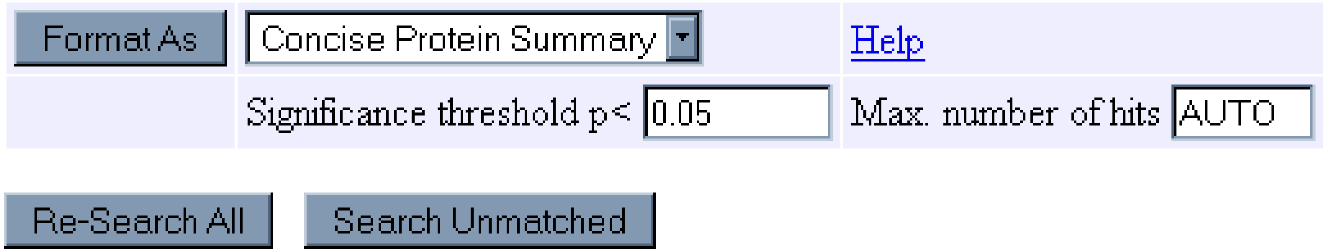
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 69 are significant ($p<0.05$).



Database	**SwissProt**
Taxonomy	**All taxa**
Missed Cleavages	**1**
Fixed Mods	**0**
Variable Mods	**0**
Peptide tol	**1000**

Concise Protein Summary Report

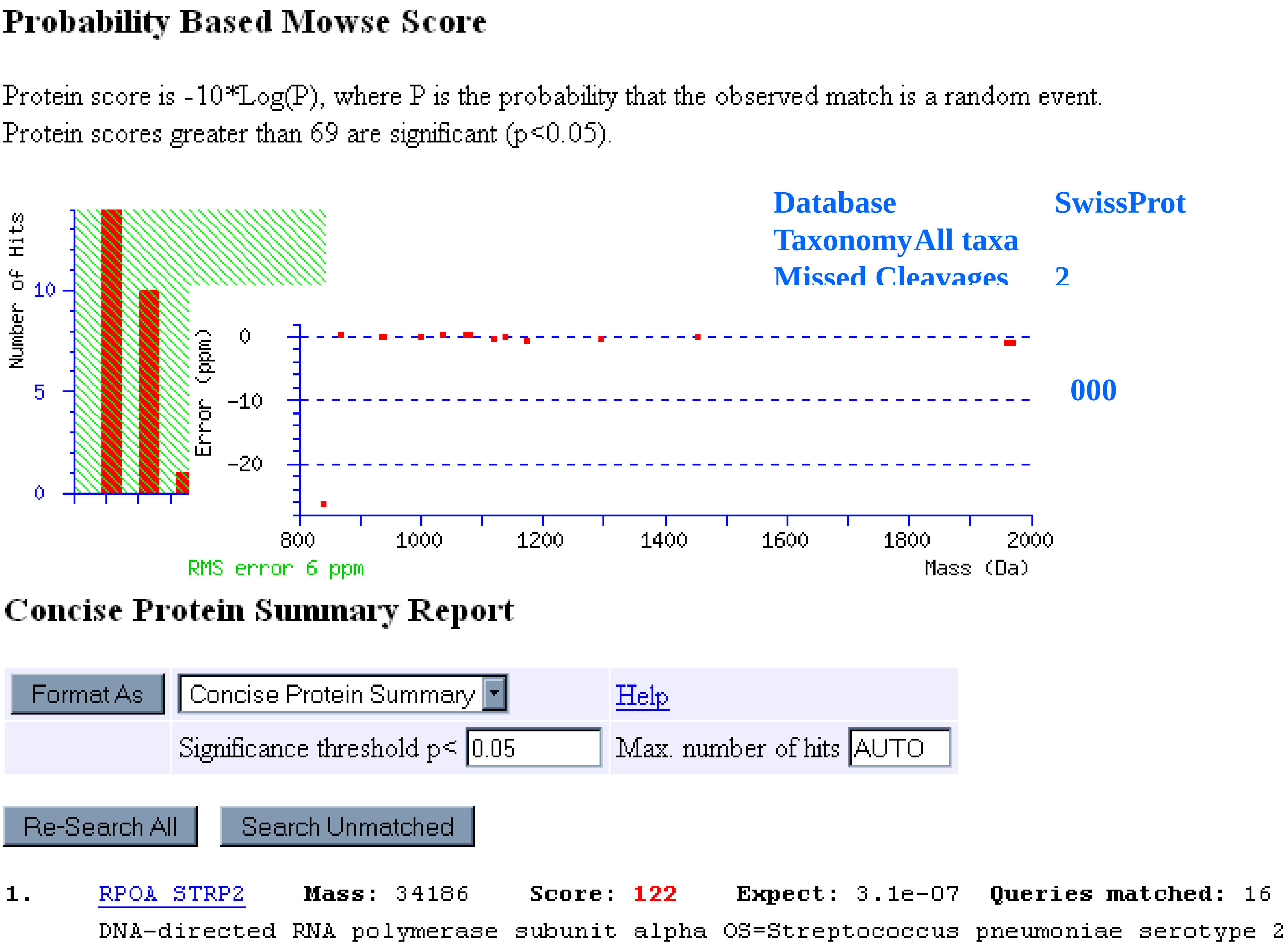


Significance threshold p< 0.05

Max. number of hits AUTO

```
1.      RPOA_STRP2      Mass: 34186     Score: 123     Expect: 2.5e-07   Queries matched: 15
        DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2
```

Peptidové mapování - vliv parametrů



Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 69 are significant (p<0.05).

Database **SwissProt**
Taxonomy **All taxa**
Missed Cleavages **2**

000

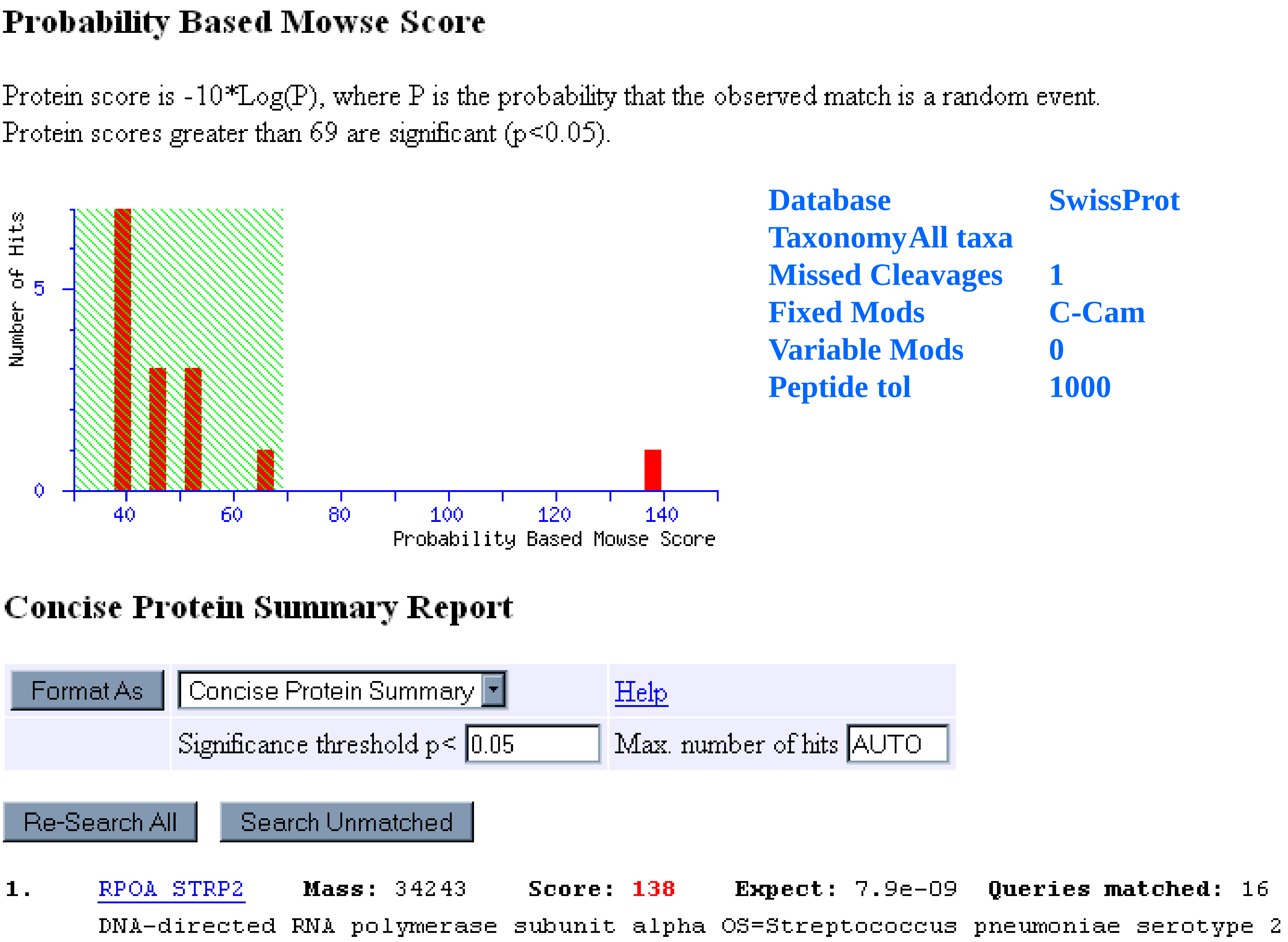
Concise Protein Summary Report

Significance threshold p< 0.05

Max. number of hits AUTO

1. **RPOA_STRP2** **Mass:** 34186 **Score:** 122 **Expect:** 3.1e-07 **Queries matched:** 16
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2

Peptidové mapování – vliv parametrů



Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 69 are significant (p<0.05).

Database	**SwissProt**
Taxonomy	**All taxa**
Missed Cleavages	**1**
Fixed Mods	**C-Cam**
Variable Mods	**0**
Peptide tol	**1000**

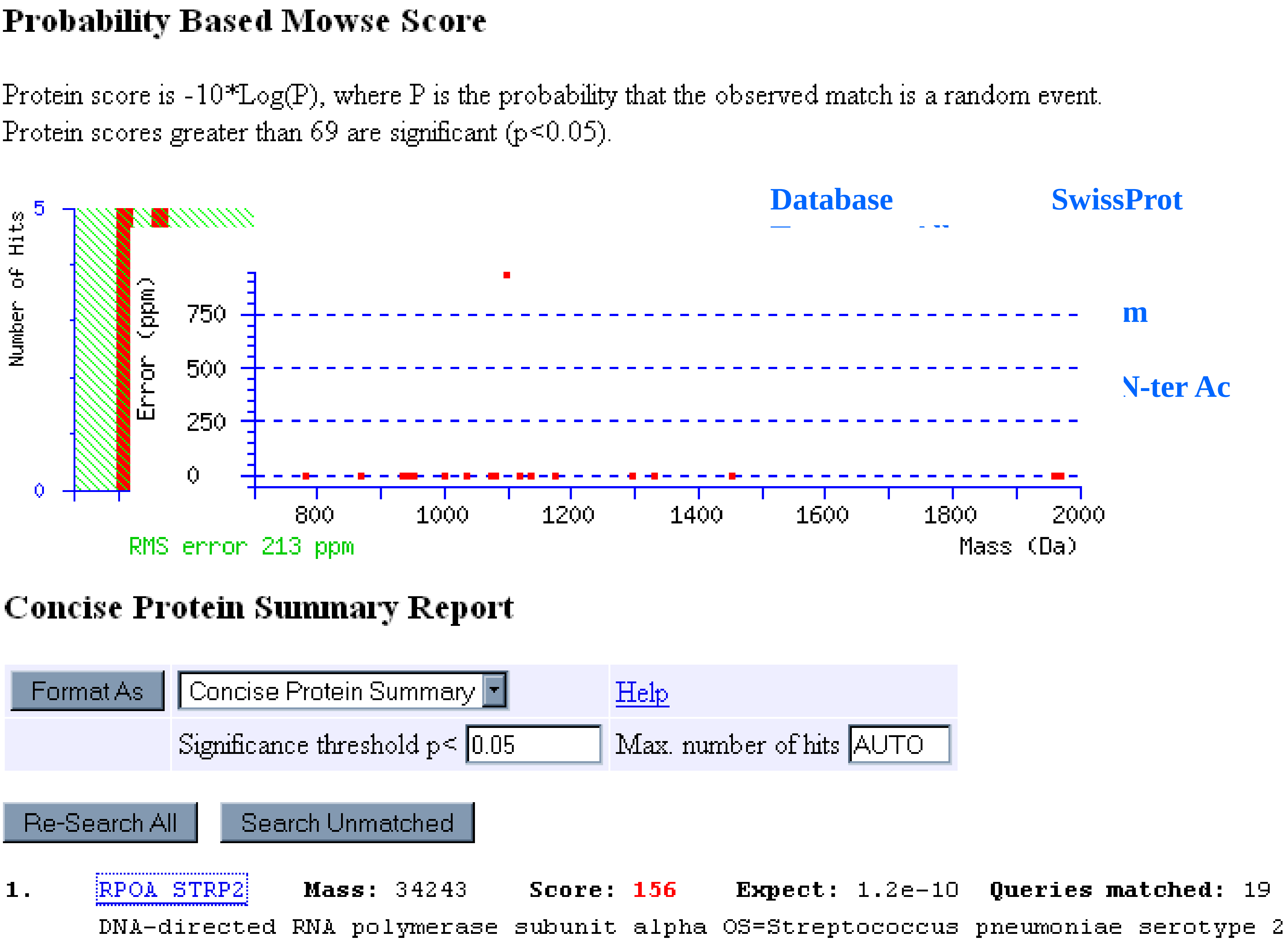
Concise Protein Summary Report

Significance threshold p< 0.05

Max. number of hits AUTO

1. RPOA_STRP2 **Mass:** 34243 **Score:** 138 **Expect:** 7.9e-09 **Queries matched:** 16
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2

Peptidové mapování – vliv parametrů



Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 69 are significant (p<0.05).

Database **SwissProt**

m

N-ter Ac

Concise Protein Summary Report

Significance threshold p< 0.05

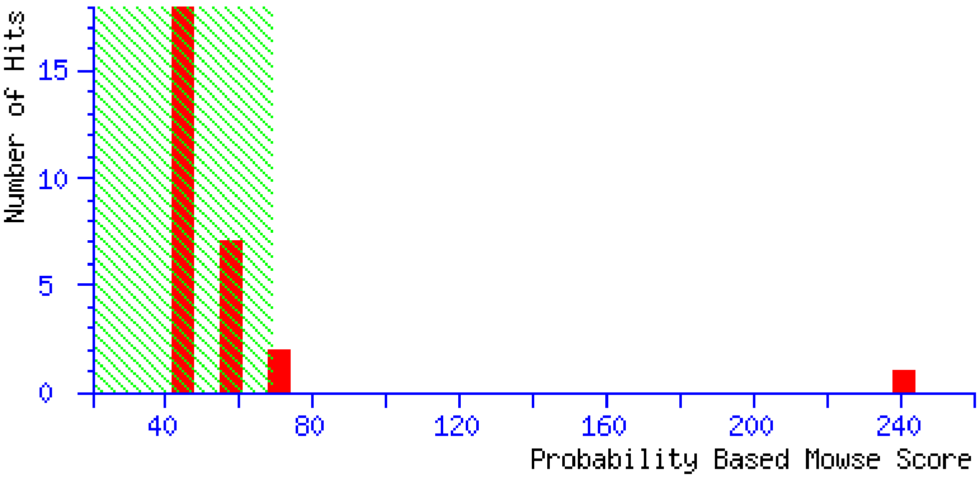
Max. number of hits AUTO

1. RPOA_STRP2 **Mass:** 34243 **Score:** 156 **Expect:** 1.2e-10 **Queries matched:** 19
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2

Peptidové mapování – vliv parametrů

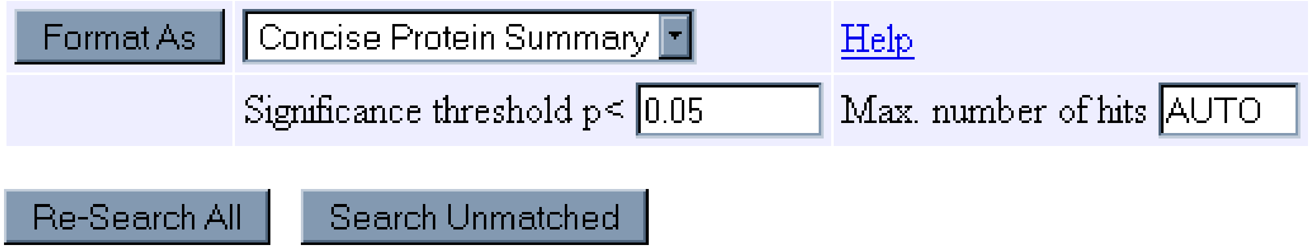
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 69 are significant (p<0.05).



Database	**SwissProt**
Taxonomy	**All taxa**
Missed Cleavages	**1**
Fixed Mods	**C-Cam**
Variable Mods	**M-ox** **Prot N-ter Ac**
Peptide tol	**100**

Concise Protein Summary Report



Significance threshold p< 0.05

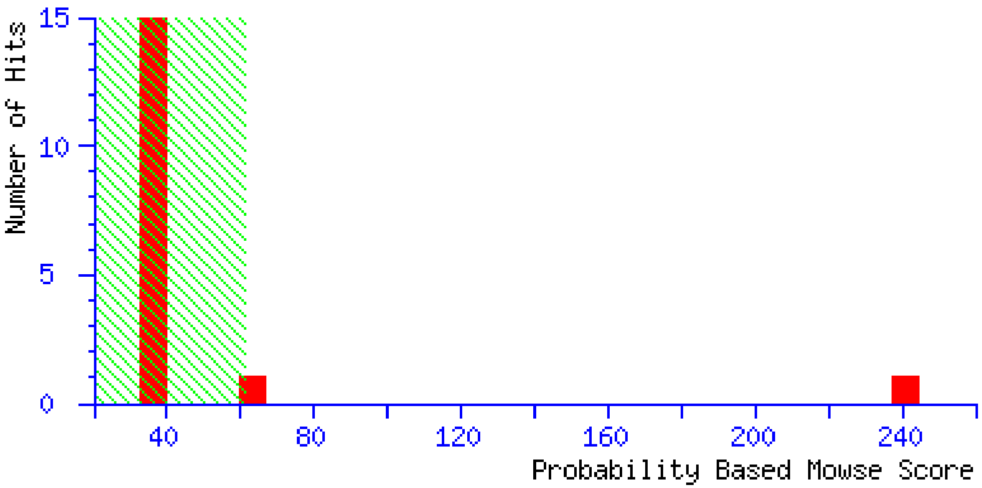
Max. number of hits AUTO

1. RPOA_STRP2 **Mass:** 34243 **Score:** **241** **Expect:** 3.9e-19 **Queries matched:** 18
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2

Peptidové mapování – vliv parametrů

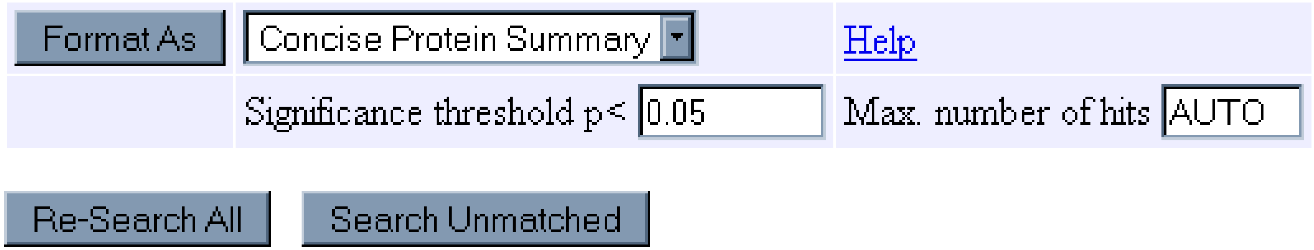
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 61 are significant (p<0.05).



Database	**SwissProt**
Taxonomy	**Firmicutes**
Missed Cleavages	**1**
Fixed Mods	**C-Cam**
Variable Mods	**M-ox**
	Prot N-ter Ac
Peptide tol	**100**

Concise Protein Summary Report



Concise Protein Summary

Significance threshold p< 0.05

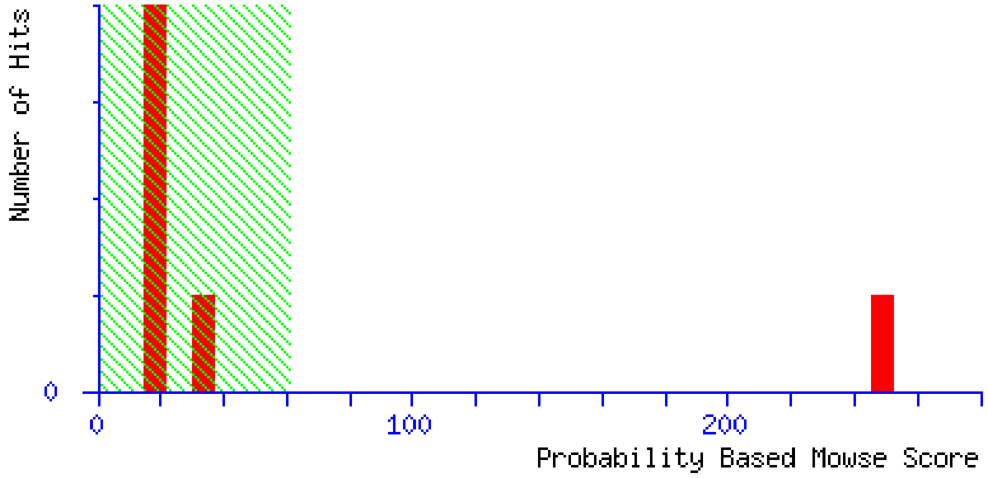
Max. number of hits AUTO

```
1.     RPOA_STRP2     Mass: 34243     Score: 241     Expect: 5.1e-20   Queries matched: 18
       DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2
```

Peptidové mapování – vliv parametrů

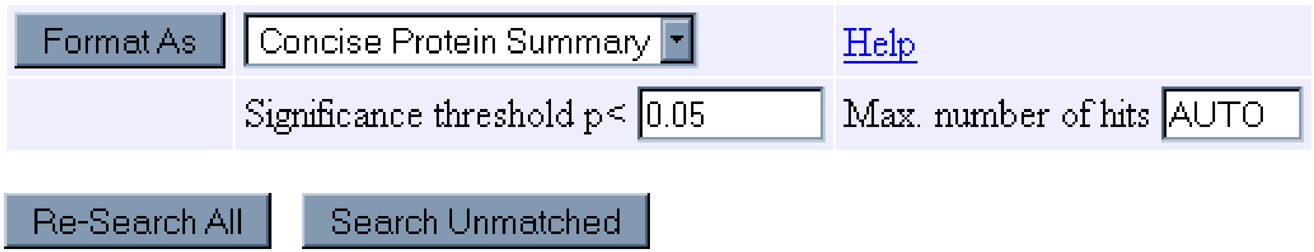
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 61 are significant ($p<0.05$).



Database	**SwissProt**
Taxonomy	**Firmicutes**
Missed Cleavages	**1**
Fixed Mods	**C-Cam**
Variable Mods	**M-ox** **Prot N-ter Ac**
Peptide tol	**10**

Concise Protein Summary Report



Significance threshold p< 0.05

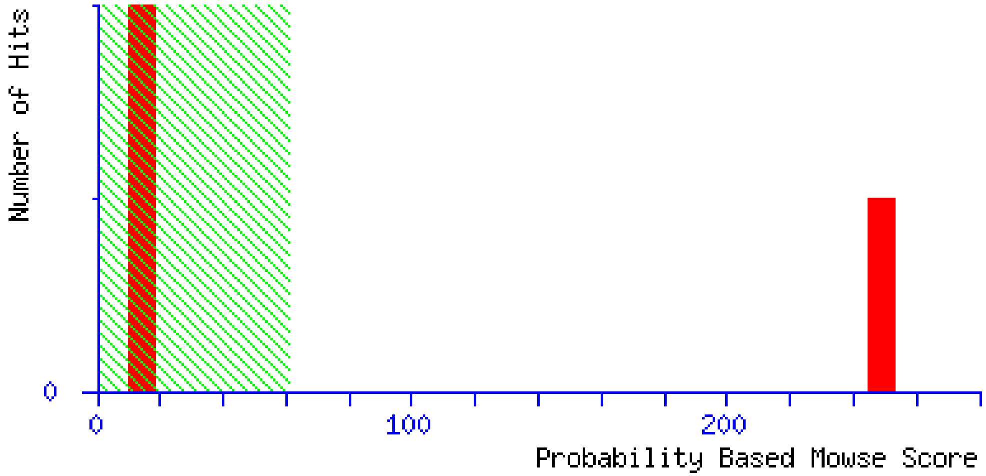
Max. number of hits AUTO

1. RPOA_STRP2 **Mass:** 34243 **Score:** **249** **Expect:** 8e-21 **Queries matched:** 18
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2

Peptidové mapování – vliv parametrů

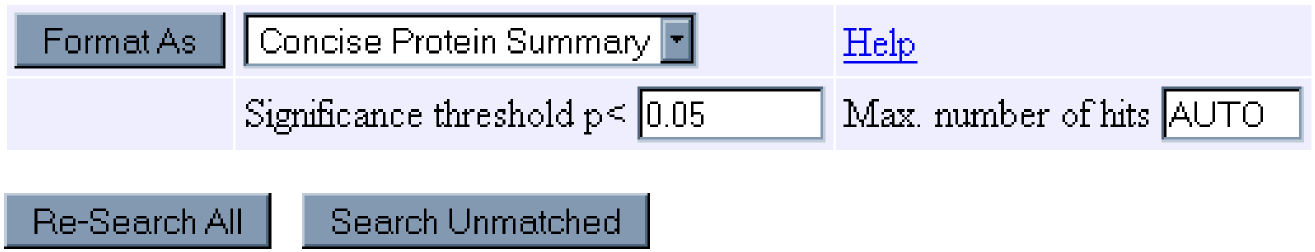
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 61 are significant ($p<0.05$).



Database	**SwissProt**
Taxonomy	**Firmicutes**
Missed Cleavages	**1**
Fixed Mods	**C-Cam**
Variable Mods	**M-ox** **Prot N-ter Ac**
Peptide tol	**1**

Concise Protein Summary Report



Significance threshold p< 0.05

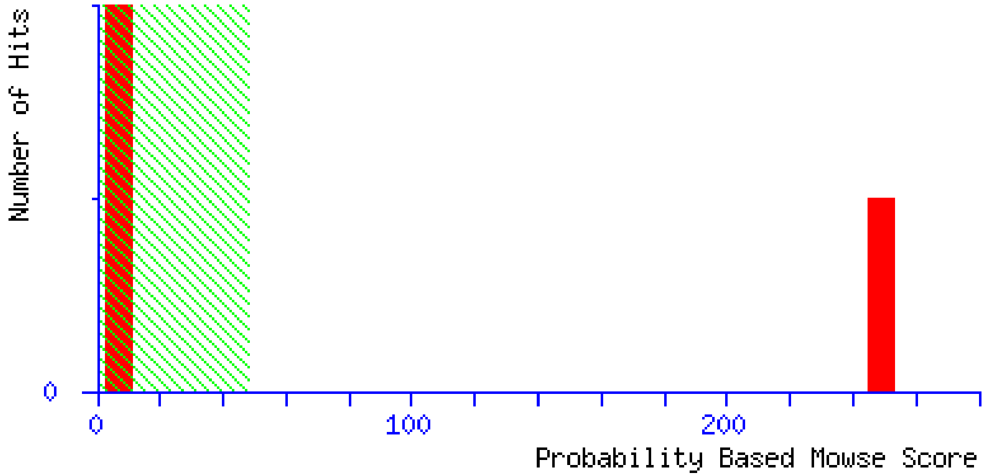
Max. number of hits AUTO

1. RPOA_STRP2 **Mass:** 34243 **Score:** 249 **Expect:** 8e-21 **Queries matched:** 18
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2

Peptidové mapování – vliv parametrů

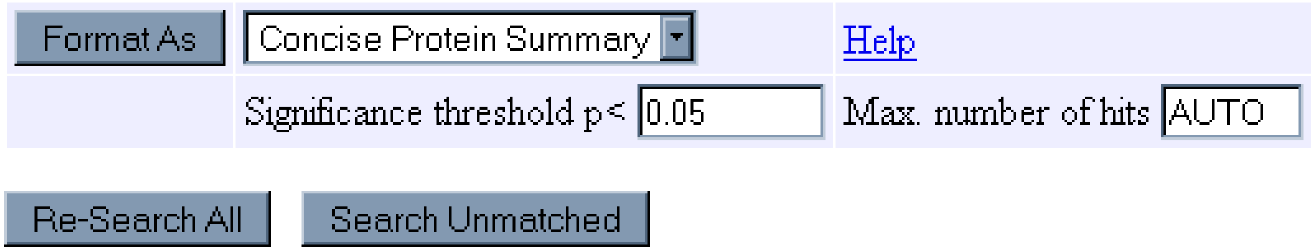
Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 48 are significant (p<0.05).



Database	**SwissProt**
Taxonomy	***S. penumoniae***
Missed Cleavages	**1**
Fixed Mods	**C-Cam**
Variable Mods	**M-ox** **Prot N-ter Ac**
Peptide tol	**1**

Concise Protein Summary Report



Significance threshold p< 0.05
Max. number of hits AUTO

```
1.     RPOA_STRP2     Mass: 34243     Score: 249     Expect: 4.5e-22   Queries matched: 18
       DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2
```

Peptidové mapování – vliv parametrů

Protein View

```
Match to: RPOA_STRP2 Score: 249 Expect: 4.5e-22
DNA-directed RNA polymerase subunit alpha OS=Streptococcus pneumoniae serotype 2 (strain D39 / NCTC 7466) GN=rpoA PE=3 SV=1

Nominal mass (Mr): 34243; Calculated pI value: 4.64
NCBI BLAST search of RPOA_STRP2 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Streptococcus pneumoniae

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Acetyl (Protein N-term),Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: 27
Number of mass values matched: 18
Sequence Coverage: 53%

Matched peptides shown in Bold Red

    1 MIEFEKPNIT KIDENKDYGK FVIEPLERGY GTTLGNSLRR VLLASLPGAA
   51 VTSINIDGVL HEFDTVPGVR EDVMQIILNI KGIAVKSYVE DEKIIELDVE
  101 GPAEVTAGDI LTDSDIEIVN PDHYLFTIGE GSSLKATMTV NSGRGYVPAD
  151 ENKKDNAPVG TLAVDSIYTP VTKVNYQVEP ARVGSNDGFD KLTLEILTNG
  201 TIIPEDALGL SARILTEHLD LFTNLTEIAK STEVMKEADT ESDDRILDRT
  251 IEELDLSVRS YNCLKRAGIN TVHDLTEKSE AEMMKVRNLG RKSLEEVKLK
  301 LIDLGLGLKD K
```

```
 Start - End     Observed   Mr(expt)   Mr(calc)      ppm   Miss Sequence
  260 - 265     784.3656   783.3583   783.3585       -0      0  R.SYNCLK.R
   87 - 93      869.3889   868.3816   868.3814        0      0  K.SYVEDEK.I
  136 - 144     936.4567   935.4494   935.4495       -0      0  K.ATMTVNSGR.G
  183 - 191     938.4214   937.4141   937.4141       -0      0  R.VGSNDGFDK.L
  301 - 309     941.6030   940.5957   940.5957        0      0  K.LIDLGLGLK.D
  136 - 144     952.4515   951.4442   951.4444       -0      0  K.ATMTVNSGR.G  Oxidation (M)
   21 - 28     1002.5619  1001.5546  1001.5546        0      0  K.FVIEPLER.G
  237 - 245    1037.4020  1036.3947  1036.3945        0      0  K.EADTESDDR.I
  174 - 182    1075.5534  1074.5461  1074.5458        0      0  K.VNYQVEPAR.V
   12 - 20     1081.5163  1080.5090  1080.5087        0      1  K.IDENKDYGK.F
  145 - 154    1120.5630  1119.5557  1119.5560       -0      1  R.GYVPADENKK.D
   29 - 39     1138.5850  1137.5777  1137.5778       -0      0  R.GYGTTLGNSLR.R
  250 - 259    1174.6306  1173.6233  1173.6241       -1      0  R.TIEELDLSVR.S
  267 - 278    1297.6743  1296.6670  1296.6674       -0      0  R.AGINTVHDLTEK.S
   71 - 81     1331.7240  1330.7167  1330.7166        0      0  R.EDVMQIILNIK.G  Oxidation (M)
  266 - 278    1453.7759  1452.7686  1452.7685        0      1  K.RAGINTVHDLTEK.S
  155 - 173    1961.0208  1960.0135  1960.0153       -1      0  K.DNAPVGTLAVDSIYTPVTK.V
  214 - 230    1971.0781  1970.0708  1970.0724       -1      0  R.ILTEHLDLFTNLTEIAK.S

No match to: 842.5097, 870.5407, 1016.5771, 1024.5436, 1089.5689, 1097.5338, 1128.5281, 1188.6459, 1308.6543
```

Unmatched masses:
842.50 – Trypsin
870.54 – very tiny peak
1016.57 - methylation of 1002.56
1024.54 – sodium adduct of 1002.56
1089.57 - methylation of 1075.55
1097.53 - sodium adduct of 1075.55
1128.53 – very tiny peak
1188.65 – methylation of 1174.63
1308.65 - very tiny peak

Peptidové mapování – prezentace výsledků

Nutno uvést

- instrument a jeho nastavení, typ kalibrace (interní, externí, post-kalibrace)
- program/algoritmus použitý pro označení píků, dekonvoluci, vyhlazení/korekce
- program a všechny parametry použité pro identifikaci
- Skóre (+limit), pokrytí sekvence, počet použitých/přiřazených píků

MALDI FT-ICR (Apollo II Dual source Apex Qe 9.4T)

externí kalibrace na Bruker Peptide Mix

DataAnalysis 4.0, unprocessed, SNAP v 2.0

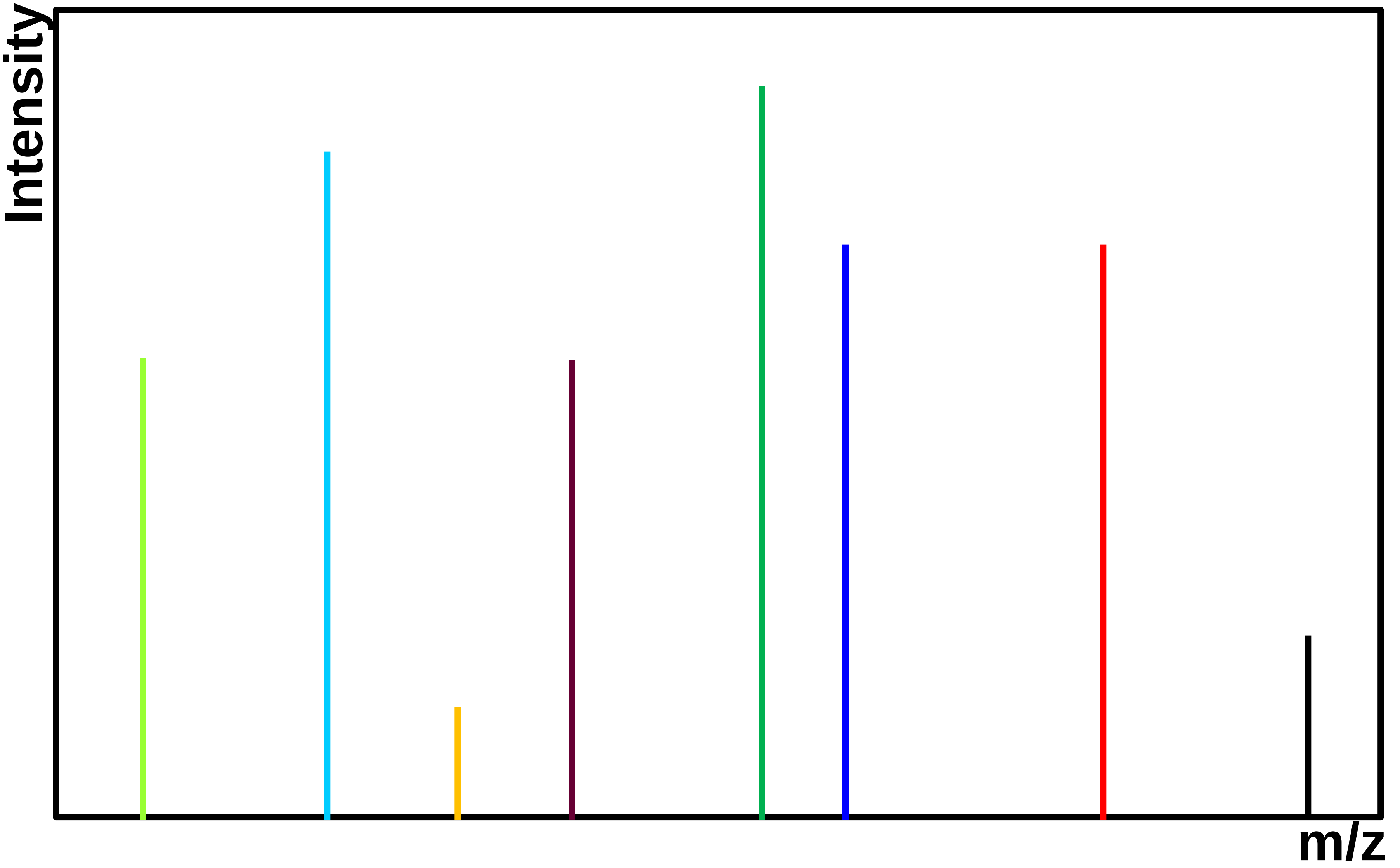
MASCOT v 2.2 – Swiss-Prot (verze/datum), S. pneumoniae, Trypsin, m.c. - 1, f.m. - C-Cam, v.m. – Prot N-term acetylated + M-ox, p.t. 1ppm

Skóre 249 (48 limit), pokrytí sekvence 53%, 23/26 píků přiřazeno

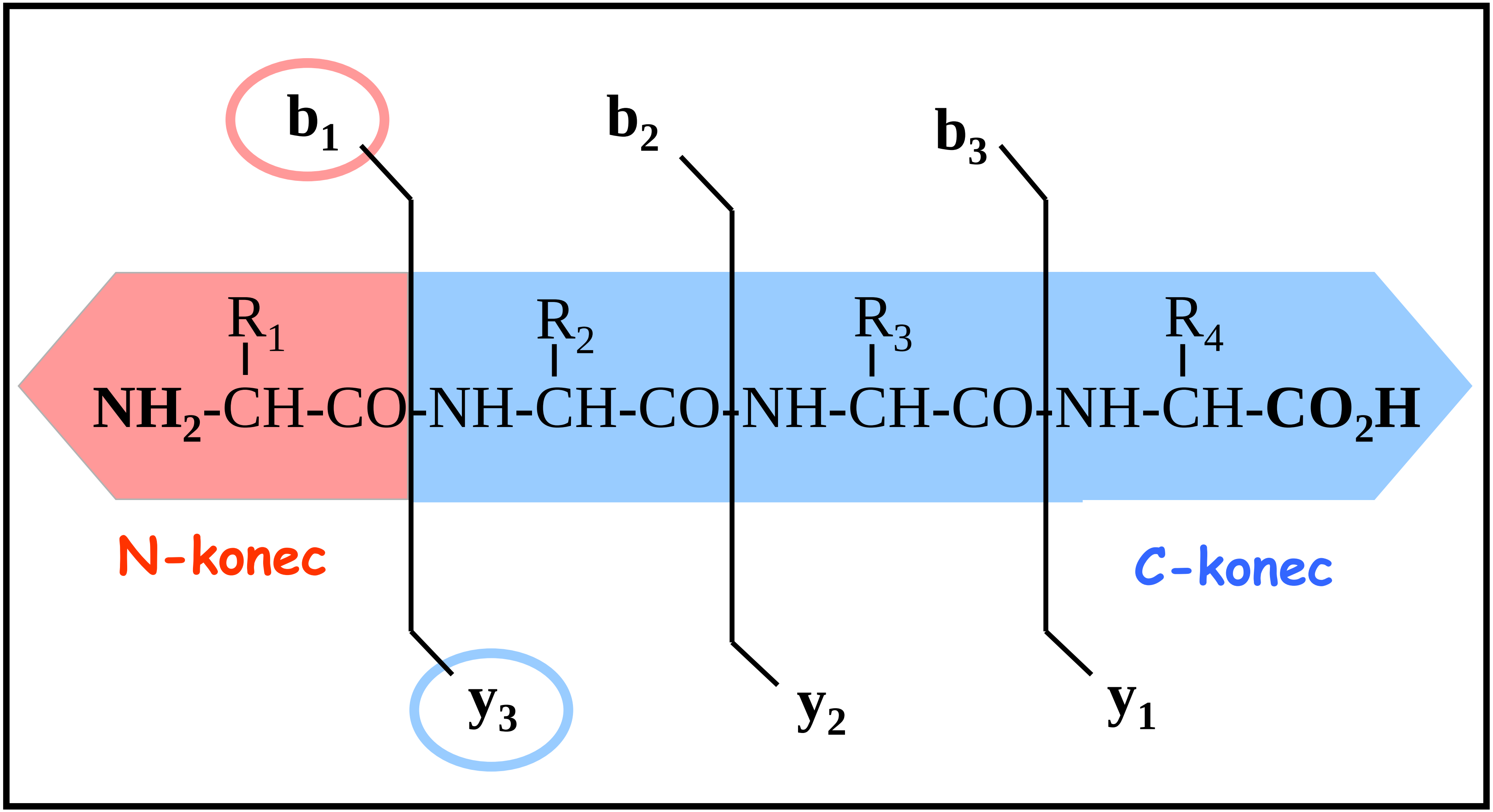
http://www.mcponline.org/site/misc/PhialdelphiaGuidelinesFINALDRAFT.pdf

Identifikace proteinů ve směsi

ADSMK EETSLINVCPWK VNMGGHDSIYTLR EEIYTLK
MLIILLR TYSHEEDK EWQIDSLAEIR IQPLPMNVSA



Fragmentace peptidu



b_1 b_2 b_3

$$\mathbf{NH_2}\text{-}\overset{\displaystyle R_1}{\overset{|}{\mathrm{C}}}\mathrm{H}\text{-}\mathrm{CO}\text{-}\mathrm{NH}\text{-}\overset{\displaystyle R_2}{\overset{|}{\mathrm{C}}}\mathrm{H}\text{-}\mathrm{CO}\text{-}\mathrm{NH}\text{-}\overset{\displaystyle R_3}{\overset{|}{\mathrm{C}}}\mathrm{H}\text{-}\mathrm{CO}\text{-}\mathrm{NH}\text{-}\overset{\displaystyle R_4}{\overset{|}{\mathrm{C}}}\mathrm{H}\text{-}\mathbf{CO_2H}$$

N-konec

C-konec

y_3 y_2 y_1

Přehled aminokyselin v číslech

AK		Hmotnost	Boční řetězec	Immoniové ionty
G	Gly	57.02	1	30
A	Ala	71.08	15	44
S	Ser	87.03	31	60
P	Pro	97.05	41	70
V	Val	99.07	43	72
T	Thr	101.05	45	74
C	Cys	103.01	47	76
L	Leu	113.08	57	86(72)
I	Ile	113.08	57	86(72)
N	Asn	114.04	58	87(70)
D	Asp	115.03	59	88
Q	Gln	128.06	72	101(84, 129)
K	Lys	128.09	72	101(129, 112, 84, 70)
E	Glu	129.04	73	102
M	Met	131.04	75	104(61)
H	His	137.06	81	110(166, 138, 123, 121, 82)
F	Phe	147.07	91	120(91)
R	Arg	156.10	100	129(112, 100, 87, 73, 70, 59)
Y	Tyr	163.06	107	136
W	Trp	186.08	130	159

Fragmentace peptidu

S-P-A-F-D-S-I-M-A-E-T-L-K MH$^+$ = 1410.6

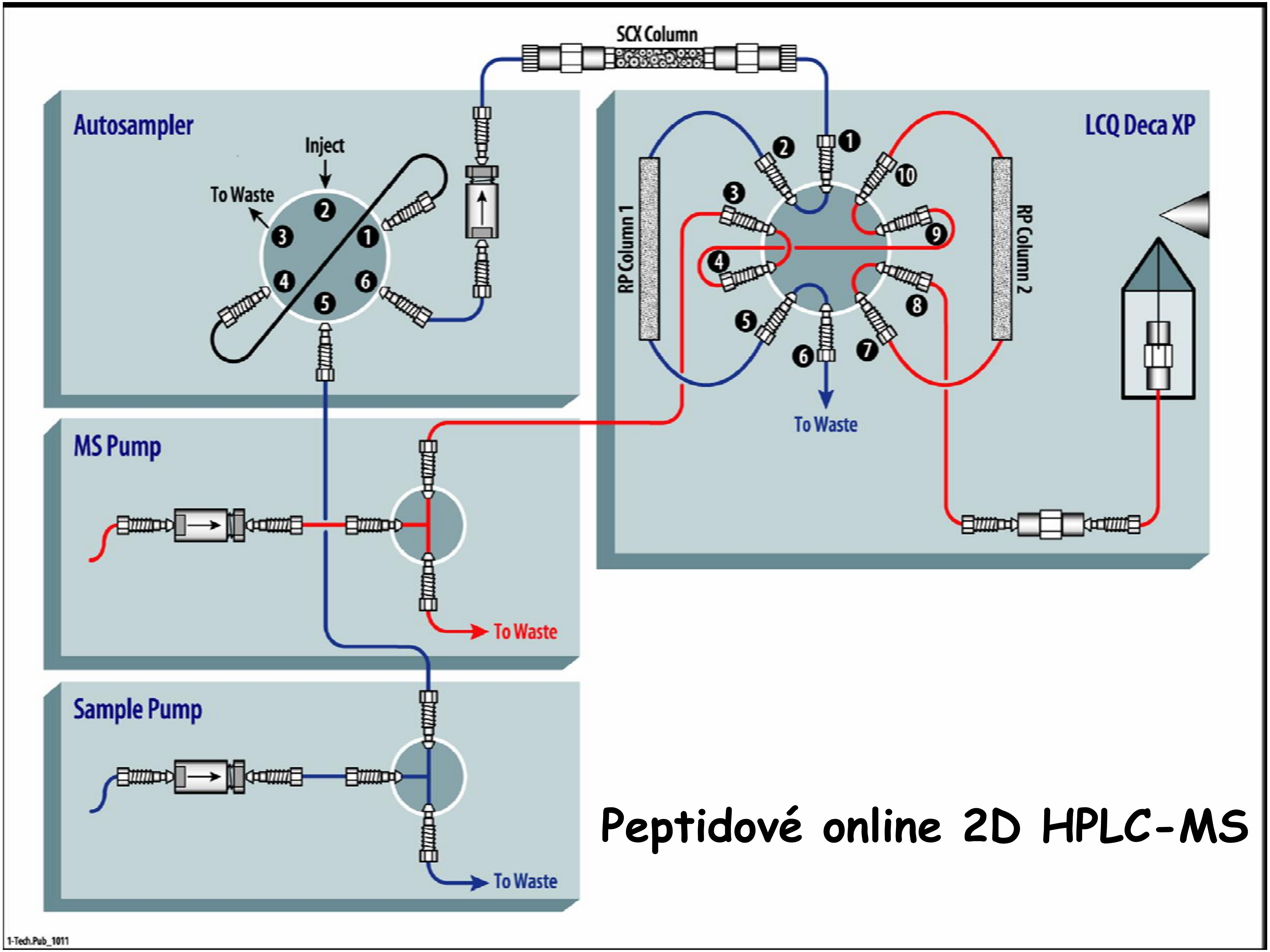
b-ions$^+$			y-ions$^+$
88.1	S	PAFDSIMAETLK	1323.6
185.2	SP	AFDSIMAETLK	1226.4
256.3	SPA	FDSIMAETLK	1155.4
403.5	SPAF	DSIMAETLK	1008.2
518.5	SPAFD	SIMAETLK	893.1
605.6	SPAFDS	IMAETLK	806.0
718.8	SPAFDSI	MAETLK	692.3
850.0	SPAFDSIM	AETLK	561.7
921.1	SPAFDSIMA	ETLK	490.6
1050.2	SPAFDSIMAE	TLK	361.5
1151.3	SPAFDSIMAET	LK	260.4
1264.4	SPAFDSIMAETL	K	147.2

Shotgun proteomics – Analýza proteinových směsí

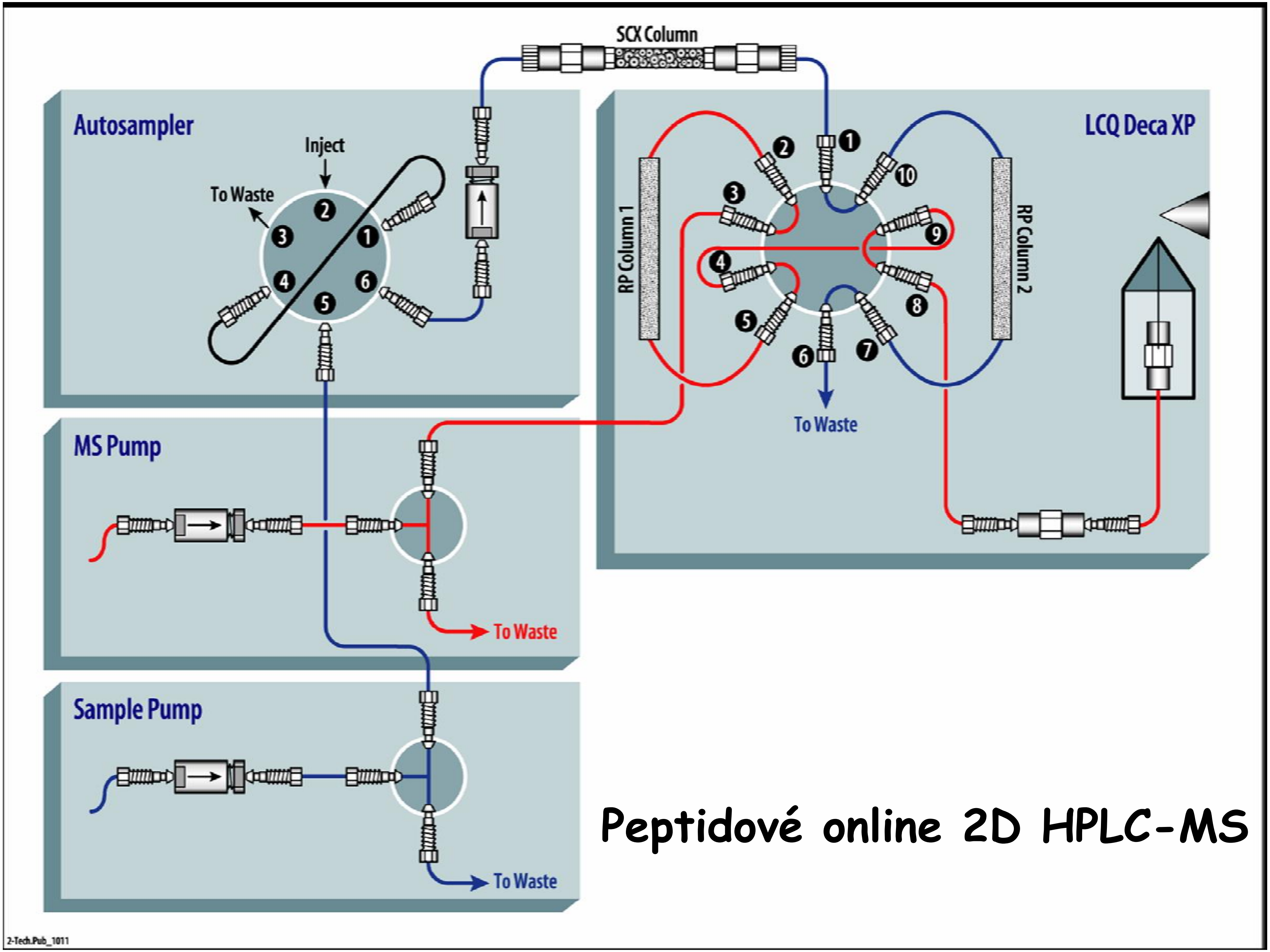
- Bez gelů - 1D nebo 2D – ztráta informace o PTM
- Komplexní směsi (organely, celé buňky,…)
- HPLC – 1D – RP C18
-   2D – SCX/SAX + RP C18
- 3D – GF + SCX/SAX + RP C18
- 2D or 3D + afinitní purifikace
- data dependentní analýza
- “frakcionace v plynné fázi”

Online spojení pro separaci peptidů

- Afinitní purifikace se využívají při cílené analýze určitého typu peptidů
  - glykopeptidy – lektinové kolony, diolové kolony, aktivní uhlí
  - fosfopeptidy = IMAC (Ga, Fe), TiO2, ZrO2
- Pro urychlení analýzy lze použít kolonku s imobilizovaným trypsinem. Více odolný a digesce je rychlejší kvůli velkému nadbytku enzymu.
- Při přímém spojení s MS musíme dbát na kompatibilitu solventů s MS. Zároveň musí být počáteční solvent kompatibilní s druhou kolonou. Vyžíváme předkolny pro zakoncentrování a pro transfer z jednoho eluentu do druhého.
- Při spojení RP s ESI musí počáteční fáze obsahovat alespoň 3-5% organického solventu – při míchání v pumpách hrozí nebezpečí vzniku bublinek a také v čisté vodě není stabilní elektrosprej.

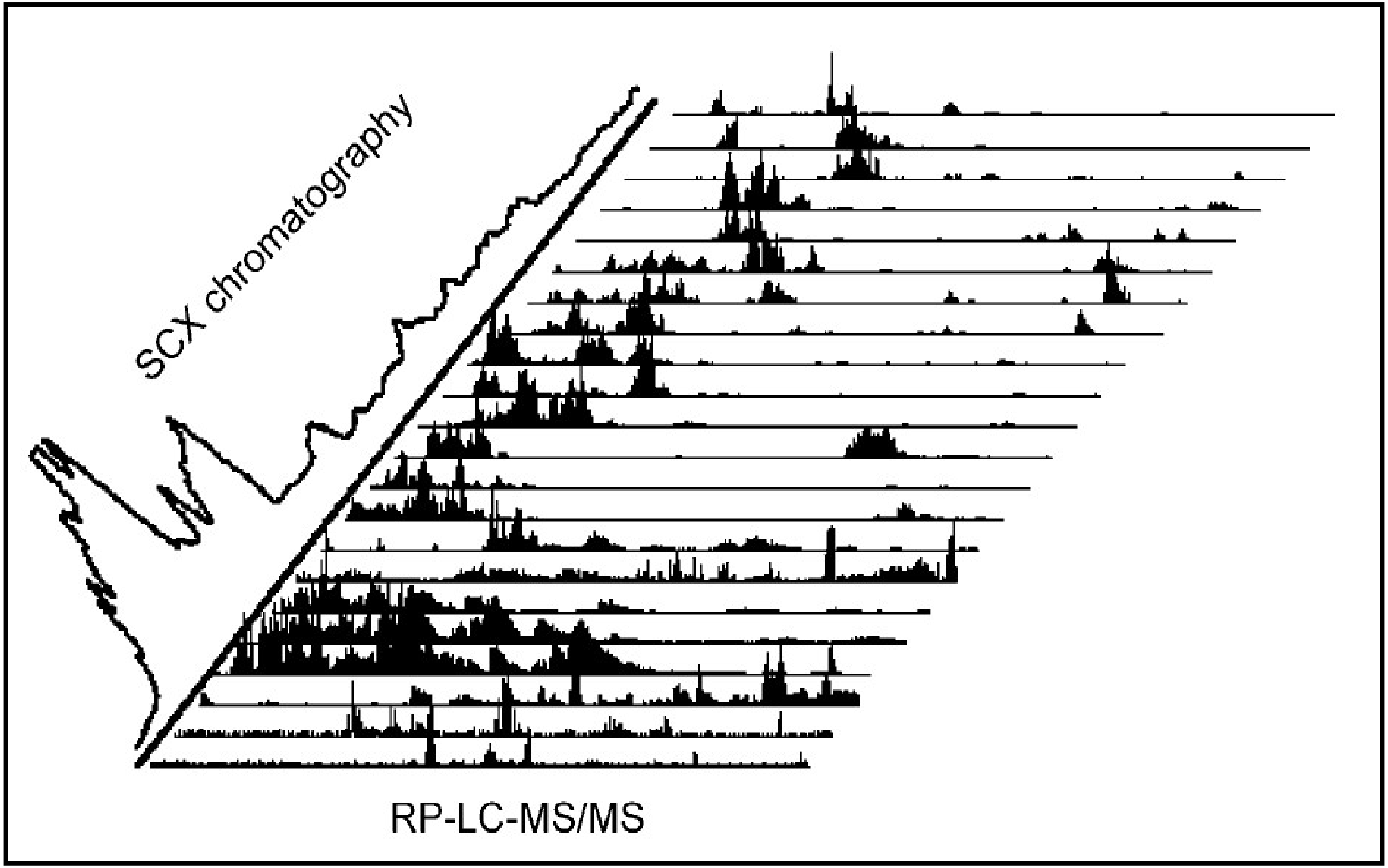


Peptidové online 2D HPLC-MS



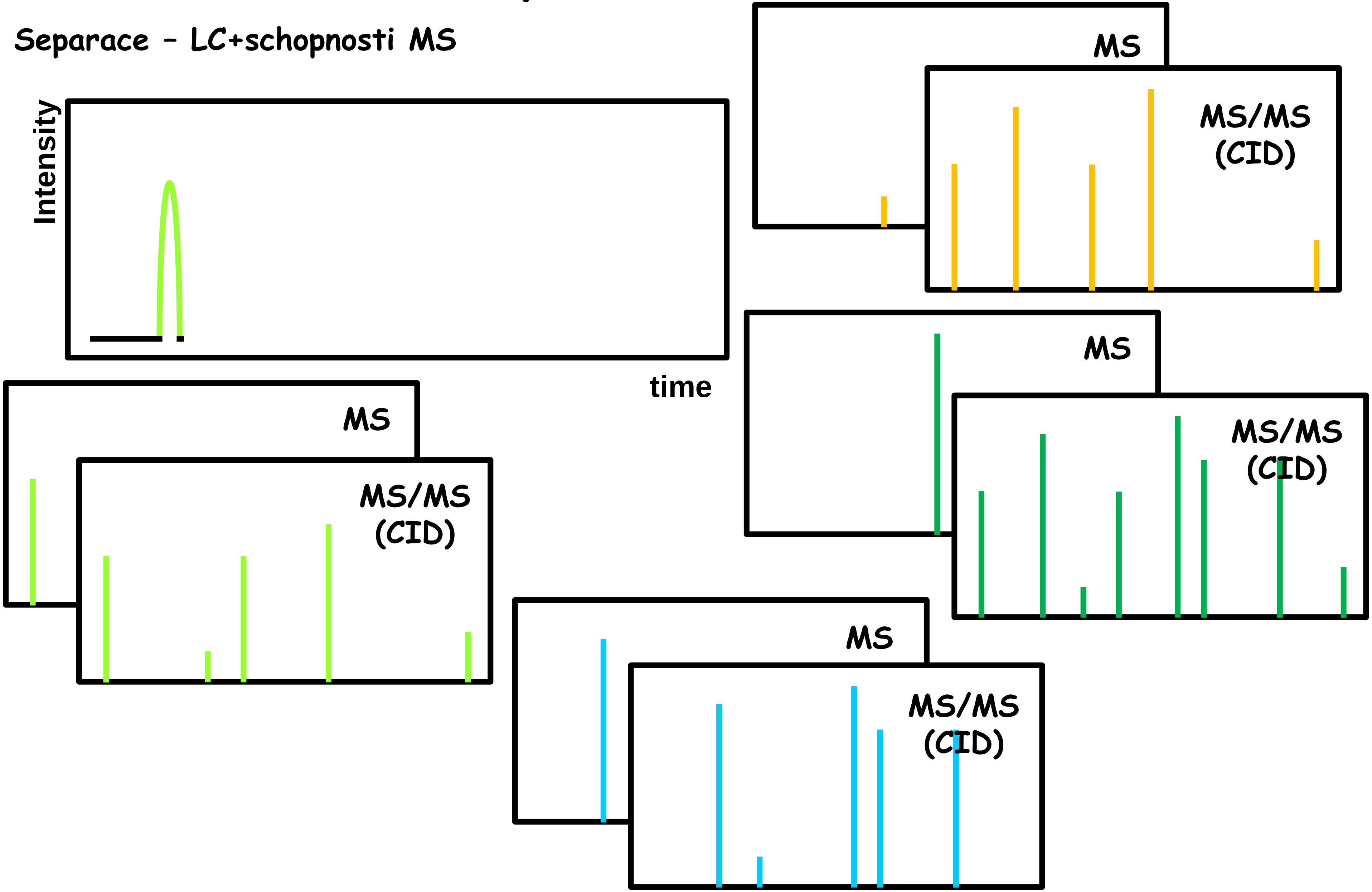
Peptidové online 2D HPLC-MS

Result of 22 step 2D HPLC run



Identifikace proteinů ve směsi
- separace a MS/MS

Separace – LC+schopnosti MS

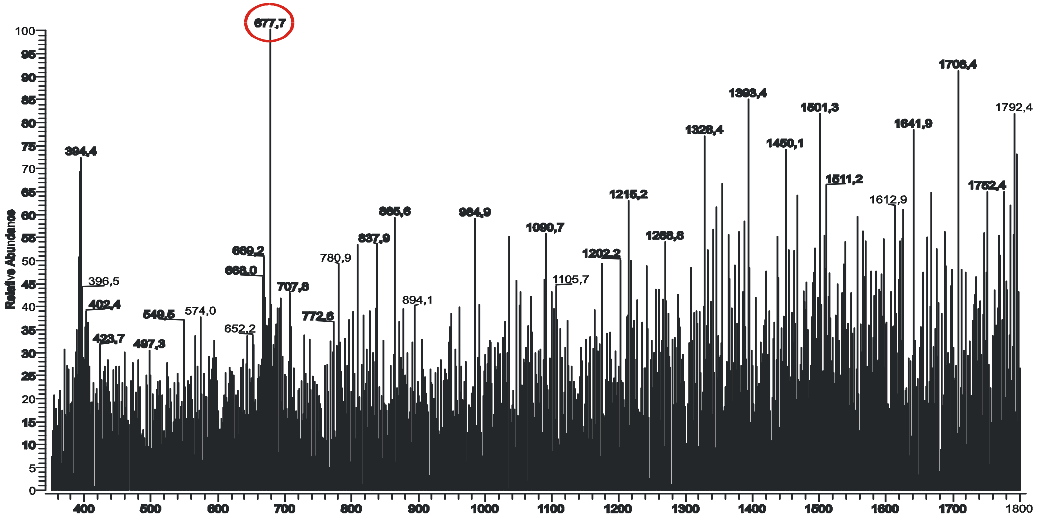


Jak MS pracuje v reálném čase…

data-dependentní analýza

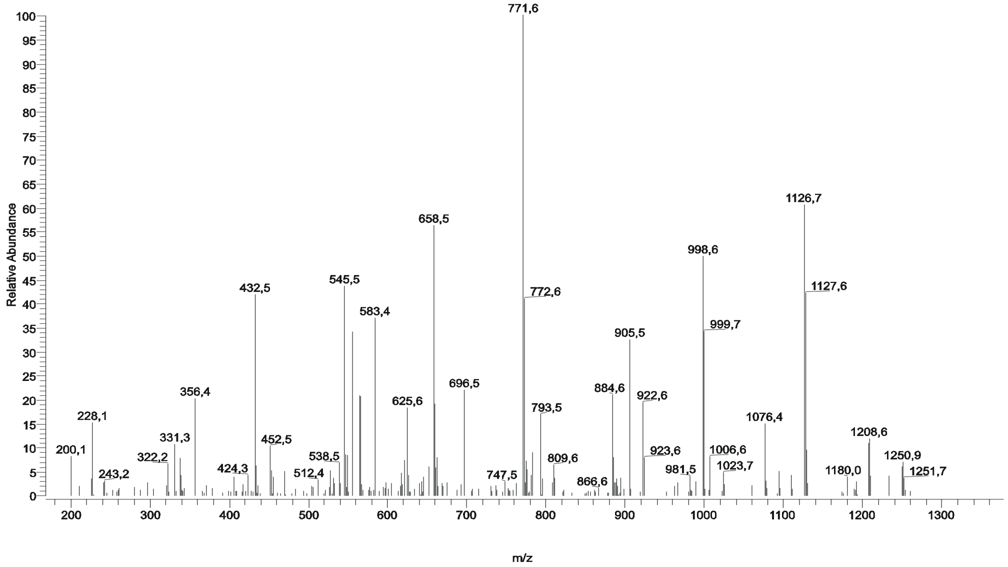
MS

Rozsah (350-1800)
Sken 1479



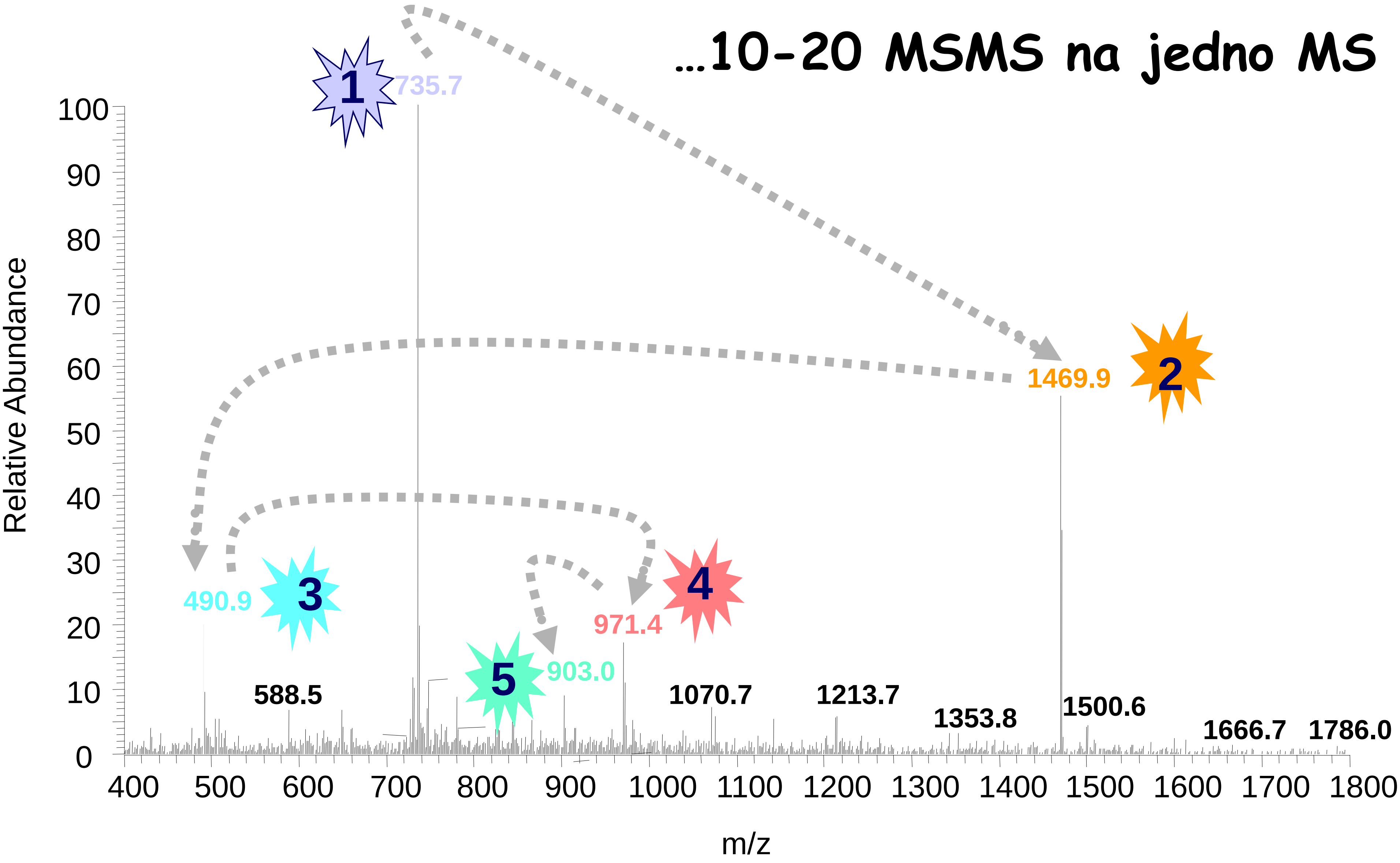
MS/MS (CID)

Sken 1480
Prekurzor 667.7

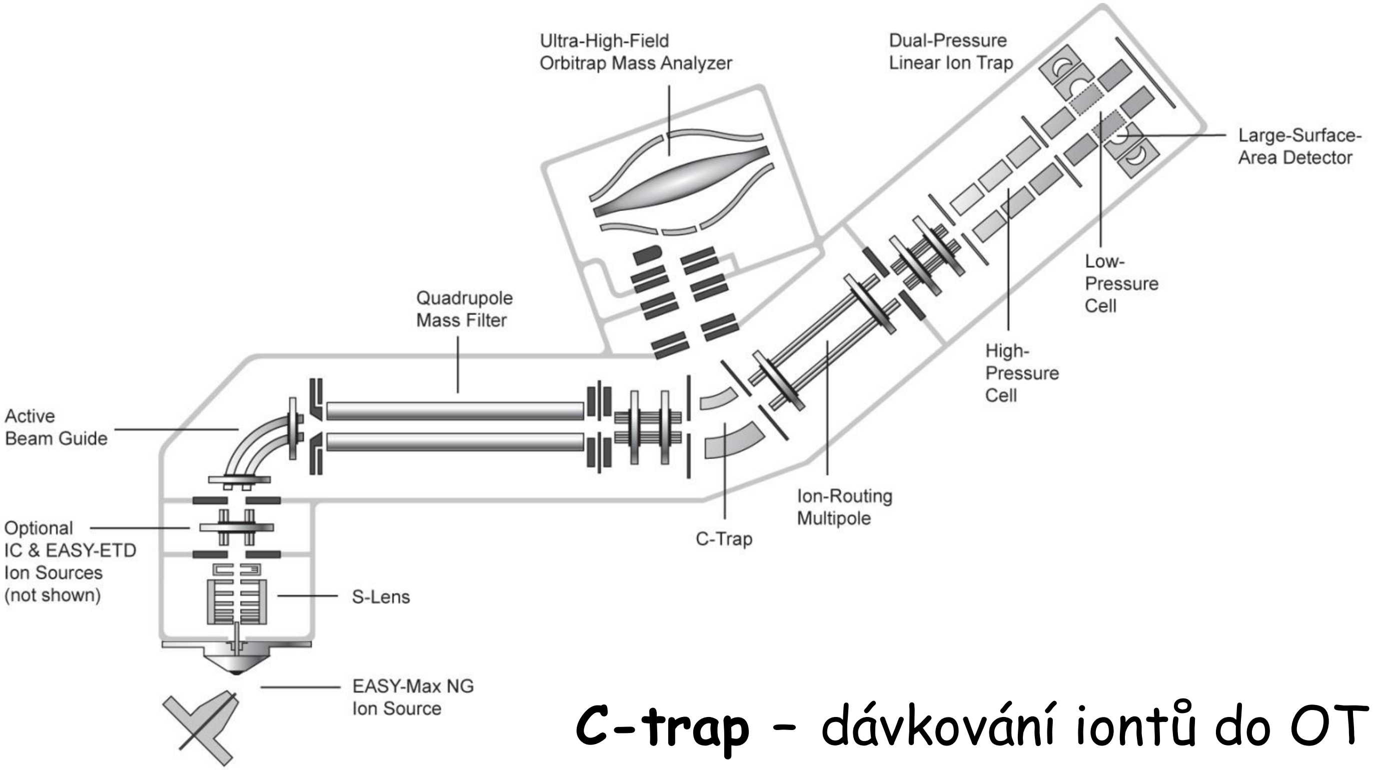


Data-Dependentní MS & Dynamická Exkluze

…10-20 MSMS na jedno MS



ORBITRAP + dual linear trap

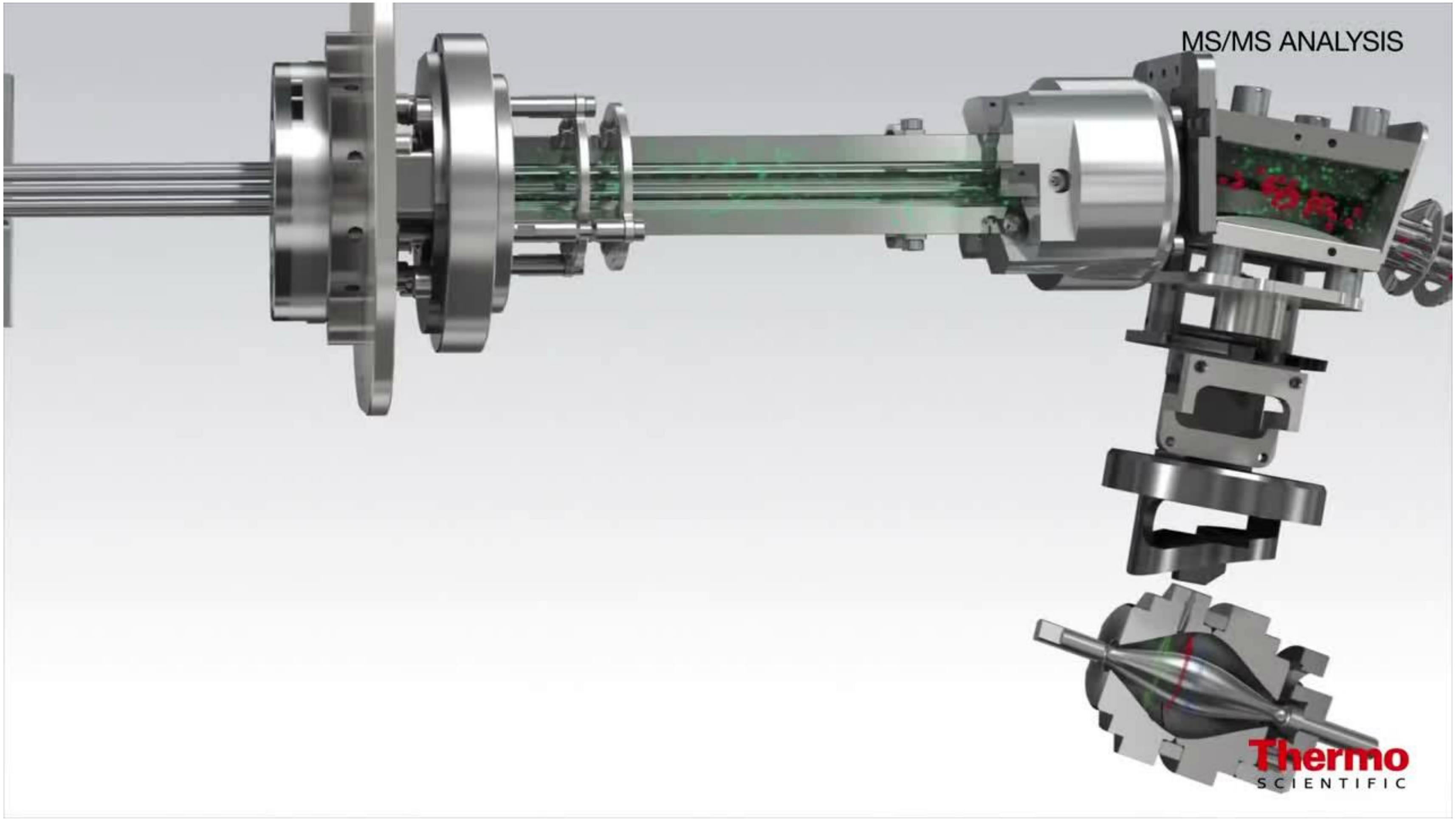


C-trap – dávkování iontů do OT + fragmentace

HCD collision cell - Higher Energy Collisional Dissociation fragmentace mimo C-trap (+ nemá 1/3 omezení – nižší RF než u pastí)

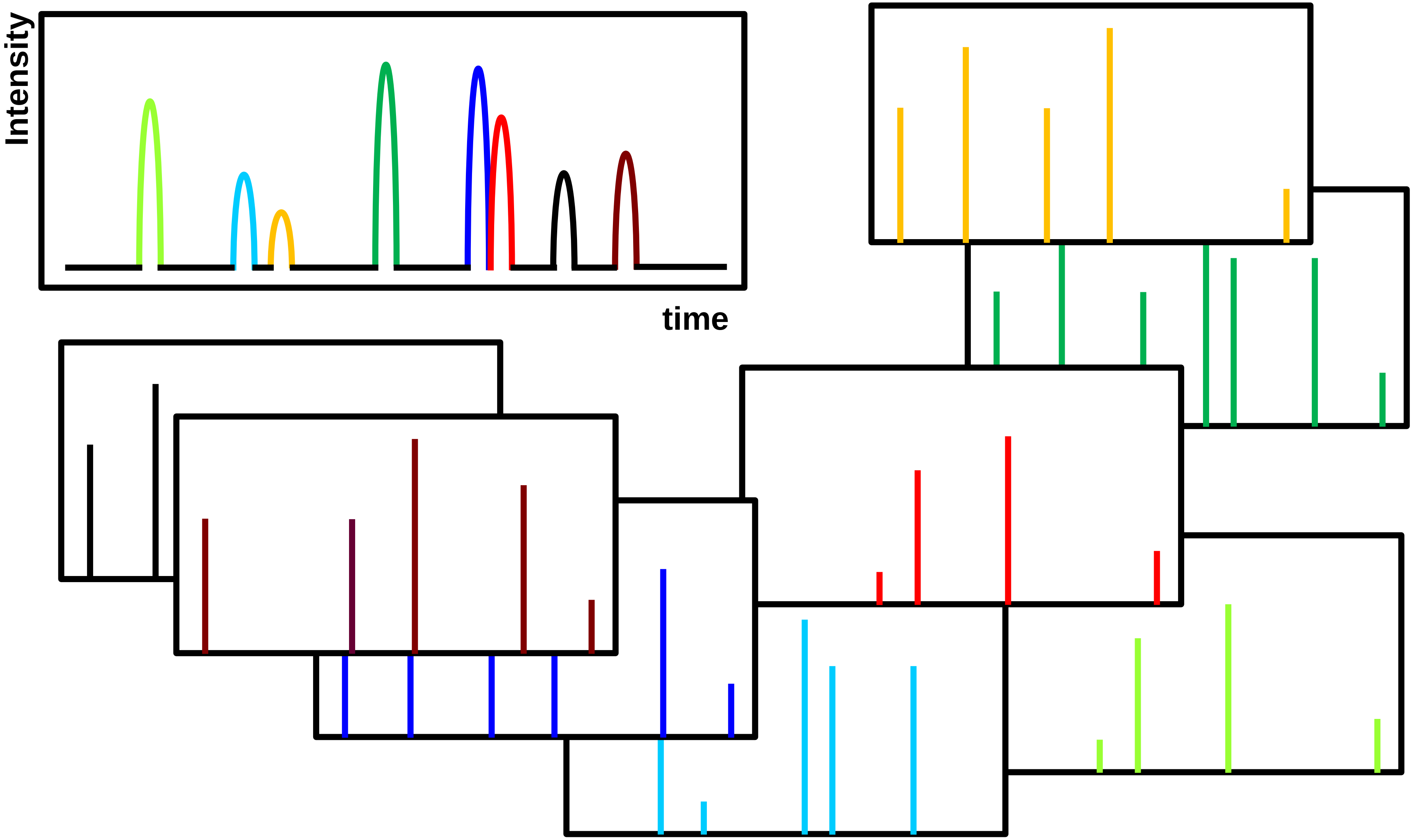
High and low pressure cell (linear trap) - vyšší tlak kolizního plynu je lepší pro záchyt iontů (60->90%) a má vyšší výtěžnost fragmentace (68->80%); nižší tlak kolizního plynu vhodnější pro skenování – lepší rozlišení a rychlejší sken

Hybridní MS – výhoda současné analýzy na více místech

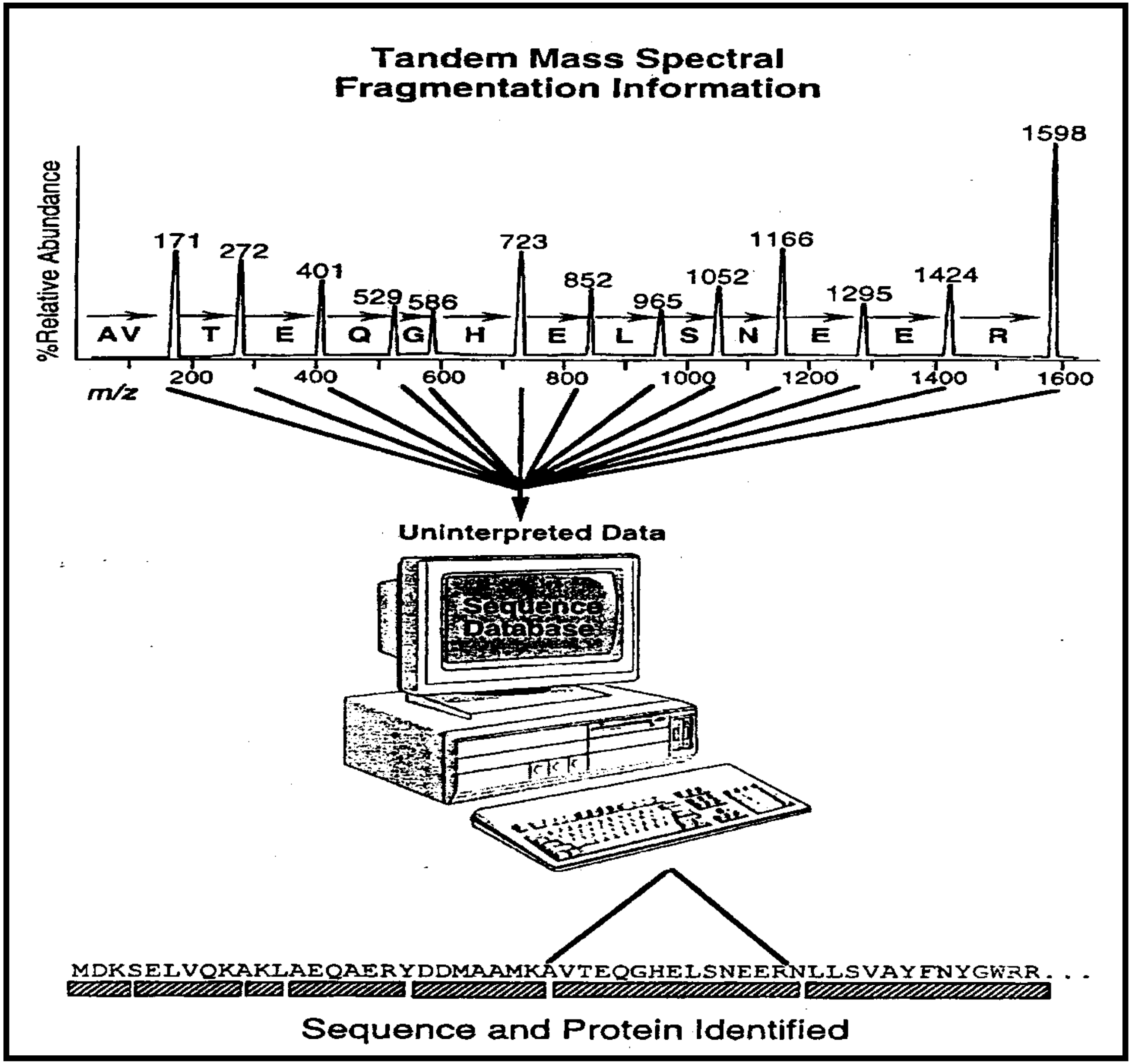


Identifikace proteinů ve směsi - separace a MS/MS

LC-MS/MS – list m/z celých peptidů a MS/MS fragmnetace každého z nich



Peptidové sekvenování



Interpretace MS/MS Spekter

MASCOT MS/MS SEARCH - *Matrix Science*

http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=MIS

MS/MS SONAR - *Rockefeller University*

http://hs2.proteome.ca/prowl/sonar/sonar_cntrl.html

MyriMatch

http://fenchurch.mc.vanderbilt.edu/software.php

OMSSA

http://pubchem.ncbi.nlm.nih.gov/omssa/omssacgi.cgi?searchsettings=iontrap.xml

!XTANDEM - *Rockefeller University*

http://www.thegpm.org/

PepFrag –*Rockefeller University*

http://prowl.rockefeller.edu/prowl/pepfrag.html

Protein Guru

http://www.proteinguru.com/MassSpec/OLMAT/login

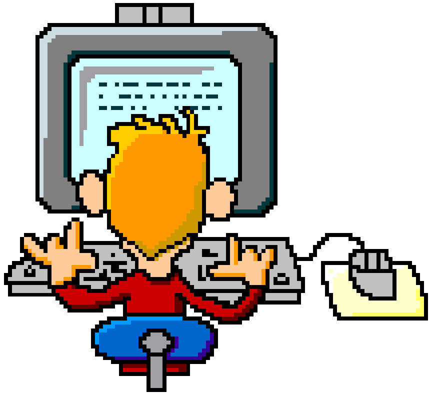
MS Seq, MS Tag - **Protein Prospector**, *UCSF*

http://prospector.ucsf.edu/cgi-bin/msform.cgi?form=msseq
http://prospector.ucsf.edu/cgi-bin/msform.cgi?form=mstagstandard

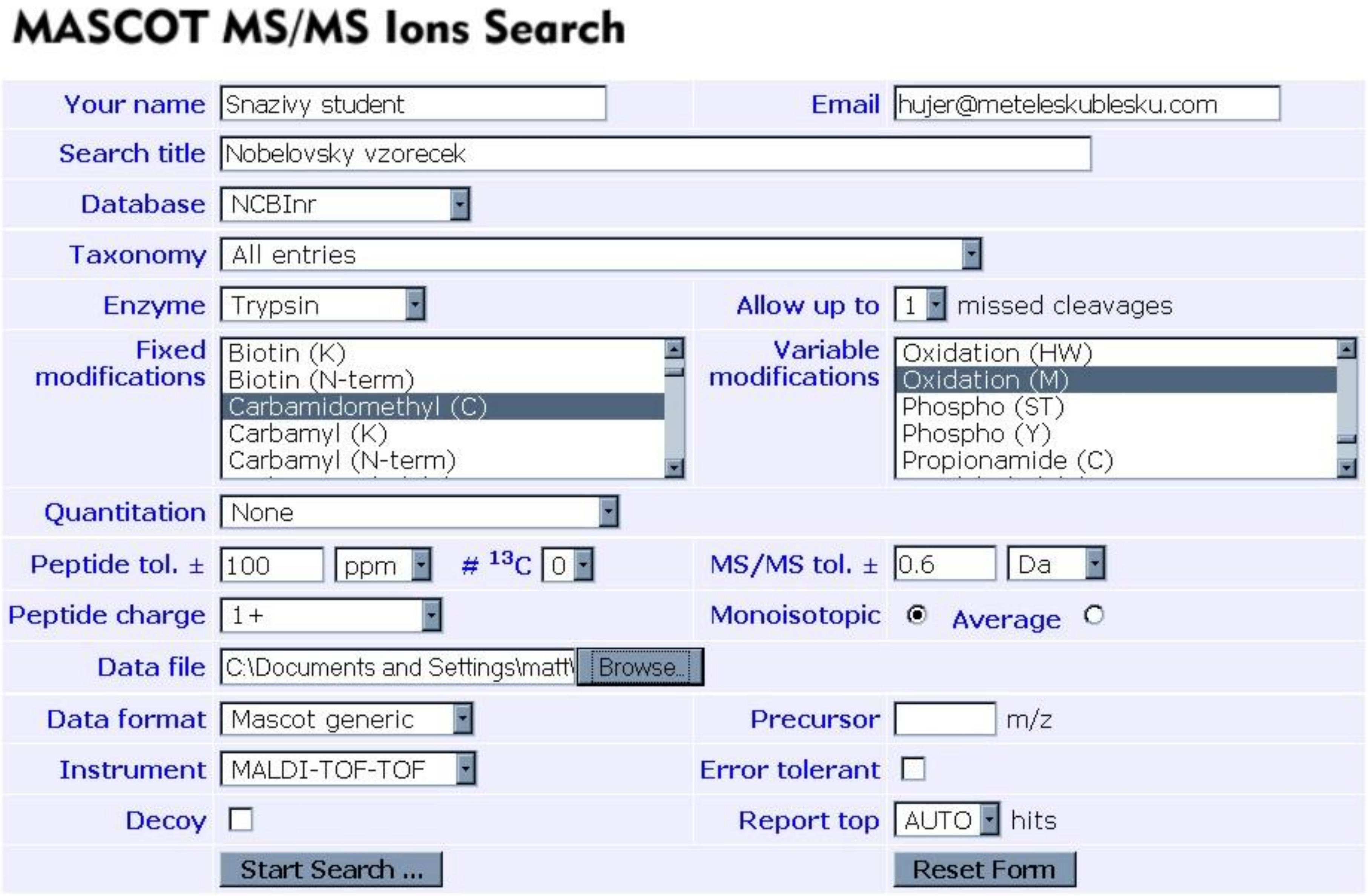
SEQUEST – Scripps, ThermoFisher

Protein Pilot – ABI Sciex

Protein Lynx - Waters



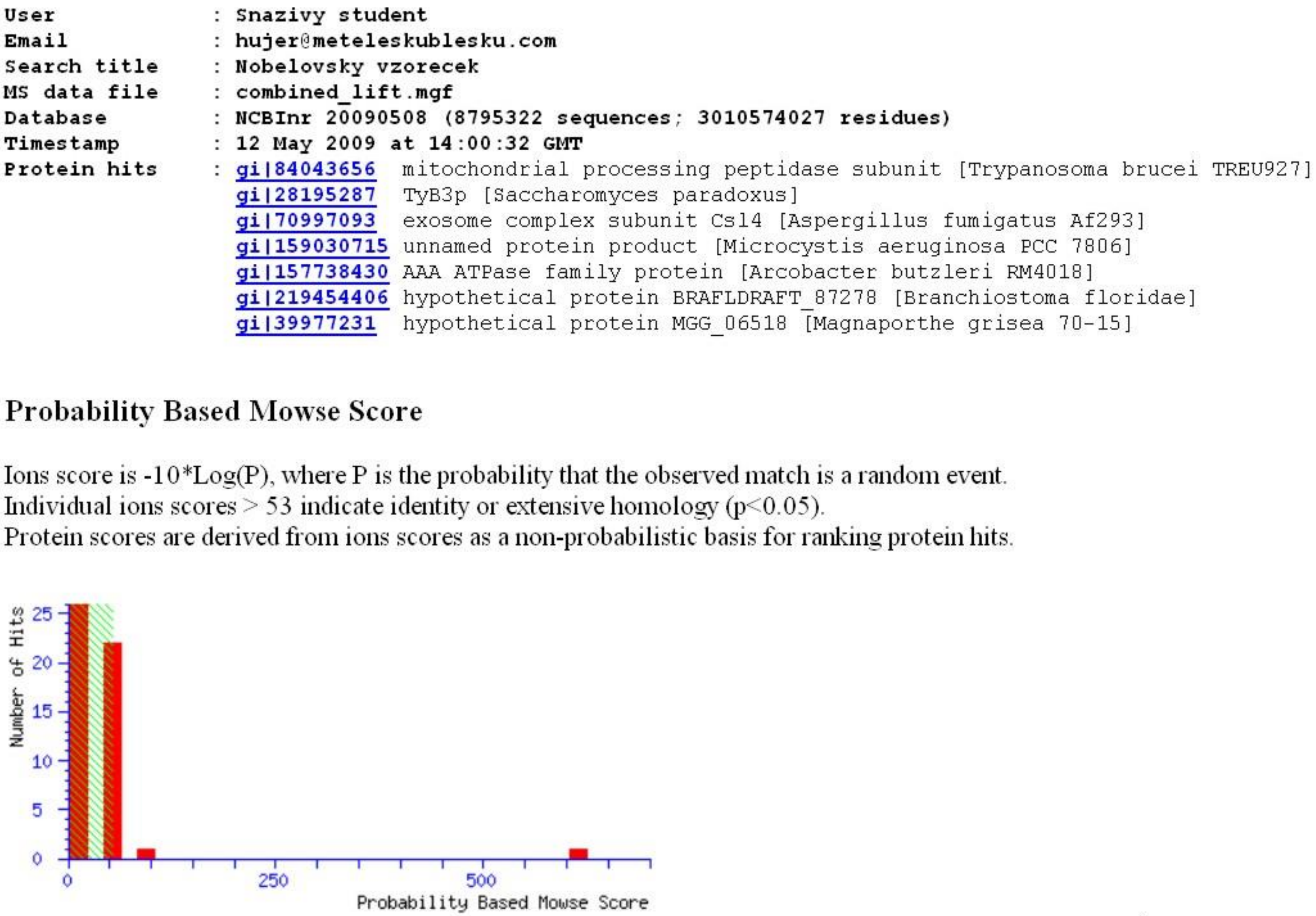
MASCOT – MS/MS



MASCOT MS/MS Ions Search

Your name	Snazivy student	Email	hujer@meteleskublesku.com
Search title	Nobelovsky vzorecek		
Database	NCBInr		
Taxonomy	All entries		
Enzyme	Trypsin	Allow up to	1 missed cleavages
Fixed modifications	Biotin (K) Biotin (N-term) Carbamidomethyl (C) Carbamyl (K) Carbamyl (N-term)	Variable modifications	Oxidation (HW) Oxidation (M) Phospho (ST) Phospho (Y) Propionamide (C)
Quantitation	None		
Peptide tol. ±	100 ppm # ^{13}C 0	MS/MS tol. ±	0.6 Da
Peptide charge	1+	Monoisotopic ●	Average ○
Data file	C:\Documents and Settings\matt\ Browse...		
Data format	Mascot generic	Precursor	m/z
Instrument	MALDI-TOF-TOF	Error tolerant	☐
Decoy	☐	Report top	AUTO hits
	Start Search ...		Reset Form

MASCOT – MS/MS

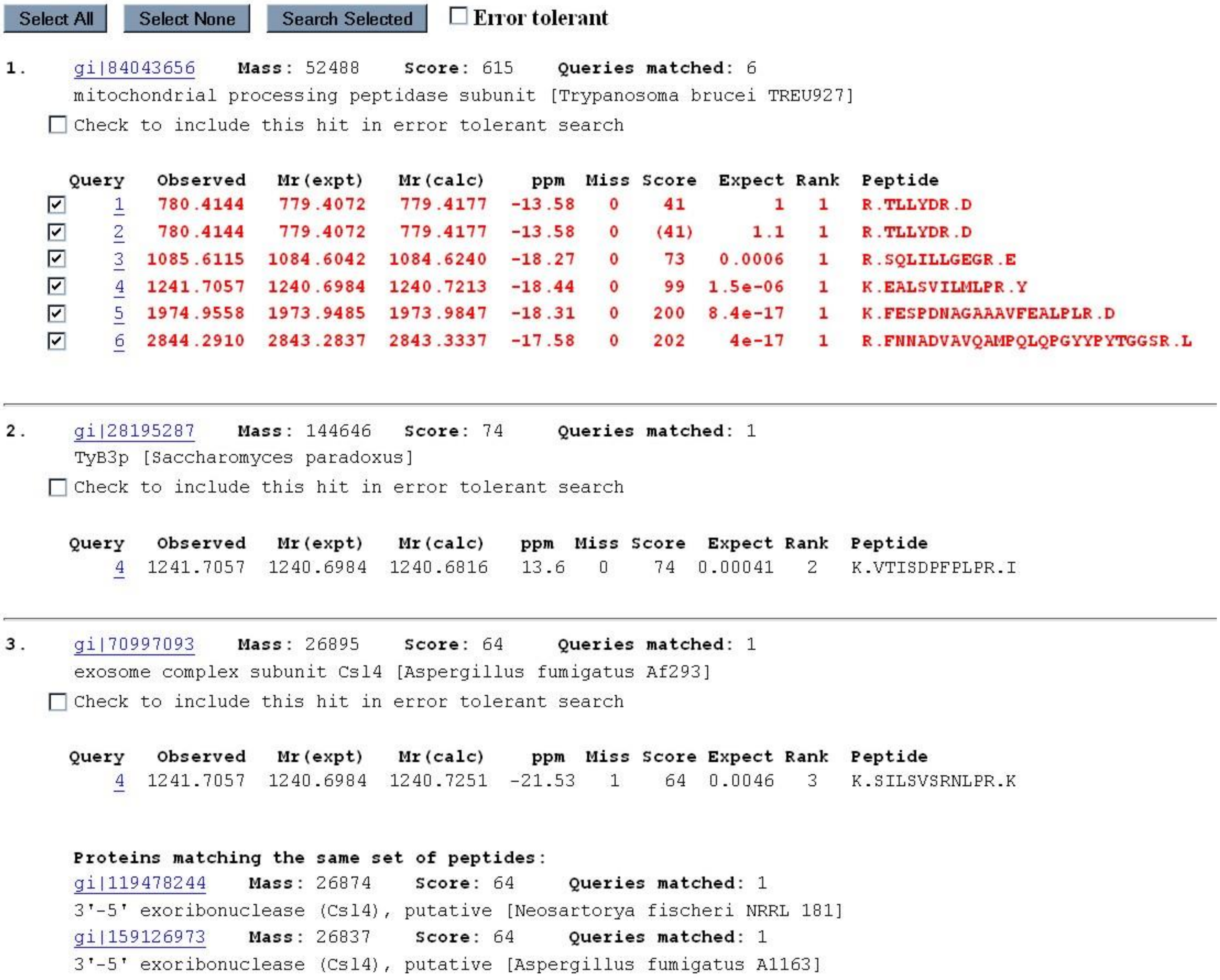


```
User             : Snazivy student
Email            : hujer@meteleskublesku.com
Search title     : Nobelovsky vzorecek
MS data file     : combined_lift.mgf
Database         : NCBInr 20090508 (8795322 sequences; 3010574027 residues)
Timestamp        : 12 May 2009 at 14:00:32 GMT
Protein hits     : gi|84043656  mitochondrial processing peptidase subunit [Trypanosoma brucei TREU927]
                   gi|28195287  TyB3p [Saccharomyces paradoxus]
                   gi|70997093  exosome complex subunit Csl4 [Aspergillus fumigatus Af293]
                   gi|159030715 unnamed protein product [Microcystis aeruginosa PCC 7806]
                   gi|157738430 AAA ATPase family protein [Arcobacter butzleri RM4018]
                   gi|219454406 hypothetical protein BRAFLDRAFT_87278 [Branchiostoma floridae]
                   gi|39977231  hypothetical protein MGG_06518 [Magnaporthe grisea 70-15]
```

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 53 indicate identity or extensive homology ($p<0.05$).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

MASCOT – MS/MS



1. gi|84043656 **Mass:** 52488 **Score:** 615 **Queries matched:** 6
mitochondrial processing peptidase subunit [Trypanosoma brucei TREU927]
☐ Check to include this hit in error tolerant search

	Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
☑	1	**780.4144**	**779.4072**	**779.4177**	**-13.58**	**0**	**41**	**1**	**1**	**R.TLLYDR.D**
☑	2	**780.4144**	**779.4072**	**779.4177**	**-13.58**	**0**	**(41)**	**1.1**	**1**	**R.TLLYDR.D**
☑	3	**1085.6115**	**1084.6042**	**1084.6240**	**-18.27**	**0**	**73**	**0.0006**	**1**	**R.SQLILLGEGR.E**
☑	4	**1241.7057**	**1240.6984**	**1240.7213**	**-18.44**	**0**	**99**	**1.5e-06**	**1**	**K.EALSVILMLPR.Y**
☑	5	**1974.9558**	**1973.9485**	**1973.9847**	**-18.31**	**0**	**200**	**8.4e-17**	**1**	**K.FESPDNAGAAAVFEALPLR.D**
☑	6	**2844.2910**	**2843.2837**	**2843.3337**	**-17.58**	**0**	**202**	**4e-17**	**1**	**R.FNNADVAVQAMPQLQPGYYPYTGGSR.L**

2. gi|28195287 **Mass:** 144646 **Score:** 74 **Queries matched:** 1
TyB3p [Saccharomyces paradoxus]
☐ Check to include this hit in error tolerant search

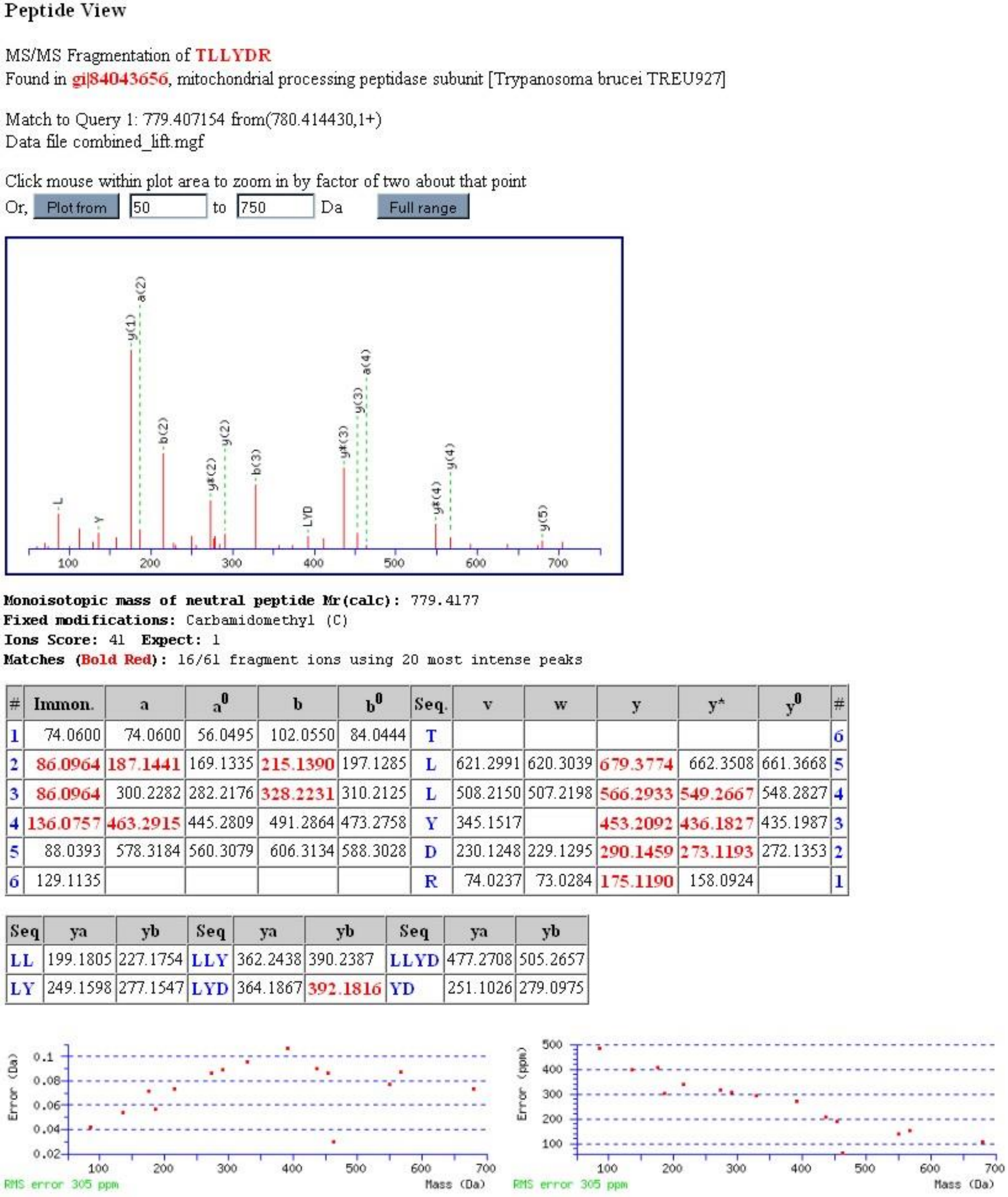
Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
4	1241.7057	1240.6984	1240.6816	13.6	0	74	0.00041	2	K.VTISDPFPLPR.I

3. gi|70997093 **Mass:** 26895 **Score:** 64 **Queries matched:** 1
exosome complex subunit Csl4 [Aspergillus fumigatus Af293]
☐ Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
4	1241.7057	1240.6984	1240.7251	-21.53	1	64	0.0046	3	K.SILSVSRNLPR.K

Proteins matching the same set of peptides:
gi|119478244 **Mass:** 26874 **Score:** 64 **Queries matched:** 1
3'-5' exoribonuclease (Csl4), putative [Neosartorya fischeri NRRL 181]
gi|159126973 **Mass:** 26837 **Score:** 64 **Queries matched:** 1
3'-5' exoribonuclease (Csl4), putative [Aspergillus fumigatus A1163]

MASCOT – MS/MS



Peptide View

MS/MS Fragmentation of **TLLYDR**
Found in gi|84043656, mitochondrial processing peptidase subunit [Trypanosoma brucei TREU927]

Match to Query 1: 779.407154 from(780.414430,1+)
Data file combined_lift.mgf

Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 50 to 750 Da Full range

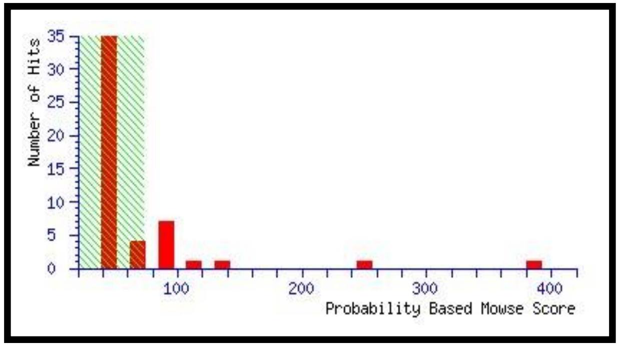
Monoisotopic mass of neutral peptide Mr(calc): 779.4177
Fixed modifications: Carbamidomethyl (C)
Ions Score: 41 **Expect:** 1
Matches (Bold Red): 16/61 fragment ions using 20 most intense peaks

#	Immon.	a	a^0	b	b^0	Seq.	v	w	y	y*	y^0	#
1	74.0600	74.0600	56.0495	102.0550	84.0444	T						6
2	**86.0964**	**187.1441**	169.1335	**215.1390**	197.1285	L	621.2991	620.3039	**679.3774**	662.3508	661.3668	5
3	**86.0964**	300.2282	282.2176	**328.2231**	310.2125	L	508.2150	507.2198	**566.2933**	**549.2667**	548.2827	4
4	**136.0757**	**463.2915**	445.2809	491.2864	473.2758	Y	345.1517		**453.2092**	**436.1827**	435.1987	3
5	88.0393	578.3184	560.3079	606.3134	588.3028	D	230.1248	229.1295	**290.1459**	**273.1193**	272.1353	2
6	129.1135					R	74.0237	73.0284	**175.1190**	158.0924		1

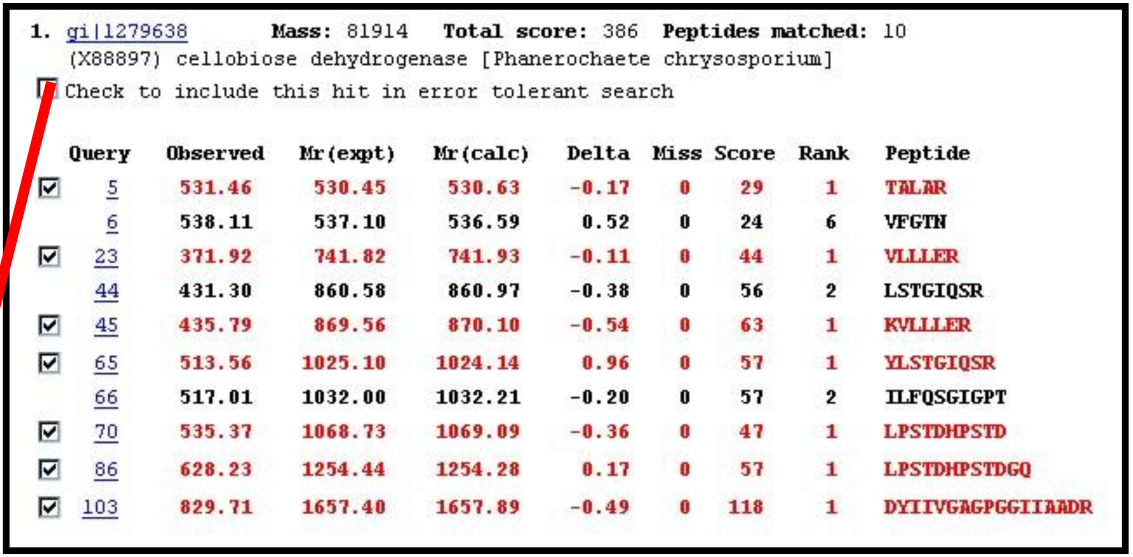
Seq	ya	yb	Seq	ya	yb	Seq	ya	yb
LL	199.1805	227.1754	LLY	362.2438	390.2387	LLYD	477.2708	505.2657
LY	249.1598	277.1547	LYD	364.1867	**392.1816**	YD	251.1026	279.0975

MASCOT – MS/MS výsledek

Skóre

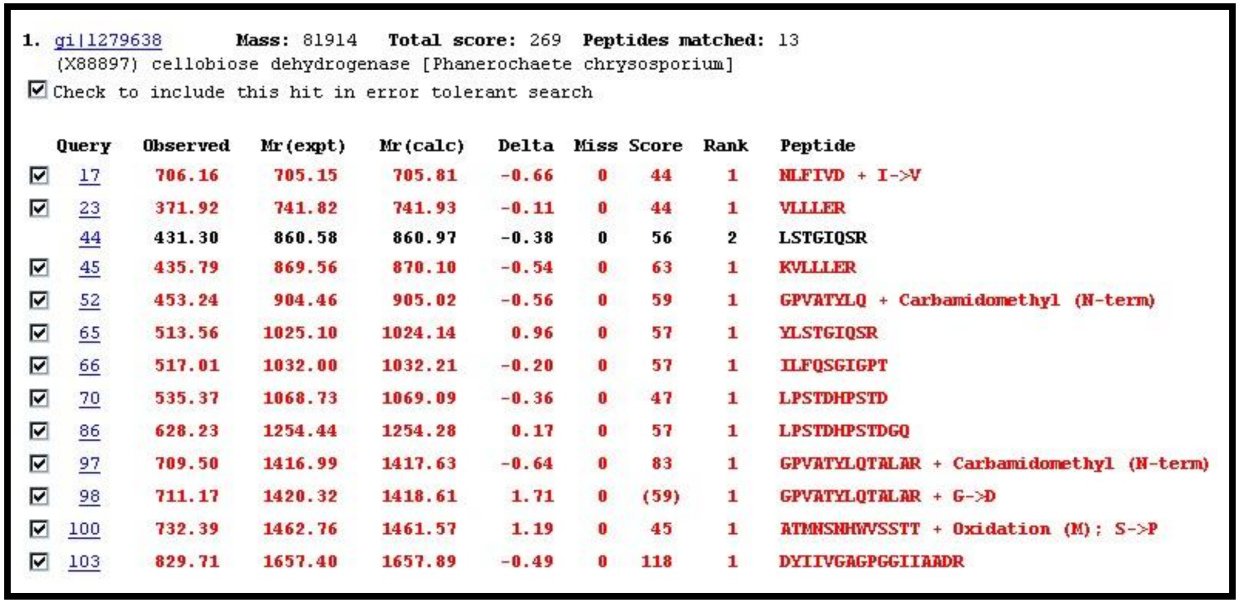


Report



1. gi|1279638 **Mass:** 81914 **Total score:** 386 **Peptides matched:** 10
(X88897) cellobiose dehydrogenase [Phanerochaete chrysosporium]
☐ Check to include this hit in error tolerant search

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
☑	5	531.46	530.45	530.63	-0.17	0	29	1	TALAR
	6	538.11	537.10	536.59	0.52	0	24	6	VFGTN
☑	23	371.92	741.82	741.93	-0.11	0	44	1	VLLLER
	44	431.30	860.58	860.97	-0.38	0	56	2	LSTGIQSR
☑	45	435.79	869.56	870.10	-0.54	0	63	1	KVLLLER
☑	65	513.56	1025.10	1024.14	0.96	0	57	1	YLSTGIQSR
	66	517.01	1032.00	1032.21	-0.20	0	57	2	ILFQSGIGPT
☑	70	535.37	1068.73	1069.09	-0.36	0	47	1	LPSTDHPSTD
☑	86	628.23	1254.44	1254.28	0.17	0	57	1	LPSTDHPSTDGQ
☑	103	829.71	1657.40	1657.89	-0.49	0	118	1	DYIIVGAGPGGIIAADR



1. gi|1279638 **Mass:** 81914 **Total score:** 269 **Peptides matched:** 13
(X88897) cellobiose dehydrogenase [Phanerochaete chrysosporium]
☑ Check to include this hit in error tolerant search

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
☑	17	706.16	705.15	705.81	-0.66	0	44	1	NLFIVD + I->V
☑	23	371.92	741.82	741.93	-0.11	0	44	1	VLLLER
	44	431.30	860.58	860.97	-0.38	0	56	2	LSTGIQSR
☑	45	435.79	869.56	870.10	-0.54	0	63	1	KVLLLER
☑	52	453.24	904.46	905.02	-0.56	0	59	1	GPVATYLQ + Carbamidomethyl (N-term)
☑	65	513.56	1025.10	1024.14	0.96	0	57	1	YLSTGIQSR
☑	66	517.01	1032.00	1032.21	-0.20	0	57	1	ILFQSGIGPT
☑	70	535.37	1068.73	1069.09	-0.36	0	47	1	LPSTDHPSTD
☑	86	628.23	1254.44	1254.28	0.17	0	57	1	LPSTDHPSTDGQ
☑	97	709.50	1416.99	1417.63	-0.64	0	83	1	GPVATYLQTALAR + Carbamidomethyl (N-term)
☑	98	711.17	1420.32	1418.61	1.71	0	(59)	1	GPVATYLQTALAR + G->D
☑	100	732.39	1462.76	1461.57	1.19	0	45	1	ATMNSNHWVSSTT + Oxidation (M); S->P
☑	103	829.71	1657.40	1657.89	-0.49	0	118	1	DYIIVGAGPGGIIAADR

„Error tolerant search“ report

Jak ověřovat MS a MS/MS data?

- Nezávislý experiment – biologický

- Manuální procházení MSMS spekter a kontrola přiřazení – expert mode, časově náročné, nemusí být 100%ní

- Statistické ověření pomocí tzv. rozpadových databází – decoy search

Jak ověřovat MS a MS/MS data?

Decoy – dvě strategie
- invertovaná databáze
- zpřeházené sekvence

Kolik procent dobrých identifikací získám v „nesmyslné“ databázi – FDR (False Discovery Rate) FDR = FP / (FP + TP)   FP – false positive, TP – true postive

http://www.mcponline.org/site/misc/PhialdelphiaGuidelinesFINALDRAFT.pdf

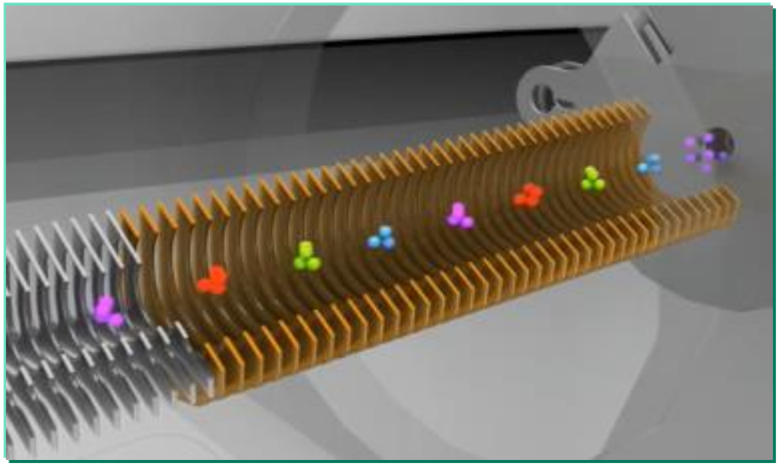
False Discovery Rates at Peptide and Protein levels : For large scale experiments, the results of any additional statistical analyses that estimate a measure of identification certainty for the dataset, or allow a determination of the false discovery rate, e.g., the results of decoy searches or other computational approaches.

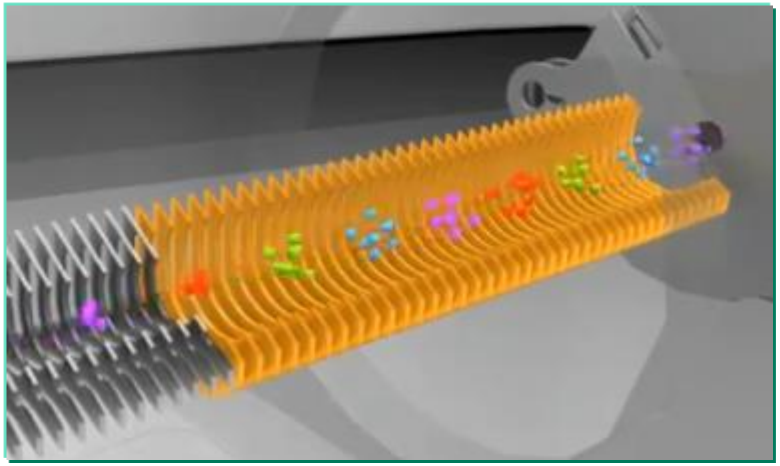
Nové přístupy = MS^e
DIA – Data Independent Acquisition

MSe – cyklické přepínání mezi normálním MS skenem a MSMS skenem kde je fragmentováno vše najednou – potřebujeme vysokou přesnost, rozlišení a velmi dobrou chromatografii (UPLC).
MSMS spektrům nerozumíme - data jsou dodatečně rozklíčována – fragmentové ionty přiřazeny prekurzorům.

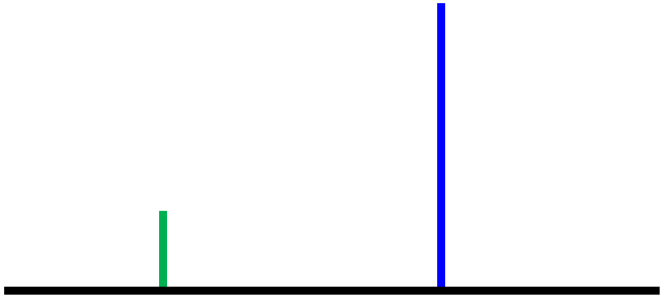
Rychlejší, větší dynamický rozsah, větší citlivost (odpadá ztráta při izolaci)

Waters

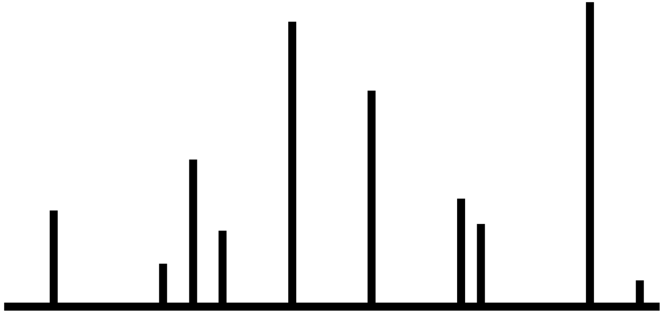




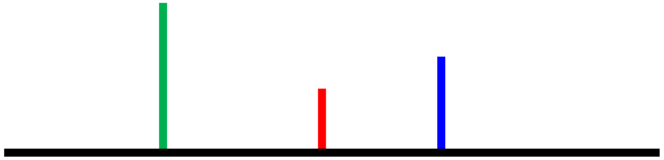
MS no 1



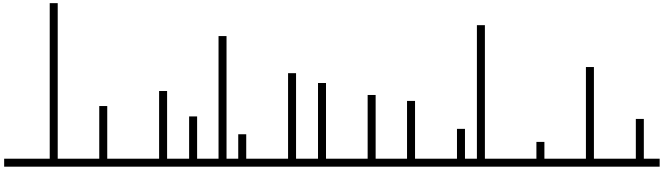
MS/MS all no 1



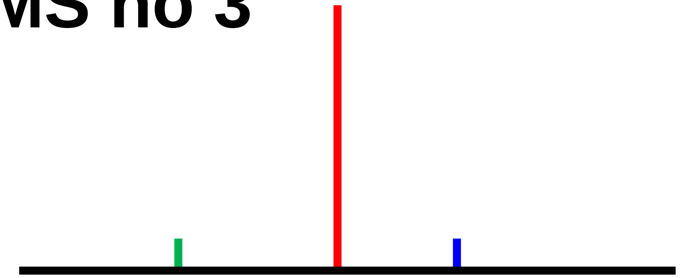
MS no 2



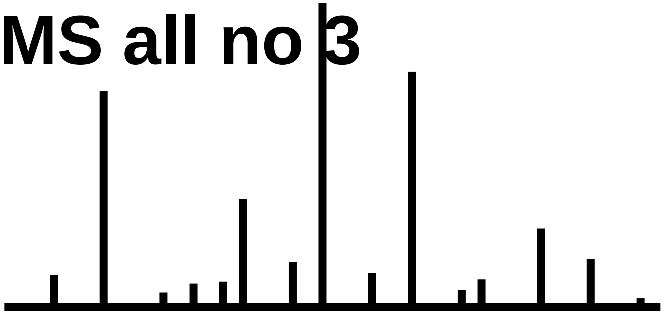
MS/MS all no 2



MS no 3

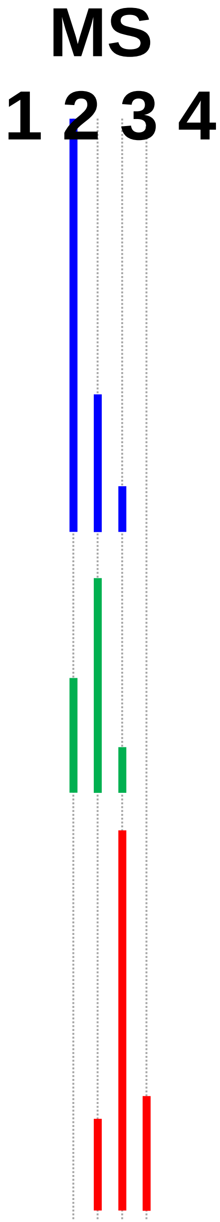


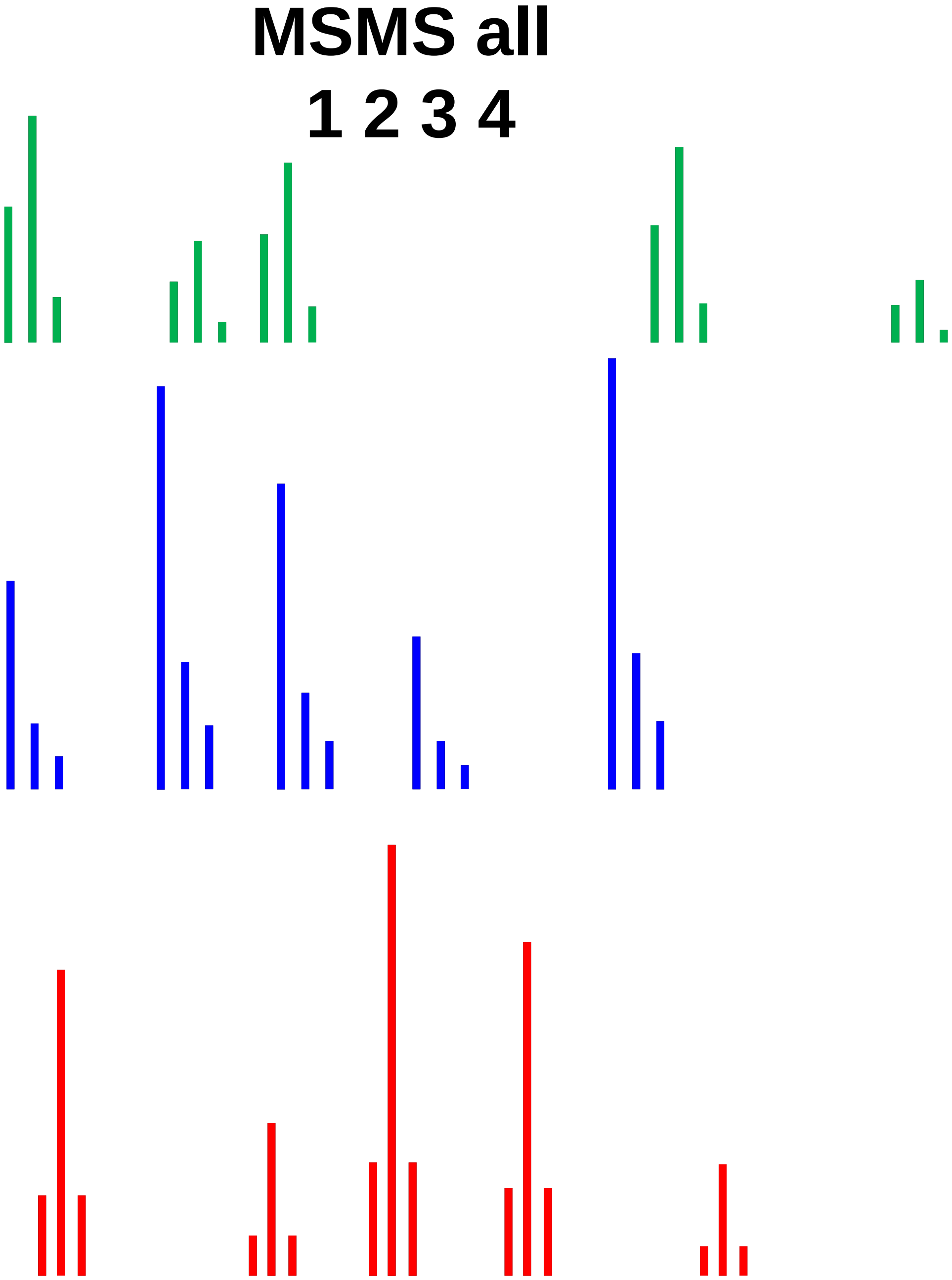
MS/MS all no 3



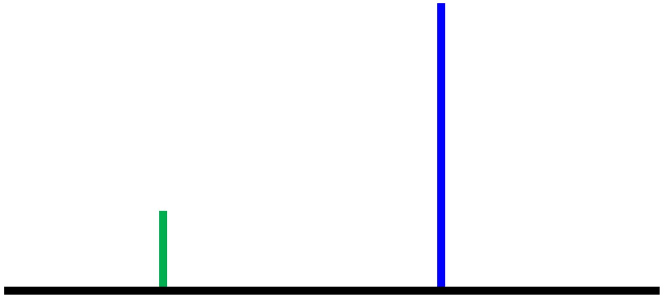
MS no 4

MS/MS all no 4

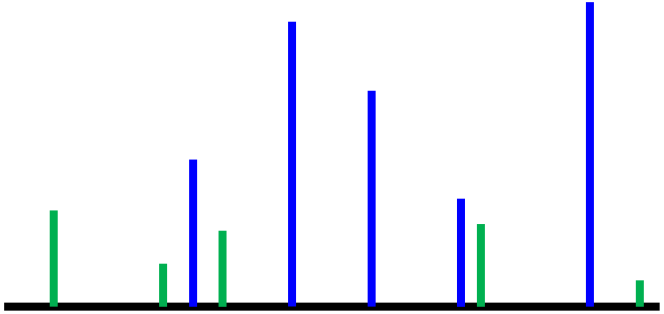

MS
1 2 3 4


MSMS all
1 2 3 4

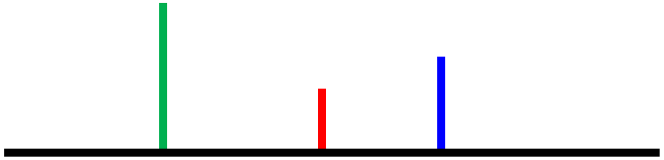
MS no 1



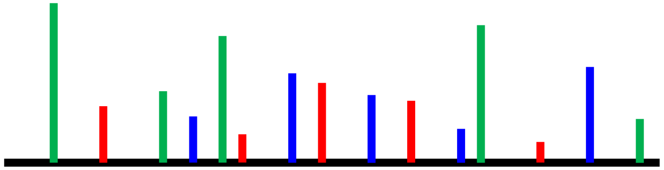
MS/MS all no 1



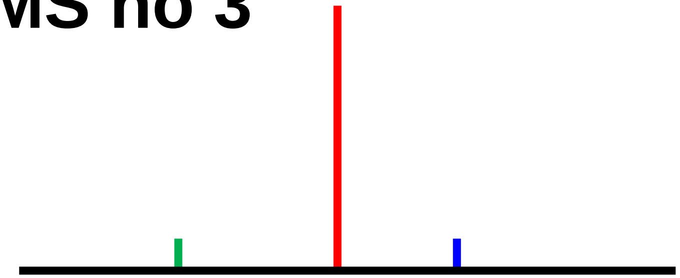
MS no 2



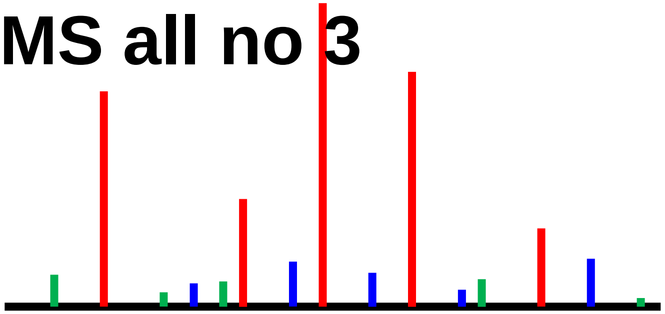
MS/MS all no 2



MS no 3



MS/MS all no 3



MS no 4

MS/MS all no 4

Post-Translační Modifikace

!!! změna MW !!!

- Fosforylace (+80 Da)
- Glykosylace (+162,203,146 Da)
- S-S můstky (-2 Da)
- Deamidace (+1 Da)
- Gln -> pyro-Glu (-17 Da)
- Oxidace (+16 Da)
- Acetylace (+42 Da)
- Acylace (+238,210 Da)
- Ztráta N-konc.Met (-131 Da)
- Zkrácení proteinu (??? Da)

Fosforylace

Signalizace, regulace proteinové aktivity, lokalizace, interakce. . .

reversibilní proces: kinasy vs. fosfatasy

Fosforylační motivy jsou různé pro jednoltivé kinasy

S (90%), T (10%), Y (0.05%),. . . . .H, K, D, E, C

1/3 eukaryotických proteinů je fosforylována

P skupina = 80 Da (PO_3)

Fosforylace – potíže…

- Nízká stechiometrie
- Většinou jsou fosfoproteiny v malém množství
- Heterogení
- Ztráta během zpracování vzorku (fosfatasy,…) a separace/purifikace
- P-peptidy se hůře ionizují, ztrácejí P při MSMS

Fosforylace – cíle…

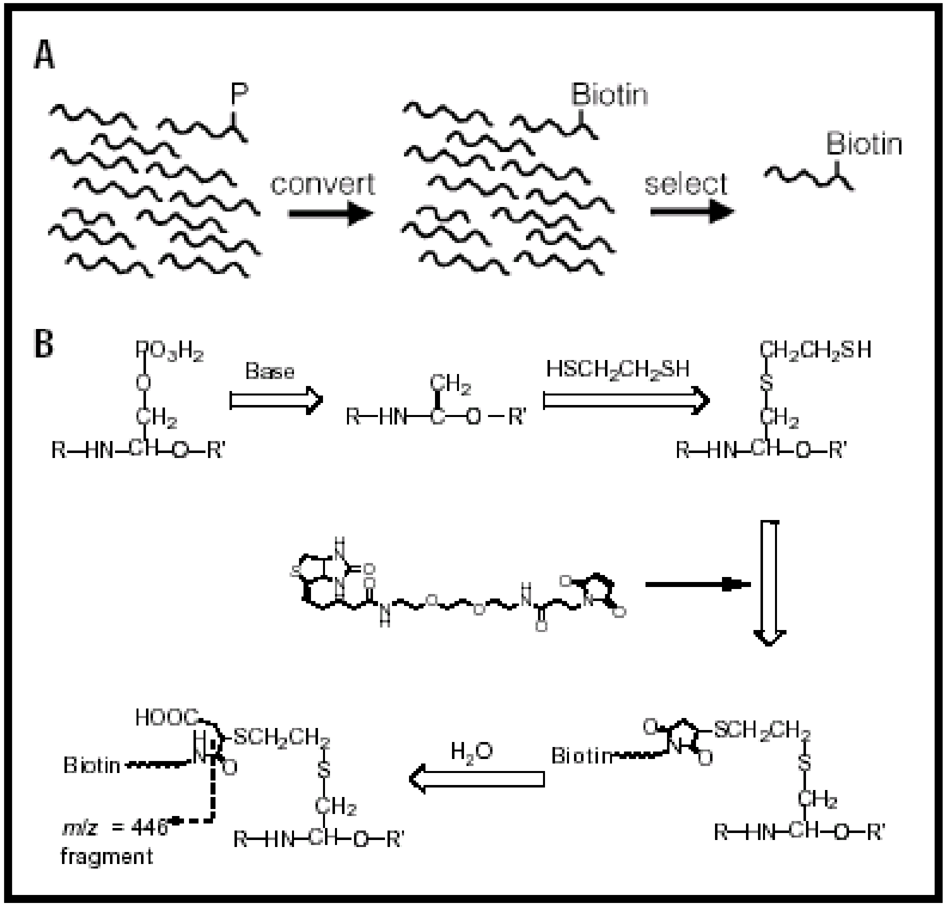
I. Identifikace/nabohacení fosfopeptidů

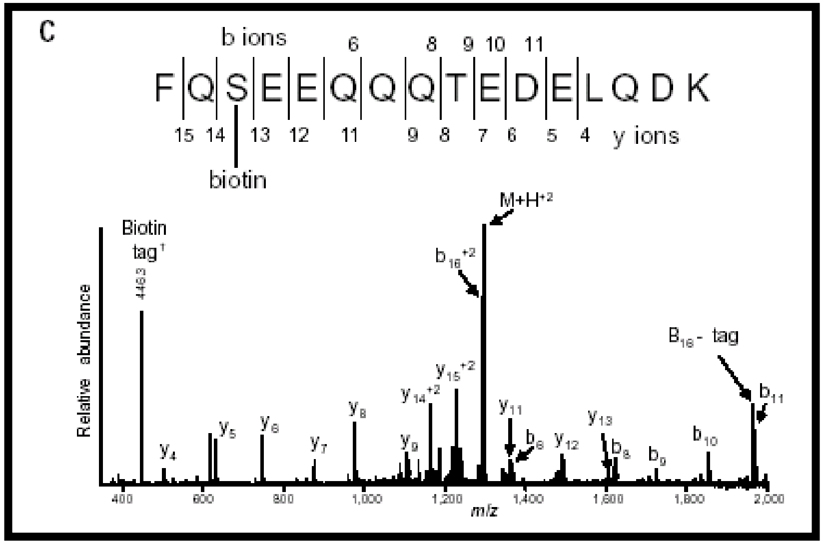
II. Lokalizace fosforylace

III. Kvantifikace fosforylací

Lovení fosfopeptidů I

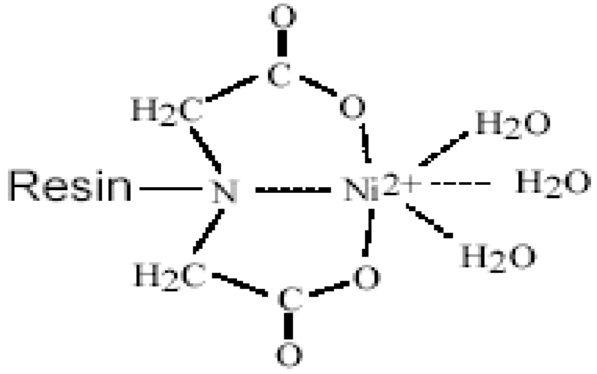
- Oxidace Cys
- Eliminace P
- Adice ethanditholu
- Biotinylace
- Afinitní purifikace
- MS, MS/MS
- Pouze pro pS, pT !!!



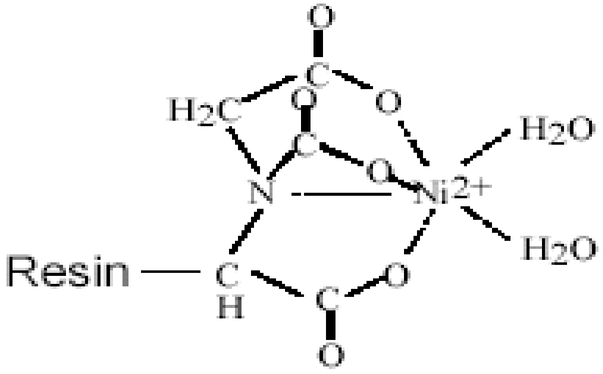


Selektivní vychytání fosfopeptidů pomocí chromatografie IMAC – Immobilized Metal Ion Affinity Chromatography

- IDA (iminodiacetic acid)
  NTA (nitrilo-triacetic acid)
- Ionty – Ga^{3+}, Fe^{3+}
- Eluce NH_4OH (nutná neutralizace)
- Odsolení na RP
- Nespecifická vazba COOH – esterifikace (potíže s deamidací…)



Nickel-IDA (Imino-DiAcetic acid) resin



Nickel-NTA (Nitrilo-TriAcetic acid) resin

Obohacení fosfopeptidů pomocí TiO_2

Digesce v gelu

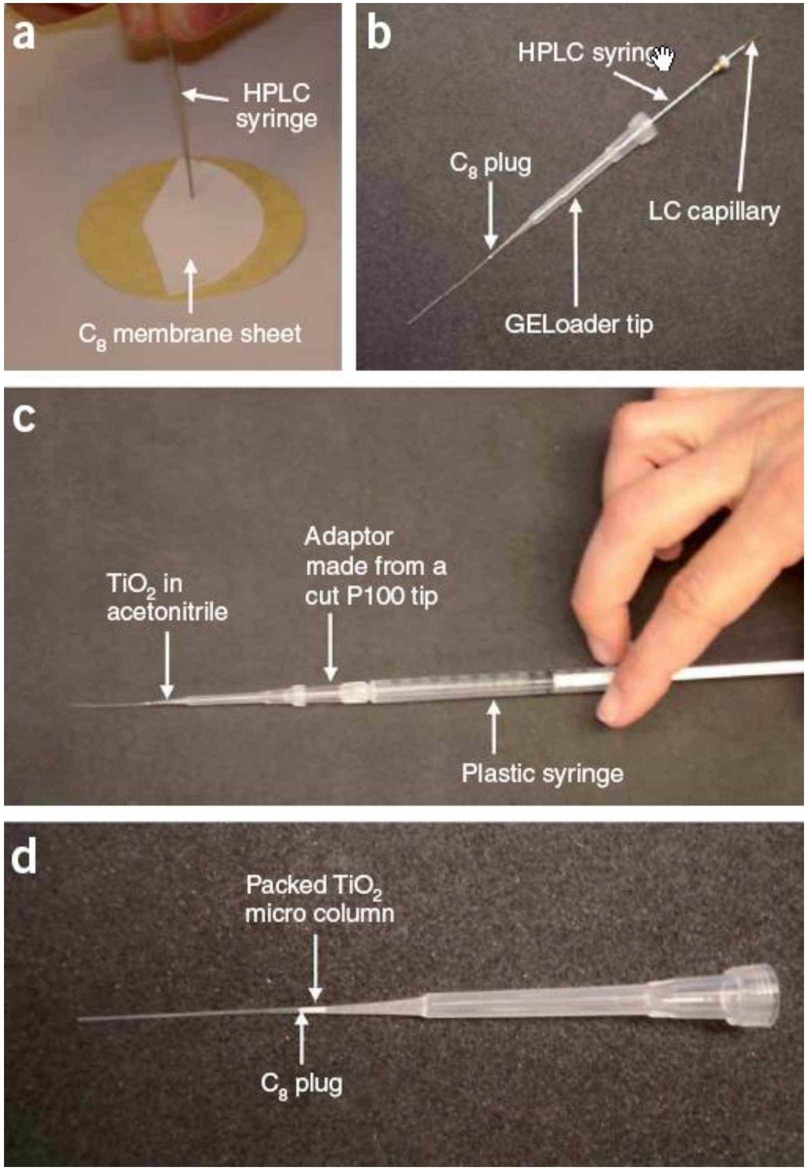
Zachycení P-peptidů na TiO_2 mikrokolonce kyselé prostředí a přídavek DHB blokuje nespecifické interakce

Eluce P-peptidů (pH 10.5)

Okyselení a purifikace na RP

Eluce matricí DHB přímo na MALDI desku

Může pevněji vázat vícenásobně fosforylované peptidy



Obohacení fosfopeptidů pomocí TiO_2

Směs tryptických peptidů z 12ti proteinů

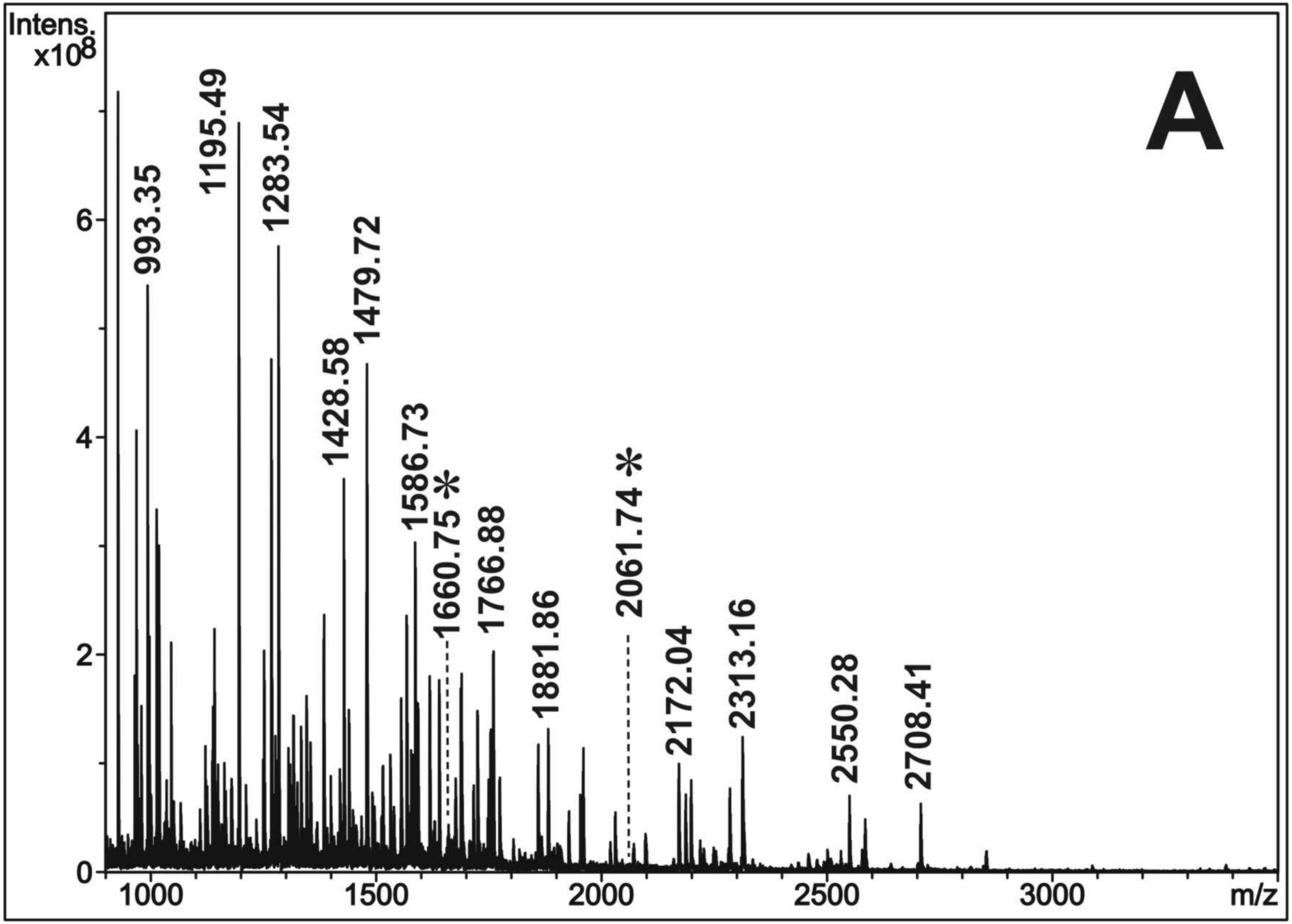
3 fosfoproteiny
Kasein alfa (S1 + S2) a beta
Ovalbumin

> 300 peptidů

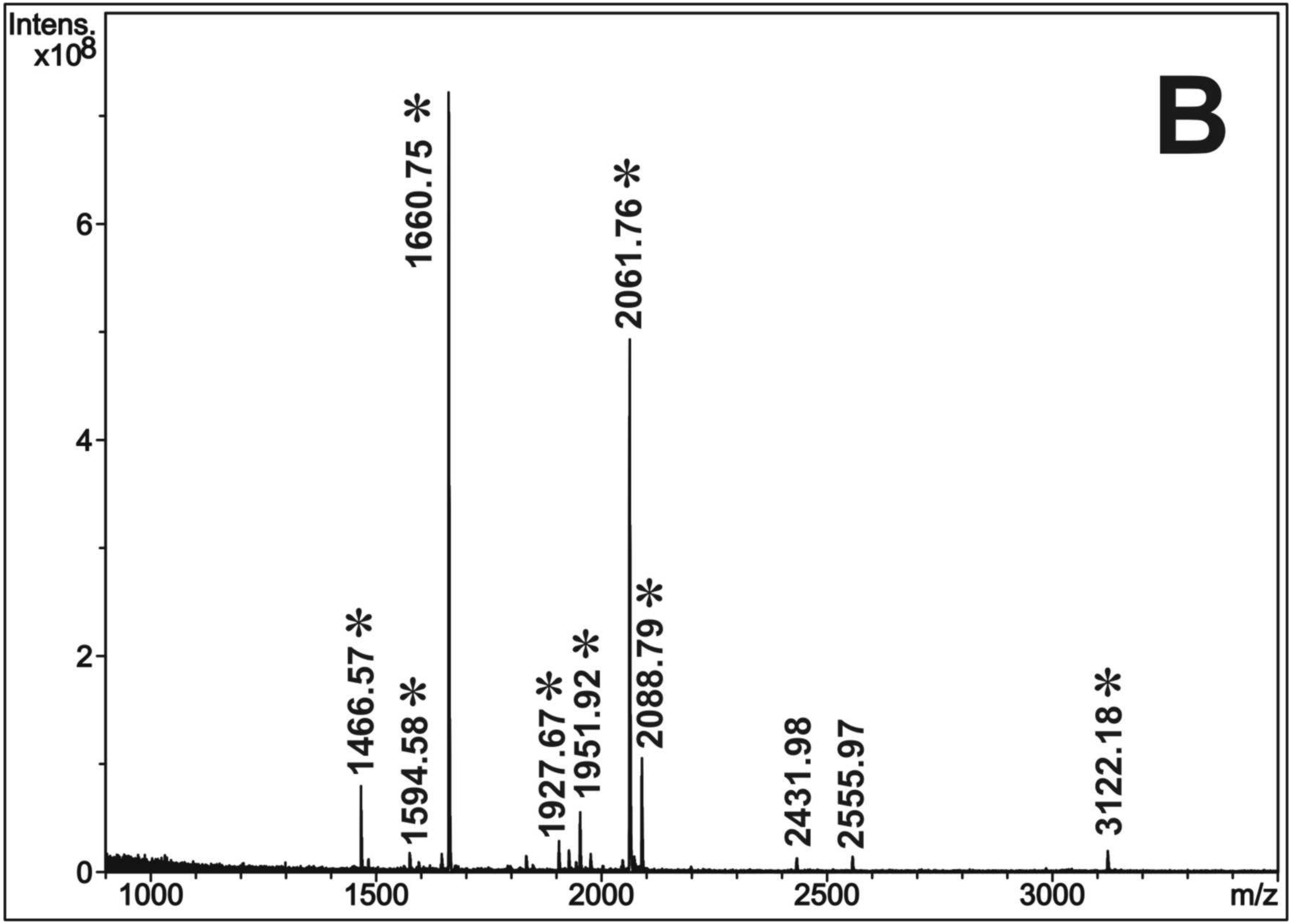
> 20 P-peptidů

Peptide sequence	Number of phosphate groups	$(M+H)^+$ Da
EVVGSAEAGVDAA (Ov. 340-352)*[1]	1	1254.52
EQLSTSEENSK (α-S2. 141-151)	1	1331.53
EQLSTSEENSK (α-S2. 141-151)	2	1411.50
TVDMESTEVFTK (α-S2. 153-164)	1	1466.61
TVDMESTEVFTKK (α-S2. 153-165)	1	1594.70
VPQLEIVPNSAEER (α-S1. 121-134)	1	1660.79
YLGEYLIVPNSAEER (α-S1)*[2]	1	1832.83
DIGSESTEDQAMEDIK (α-S1. 58-73)	1	1847.69
DIGSESTEDQAMEDIK (α-S1. 58-73)	2	1927.69
YKVPQLEIVPNSAEER (α-S1. 119-134)	1	1951.95
FQSEEQQQTEDELQDK (β-C. 33-48)	1	2061.83
EVVGSAEAGVDAASVSEEFR (Ov. 340-359)	1	2088.89
NVPGEIVESLSSSEESITR (β-C. 7-25)*[3]	4	2352.85
NTMEHVSSSEESIISQETYK (α-S2. 17-36)	4	2619.04
VNELSKDIGSESTEDQAMEDIK (α-S1. 52-73)	3	2678.01
QMEAESISSSEEIVPNSVEAQK (α-S1. 74-94)	5	2720.91
NTMEHVSSSEESIISQETYKQ (α-S2. 17-37)*[4]	4	2747.10
EKVNELSKDIGSESTEDQAMEDIK (α-S1. 50-73)	3	2935.15
ELEELNVPGEIVESLSSSEESITR (β-C. 17-40)	4	2966.16
NANEEEYSIGSSSEESAEVATEEVK (α-S2. 61-85)	4	3008.01
NANEEEYSIGSSSEESAEVATEEVK (α-S2. 61-85)	5	3087.99
RELEELNVPGEIVESLSSSEESITR (β-C. 16-40)	4	3122.27
KNTMEHVSSSEESIISQETYKQEK (α-S2. 16-39)*[3]	4	3132.20

Obohacení fosfopeptidů pomocí TiO_2



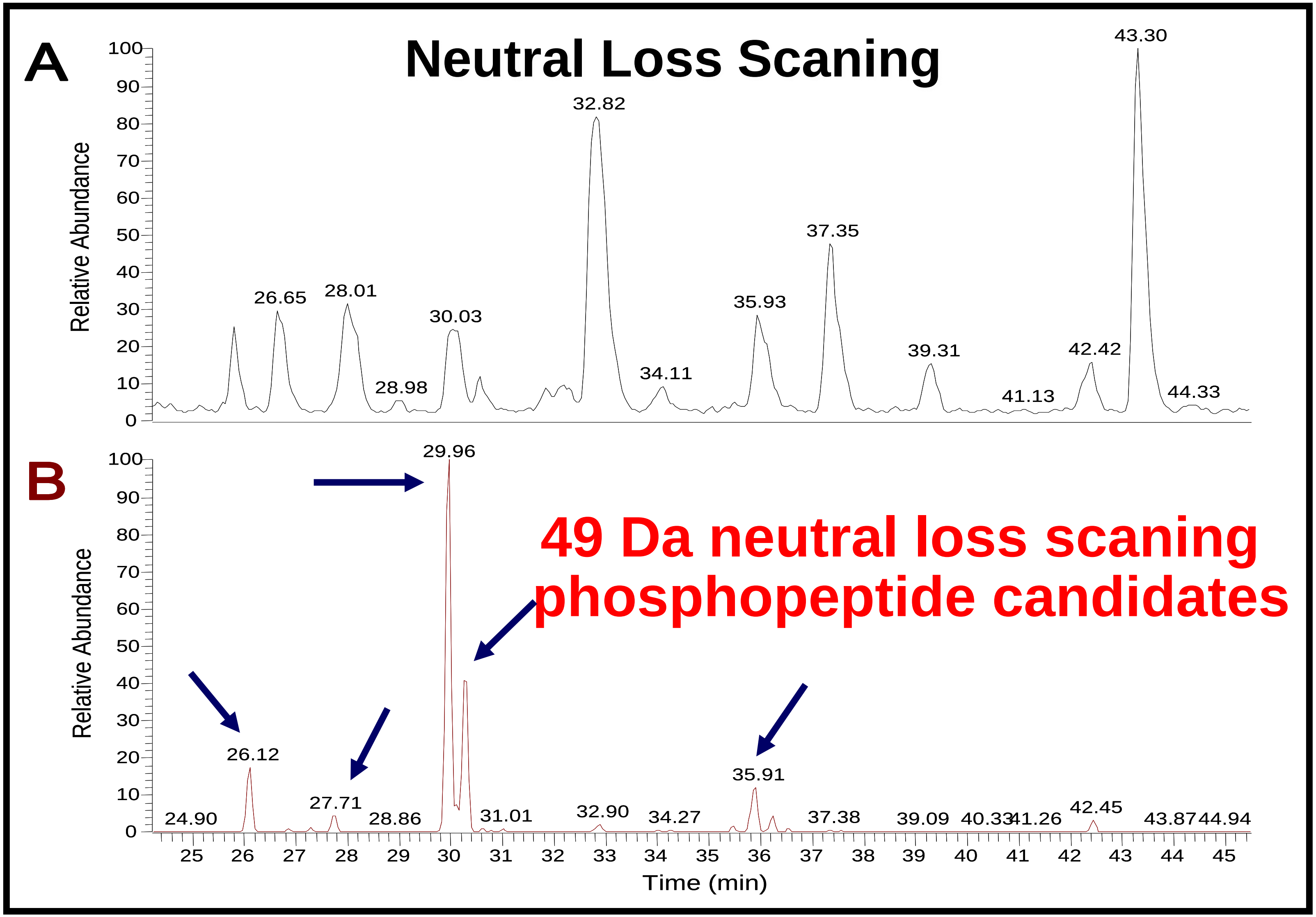
Obohacení fosfopeptidů pomocí TiO_2



Identifikace P-peptidů pomocí MS

- pY $[MH-80]^+$, pS,pT $[MH-98]^+$
- Neutrální ztráta (pS, pT: 98, 49, 33)
- P-peptidy spouštěné MS^3 (MS^2 – je-li ztráta P -> MS^3 na hlavním píku v MS^2)
- Prekurzorový sken +ní mód (pY m/z 216.043)
- Prekurzorový sken -ní mód(pS, pT, pY: m/z 79)

Identifikace P-peptidů pomocí MS



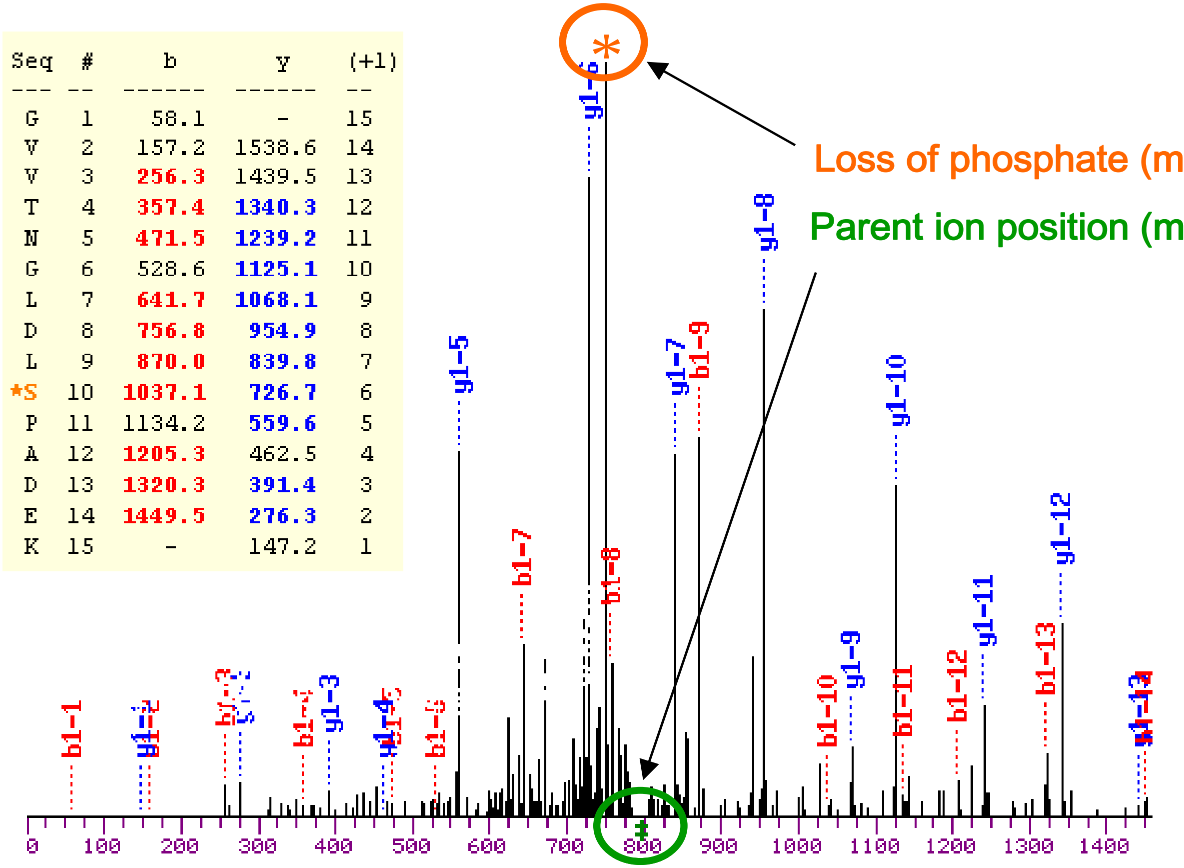
Neutral Loss Scaning

49 Da neutral loss scaning phosphopeptide candidates

Phosphoproteomics: Methods III
localization of phosphorylation site(s)

A. tandem mass spectrometry MS/MS

GVVTNGLDLSPADEK



Seq	#	b	y	(+1)
G	1	58.1	-	15
V	2	157.2	1538.6	14
V	3	256.3	1439.5	13
T	4	357.4	1340.3	12
N	5	471.5	1239.2	11
G	6	528.6	1125.1	10
L	7	641.7	1068.1	9
D	8	756.8	954.9	8
L	9	870.0	839.8	7
*S	10	1037.1	726.7	6
P	11	1134.2	559.6	5
A	12	1205.3	462.5	4
D	13	1320.3	391.4	3
E	14	1449.5	276.3	2
K	15	-	147.2	1

Loss of phosphate (m/z749)

Parent ion position (m/z 798)

Glykosylace

- Mnoho důležitých funkcí: stabilizace struktury, dobré poskládání, ochrana před degradací, interakce ligand-receptor, targeting, modulace aktivity

- N-glykosylace
  na sekvenci: NXT/S/C – v β-strukturách nebo smyčkách
  sacharid: větvený, 3 typy: mannosový, komplexní, hybridní

- O-glykosylace
  nemá klasický motiv
  S, T, ale také hydroxyPro, Lys, atd.
  sacharid: cokoliv od monosacharidu pro dlouhé, případně mírně větvené struktury

N-glykosylace

1. Charakterizace sacharidových struktur+určení heterogenity
2. Nalezení míst glykosylace
3. Přiřazení struktur na jednotlivá glykosylační místa

1. a) uvolnění glykanu(-ů) –
   <u>enzymaticky</u> – PNGasa F, Endo H,…
   b) charakterizace glykanů – NMR, MS, MS/MS, permethylace+MS/MS, exoglykosidasy+MS

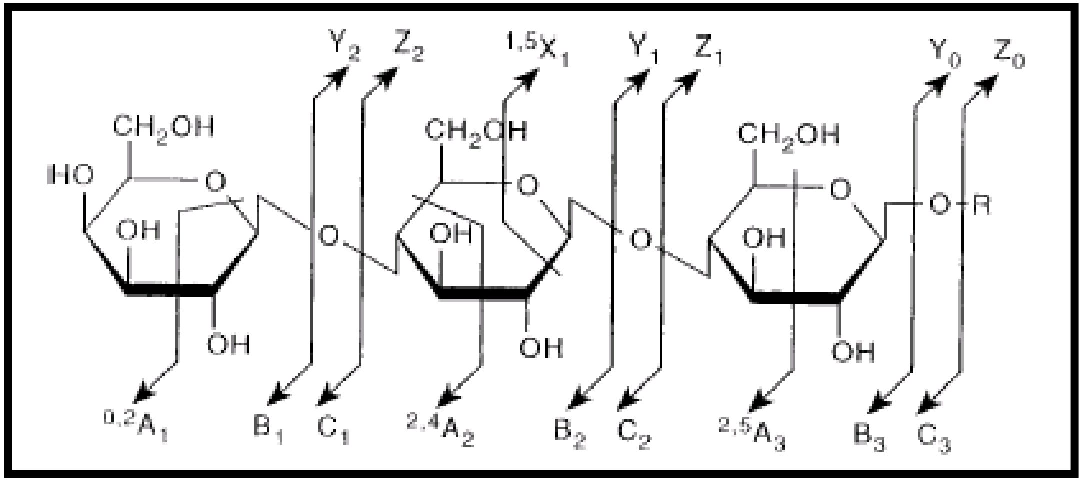
2. Deglykosylace (PNGasa F – N->D, využití $H_2{}^{18}O$) – identifikace „modifikovaných peptidů"

3. Štěpení proteinu na peptidy, separace, identifikace a nabohacení glykopeptidů – opakování kroku 1. nebo přímo MS

Určení struktury glykanu

Určení MW - MS

MS/MS – nedá dost informace o větvení (nutno modifikovat či dělat permethylaci)



Nomenklatura
Domon-Costello (1998)

MS “vidí“ hmotnosti – nerozliší např. Glc a Man - Hex (162 Da), dHex (146 Da), HexNAc (203 Da), Sia (291 Da),...

. . . naštestí víme jaké glykany očekávat, nepotřebuje znát přesnou strukturu, spíš nás zajímá např. zda je přítomna fukosylace, atd – mnohdy vystačíme s přesnou MW a hledání v dtb glykanů (Glycosuite)

Jak to lze urychlit? - přímá analýza glykopeptidů

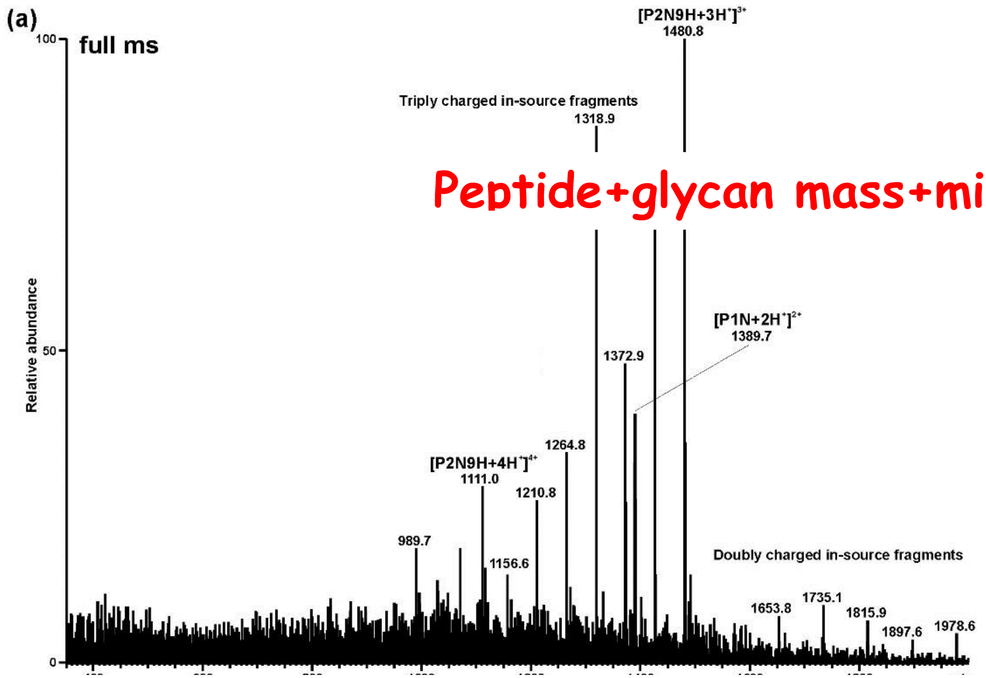
- Štěpení různými proteázami a LC-MS/MS (iontová past, q-tof, q-FT-MS) nebo LC-MS (FT-MS)

- Jak identifikovat glykopeptidy
  a) hledáme v MS záznamu diference odpovídající sacharidům (např. 162, 81, 54 pro Hexosu)
- b) hledáme oxoniové ionty – fragmenty glykanu (163, 204, 274, 292, 366)
- c) pomocí SW hledáme glykopeptidy podle velice přesných m/z (FT-ICR)

DorinM …ověření sekvence…

MLRNVCPVLILLIIGATAQDPTDVGEAFANVEWSVAELKRVL
VMGVPRDCGELFLSGQNHSGVYNIYPYKDSLLPVS
AYCDMETDGGGWTVFQRRGQFGNPVYYFYKKWA
DYAHGFGDPAKEYWLGNNVLHALTSDKAMSLRIE
MKNHSLETLTAEYSVFKVASEEEYFKINVGGYIGSK
GSDAFSIANGSMFTASDQDHDTYTNNCAVEFKGA
WWYTSCHGSNLNGLNLNGEHPSYADGIEWSARG
GSTGLYYYSYPNVEMKVRDAHFISRVADGRAS

MS, MS2, MS3



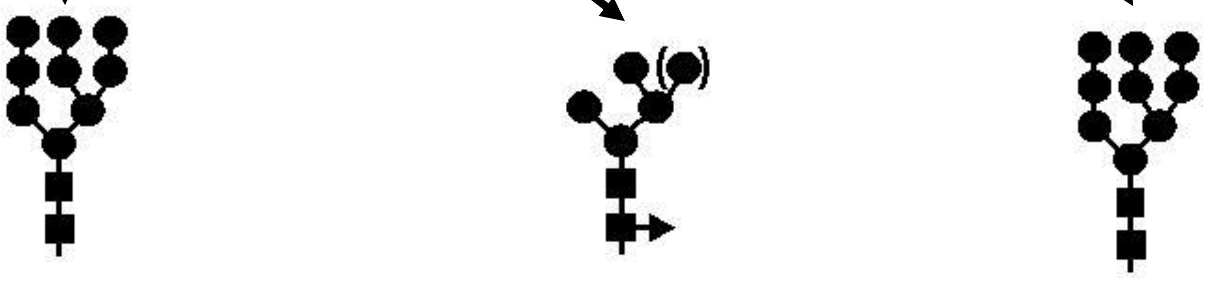
Peptide+glycan mass+microheterogeneity

Peptide sequence
&
Site of attachment

Glycan structure

… 100% pokytí a určení tří míst glykosylace (+ odhalení pyroGlu na N-konci)

MLRNVCPVLILLIIGATA**(pyro-Glu)**DPTDVGEAFANVEWSV
AELKRVLVMGVPR**DCGELFLSGQNHSGVYNIYPYK**
DSLLPVSAYCDMETDGGGWTVFQRRGQFGNPVY
YFYKKWADYAHGFGDPAKEYWLGNNVLHALTSDK
AMSLRIEMK**NHSLETLTAEYSVFK**VASEEEYFKINV
GGYIGSK**GSDAFSIANGSMFTASDQDHDTYTNNC**
AVEFKGAWWYTSCHGSNLNGLNLNGEHPSYADGI
EWSARGGSTGLYYYSYPNVEMKVRDAHFISRVAD
GRAS



O- (a N-) glykosylace = Propeptid z *Penicillium oxalicum.*

VKVNPLPAPRNITWGSSGPISITKPALHLENHHGQNQDILHHAWDRTWATITN
LEWVPAAIEAPIPSFRPFPTPA(D)

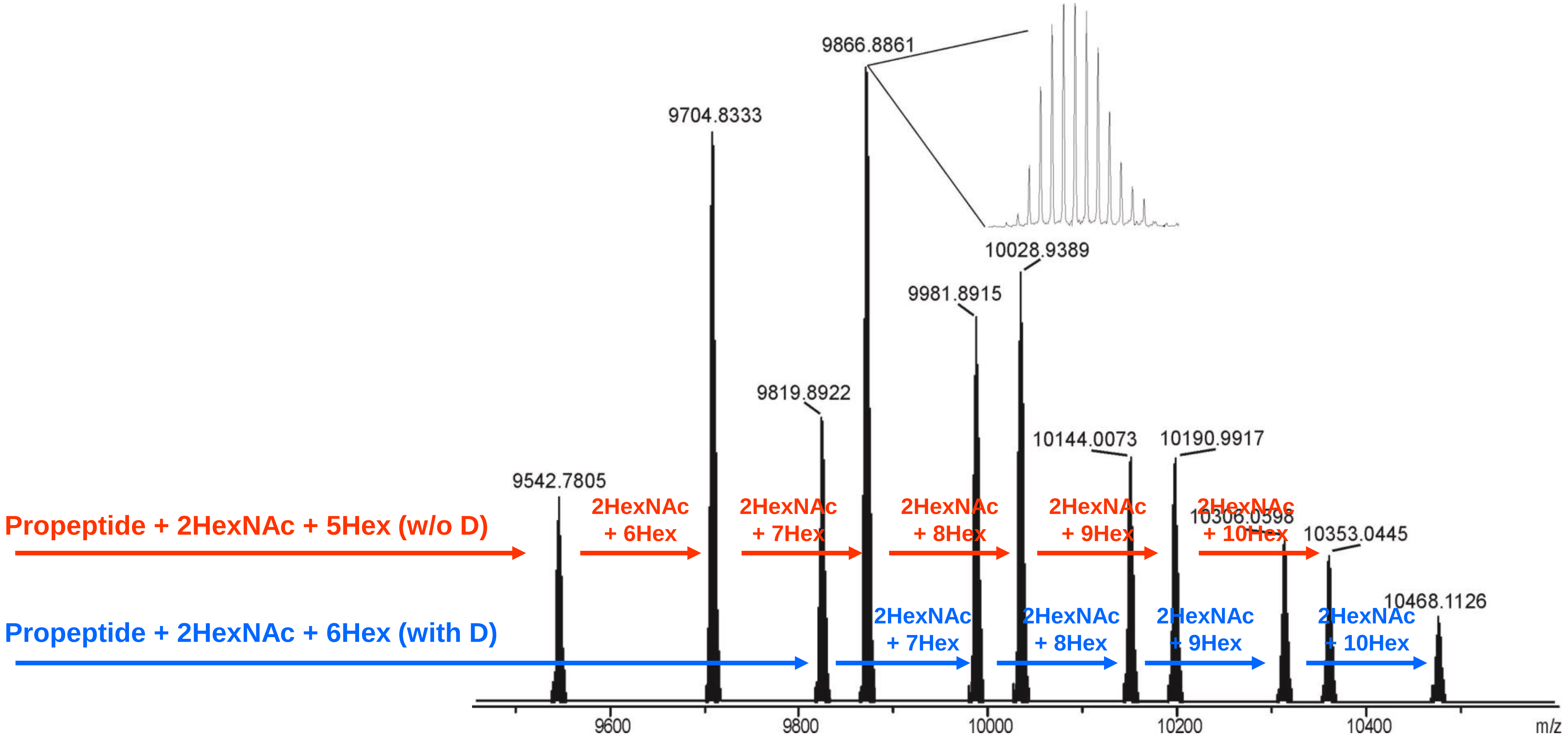
$[M+H^+]^+$ = 8326.3431 Da (without C-term D)

$[M+H^+]^+$ = 8441.3700 Da (with C-term D)

Theoretical Mass: 9866.8715

Experimental Mass: 9866.8861

Error: 1.4 ppm



Propeptide + 2HexNAc + 5Hex (w/o D)

2HexNAc + 6Hex

2HexNAc + 7Hex

2HexNAc + 8Hex

2HexNAc + 9Hex

2HexNAc + 10Hex

Propeptide + 2HexNAc + 6Hex (with D)

2HexNAc + 7Hex

2HexNAc + 8Hex

2HexNAc + 9Hex

2HexNAc + 10Hex

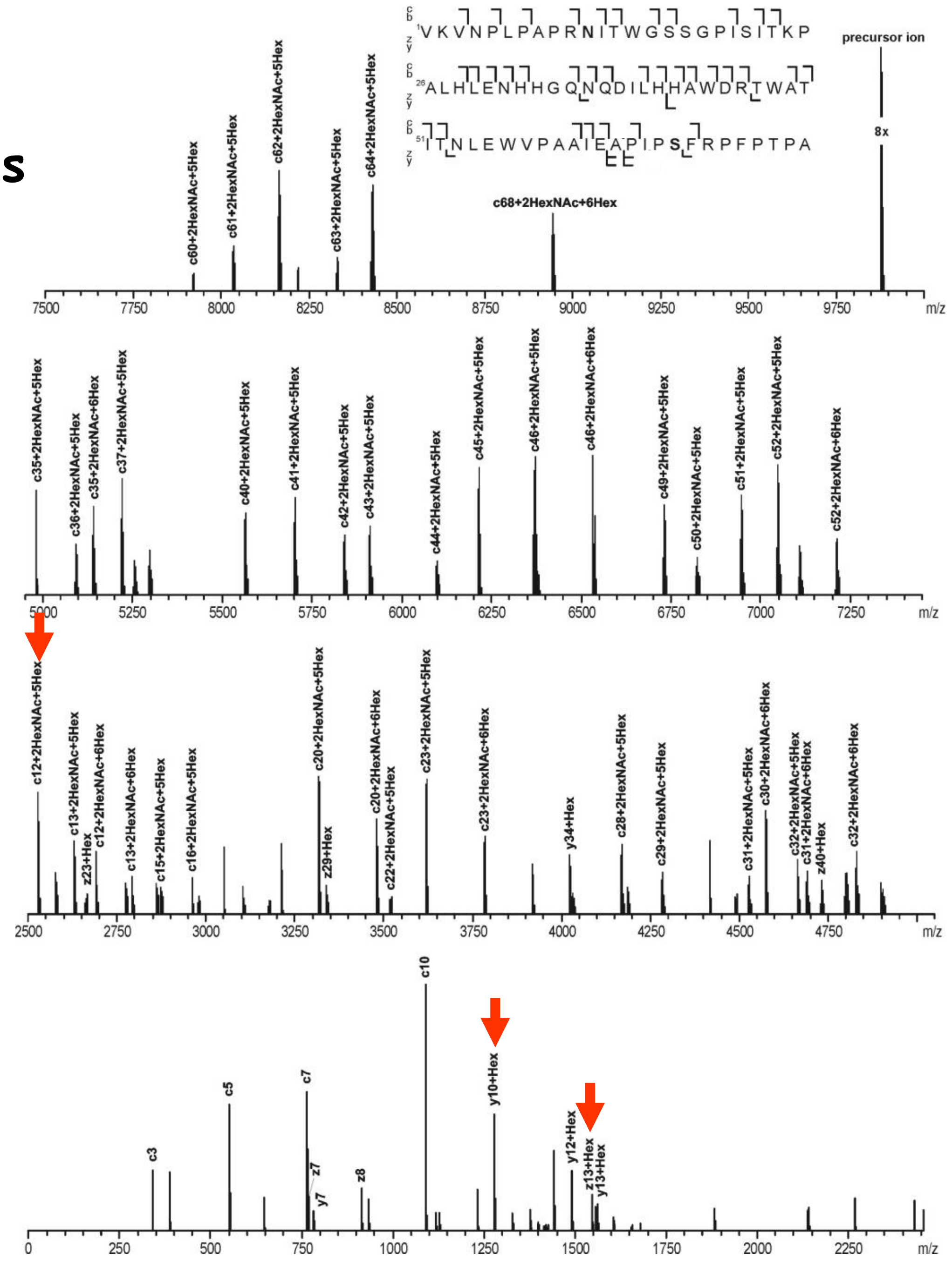
ECD - *P. oxalicum*

Propeptide w/o -D at C terminus

Observed Ions

c3
c5
c7
c9
c10
c12 + 2HexNAc + 5HEX
c12 + 2HexNAc + 6HEX
c13 + 2HexNAc + 5HEX
c13+ 2HexNAc + 6HEX
c15 + 2HexNAc + 5HEX
c16 + 2HexNAc + 5HEX
c17 + 2HexNAc + 6HEX
c17 + 2HexNAc + 5HEX
c20 + 2HexNAc + 5HEX
c20 + 2HexNAc + 6HEX
c22 + 2HexNAc + 5HEX
c23 + 2HexNAc + 5HEX
c23 + 2HexNAc + 6HEX
c28 + 2HexNAc + 5HEX
c29 + 2HexNAc + 5HEX
c30 + 2HexNAc + 5HEX
c30 + 2HexNAc + 6HEX
c31 + 2HexNAc + 5HEX
c32 + 2HexNAc + 5HEX
c32 + 2HexNAc + 6HEX

c33 + 2HexNAc + 5HEX
c35 + 2HexNAc + 5HEX
c35 + 2HexNAc + 6HEX
c36 + 2HexNAc + 5HEX
c37 + 2HexNAc + 5HEX
c40 + 2HexNAc + 6HEX
c41 + 2HexNAc + 5HEX
c42 + 2HexNAc + 5HEX
c43 + 2HexNAc + 5HEX
c44 + 2HexNAc + 5HEX
c45 + 2HexNAc + 5HEX
c45 + 2HexNAc + 6HEX
c46 + 2HexNAc + 5HEX
c46 + 2HexNAc + 6HEX
c49 + 2HexNAc + 5HEX
c50 + 2HexNAc + 5HEX
c51 + 2HexNAc + 5HEX
c51 + 2HexNAc + 6HEX
c52 + 2HexNAc + 5HEX
c52 + 2HexNAc + 6HEX
c60 + 2HexNAc + 5HEX
c61 + 2HexNAc + 5HEX
c62 + 2HexNAc + 5HEX
c63 + 2HexNAc + 5HEX
c68 + 2HexNAc + 6HEX



Other Ions: z8, z13 + HEX, z23 + HEX, z29 + HEX, z40 + HEX, y7, y10 + HEX, y13 + HEX, y34 + HEX