

**Proteomické aplikace a  
experimenty  
v onkologickém výzkumu**

Pavel Bouchal

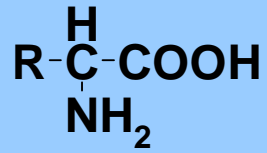
# Analýza genové exprese: mRNA nebo protein?

**Protein: „proteom“** (PROTEin complement expressed by a genOME)

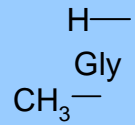
- informace o skutečných efektorech buněčných procesů – (což není případ mRNA)
- proteiny jsou oproti mRNA chemicky velmi heterogenní – různá hydrofobicita aminokyselin – stanovení genové exprese na úrovni proteinu je technicky složitější

**mRNA: „transkriptom“**

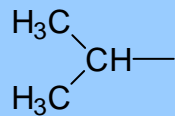
- hladiny mRNA neodpovídají plně hladinám proteinů (degradace mRNA-čas, splice varianty, posttranslační modifikace proteinů)
- stanovení genové exprese je snadnější ve srovnání s proteiny



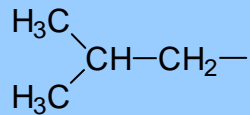
## PŘEHLED AMINOKYSELIN



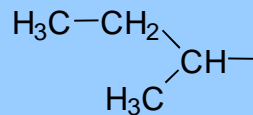
Ala



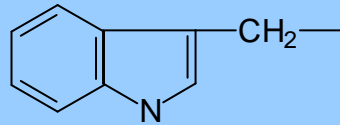
Val



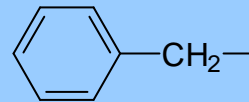
Leu



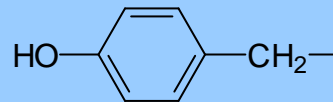
Ile



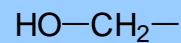
Trp



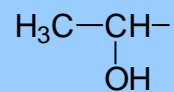
Phe



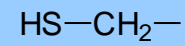
Tyr



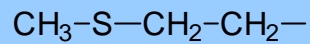
Ser



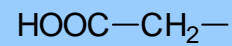
Thr



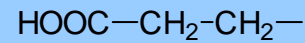
Cys



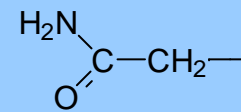
Met



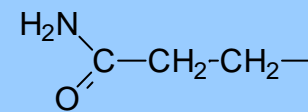
Asp



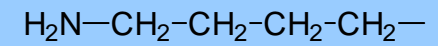
Glu



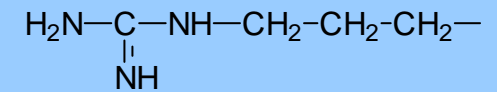
Asn



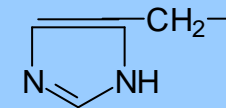
Gln



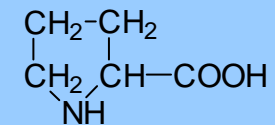
Lys



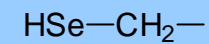
Arg



His



Pro



SeCys

# Analýza genové exprese: metody k dispozici

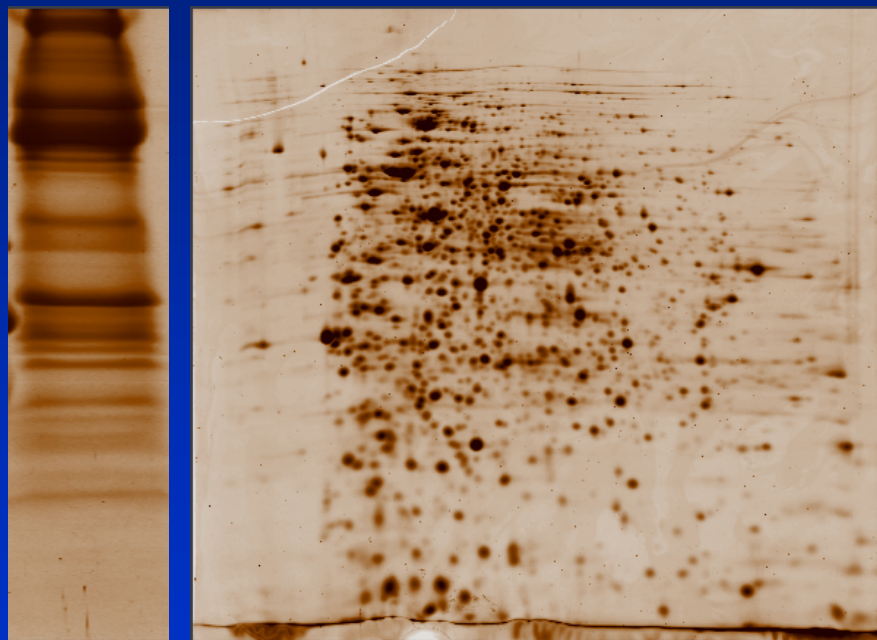
## mRNA metody

- nejprve přepis mRNA → cDNA (reverzní transkriptáza)
- qRT-PCR – analýza exprese 1 genu – PCR amplifikace
- více genů: čipové metody (expresní čipy-i pro celý genom): na principu hybridizace nebo se zahrnutím PCR (citlivější)

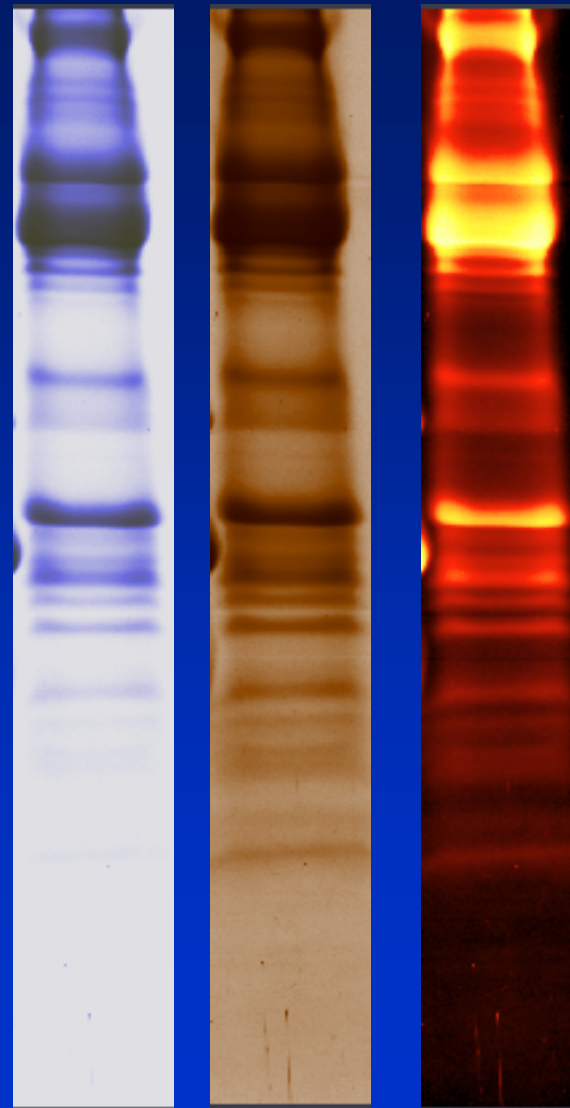
## Proteinové metody

- Separace proteinů (elektroforéza, chromatografie) nebo peptidů-před MS
- Identifikace proteinů: hmotnostní spektrometrie (MS), nebo imunochemie - protilátky
- Protilátky: proti konkrétnímu proteinu (variantě), velmi specifické (ale někdy crossreaktivita!), velmi citlivé
- MS: kvantifikace i několika tisíc proteinů v 1 analýze, méně citlivé

# Proteinové metody: rozlišení a citlivost



Rozlišení →



Citlivost →

# Dynamický rozsah koncentrací proteinů v proteomice

- Jednotlivé proteiny v celobuněčném vzorku se liší až o 12 řádů v rozsahu koncentrací
- Více koncentrované proteiny zákonitě překrývají ty méně koncentrované
- Málo koncentrované proteiny jsou pod limitem detekce

Náznak řešení:

- Frakcionace biologických vzorků
- Odstranění nejkonzentrovanějších proteinů (např. albuminu a dalších z krevní plasmy)
- Použití imunochemie namísto MS (+/-)
- Kvantifikace exprese na úrovni transkriptu (+/-)

- **Proteomické workflow**

⇒ z jakých přístupů můžeme vybírat?

- **Experimenty v onkologickém výzkumu**

Funkční studie

Protein-proteinové interakce

Strukturní charakterizace proteinů

Vyhledávání biomarkerů

Verifikace a validace biomarkerů

⇒ jak sestavit svůj experiment a jaké přístupy zvolit?

- **Statistická analýza a validace na dalších biologických úrovních**

⇒ nejdůležitější poznámky ke statistice

⇒ databáze použitelné při interpretaci výsledků

# I. Nejpoužívanější proteomické přístupy a jejich kombinace

- Dvourozměrná gelová elektroforéza (2-DE)
- Hmotnostní spektrometrie a její možnosti
- SELDI-TOF-MS
- SILAC-LC-MS
- iTRAQ-(2D)LC-MS/MS
- Cílená proteomika (SRM)
- Protilátkové arrays (cell arrays, tissue arrays)



# Klasické proteomické workflow (2-DE-MS)

## 1. Rozměr 2-DE: isoelektrická fokusace

**IPG-IEF** =proužky s imobilizovaným pH gradientem) –  
dobrá reprodukovatelnost, napětí až 10000 V,  
současný standard, komerčně dostupné

## • 2. Rozměr 2-DE: SDS-PAGE

- Ekvilibrace IPG proužků před SDS-PAGE: obalení proteinů SDS , redukce (DTT) a alkylace (jodacetamid)

- SDS-PAGE (malý i velký formát)

10 až 12 % (homogenní)  $M_r$  15000-100000

8-16% (porozitní gradient)  $M_r$  10000-150000

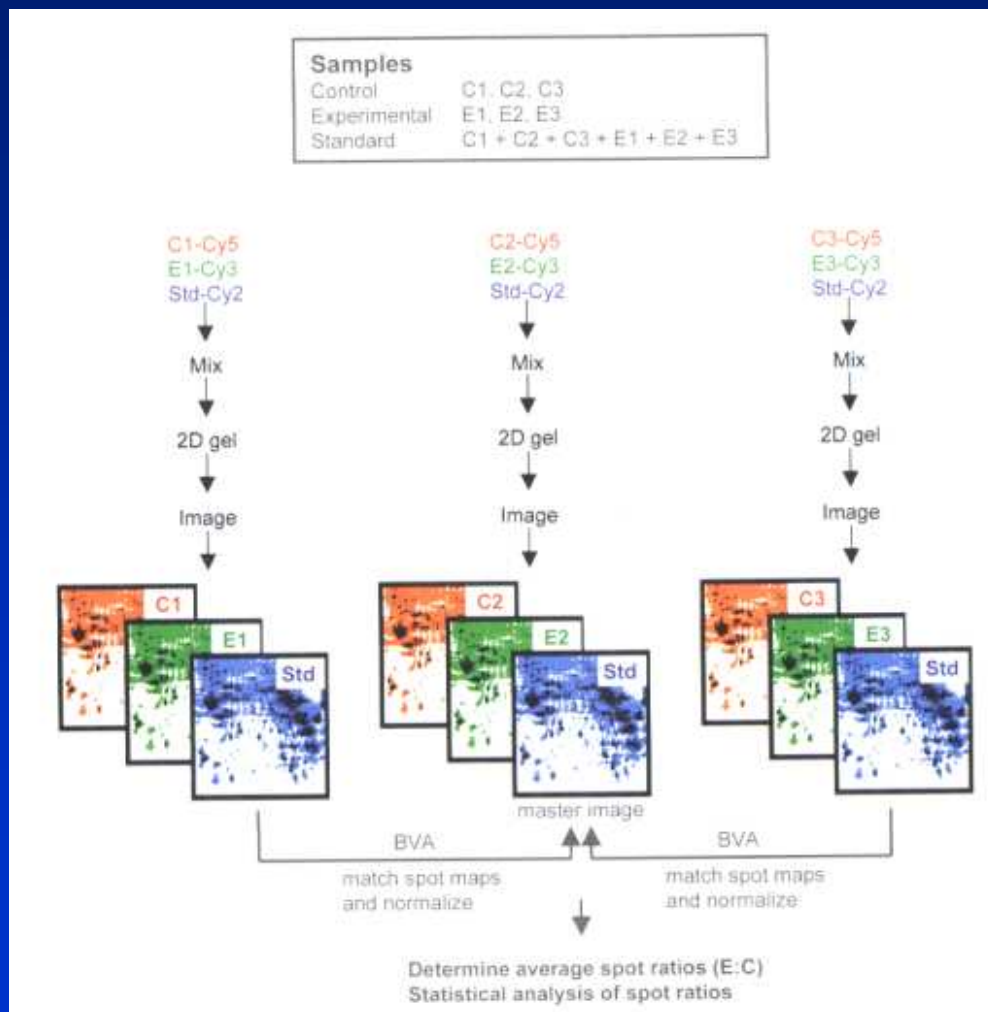
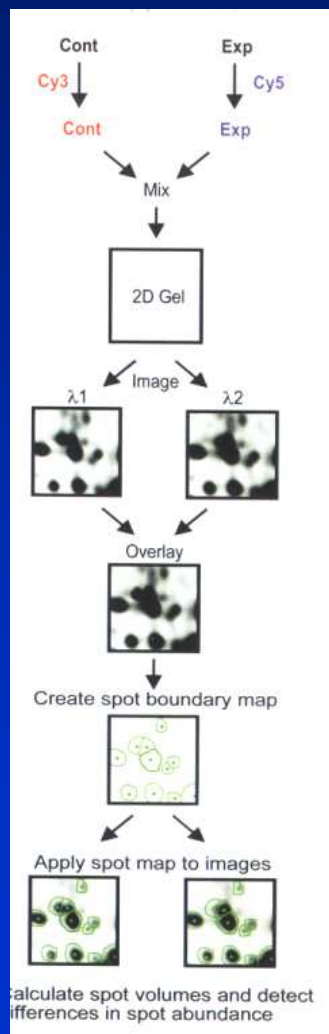
reprodukovatelnost!

Barvení gelů

Obrazová analýza

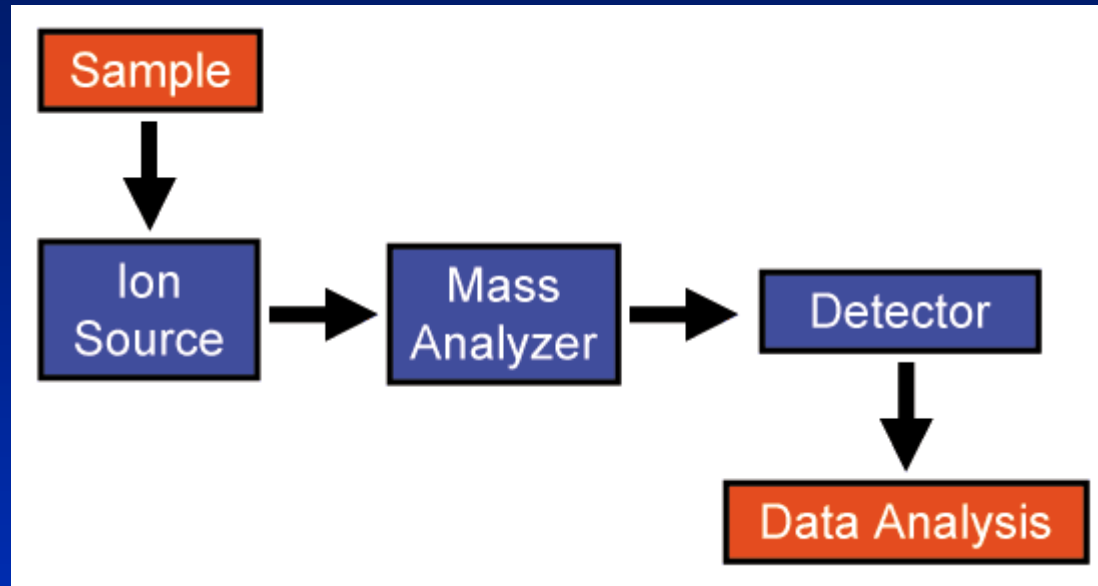
Identifikace proteinů hmotnostní spektrometrií

# 2-D DIGE (diferenční 2-D elektroforéza)

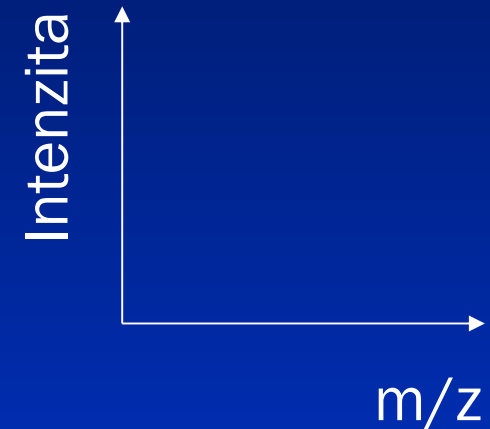




# Hmotnostní spektrometrie



MS spektrum



## Iontové zdroje:

MALDI (matrix-assisted laser desorption-ionization)

ESI (electrospray ionization)

## Analyzátoři:

TOF (time-of-flight)

Q3 (trojitý kvadrupól)

IT (ion trap-iontová past)

Orbitrap

Kombinace analyzátorů -  
>hybridní systémy

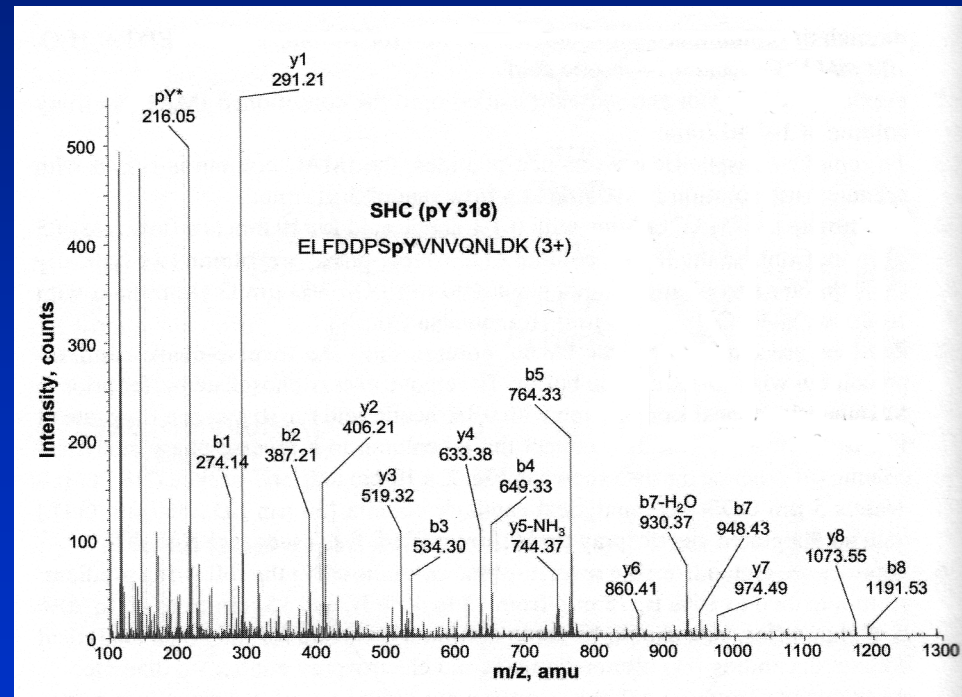
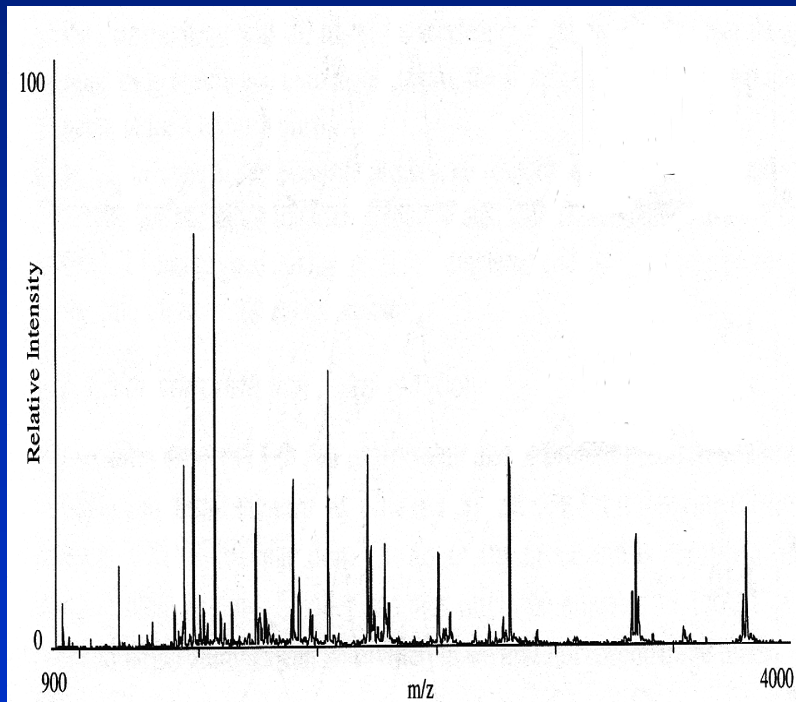
# Co umožňuje hmotnostní spektrometrie

- Identifikace proteinů z gelu a roztoku:  
Štěpení trypsinem na peptidy → MS spektrum → peptide mass fingerprinting v **MS** módu  
peptidy ve hmotnostním spektrometru lze dále fragmentovat →  
částečná sekvenace peptidů → peptide mass fingerprinting v  
**MS/MS** módu (přesnější identifikace)  
identifikace proteinů na základě srovnání s databázemi
- Kvantifikace proteinů a peptidů:  
kvantifikace proteinů: **SELDI**-TOF MS, MALDI imaging  
kvantifikace peptidů v MS módu: **SILAC**, label-free kvantifikace  
(**LFQ**)  
kvantifikace peptidů v MS/MS módu: **iTRAQ**, **SRM**

# MS

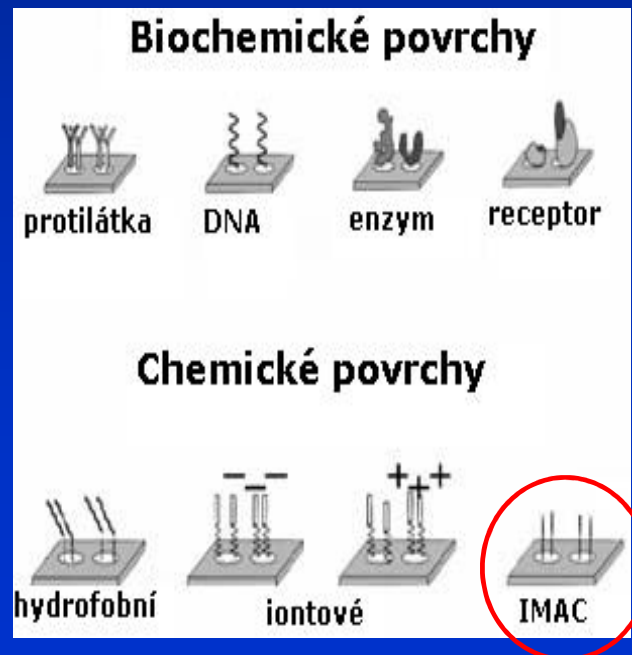
# versus

# MS/MS



# SELDI-TOF MS

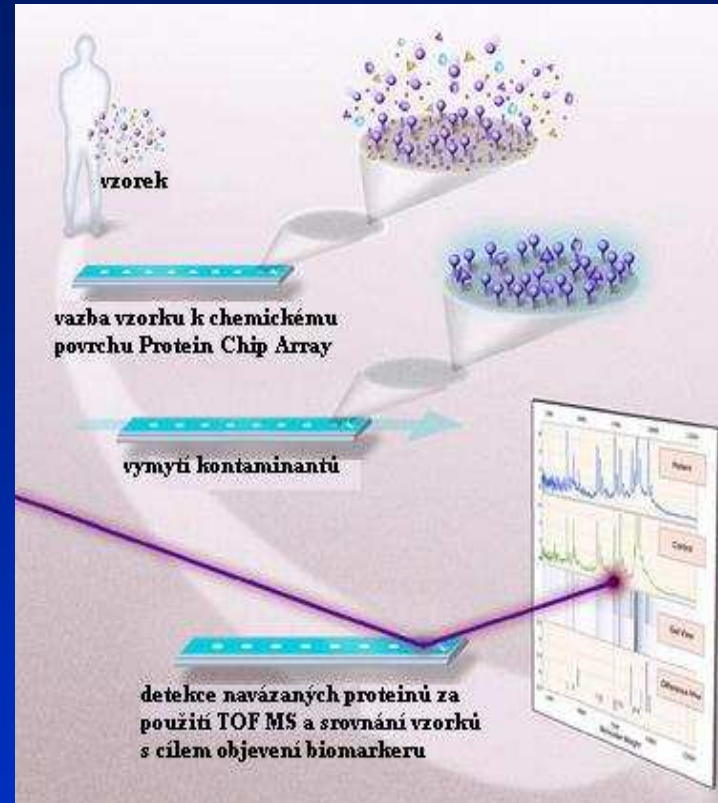
- Surface – enhanced laser desorption-ionization time-of-flight mass spectrometry
- Příprava vzorku: podobně jako pro 2-DE
- Surface=povrch čipu, na nějž se proteiny vážou na základě chemické nebo biochemické interakce



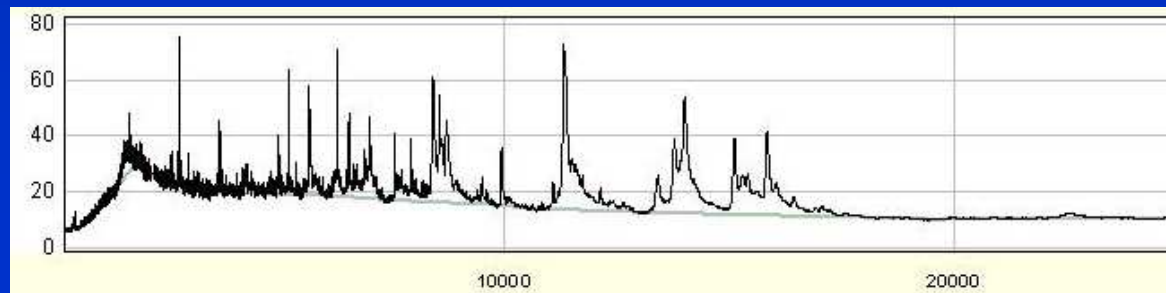
$\text{Cu}^{2+}$



# SELDI – TOF MS



SELDI-TOF MS pracuje pouze v MS módu, detekuje celé intaktní proteiny

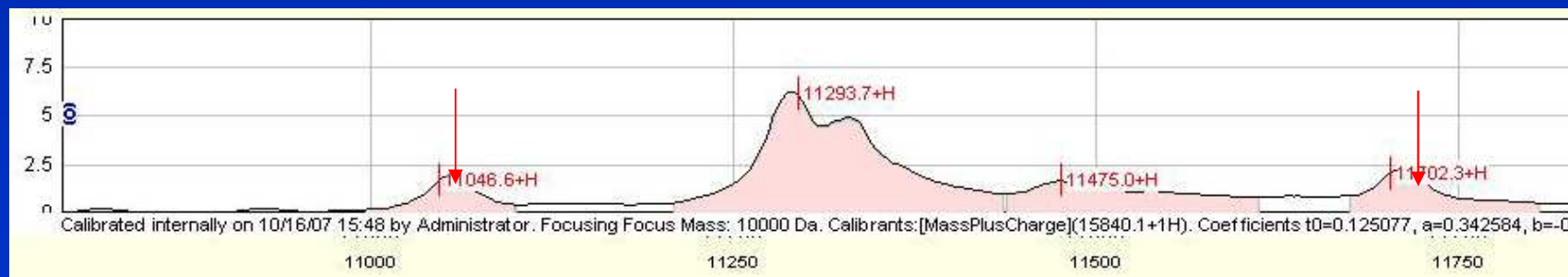
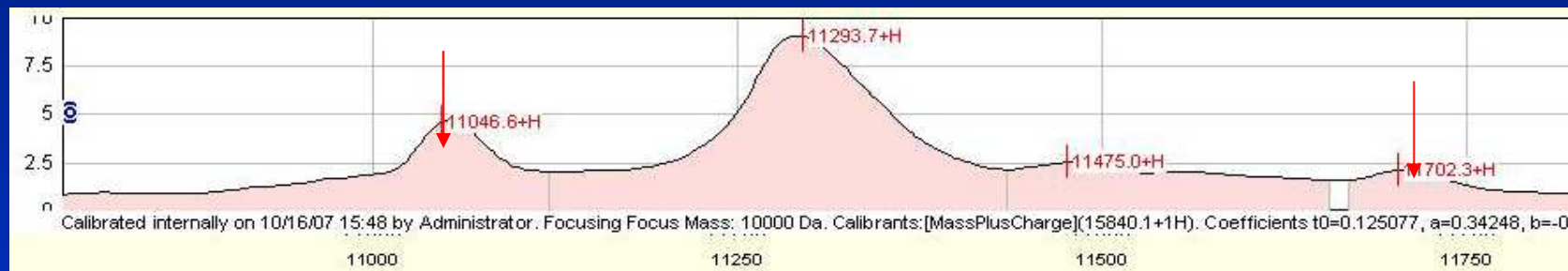
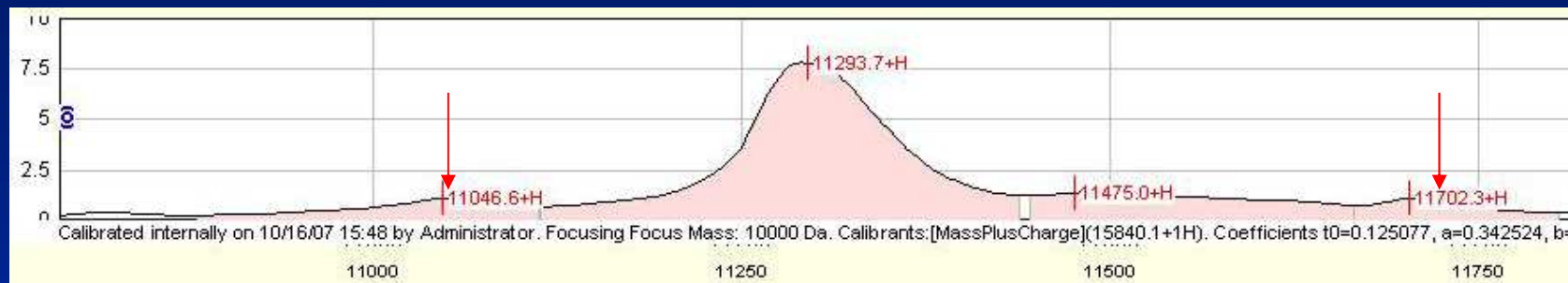


SELDI TOF MS spektrum



# Kvantifikace pomocí SELDI-TOF MS

Princip: Srovnání ploch píků mezi spektry - např. vzorky různých pacientů



Identifikace proteinů=problém!

kombinací separačních přístupů a externího MS/MS – zdlouhavá, často se nedaří. Je třeba mít štěstí

## II. Proteomické workflow LC-MS

- Návrh experimentu
- Extrakce proteinů/příprava vzorku
- Štěpení proteinů na peptidy
- Frakcionace
- LC-MS/MS analýza
- Analýza dat: Identifikace a kvantifikace proteinů
- Statistická analýza dat

# Extrakce proteinů/příprava vzorku pro LC-MS

**Mechanická homogenizace** v lyzačních pufrech na různé bázi:

## Denaturant

- močovina 2-8 M
- SDS až 4% (CMC~0.5%)
- Rapigest  
výhody a nevýhody?

## Redukční činidlo

- dithiothreitol (DTT)
- tris(carboxyethyl)fosfin (TCEP)

# Štěpení proteinů

- In-solution
- In-gel
- Filter aided sample preparation (FASP)
  
- Precipitace proteinů - optional
- Redukce S-S můstků
- Alkylace –SH skupin (jodacetamid, MMTS)
  
- Trypsin (C-strana Lys, Arg pokud nenásleduje Pro)
- Lys-C (C-strana Lys i když následuje Pro)

# Frakcionace

**Snížení komplexity u komplexních vzorků** (buněčné a tkáňové lyzáty, séra) pro necílené analýzy => zvýšení pokrytí proteomu

## Frakcionace proteinů

- SDS-PAGE (SEC), pak in-gel digest

## Frakcionace peptidů

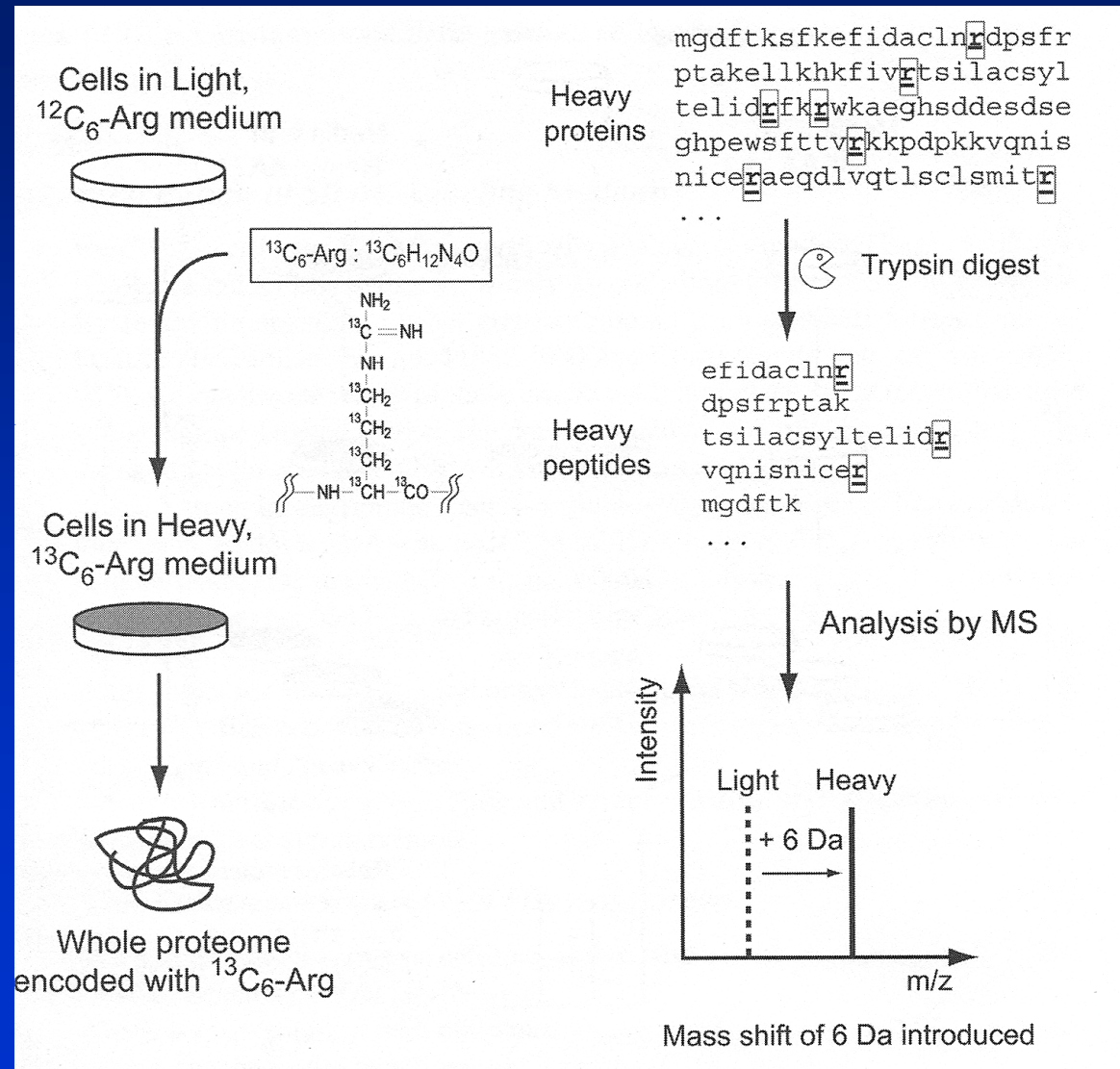
- SCX-strong cation exchange
- RP-reverzní fáze v alkalickém pH
- HILIC-hydrofilní chromatografie

Orthogonalita separace vůči následné LC-MS/MS analýze (RP v kyselém pH)

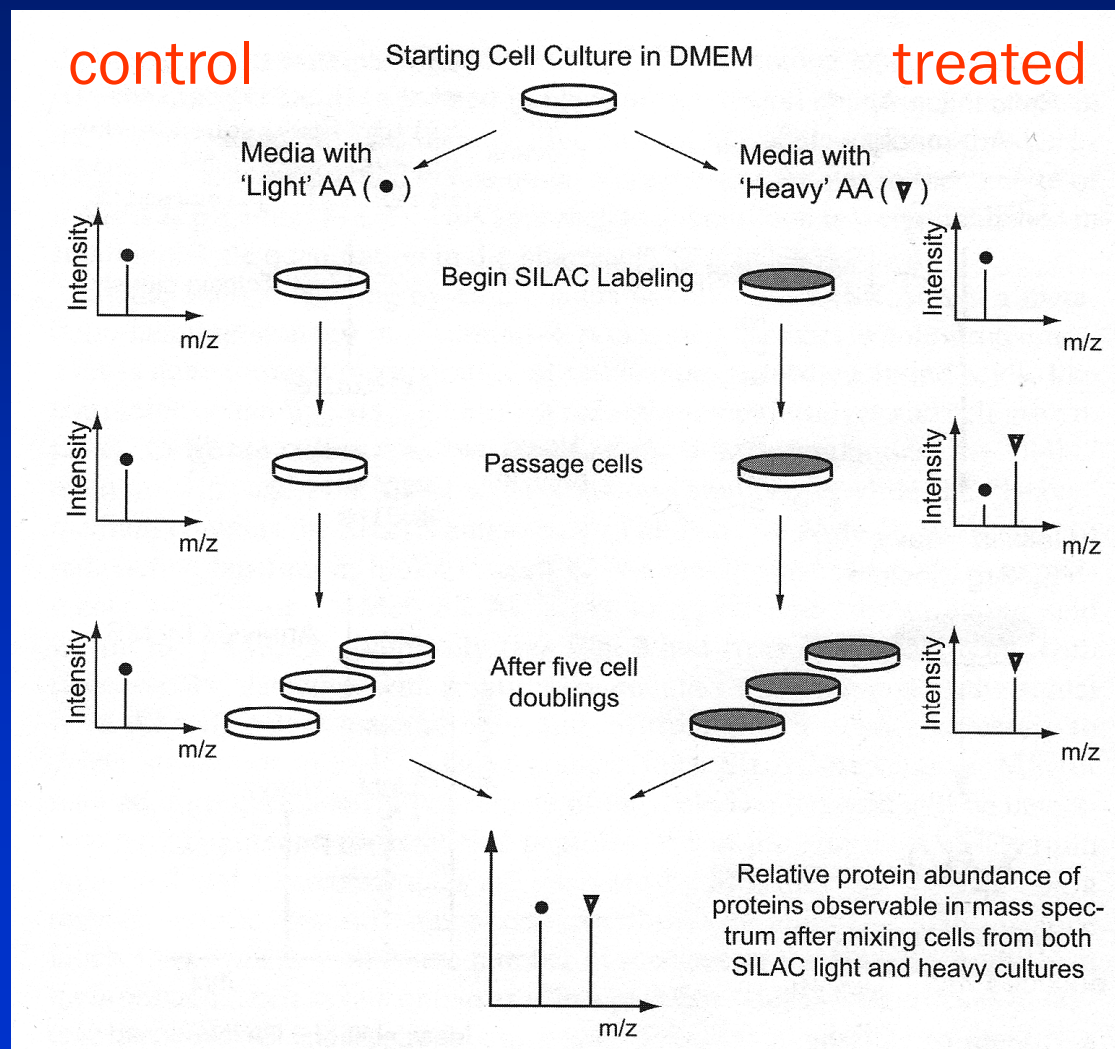
# Kvantifikace s použitím SILAC značení

SILAC=  
Stable isotope labeling  
by amino acids  
in cell culture

Značené AK k dispozici:  
Lys, Arg



# Kvantifikace s použitím SILAC značení



# iTRAQ – (2D)LC-MS/MS

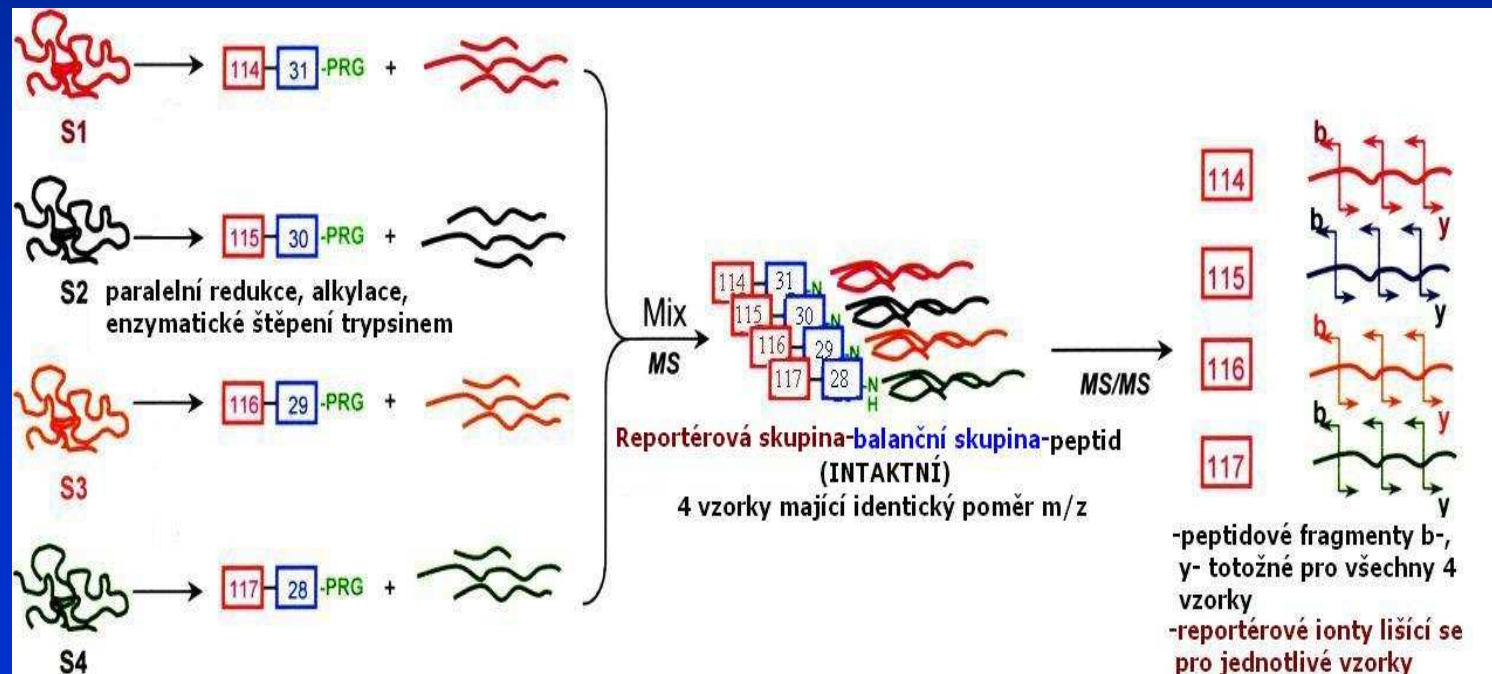
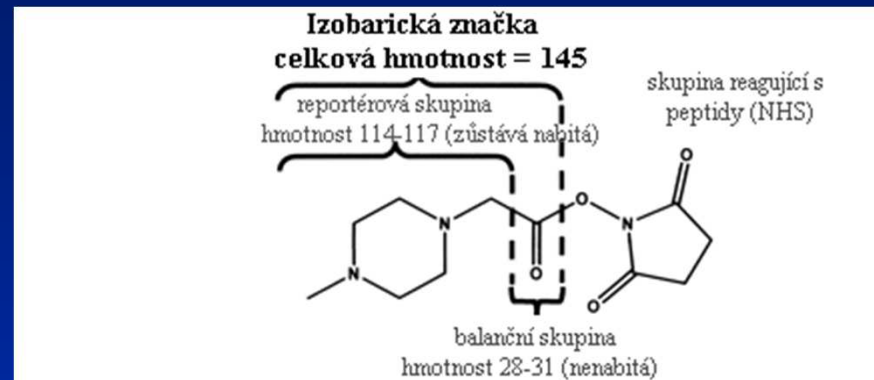
iTRAQ = Isobaric tags for relative and absolute quantitation

Postup:

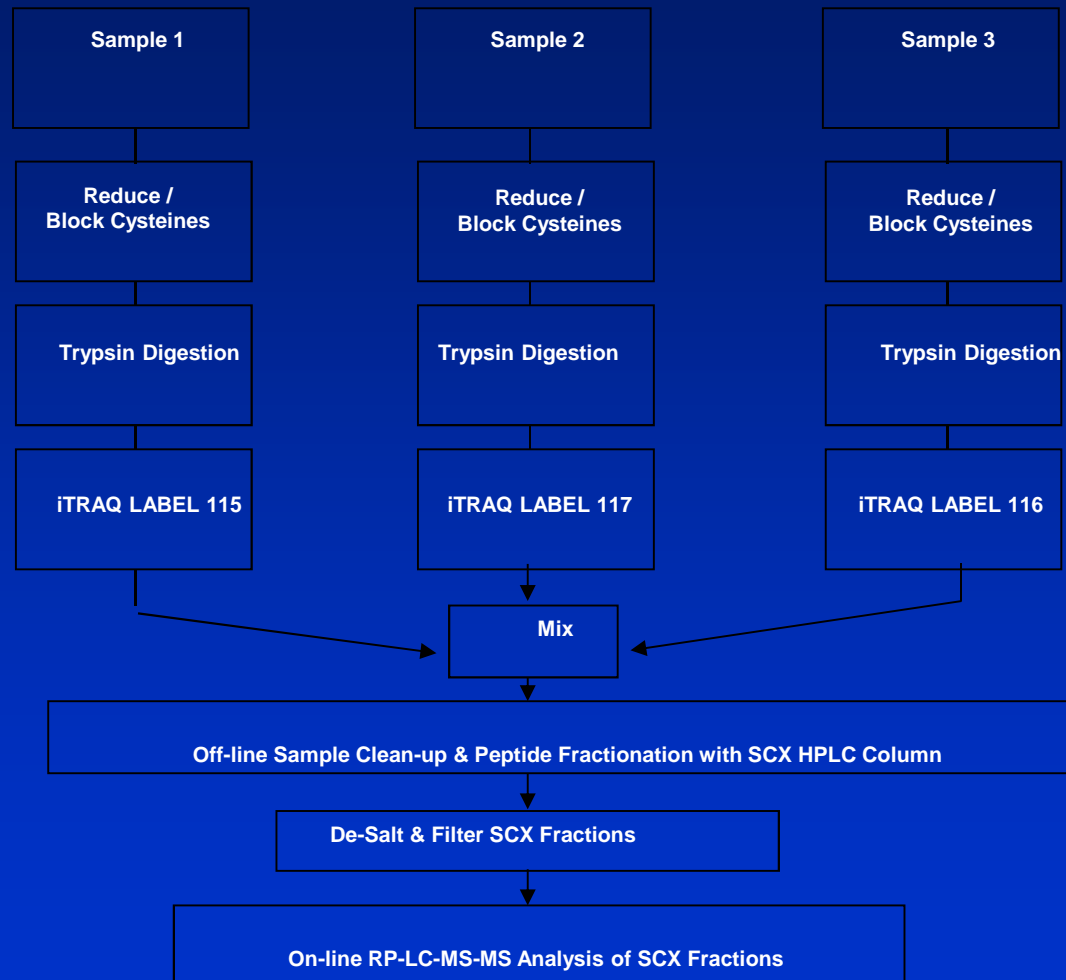
- Příprava vzorku (např. 0.1% SDS) – až 8 kvantitativně srovnávaných vzorků v 1 analýze
- Redukce a alkylace
- Digeste trypsinem
- Značení (iTRAQ značky: 113, 114, 115, 116, 117, 118, 119, 121)
- Smíchání 8 značených digestů
- Frakcionace peptidů (SCX, RPalk., IPG, HILIC, ...)
- Kvantifikace a identifikace proteinů (LC-MS/MS nebo MALDI-MS/MS)



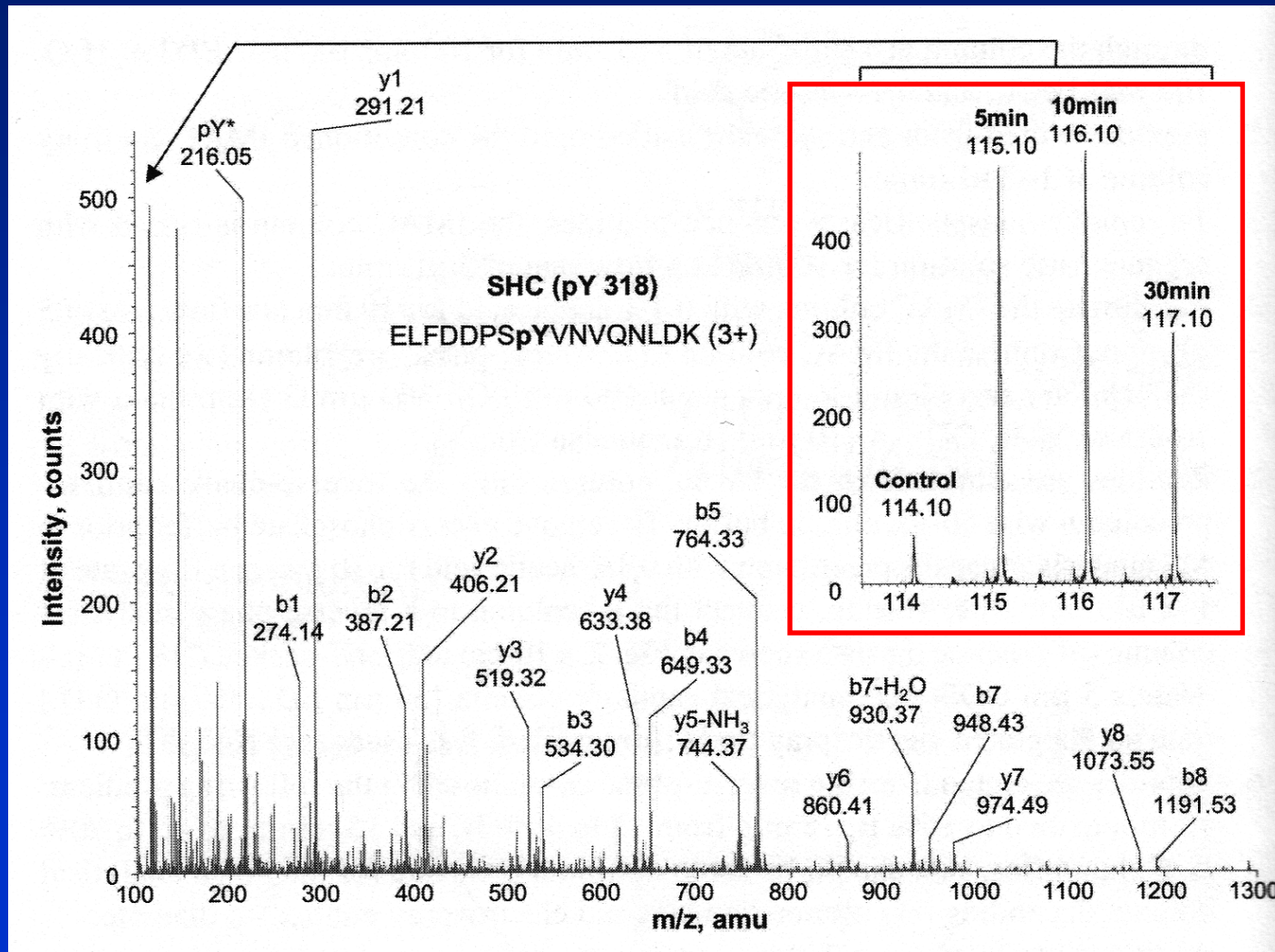
# iTRAQ – (2D)LC-MS/MS



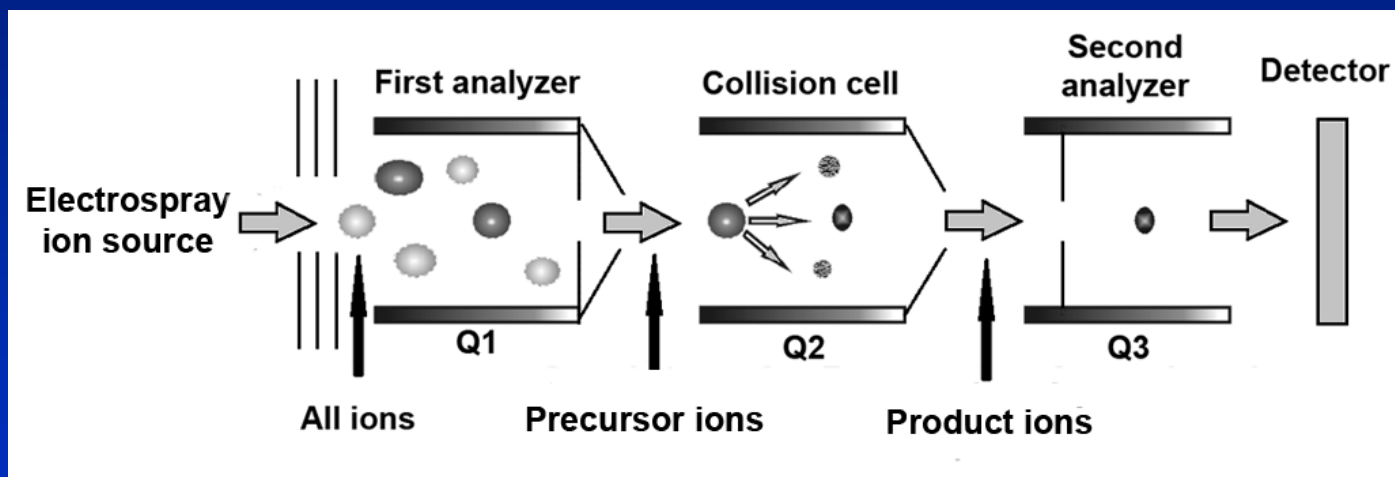
# Schéma typického iTRAQ – 2DLC-MS/MS experimentu (zde 3-plex, nyní je možný až 8-plex)



# MS/MS spektrum peptidu s iTRAQ reportérovými ionty



# Cílená proteomika a selected reaction monitoring (SRM)



# SRM a cílená proteomika

- SRM=multiple reaction monitoring (=MRM=multiple reaction monitoring)
- Metoda kvantifikace peptidů z vybraného proteinu na základě tzv. přechodů (transitions)
- Metoda vysoce selektivní, citlivá, používaná v analýze nízkomolekulárních látek
- Velký potenciál pro validaci výsledků z proteomických studií jako „HMOTNOSTNĚ-SPEKTROMETRICKÝ WESTERN BLOT“ – místo přípravy protilátky by mohl stačit návrh „přechodu“ pomocí software
- Uvádí se postupně do praxe, vyžaduje vývoj metody na konkrétní proteiny

# LC-MS/MS analýza

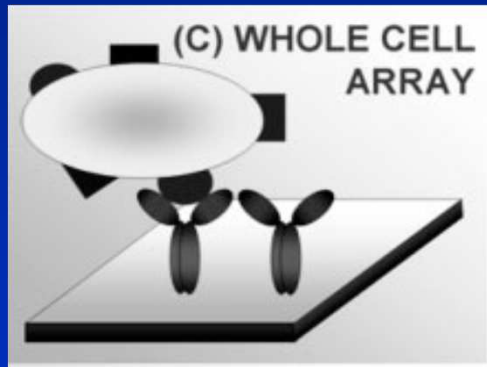
*(instrumentace v RECAMO)*

- LC: RP v kyselém pH (formic acid (FA))  
on-line spojena s MS

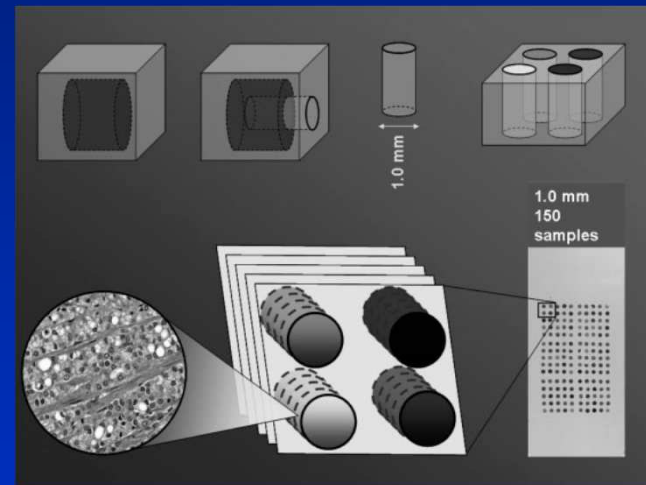
- MS/MS analýza:  
**Orbitrap ELITE** (Thermo)  
CID/HCD/ETD

**TripleTOF 5600+** (ABSCIEX)  
necílená i cílená kvantifikace, SWATH

# Protilátkové arrays



Cell array



Tissue array  
(imunohistochemie)

# Srovnání proteomických metod: Výhody a nevýhody

(úkol pro studenty na přednášce k doplnění)

2-DE SELDI SILAC iTRAQ-(2D)LC-MS/MS SRM

+

-



- **Proteomické workflow**

⇒ z jakých přístupů můžeme vybírat?

- **Experimenty v onkologickém výzkumu**

Funkční studie

Protein-proteinové interakce

Strukturní charakterizace proteinů

Vyhledávání biomarkerů

Verifikace a validace biomarkerů

⇒ jak sestavit svůj experiment a jaké přístupy zvolit?

- **Statistická analýza a validace na dalších biologických úrovních**

⇒ nejdůležitější poznámky ke statistice

⇒ databáze použitelné při interpretaci výsledků

# III. Proteomické experimenty v onkologickém výzkumu

- Funkční studie
- Protein-proteinové interakce
- Struktura proteinů a proteinových komplexů
- Vyhledávání biomarkerů
- Verifikace a validace biomarkerů

# Návrh experimentu

- Volba vhodného materiálu – dáno typem experimentu  
buněčné linie, xenografty, tkáňové lyzáty, archivní FFPE materiál, mikrodisekované (LCM) buňky z něj, kostní dřeň, sekretované proteiny, krevní sérum/plasma
- Volba kvantifikačního postupu  
2-DE/SELDI/SILAC/iTRAQ/LFQ/SRM
- Design experimentu  
Biologické replikáty  
Analytické replikáty  
Randomizace  
Blocking

## A. Funkční studie

- Buněčné linie, případně *in vivo* modely
- Umlčení exprese genu v buněčné linii přirozeně produkující daný protein: siRNA, TALEN (vs. kontrola)
- Mutantní forma proteinu (vs. wild type)
- Navození exprese genu v buněčné linii přirozeně neprodukující daný protein: (Stabilně) transfekovaná linie
- Funkční charakterizace na buněčné úrovni (např. rozdíly v invazivitě, migraci, proliferaci, expresi markerových proteinů signálních drah, ....)
- Proteomika jako systémově biologický nástroj může pomoci odhalit globální molekulární odpověď na (ne)přítomnost sledovaného proteinu a tedy souvislost se změnami hladin dalších proteinů – **funkčních partnerů**
- Sledujeme změny hladin všech **detekovaných** proteinů

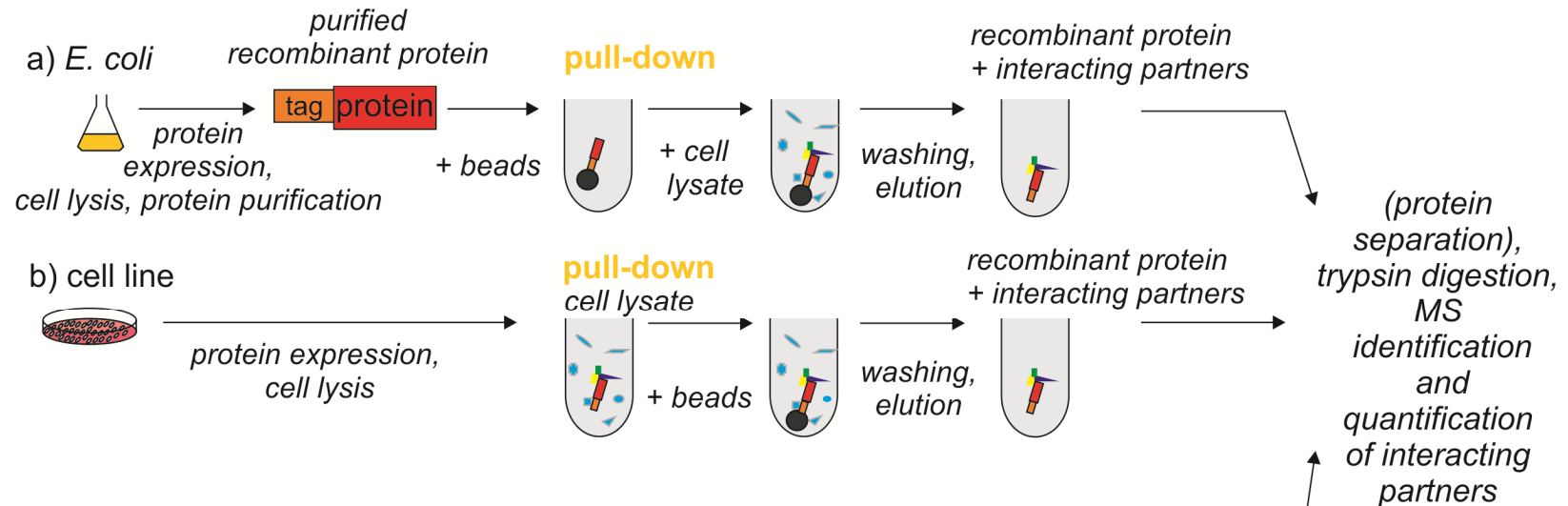
# Workflow: Funkční studie

- Návrh experimentu?
- Volba kvantifikačního přístupu?
- Extrakce proteinů/příprava vzorku
- Štěpení proteinů na peptidy
- Frakcionace?
- LC-MS/MS analýza
- Analýza dat: Identifikace a kvantifikace proteinů
- Statistická analýza dat?

# B. Analýza protein-protein interakcí

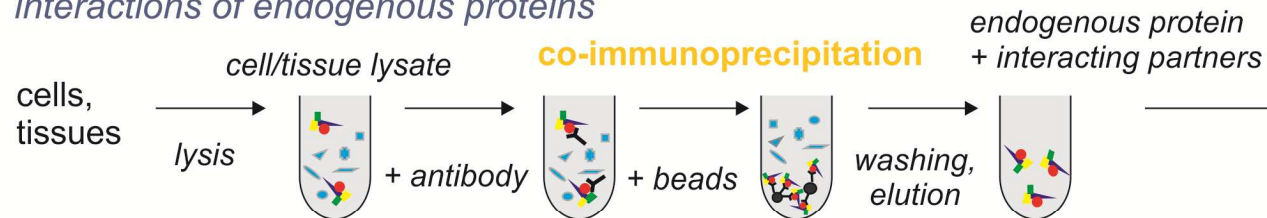
## (1) Pull-down-MS approach

*interactions of tag-fused recombinant proteins*



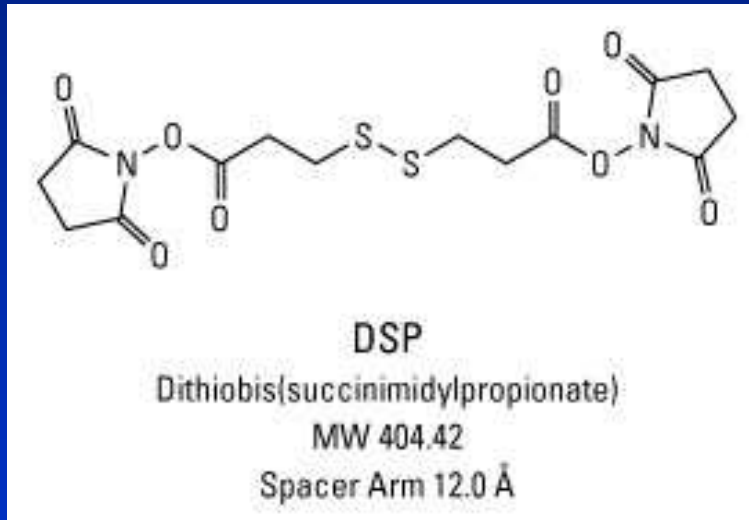
## (2) Co-immunoprecipitation-MS approach

*interactions of endogenous proteins*



Symbols: ● protein whose interacting partners are to be identified    ■ interacting partners    Y antibody    ■ non-interacting proteins in the cell lysate

# Cross-linkery při identifikaci protein-protein interakčních partnerů



# Workflow: analýza protein-protein interakcí

- Návrh experimentu?
- Volba kvantifikačního přístupu?
- Extrakce proteinů/příprava vzorku
- Štěpení proteinů na peptidy
- Frakcionace?
- LC-MS/MS analýza
- Analýza dat: Identifikace a kvantifikace proteinů
- Statistická analýza dat?



## C. Proteomika a strukturní charakterizace proteinů

- Top-down approach (=analýza intaktních proteinů bez štěpení proteázou, ETD fragmentace) (vs. bottom-up)
- H/D exchange – protein-protein interakce v čase
- Posttranslační modifikace (PTM) – fosforylace, glykosylace, acetylace
  - hladina proteinu vs. změna PTM!

## D. Proteomika při vyhledávání biomarkerů

- Značný klinický zájem
- Úspěšnost uplatnění v klinické praxi zatím malá – celá řada důvodů
- „Biomarkerová krize“ 2008-9
- Poté posun v pokrytí proteomu díky masovějšímu uplatnění analyzátoru Orbitrap
- Nutno věnovat zásadní pozornost
  - správně položené klinické otázce
  - výběru vzorků
  - experimentálnímu designu
  - kvalitě klinických vzorků
  - statistické analýze
  - verifikaci proteomických dat

# Jaký biologický materiál analyzovat?

- Modelové systémy - buněčné linie
- Tkáně
- Buňky mikodisekované z tkání (LCM – laser captured microdissection)
- Kostní dřeň – leukémie
- Plasma – sérum

?

**Vzorky od pacientů - etická pravidla!!**

# Workflow: Vyhledávání biomarkerů

- Návrh experimentu?
- Volba kvantifikačního přístupu?
- Extrakce proteinů/příprava vzorku?
- Štěpení proteinů na peptidy
- Frakcionace?
- LC-MS/MS analýza
- Analýza dat: Identifikace a kvantifikace proteinů
- Statistická analýza dat?

# E. Verifikace a validace potenciálních biomarkerů

Screeningové metody:

- 2-DE, SELDI-TOF MS, SILAC-LC-MS, iTRAQ-(2D)LC-MS/MS, LFQ-MS

Validační metody na proteinové úrovni:

- western blotting, SRM

Validační metody na úrovni mRNA (???):

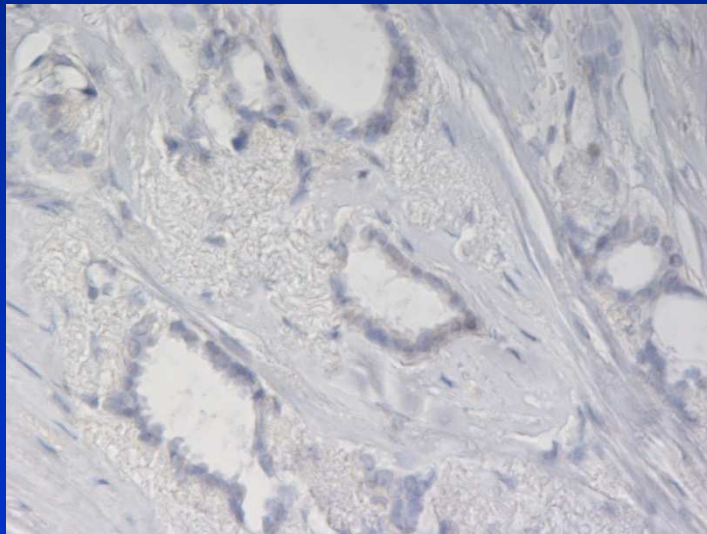
- qRT-PCR, (expresní čipy)

Validační metoda na proteinové úrovni – tkáňové řezy:

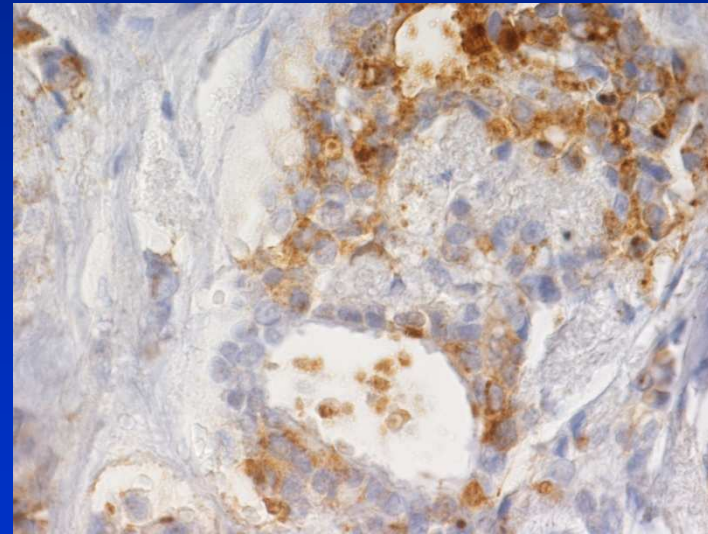
- Imunohistochemie !!

# Verifikace a validace potenciálních biomarkerů

- kvantitativní vs. funkční verifikace
- Imunohistochemie (IHC)



negativní



pozitivní

- IHC vs. cílená proteomika

# Verifikace a validace potenciálních biomarkerů pomocí cílené proteomiky

## Návrh SRM metod

- výběr detekovatelných proteotypických peptidů (na základě vlastních naměřených dat nebo knihovny spekter – např. SRM atlas)
- výběr vhodných produktových iontů
- testování na základě kvantitativních parametrů

# Workflow: Verifikace a validace potenciálních biomarkerů pomocí cílené proteomiky

- Návrh experimentu? *Training + validation sample set*
- Volba kvantifikačního přístupu?
- Extrakce proteinů/příprava vzorku
- Štěpení proteinů na peptidy
- Frakcionace?
- LC-MS/MS analýza?
- Analýza dat: Identifikace a kvantifikace proteinů?
- Statistická analýza dat?

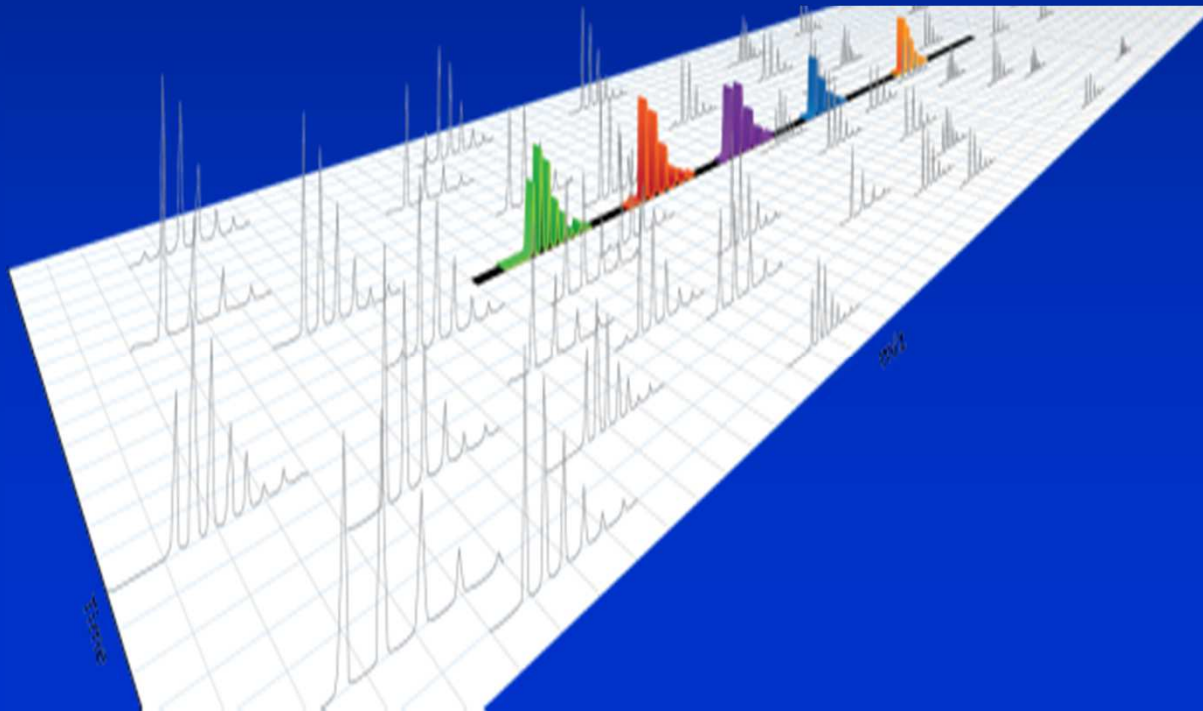
SRM lze využít rovněž v základním onkol. výzkumu pro kvantifikaci proteinů jako alternativa western-blotů či ELISY

SRM se používá jako rutinní kvantifikační metoda malých molekul



# SWATH koncept v onkologickém výzkumu

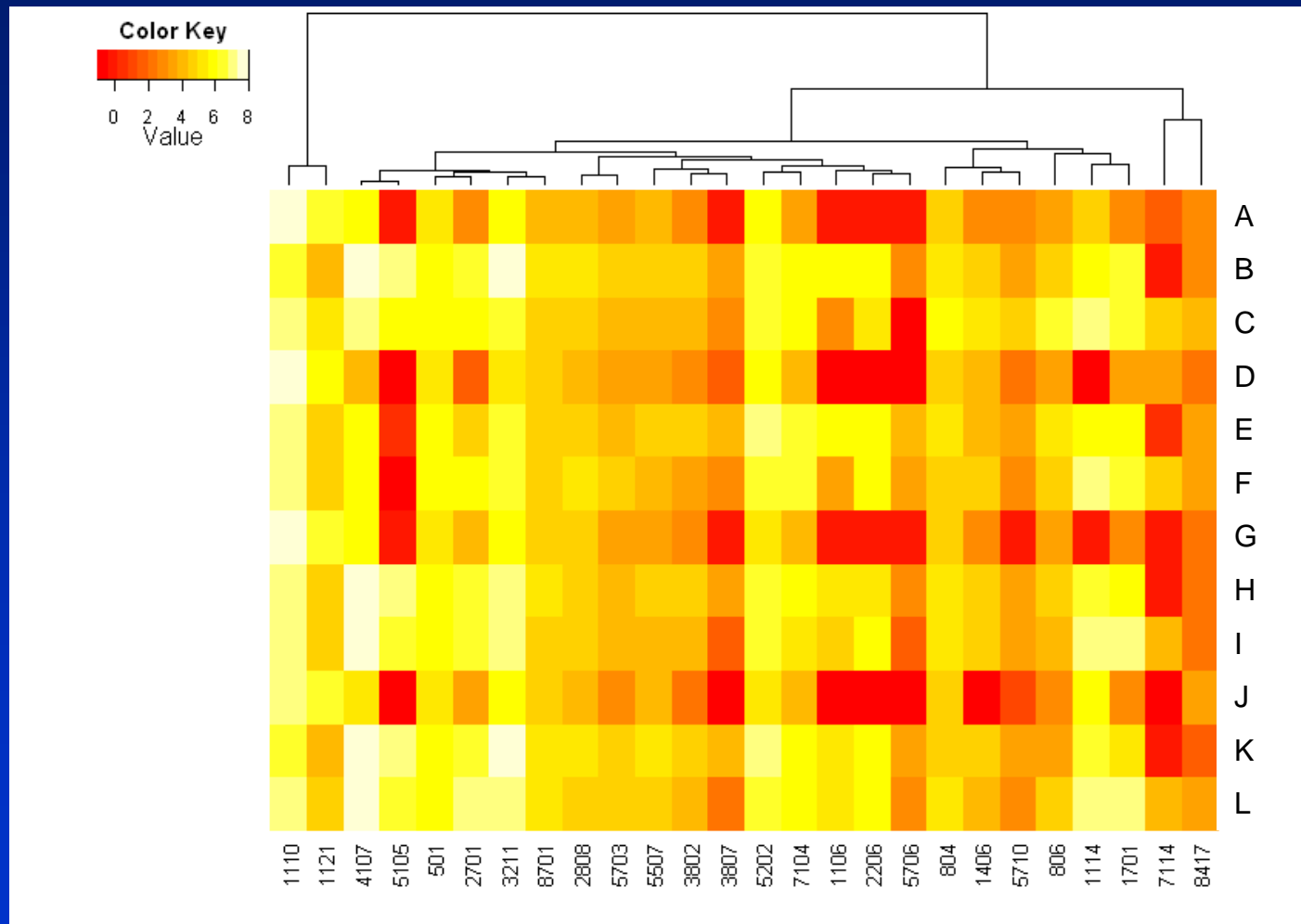
- myšlenka: digitální fingerprinty, z nichž je možné extrahovat kvantitativní proteomická data
- digitized biobanking



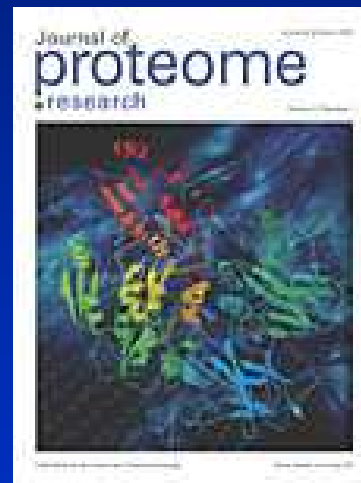
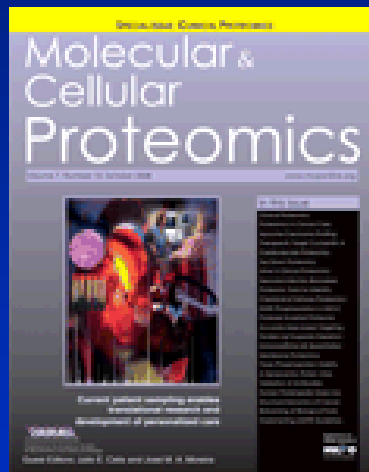
## IV. Statistika

- Při srovnání exprese nutnost paralelní analýzy **minimálně 3+3 replikátů (analytické vs. biologické replikáty)**
- **Normalizace** dat
- Kvantitativně vs. statisticky významný rozdíl: **>2x vs.  $p$ -hodnota** (Student t-test, Mann-Whitney test);
- **Korelační analýza** – nutno přizpůsobit design studie
- Korekce na **mnohonásobné testování** (FDR-korekce)
- při srovnávání více než 2 vzorků vhodná **konzultace se statistikem** – často nutnost aplikace pokročilých statistických metod
- Analyzujeme-li vzorky pacientů, soubor musí být dostatečně reprezentativní a klinicky (patologicky) charakterizovaný – **spolupráce s patologem, s lékaři**

# Heatmapa – korelační analýza



# Další zdroje informací – reviews v předních světových časopisech



+ Quantitative Proteomics by Mass Spectrometry. Methods in Molecular Biology series, No.359, Sechi, S. (Ed.). Humana Press 2007

Děkuji za pozornost