

Metode za analizo proteoma

Igor KRIŽAJ

Odsek za molekularne in biomedicinske znanosti, Inštitut Jožef Stefan
Jamova cesta 39, 1000 Ljubljana, Slovenija

Izvleček:

Proteini so nosilci večine bioloških funkcij. Da bi razumeli, kako celice delujejo, je potrebno raziskovati kateri proteini so prisotni, kje se nahajajo, kako interagirajo z drugimi molekulami in katere funkcije opravljajo tekom razvoja ali v različnih bolezenskih stanjih. Pojem proteom zajema celokupen nabor proteinov v določeni celici, tkivu ali organizmu v določenem trenutku. Pričujoči prispevek podaja pregled metodološkega napredka in bodočih usmeritev na področju, ki se ukvarja z analizo kompleksnih proteinskih mešanic v bioloških vzorcih in področju, ki se ukvarja z obširno identifikacijo proteinov, kvalitativno in kvantitativno, z visoko občutljivostjo in zanesljivostjo.

Ključne besede: proteom, separacija proteinov, masna spektrometrija, identifikacija proteinov, kvantifikacija proteinov

Methods for proteome analysis

Abstract:

Encoded proteins carry out most biological functions, and to understand how cells work, one must study what proteins are present, where they are located, how they interact with other molecules and what they do during development or in response to disease. The term proteome defines the entire protein complement in a given cell, tissue or organism at particular time. The present review reflects the methodological progress made and future developments needed to resolve complex protein mixtures in biological samples and subsequently comprehensively identify proteins, qualitatively and quantitatively, with both high sensitivity and fidelity.

Key words: proteome, protein separation, mass spectrometry, protein identification, protein quantification

1 UVOD

Informacije pridobljene z določitvijo nukleotidnega zaporedja celotnih genomov so odprle poti in omogočile nove globalne pristope k študiju proteinskih interakcijskih mrež v organizmih. Proteini so namreč tiste ključne molekule, ki so v celicah odgovorne za izvajanje fizioloških in patoloških funkcij. Tehnologija DNA mikromrež se je razvila do točke, ki omogoča celostno spremljanje izražanja genov pri vedno večjem številu organizmov. Take vrste poskusi postrežejo s podatki o časovnem poteku prepisovanja ali transkripcije genov (transkriptomika), niso pa vedno v korelaciji z izražanjem genov na proteinskem nivoju, niti ne vsebujejo nekaterih zelo pomembnih sporočil, npr. o tem, kam v celici je po sintezi nek protein namenjen in kje bo izražal svojo aktivnost. Po zadnjih ocenah se v človeškem genomu nahaja okoli 23,000 genov. Teh 23,000 genov se v procesih transkripcije DNA in prevajanja ali translacije RNA prevede v precej večje število proteinskih produktov. Vzroki za bistveno večjo število proteinov v celicah kot je genov v genomu so procesi alternativnega izrezovanja RNA med transkripcijo in posttranslacijske modifikacije proteinov. Kompleksnost proteinskega komplementa genoma še dodatno močno povečata časovna in prostorska organizacija proteinov v celičnem sistemu, ki sta bistvenega pomena za njegovo delovanje. Proteinski profil celičnega sistema je dinamičen. Vsakozi se spreminja in je odvisen je od fiziološkega in patološkega stanja sistema. Za popoln opis delovanja neke celice, tkiva ali pa celotnega organizma je torej potrebno poznati vse igralce (proteine), njihove vloge (encimske aktivnosti), postavitve na "odru" (lokalizacijo), "dialoge" med njimi (interakcije) in to v vsakem trenutku. Rešitev te izjemno kompleksne naloge se kljub skokovitemu napredku znanosti še vedno zdi zelo daleč.

2 OPREDELITEV PROTEOMSKIH METOD

Pojem proteom je definiran kot celokupni nabor proteinov v določeni celici, tkivu ali organizmu. Proteomika je veda, ki se ukvarja z globalno analizo funkcije genov na proteinskem nivoju. Proteomske analize izvajamo z vrsto različnih pristopov, ki vključujejo celično-lokalizacijske študije s pomočjo svetlobne in elektronske mikroskopije, interakcijske in aktivnostne študije v obliki nizov na mikroploščicah (čipih) in metodologije na genetski osnovi, kot je npr. kvasna dvohibridna metoda. Zelo pomemben raziskovalni pristop v proteomiki je visoko-pretočna strukturna analiza/identifikacija proteinov. Posebna izziva temu pristopu sta visoka kompleksnost naborov proteinov, zelo nizka vsebnost večine komponent in več kot 10 redov velikosti velik koncentracijski razpon med najbolj bogato zastopanimi proteini in tistimi, ki jih je v vzorcu najmanj.

Proteomika v klasičnem pomenu pozna dva ključna koraka, separacijo proteinov iz celičnega ali tkivnega vzorca in njihovo identifikacijo. To poglavje je posebej posvečeno metodam za ločevanje kompleksnih zmesi proteinov in metodam za strukturno karakterizacijo polipeptidov v sledovih na globalni, visoko-pretočni način.

3 METODE LOČEVANJA PROTEINOV

Zaenkrat še ne poznamo metode s katero bi lahko na preprost način, v enem koraku identificirali in kvantificirali vse componentne kompleksnih proteinskih zmesi. Tako je potrebno proteinski vzorec najprej ločiti na posamezne komponente, šele nato sledi faza identifikacije in kvantifikacije. Pri separaciji kompleksnih proteinskih vzorcev sta se v množici postopkov v proteomiki najbolj uveljavila dva: poliakrilamidna gelska elektroforeza (PAGE) in tekočinska kromatografija (LC). Posebej velja omeniti afinitetno kromatografijo, ki je zelo pomembna metoda posebej v povezavi s karakterizacijo posttranslacijskih modifikacij v proteinih (Jensen, 2004) ter proteinskih kompleksov in njihove dinamike (Warnock in sod., 2004; Gingras in sod., 2007).

3.1 POLIAKRILAMIDNA GELSKA ELEKTROFOREZA (PAGE)

Proteine ločujemo s pomočjo električnega polja v poliakrilamidnih gelih na osnovi njihovega naboja z nativno elektroforezo, izoelektričnim fokusiranjem in elektroforezo v prisotnosti anionskega detergenta natrijevega dodecil sulfata (NaDS). Tem separacijskim postopkom z eno besedo pravimo eno-dimenzionalni (1D-PAGE) (Westermeier, 1997).

3.1.1 Nativna PAGE

Nativna PAGE temelji na potovanju proteinskih molekul zaradi njihovega naboja v električnem polju. Molekule z večjim nabojem pri določenih pogojih potujejo hitreje kot tiste z manjšim. Na mobilnost proteinskih molekul vpliva do določene mere tudi njihova velikost in oblika, saj se morajo na svoji poti proti ustrezni elektrodi "prebivati" skozi pore poliakrilamidnega gela. Pri manj zamreženih gelih je ta vpliv manjši, pri bolj zamreženih pa večji. Resolucija metode se povečuje z daljšanjem gelov.

3.1.2 Izoelektrično fokusiranje

Izoelektrično fokusiranje se od nativne elektroforeze loči po tem, da se ne izvaja v okolju z enoznačno določenim pH, temveč v okolju, kjer je pH vrednost odvisna od položaja na gelu. Običajno se pH v gelu linearno spreminja od enega konca gela do drugega (linearni pH gradient), lahko pa je sprememba gradienta pH tudi drugačna, npr. eksponentna. Protein v električnem polju potuje toliko časa dokler ne dospe v področje gela, kjer je pH natančno enak vrednosti njegove izoelektrične točke (pI). Takrat postane število pozitivnih nabojev, ki jih protein nosi, enako številu negativnih nabojev, navzven postane električno nevtralen. Sila električnega polja nanj ne deluje več, zato v gelu obmiruje – se fokusira. Resolucija metode se poveča, če določen pH gradient pripravimo v daljšem gelu.

3.1.3 PAGE v prisotnosti natrijevega dodecil sulfata (NaDS-PAGE)

Pri NaDS-PAGE proteinski vzorec najprej obdelamo z anionskim detergentom NaDS v prisotnosti reducenta ali pa brez njega. Proteinske molekule se

razvijejo in se enakomerno obdajo z NaDS. To pomeni, da postane njihov negativni naboj na masno enoto enak. Ker molekule zavzamejo tudi enako paličasto obliko, postane njihova mobilnost v poliakrilamidnem gelu odvisna izključno od njihove mase (velikosti). Lažje molekule se lažje in hitreje »prebivajo« skozi pore gela, težje molekule pa težje, zato v gelu potujejo počasneje. S stopnjo zamreženosti gela lahko vplivamo na hitrost potovanja molekul in tako določamo resolucijo v različnih masnih področjih. Na resolucijo vpliva tudi dolžina gela.

3.1.4 Dvo-dimenzionalna PAGE (2D-PAGE)

Metoda z izredno resolucijsko močjo je dvo-dimenzionalna PAGE (2D-PAGE), pri kateri mešanico proteinov, ekstrahirano iz celic ali tkiva, ločimo v poliakrilamidnem gelu (Rabilloud, 2000). Proteine, ki jih je v zmesi lahko nekaj deset ali pa več tisoč, ločimo v gelu postavljenem v električno polje v dveh korakih. V prvem koraku jih ločujemo glede na njihov električni naboj z izoelektričnim fokusiranjem (glej 3.1.2). V drugem koraku jih ločujemo glede na njihovo molekulsko maso z NaDS-PAGE (glej 3.1.3). Separacija glede na molekulsko maso poteka v smeri, ki je pravokotna na smer izvedbe izoelektričnega fokusiranja. Proteine v gelu lahko detektiramo na različne načine. Klasičen način je barvanje z barvilom Coomassie modrim. Okoli sto-krat bolj občutljiva načina detekcije sta barvanje s koloidnim srebrom in fluorescenčnimi barvili (Unlu in sod., 1997; Rabilloud, 2002). Še bolj občutljiva metoda detekcije proteinov v gelu je avtoradiografija radioaktivno označenih proteinov. Običajno lahko z 2D-PAGE na gelu zanesljivo ločimo okrog 2000 proteinov, v optimalnih pogojih pa lahko s to tehniko na enem gelu ločimo celo do 11.000 proteinov.

2D-PAGE je tipična proteomska separacijska tehnika, s katero lahko naenkrat opazujemo celoten proteom (ali njegov del) dane celice, tkiva ali organizma. Je osnovna tehnika t.i. diferencialne proteomike, kjer primerjamo proteoma dveh sistemov, npr. dveh celičnih linij, zdrave in rakaste. Klasična diferencialna proteomika je temeljila na ločeni 2D-PAGE analizi primerjanih sistemov pri karseda enakih pogojih. Kljub temu je pri primerjanju 2D-PAGE gelov pogosto prihajalo do napak, kot posledice slabe ponovljivosti izvedbe separacije in detekcije vzorcev. Število napak se je precej zmanjšalo ob zamenjavi tekočih amfolitov za pripravo pH-gradienta za izoelektrično fokusiranje v prvi dimenziji z imobilizacijo pH gradienta v gelu samem (IPG). Napake je kasneje v veliki meri odpravila 2D-PAGE tehnika, kjer primerjana vzorca (lahko pa tudi tri) pred analizo označimo z različnima fluorescenčnima barviloma (ponavadi z barvili iz serije Cy: Cy2, Cy3 in Cy5), ju združimo in analiziramo skupaj. Tej metodi pravimo diferencialna gelska elektroforeza v gelu – DIGE ("Differential In-Gel Electrophoresis"). Za primerjalno analizo 2D-PAGE vzorcev obstajajo različni sofisticirani primerjalni algoritmi (npr. Hoogland in sod., 2004). Namen primerjanja je identifikacija proteinov, ki se med dvema sistemoma razlikujejo v stopnji izražanja.

Strukturno identifikacijo proteinov ločenih z eno od omenjenih PAGE tehnik ponavadi v proteomiki izvedemo z masno spektroskopijo (MS), lahko pa tudi z Edmanovo degradacijo (Kamp in sod., 1997). Vzorec je pred MS analizo potrebno ustrezno pripraviti. V gelu ga razgradimo na peptide, kar storimo z eno od specifičnih proteaz, ponavadi tripsinom, nato peptide iz gela izločimo z

ekstrakcijo, jih neposredno analiziramo z MS ali pa pred MS analizo še dodatno očistimo s tekočinsko kromatografijo (LC). Izrezovanje proteinskih lis iz gelov, razgradnja proteinov v gelih s proteazami in ekstrakcija nastalih peptidov so v proteomskih postopkih robotizirani. Peptidne vzorce bodisi nanesimo na MALDI tarčo za ionizacijo z lasersko desorpcijo (glej 4.1.1.1), v primeru dodatnega čiščenja z LC, pa je ponavadi tekočinski sistem neposredno sklopljen z elektrosprej (ESI) ionizatorjem (glej 4.1.1.2).

3.2 TEKOČINSKA KROMATOGRAFIJA (LC)

Tekočinska kromatografija predstavlja idealen način priprave vzorca za MS saj jo je mogoče neposredno sklopiti s kapilaro elektrosprej ionizatorja (glej 4.1.1.2), kar omogoča visoko-pretočno identifikacijo vzorcev. Problem eno-dimenzionalne LC je pogosto nezadostna resolucija zelo kompleksnih peptidnih zmesi, ki jih npr. generiramo s proteolizo množice proteinov iz celotnega celičnega ekstrakta. Poleg tega je količina v take vrste analizah zbranih podatkov ogromna in njihova interpretacija zelo zahtevna in zamudna. Za poenostavitev vzorcev za MS analizo se razvijajo več-dimenzionalne separacijske platforme, ki združujejo različne kombinacije postopkov za separacijo proteinov in peptidov. Med najbolj uporabljane sodi kombinacija ionske kromatografije na močnem kationskem izmenjevalcu in kromatografije na reverzni fazi (dvo-dimenzionalni separacijski postopek) ali pa kombinacija ionske kromatografije na močnem kationskem izmenjevalcu, afinitetne kromatografije na avidinskem nosilcu in kromatografije na reverzni fazi (tri-dimenzionalni separacijski postopek). Pri zelo kompleksnih polipeptidnih vzorcih se pogosto pred več-dimenzionalno LC analizo izvede tudi predhodna osnovna separacija z 1D-PAGE. Izraz MudPIT ("Multidimensional Protein Identification Technology") je bil vpeljan za identifikacijske tehnike, ki bazirajo na več-dimenzionalnih separacijskih postopkih čiščenja proteinov in peptidov pred MS analizo (Kislinger in Emili, 2005).

4 METODE IDENTIFIKACIJE PROTEINOV

4.1 MASNA SPEKTROMETRIJA (MS)

MS je glavna metoda za strukturno analizo kompleksnih proteinskih zmesi. Meritve se izvajajo z ioniziranimi molekulami v plinski fazi. Glavni sestavni deli masnega spektrometra so ionizator, masni analizator in detektor. V ionizatorju se tvorijo ioni, v masnem analizatorju se ioni ločijo glede na razmerje med maso in nabojem (m/z), detektorji pa izmerijo ionski tok pri posamezni vrednosti m/z . Rezultat meritve, masni spekter, je prikaz absolutne ali relativne intenzitete ionskega toka v odvisnosti od m/z . Zelo pomemben korak, ki sledi, je interpretacija masnega spektra.

4.1.1 Metode za ionizacijo proteinov

Za najbolj pomembno odkritje, rešitev osnovnega problema MS bioloških molekul, kako prevesti visoko polarne, popolnoma nehlapne molekule z masami več deset kDa v plinasto stanje, ne da bi jih pri tem uničili, sta si leta 2002 John B. Fenn in Koichi Tanaka delila Nobelovo nagrado za kemijo. Odkritje in razvoj

postopkov za ionizacijo proteinov je omogočil hiter razvoj na MS-temelječe proteomike.

Metod za volatalizacijo in ionizacijo bioloških molekul je več, npr. ionizacija z elektroni (EI), kemijska ionizacija (CI) in ionizacija s hitrimi atomi (FAB), najpogosteje pa se uporabljata dve, elektrosprej ionizacija (ESI) in ionizacija s pomočjo laserske svetlobe (MALDI). Principa slednjih sta podrobneje opisana v nadaljevanju, medtem ko opise principov EI, CI in FAB ionizacije najdete npr. v knjigi Chhabila Dassa (2001).

4.1.1.1 Ionizacija v nosilcu z lasersko desorpcijo (MALDI)

MALDI je okrajšava za angleški izraz: "matrix-assisted laser desorption/ionization". Po slovensko bi izraz lahko prevedli kot ionizacija v nosilcu (matriksu) z lasersko desorpcijo. MALDI tehnika ionizacije proteinskih molekul sodi med t.i. "mehke", nedestruktivne ionizacijske tehnike (Karas in Hillenkamp, 1988). Pri MALDI tehniki protein vmešamo v visok prebitek nosilca, običajno aromatske kisline nizke molekulske mase z visoko absorptivnostjo UV svetlobe. Obsevanje z usmerjenim žarkom laserske svetlobe primerne valovne dolžine povzroči sublimacijo molekul nosilca, z njim pa tudi prenos vmešanih nehlapnih molekul polipeptida v plinsko fazo. Pri številnih trkih med ioni in molekulami v gosti plinski fazi, se tvorijo v glavnem enkrat protonirani ioni polipeptida, ki se v visokem električnem polju pospešijo v smeri masnega analizatorja.

4.1.1.2 Ionizacija z elektrosprejem (ESI)

Tudi ionizacija proteinskih molekul z elektrosprejem je "mehak", nedestruktiven način ionizacije (Fenn in sod. 1989). Pri elektrospreju se vzorec v ionizator dovaja v raztopini skozi kovinsko kapilaro, ki se nahaja na visokem, običajno pozitivnem električnem potencialu nekaj kV glede na vhod v masni analizator. Ob izhodu iz kapilare se raztopina polipeptida zaradi elektrostatske disperzije razprši v zelo fine, pozitivno nabite kapljice. V evakuiranem prostoru ionizatorja, topilo odpareva, kapljice se manjšajo in nazadnje ostanejo le nesolvatirani, pozitivno nabiti ioni polipeptida, ki se v visokem električnem polju pospešijo v smeri masnega analizatorja. Pri tem načinu ionizacije dobimo ione tudi z deset in več naboji (protoni). Način dovajanja vzorca v ionizator omogoča enostavno sklopitev različnih vrst tekočinske kromatografije z MS analizo (LC-MS in LC-MS/MS).

4.1.2 Masni analizatorji

Masni analizator je centralni element na MS-osnovanih proteomskih tehnologij. Glavni parametri, ki označujejo masne analizatorje so njihova občutljivost (to je najmanjša množina vzorca, ki jo lahko ob določeni zanesljivosti še detektiramo), ločljivost (to je sposobnost ločbe dveh sosednjih masnih vrhov) in območje delovanja (to je največja masa, ki jo še lahko zanesljivo določimo). Poznamo pet osnovnih vrst masnih analizatorjev: magnetni sektorski analizator, analizator na čas preleta ionov, kvadrupolni analizator, analizator z ionsko pastjo in ionsko ciklotronski resonančni analizator. Po principih določanja mase ionov, obliki in zgoraj omenjenih parametrih se med seboj precej ločijo. Vsak ima svoje dobre

in šibke plati. Analizatorje uporabljamo ločeno, lahko pa jih tudi združujemo in na tak način izkoristimo dobre plati vsakega izmed njih. Sledi kratek opis principov delovanja posameznih vrst masnih analizatorjev in nekaterih njihovih lastnosti.

4.1.2.1 Magnetni sektorski analizator

Najstarejši tip masnega analizatorja je magnetni sektorski analizator. V praksi se sicer še uporablja, vendar ga hitro izpodrivajo bolj zmogljivi analizatorji. Ko prileti linearno gibajoči se ion v področje magnetnega polja, se njegova pot ukloni v ravnini pravokotni na smer magnetnega polja. Uklon iona (r) je odvisen od gostote magnetnega polja (B), mase (m) in naboja (z) iona v skladu z enačbo: $r = (1/B) \times (2mV/z)^{1/2}$. V procesu meritve ponavadi napetost (V), s katero pospešujemo ione pred vstopom v masni analizator ohranjamo konstantno, zvezno pa spreminjamo B . Detektorski sistem meri ionski tok na fiksno določenem mestu ($r = \text{konstanta}$). V odvisnosti od B je m/z enoznačno določen z enačbo: $m/z = (B^2 r^2)/2V$.

4.1.2.2 Analizator na čas preleta ionov

Analizator na čas preleta ionov (TOF iz angleškega "time of flight") je eden najpreprostejših masnih analizatorjev. Ione ločuje na osnovi razlike v hitrosti njihovega potovanja. Osnovna zahteva za natančno delovaje TOF analizatorja je, da ioni istočasno vstopijo v merilno cev in, da so ob vstopu vsi pospešeni do enake kinetične energije (zV), določene z nabojem iona (z) in pospeševalne napetosti (V). Ob izpolnjevanju teh pogojev je hitrost potovanja ionov (v) v merilni cevi obratno sorazmerna z maso ionov. Lažji ioni potujejo hitreje, težji pa počasneje. Čas (t), ki je potreben, da ion preleti pot (L) od vstopa v merilno cev, do detektorja, ki ga lahko zelo natančno izmerimo, je z m/z v sledeči zvezi: $t = L/v = L(m/2zV)^{1/2}$. Energijsko disperzijo ionov na vstopu v TOF analizator lahko dodatno zmanjšamo s t.i. reflektrom. Ioni z večjo kinetično energijo od povprečne se odbijejo od "elektrostatskega zrcala" z zaostankom, medtem ko se ioni z nižjo kinetično energijo od povprečne odbijejo od reflektora v smeri detektorja pred ostalimi ioni, kar kompenzira energijske razlike med njimi in zmanjša napako meritve. S podaljšanjem poti potovanja ionov se poveča tudi natančnost meritve časa potrebnega za prelet merilne cevi in s tem natančnost meritve m/z . Merilne cevi imajo lahko tudi več reflektromov.

4.1.2.3 Kvadrupolni analizator

Kvadrupolni masni analizator je najpogosteje uporabljan masni analizator, ker je pri njem razmerje med ceno in zmogljivostjo najugodnejše. Masna separacija pri tej vrsti analizatorja temelji izključno na uporabi električnih polj. Kvadrupolni analizatorji so dinamični analizatorji – trajektorije ionov nadzorujemo z vrsto časovno-odvisnih sil, ustvarjenih s konstantno usmerjeno napetostjo (U), ki ji je prištet radio-frekvenčno oscilirajoči potencial ($V \cos \omega t$) na štirih paroma vzporednih paličastih elektrodah. V poenostavljenem približku, so elektrode v prostoru razporejene tako, da tvorijo stranice kvadra. Diagonalne elektrode so med seboj električno sklopljene. Ioni priletijo med elektrode v daljši smeri

elektrod. Separacija različnih m/z temelji na stabilnosti poti ionov skozi kvadrupolno električno polje, ki se tvori v prostoru med elektrodami. Kvadrupol je pravzaprav ionski filter. Pri določenih pogojih napetosti na elektrodah lahko skozi ta filter do detektorja prodre samo ion s točno določeno m/z , ostali pa končajo na elektrodah. Kvadrupol se uporablja tudi kot element masnega spektrometra za pripravo sekundarnih ionov s fragmentacijo izbranega primarnega iona.

4.1.2.4 Analizator z ionsko pastjo

Masni analizatorji z ionsko pastjo delujejo na drugačnem principu kot do sedaj opisani analizatorji, ki analizirajo ione v snopih. Kot že samo ime pove, ione pri tej vrsti analizatorjev najprej ujamemo v past in jih šele nato obdelujemo. Zelo popularna in dostopna vrsta tovrstnih instrumentov so ionske pasti, ki so pravzaprav tridimenzionalni analogi kvadrupolnih masnih filtrov (glej 3.4.1.2.3.). Past sestavljajo toroidna centralna elektroda in elektrodi, ki kot kapi zapirata z zgornje in spodnje strani prostor v katerem se z oscilirajočo električno napetostjo na vseh elektrodah ustvarja spreminjajoče se električno polje. Ena od elektrod ima drobno luknjico skozi katero se uvajajo ioni v prostor med elektrodami in kjer jih ujamemo s primerno kombinacijo enosmerne in radio-frekvenčno oscilirajoče napetosti ($U + V \cos \omega t$) na elektrodah v električno polje, ki tvori t.i. ionsko past. Druga elektroda ima običajno več perforacij skozi katere potem prehajajo ioni do detektorja. Ujeti ioni krožijo v ionski pasti s frekvenco, ki je odvisna od njihove m/z . Masni spekter ujetih ionov dobimo s spreminjanjem amplitude obeh komponent napetosti in/ali frekvence oscilirajočega dela potenciala na elektrodah. Kroženje ionov z določeno m/z postane namreč pri določenih pogojih v pasti nestabilno (masno-selektivna nestabilnost) in ioni "pobegnejo" iz pasti na detektor. Ionske pasti so robustni, občutljivi in relativno poceni analizatorji. Večina strukturnih podatkov v proteomski literaturi je bila pridobljena s to vrsto aparatur. Glavna pomanjkljivost ionskih pasti je njihova relativno nizka natančnost določanja mase ionov, ki je v določeni meri posledica omejenega števila ionov, ki se še lahko akumulirajo v ionski pasti, preden pride do prevelike motnje v njihovem gibanju, posledice medsebojnega odboja. Izboljšanja v smeri večje natančnosti, ločljivosti in občutljivosti določanja mase ionov si prizadevajo doseči s t.i. "linearnimi" ali "dvo-dimenzionalnimi" ionskimi pastmi (Hager, 2002; Schwartz in sod., 2002), pri katerih se ioni akumulirajo v cilindričnem volumnu, ki je precej večji kot volumen v klasični tri-dimenzionalni ionski pasti. V ionski pasti lahko z ustreznim potencialom na elektrodah dosežemo fragmentacijo primarnih, pa tudi naslednjih generacij ionov, ki jih potem tudi analiziramo. Tako lahko z ionskimi pastmi generiramo masne spektre prve, druge in višjih (n) generacij.

4.1.2.5 Fourier transform ionsko ciklotronsko resonančni analizator

Tudi ionsko ciklotronski resonančni analizator sodi med analizatorje z ionsko pastjo, vendar za razliko analizatorjev z ionsko pastjo, ione zajema v območju visokega vakuumu in zelo visokega magnetnega polja. Zajeti ioni v pasti krožijo z določeno stabilno frekvenco. Poleg tega ioni oscilirajo s karakteristično frekvenco, tako v radialni kot v aksialni smeri (sekularna frekvenca). Na to naravno oscilacijo ionov, ki je odvisna od njihove m/z , lahko močno vplivamo z

dodatnim radio-frekvenčnim signalom. Tisti ioni, katerih frekvenca precesiranja pride v resonanco z dovedenim radio-frekvenčnim vzbujevalnim pulzom absorbirajo njegovo energijo. Resonančno vzbujeni ioni s kroženjem v neposredni bližini t.i. "sprejemnih" elektrod v le-teh inducirajo električni tok, ki ga merimo. Matematična analiza (Furierova transformacija) zabeleženega induciranelega električnega toka nam v končni fazi poda vrednosti m/z ionov v merilni celici. Te vrste analizatorjev odlikujejo visoka občutljivost, natančnost določanja mase, ločljivost in široko dinamično območje delovanja (Martin in sod., 2000). Izjemno visoka ločljivost temelji na specifičnem principu delovanja – frekvenco vzbujevalnega radio-signala je namreč mogoče izjemno natančno določiti. Omejitev za širšo uporabo teh analizatorjev pa je zaenkrat še vedno predvsem zahtevnost dela, nizka učinkovitost fragmentacije ionov in zelo visoka cena.

4.1.3 Načini identifikacije proteinov z MS

4.1.3.1 Peptidno mapiranje

V literaturi se za omenjeni način strukturne identifikacije proteinov uporablja v glavnem izraz "peptide-mass fingerprinting", najdemo pa tudi izraza "peptide mapping" in "peptide-mass mapping". Postopek so leta 1993 predlagali Henzel in sod. 1993 in temelji na razgradnji homogenega proteinskega vzorca s specifično proteazo, najpogosteje tripsinom ali endoproteinazo Lys-C, ter masno analizo dobljene peptidne zmesi. Na ta način dobimo vsakemu proteinu karakterističen masni spekter ali "prstni odtis" (od tod tudi izraz "peptide-mass fingerprinting"). Identifikacija proteina temelji na matematični primerjavi prekrivanj peptidnega masnega spektra z bazo podatkov v kateri so bodisi eksperimentalno določeni "prstni odtisi" proteinov ali pa teoretično generirani "prstni odtisi" proteinov. Za pripravo peptidov se namreč uporabljajo izključno visoko specifične proteaze, za katere se natančno ve, katere vezi v proteinih cepijo, tako da lahko napovemo fragmentacijski vzorec vsakemu proteinskemu (neposredno določenemu ali prevedenemu iz nukleotidnega zaporedja) zaporedju. Računalniška algoritma, s katerima po omenjenih principih iz masnega spektra napovemo najverjetnejšo identiteto analiziranega proteina sta npr. algoritem Mascot (Perkins in sod., 1999) in algoritem Sequest (Eng in sod., 1994). Pregled ostalih algoritmov je predstavljen v delu Nesvizhskijskega in sodelavcev (2007). Ker metoda zahteva v izhodišču homogen neznani protein, se ponavadi uporablja v povezavi s predhodnim čiščenjem vzorca z 1D- in 2D-PAGE.

4.1.3.2 "De novo" sekveniranje – tandemsko MS (MS/MS)

Bolj specifična metoda za identifikacijo proteina z MS od zgoraj omenjenega peptidnega mapiranja je "de novo" sekveniranje proteina (Steen in Mann, 2004). V ta namen se poslužujemo tehnike tandemске MS (MS/MS). Tandemski masni spektrometer se od običajnega razlikuje po tem, da ima zaporedno vezana dva masna analizatorja, med njima pa je nameščena t.i. kolizijska celica (ioni fragmentirajo ob trkih z atomi ali molekulami razredčenega plina v celici) oziroma nek drug element za fragmentacijo ionov (npr. kvadrupol). V prvem masnem analizatorju poteka analiza ionov, ki izvirajo neposredno iz ionizatorja

(primarni ioni). S specifičnimi pogoji v prvem masnem analizatorju lahko izberemo primarni ion s točno določeno m/z in ga vodimo v sektor, kjer na različne načine (Dass, 2001) izzovemo njegov razcep na t.i. sekundarne ione. Sekundarne ione potem analiziramo v drugem masnem analizatorju in dobimo sekundarni ali CID (od "Collision-Induced Dissociation") masni spekter. Fragmentacija primarnih ionov poteka v odvisnosti od količine energije, ki jo dovedemo ionom, vendar je vrstni red cepitve kovalentnih vezi v peptidu določen z energijo posameznih vezi. Najšibkejša je peptidna vez, zato se ta najprej cepi. Poimenovanje in številčenje sekundarnih ionov je standardizirano z Roepstorff–Fohlmann–Biemannovo nomenklaturjo (Roepstorff in Fohlman, 1984; Biemann, 1992). Odvisno od tega, na katerem razpadnem produktu peptida ostane pozitiven naboj dobimo ione serije b (naboj ostane na N-koncu) in serije y (naboj ostane na C-koncu). Če se cepi vez med alfa-ogljikovim atomom in karbnilno skupino v osnovni verigi peptida, dobimo ione serije a (naboj ostane na N-koncu) in serije x (naboj ostane na C-koncu), če pa popusti vez med alfa-ogljikovim atomom in dušikom v polipeptidni verigi, pa dobimo ione serije c (naboj ostane na N-koncu) in serije z (naboj ostane na C-koncu). Obstajajo različni računalniški algoritmi, ki v prvem koraku identificirajo vsaj eno od serij sekundarnih ionov (a, b, c x, y ali z) nato pa z odštevanjem mas sosednjih ionov znotraj posamezne serije identificirajo aminokislinsko zaporedje (Nesvizhskii in sod., 2007). Na osnovi primerjave določenega aminokislinskega zaporedja z bazami proteinskih zaporedij potem (z določeno verjetnostjo – glej Rappsilber in Mann, 2002) identificiramo analizirani protein (Nesvizhskii in sod., 2007).

4.1.3.3 Metoda "peptidne značke" ("Peptide Sequence Tag")

Fragmentacijski spektri ponavadi vsebujejo nekaj jasno rešljivih, vsaj dve aminokislini dolgih, aminokislinskih zaporedij. Algoritem imenovan "Peptide Sequence Tags" (Mann in Wilm, 1994) uporablja za identifikacijo proteina kratko aminokislinsko zaporedje, poleg tega pa še maso celotnega iona pred fragmentacijo in maso najlažjega iona v seriji po fragmentaciji. S pomočjo teh treh podatkov algoritem izračuna oddaljenosti identificiranega kratkega aminokislinskega zaporedja v Da do enega in drugega konca peptida. Peptidna značka, s katero algoritem potem preiskuje bazo proteinskih zaporedij (Nesvizhskii in sod., 2007), sestoji torej iz treh delov – mase zaporedja od znanega zaporedja do N-konca peptida, kratkega znanega zaporedja in mase od znanega zaporedja do C-konca peptida.

4.1.4 Bolj pogoste konfiguracije masnih spektrometrov

Masni spektrometri se med seboj razlikujejo predvsem v kombinaciji načina ionizacije in analizne metode, ki jo izvajajo. Glede na to se masni spektrometri ločijo po namembnosti. Način detekcije ionov je pri različnih masnih spektrometrih v glavnem zelo podoben (Dass, 2001). Tako je zelo pogosto uporabljena kombinacija MALDI ionizacije in TOF analizatorja, ki je zelo primerna za določanje mase intactnih polipeptidov, medtem ko je kombinacija ESI, analizatorja z ionsko pastjo in trojnega kvadrupola zelo uporabna za generiranje spektrov sekundarnih ionov in s tem za "de novo" sekveniranje peptidov (Aebersold in Goodlett, 2001). Uveljavljajo se tudi druge kombinacije.

Za fragmentacijo MALDI-generiranih primarnih ionov, so MALDI ionizator sklopili s kvadrupolom in analizatorjem z ionsko pastjo (Krutchinsky in sod., 2001). MALDI ionizator so sklopili tudi z dvema TOF analizatorjema (t.i. TOF-TOF instrument). Za prvim TOF analizatorjem se nahaja kvadrupol, kjer izbrani ion fragmentiramo, sekundarni ioni pa se potem analizirajo v drugem TOF sektorju (Medzihradszky in sod., 2000). Znan je tudi kombinacija, kjer ionizatorju sledita dva kvadrupola, drugi služi za fragmentacijo primarnih ionov, in pa TOF analizator (Loboda in sod., 2000). Primarni ion določene m/z izberemo v prvem masnem analizatorju, ki je bodisi TOF ali kvadrupol, fragmentiramo s kvadrupolom ali pa v kolizijski celici, sekundarne ione pa potem analiziramo s TOF analizatorjem. Masni spektrometri opisanih konfiguracij so visoko občutljivi, odlikuje jih visoka ločljivost in zanesljivost določitve mas. Pri kvadrupol TOF instrumentih lahko uporabljamo tako MALDI kot ESI način tvorbe primarnih ionov. Rezultirajoči spektri sekundarnih ionov so ponavadi bolj informativni kot sekundarni spektri pridobljeni z analizatorji z ionsko pastjo. Čeprav TOF, analizator z ionsko pastjo in hibridni TOF instrumenti trenutno dominirajo v proteomiki, se pojavljajo tudi druge konfiguracije, npr. z linearnimi ionskimi pastmi in ionsko ciklotronskimi resonančnimi analizatorji. Zaradi enostavnosti, visoke natančnosti, ločljivosti in občutljivosti se pri peptidnem mapiranju proteinov najpogosteje uporablja MALDI-TOF kombinacija.

4.1.5 Kvantifikacija proteinov z MS

Tako pri MALDI- kot pri ESI-MS je zveza med signalom in množino analita v vzorcu kompleksna in zaenkrat še nejasna. Masni spektrometri niso aparature s katerimi bi iz vzorca neposredno lahko pridobili kvantitativne rezultate. Za kvantitativno MS analizo je potrebno vzorec predhodno označiti z enim od stabilnih izotopov: ^{13}C , ^{15}N , ^{17}O ali ^{18}O . V vzorcu imamo po označevanju proteine, ki vsebujejo v naravi prevladujoče izotope: ^{12}C , ^{14}N in ^{16}O ter izotopsko označene proteine, ki so od prvih le nekoliko težji, sicer pa so njihove lastnosti prvim popolnoma identične. Take pare molekul lahko diferencialno analiziramo z MS in iz razmerja intenzitet m/z signalov natančno razberemo njihovo kvantitativno razmerje v vzorcu. Torej, če se izražanje določenega proteina ali skupine proteinov v danih pogojih poveča ali zmanjša, se to jasno odrazi v razmerju m/z signala naravne in izotopsko-označene oblike proteina ali skupine proteinov (Aebersold in Mann, 2003; Ong in Mann, 2005).

Vzorci izotopsko označujemo na tri načine. Stabilne izotope lahko vgradimo v proteine z metabolnim označevanjem, kjer v celično gojišče dodamo soli ali aminokisline, ki vsebujejo težke izotope (npr. ^{13}C -označen arginin). Ta metodologija se imenuje SILAC ("Stable Isotope Labeling by Amino acids in Cell culture") metodologija. Na ta način lahko označimo v principu vse peptide v določenem celičnem sistemu brez enega samega kemijskega koraka, na zelo enostaven način, prav tako pa je metoda združljiva z vsemi nadaljnjimi postopki čiščenja.

Proteine lahko izotopsko označimo na encimski način, npr. z vgradnjo ^{18}O iz molekule težke vode ob hidrolizi (proteolizi) peptidne vezi.

Tretji način izotopskega označevanja polipeptidov je kemijski način – modifikacija polipeptidnih molekul z reagenti različne izotopske sestave (Zhang

in sod., 2004). Proteine je potrebno za te vrste označevanje najprej izolirati. Ta metoda nudi trenutno največ možnosti in se v praksi tudi največ uporablja. Z izbiro ustreznega reagenta omogoča usmerjeno in selektivno označevanje funkcionalnih skupin v polipeptidih, kar nadalje omogoča selektivno izolacijo in analizo. Obstajajo reagenti, specifične za sulfhidrilno skupino, amino skupino, serin in cistein v aktivnih mestih serinskih ali cisteinskih proteaz, fosfatne estre in N-vezane sladkorje v proteinih. Ta metoda označevanja ponuja še številne možnosti in pričakuje se razvoj novih reagentov, ki bodo omogočili kvantitativno MS analizo novih skupin proteinov – “pod-proteomov”.

5 TRENDI RAZVOJA V METODOLOGIJI ZA ANALIZO PROTEOMA

V prispevku sem se omejil na metodologije za separacijo kompleksnih proteinskih mešanic in na visoko-pretočno strukturno identifikacijo polipeptidov v proteomiki. Če ostanem pri tem, potem strokovnjaki kot najbolj perspektivne za napredek področja izpostavljajo nove pristope v pripravi bolj enostavnih vzorcev pred strukturno analizo. Precej možnosti za razvoj je še na področju bioinformatike, programskih orodij za analizo kompleksnih in obsežnih podatkovnih zbirk karakterističnih pri proteomskih raziskavah (Nesvizhskii in sod., 2007). Razpoložljivost celotnih genomskih zaporedij vedno večjega števila organizmov dodatno pospešuje na MS-temelječo identifikacijo proteinov in zmanjšuje potrebo po uporabi počasnejših identifikacijskih postopkov, kot je npr. “*de novo*” sekveniranje (Taylor in sod. 2007). Tretja smer razvoja, od katere se precej pričakuje, je razvoj instrumentacije, masnih spektrometrov, v smeri še bolj občutljivih, še bolj zanesljivih, visoko ločljivih in hitrejših sistemov (Domon in Aebersold, 2006).

Na MS-temelječa proteomika se je v nekaj letih razvila v zelo močno orodje za analizo proteinske sestave bioloških sistemov, posttranslacijskih modifikacij proteinov in interakcij med proteini. Razvoj se sedaj usmerja zlasti v obsežno in zahtevno analizo nivoja izražanja posameznih proteinov v celičnih sistemih (Cox in Mann, 2007).

6 VIRI

- Aebersold R., Goodlett D.R. 2001. Mass spectrometry in proteomics. *Chem. Rev.*, 101: 269–295.
- Aebersold R., Mann M. 2003. Mass spectrometry-based proteomics. *Nature*, 422: 198–207.
- Biemann K., 1992. Mass spectrometry of peptides and proteins. *Annu. Rev. Biochem.*, 61: 977–1010.
- Cox J., Mann M. 2007. Is Proteomics the New Genomics? *Cell*, 130: 395–398.
- Domon B., Aebersold R. 2006. Mass Spectrometry and Protein Analysis. *Science*, 312: 212–217.
- Dass C. 2001. Principles and practice of biological mass spectrometry. John Wiley & Sons, Inc.: 416 str.
- Eng J.K., McCormack A.I., Yates J.R. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.*, 5: 976–989.
- Fenn J.B., Mann M., Meng C.K., Wong S.F., Whitehouse C.M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246: 64–71.
- Gingras A.-C., Gstaiger M., Raught B., Aebersold R. 2007. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.*, 8: 645–654.

- Hager J.W. 2002. A new linear ion trap mass spectrometer. *Rapid Commun. Mass. Spectrom.*, 16: 512–526.
- Henzel W.J., Billeci T.M., Stults J.T., Wong S.C. 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc. Natl Acad. Sci. USA*, 90: 5011–5015.
- Jensen O.N. 2004. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.*, 8: 33–41.
- Kamp M.R., Choli-Papadopoulou T., Wittman-Liebold B. 1997. Protein structure analysis. Preparation, characterization, and microsequencing. Springer-Verlag: 316 str.
- Karas M, Hillenkamp F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.*, 60: 2299-2301.
- Kislinger T., Emili A. 2005. Multidimensional protein identification technology. *Expert Rev. Proteomics*, 2: 27-39.
- Krutchinsky A.N., Kalkum M., Chait B.T. 2001. Automatic identification of proteins with a MALDIquadrupole ion trap mass spectrometer. *Anal. Chem.*, 73: 5066–5077.
- Hoogland C., Mostaguir K., Sanchez J.-C., Hochstrasser D.F., Appel R.D. 2004. SWISS-2DPAGE, ten years later. *Proteomics*, 4: 2352-2356.
- Loboda A.V., Krutchinsky A.N., Bromirski, M., Ens, W., Standing, K.G.A. 2000. Tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. *Rapid Commun. Mass Spectrom.*, 14: 1047–1057.
- Mann M., Wilm M.S. 1994. Error tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.*, 66: 4390–4399.
- Martin S. E., Shabanowitz J., Hunt D. F., Marto J. A., 2000. Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 72: 4266–4274.
- Medzihradszky K.F. et al. 2000. The characteristics of peptide collision-induced dissociation using a highperformance MALDI-TOF/TOF tandem mass spectrometer. *Anal. Chem.*, 72: 552–558.
- Nesvizhskii A., Vitek O., Aebersold R. 2007. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nature Methods*, 4: 787-797.
- Ong S.-E., Mann M., 2005. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.*, 1: 252-262.
- Perkins D.N., Pappin D.J., Creasy D.M., Cottrell J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20: 3551–3567.
- Rabilloud T. 2000. Proteome Research: Two Dimensional Gel Electrophoresis and Identification Methods. Springer-Verlag: 248 str.
- Rabilloud T., 2002. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics*, 2: 3–10.
- Rappsilber J., Mann M. 2002. What does it mean to identify a protein in proteomics? *Trends Biochem. Sci.*, 27: 74-78.
- Roepstorff P., Fohlman J. 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.*, 11: 601.
- Schwartz J. C., Senko M.W., Syka, J.E. 2002. A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 13: 659–669.
- Steen H., Mann M., 2004. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* 5: 699-711.
- Taylor C.F. et al. 2007. The minimum information about a proteomics experiment (MIAPE). *Nature Biotechnol.* 25: 887-893.
- Unlu M., Morgan M.E., Minden J. S. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis*, 18: 2071–2077.
- Warnock D.E., Fahy E., Taylor S.W. 2004. Identification of protein associations in organelles, using mass spectrometry-based proteomics. *Mass Spectr. Rev.*, 23: 259–280.
- Westmermeier R., 1997. Electrophoresis in practice. VCH Verlagsgesellschaft: 331 str.
- Zhang H., Yan W., Aebersold R. 2004. Chemical probes and tandem mass spectrometry: a strategy for the quantitative analysis of proteomes and subproteomes. *Cur. Opin. Chem. Biol.*, 8: 66–75.