

# Analiza proteina spektrometrijom masa

KUI – 10/2008  
Prijelo 22. veljače 2007.  
Prihvaćeno 21. lipnja 2007.

N. Galić<sup>a</sup> i M. Cindrić<sup>b</sup>

<sup>a</sup> Zavod za analitičku kemiju, Kemijski odsjek, Prirodoslovno-matematički fakultet, Sveučilište u Zagrebu, Horvátovac 102a, 10 000 Zagreb, Hrvatska

<sup>b</sup> Mario Cindrić, Pliva – Research & Development, Prilaz baruna Filipovića 29, 10 000 Zagreb, Hrvatska

Opisana je analiza bioloških spojeva spektrometrijom masa (MS) uz ionizaciju elektroraspršenjem (ESI) i matricom pomognutu ionizaciju uz desorpciju laserskim zračenjem (MALDI). Objasnjena je karakterizacija proteinskih posttranslacijskih modifikacija i identifikacija proteina na temelju određivanja molekulske mase peptida (*engl.* peptide mass fingerprint, PMF) te identifikacija na temelju fragmentiranja peptida (kolizijom inducirana fragmentacija i poslijeionizacijska fragmentacija). Također je dan pregled metoda za kemijsku modifikaciju peptida, čime se omogućuje kvantifikacija i *de novo* sekvenciranje proteina.

Ključne riječi: *Spektrometrija masa, identifikacija proteina, identifikacija posttranslacijskih modifikacija, kvantitativna analiza proteina*

## Uvod

Od otkrića elektroraspršenja (ESI)<sup>1</sup> i matricom pomognute ionizacije uz desorpciju laserskim zračenjem (MALDI)<sup>2</sup> spektrometrija masa postala je iznimno važna tehnika u analizi bioloških makromolekula. Točno određivanje mase peptida ili proteina omogućuje identifikaciju proteina, određivanje aminokiselinskog slijeda, identifikaciju i određivanje položaja posttranslacijske modifikacije (glikozilacija, nastajanje disulfidnih mostova, fosforilacija itd.), određivanje mutacija te provjeru strukture i čistoće proteina dobivenih genetskim inženjeringom.<sup>3</sup> Nadalje, spektrometrija masa primjenjuje se i u istraživanju različitih interakcija između proteina i iona metala, malih organskih molekula i različitih bioloških spojeva. Određuje se stehiometrija nastalih nekovalentnih kompleksa, afinitet i konformacijske promjene povezane s navedenim interakcijama.<sup>4,5</sup> Spektrometrija masa široko se primjenjuje u farmaceutskoj industriji u ispitivanju malih organskih molekula kao potencijalnih lijekova ali i biofarmaceutika, većinom rekombinantnih proteina kao npr. inzulina, interferona itd.<sup>6</sup> U tu svrhu rabe se spektri masa kao i spektri dobiveni tandemnom spektrometrijom masa, MS/MS ili MS<sup>n</sup>. Ionizacijske tehnike ESI<sup>7–10</sup> i MALDI<sup>11–15</sup> detaljno su objašnjene u literaturi, te su u ovom radu samo u osnovnim crtama opisana njihova svojstva nužna za razumijevanje analize biomolekula spektrometrijom masa.

## Analiza proteina spektrometrijom masa

### Tehnike analize proteina spektrometrijom masa

Tehnikom elektroraspršenja nastaju višestruko nabijeni ioni što omogućuje njihovo detektiranje i uz klasične spektrometre masa malih vrijednosti  $m/z$  kao što su kvadrupol i ionska stupica. Tehnikom MALDI uglavnom nastaju jednostruko nabijeni ioni. Budući da se radi o pulsnoj tehnici, najčešće se povezuje s analizatorom masa koji mjeri vrijeme leta (TOF), ali je moguće i povezivanje s analizatorima

koji skladište ione, primjerice ionska stupica (*engl.* ion trap) i analizator ionsko-ciklotronske rezonancije uz Fourierovu transformaciju (*engl.* Fourier transform ion cyclotron resonance, FT ICR). Osnovne karakteristike (osjetljivost, mjerno područje, točnost masa i razlučivanje) analizatora masa koji se rabe u analizi bioloških spojeva prikazani su u tablici 1. Tablica je načinjena kombiniranjem podataka iz više literaturnih izvora.<sup>4,16,17</sup>

Granica određivanja bioloških spojeva tehnikama MALDI i ESI nalazi se između femtomola i pikomola, a ovisi o prirodi uzorka, načinu pripreme i čistoći uzorka. Biološki uzorci obično su razrijeđene otopine peptida ili proteina i onečišćenja (nehlapljivi puferi, detergentski soli itd.). Onečišćenja mogu smanjiti ili u potpunosti prigušiti intenzitet željenog iona, grade adukte s ionima analita i time utječu na osjetljivost, ali i na točnost metode ukoliko ioni adukta nisu dovoljno dobro odijeljeni od iona analita.<sup>18</sup> Treba napomenuti da je MALDI otporniji na onečišćenja teško hlapljivim solima (npr. fosfati, ftalati, kloridi) u odnosu na ESI zbog različitog načina pripreme uzorka i procesa desorpcije i ionizacije. Navedeni nedostaci u praksi se rješavaju uvođenjem separacijske tehnike. Najčešće se primjenjuje tekućinska kromatografija obrnutih faza (*engl.* liquid chromatography, LC), odnosno vezani sustav LC-MS uz ionizaciju elektroraspršenjem. Izravno povezivanje tehnika LC i MALDI nije moguće, pa se eluati iskapavanjem nanose na pločicu MALDI.<sup>19</sup> Kromatografska komponenta može se nadomjestiti primjenom tehnike Zip-Tip. Zip-Tip-om se naziva malena masa (od nekoliko miligrama do jednog grama) nepokretne faze širokih pora (C<sub>18</sub>, C<sub>4</sub> itd.), koja se nalazi na kraju plastičnog nastavka pipetora. Nanošenjem uzorka i zatim gradijentnim ispiranjem nepokretne faze dobivaju se eluati u kojima analiti nisu tako dobro odvojeni kao na kromatografskoj koloni, ali su bez soli i sadrže ukoncentrirani analit. Proteini se osim kromatografskim metodama mogu odijeliti 2-D gel-elektroforezom, a zatim analizirati spektrometrijom masa, često tehnikom MALDI-TOF.<sup>20</sup> U posljednjih desetak godina se za razdvajanje kompleksnih proteinskih smjesa

T a b l i c a 1 – Usporedba analizatora masa

T a b l e 1 – General comparison of mass analyzers

Analizator Analyzer	Osjetljivost Sensitivity	Gornja granica m/z m/z upper limit	Razlučivanje Resolution	Točnost mase (u) Mass accuracy (u)
Kvadrupol Quadrupole	femtomol femtomole	<3000	2000–3000	0,1
Sektor Sector	femtomol femtomole	<15 000	5000–100 000	0,0001
Ionska stupica Ion trap	femtomol femtomole	<5000	2000	0,1
TOF* Time-of-flight	atomol attomole	Teoretski bez ograničenja No theoretical limitation	5000–10 000 uz odgođenu ekstrakciju 5000–10 000 with delayed extraction	0,0001
Orbitrap	atomol attomole	<6000	150 000	0,0001
FT ICR**	atomol attomole	<10 000	Ovisi o magnetu, za 9,4 T > 1 000 000 Dependent on magnet, >1 000 000 at 9.4 T	0,0001–0,00001

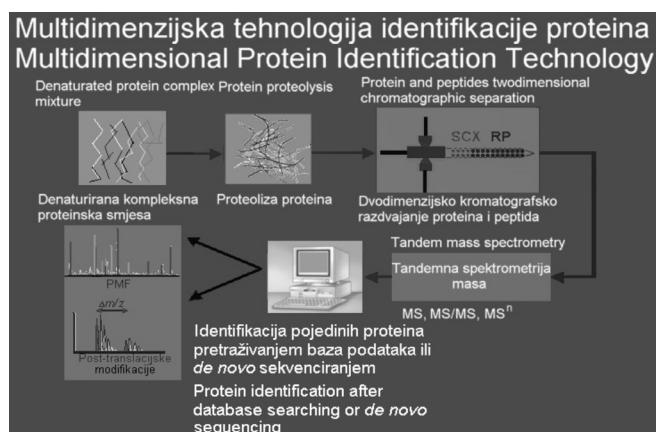
\* Analizator koji mjeri vrijeme leta

\*\* Analizator ionsko-ciklotronske rezonancije uz Fourierovu transformaciju

\*\* Fourier transform ion cyclotron resonance

(npr. smjese staničnih ekstrakata koji se sastoje od više tisuća proteina) osim primjene dvodimenzijske gel-elektroforeze služe i separacijske tehnike višedimenzijske kromatografije.<sup>21</sup> Dvodimenzijska kromatografija se može primjenjivati sukcesivno za razdvajanje proteina unutar pojedine frakcije koja je prethodno dobivena razdvajanjem kompleksne smjese na više frakcija u prvoj kromatografskoj dimenziji (npr. prva dimenzija kationski izmjenjivač, a druga dimenzija kromatografija obrnutih faza). Daljnje unapređene ovakvog pristupa prvi je put prikazano radom J. R. Yatesa III,<sup>22–24</sup> u kojem je nepokretna faza jakog kationskog izmjenjivača (*engl.* strong cation exchange, SCX) prva kromatografska dimenzija nanosena na početak kolone, a nadalje je unutar iste kolone slijedi nepokretna faza širokih pora, oktadecilsilicijev dioksid (C<sub>18</sub>).<sup>25</sup> Peptidi se nakon proteolize vežu za nepokretnu fazu jakog kationskog izmjenjivača u kiselom puferu, te se stupnjevito u frakcijama (pulsovima) ispiru prema polarosti amonijevim acetatom. Svaka frakcija nakon ispiranja s nepokretne faze jakog kationskog izmjenjivača se nadalje raščlanjuje na proteinske i peptidne komponente razdvajanjem na nepokretnoj fazi C<sub>18</sub> gradijentnim ispiranjem vodom i acetonitrilom uz dodatak trifluoroctene kiseline (TFA) i linearnim povećanjem udjela acetonitrila. Krajnji rezultat takve analize širi mogućnosti kromatografskog razdvajanja na par tisuća proteina i peptida uz mogućnost daljnje analize tehnikama MS (amonijev acetat je lako hlapljiv pufer i u koncentracijama od  $c = 250 \text{ mmol dm}^{-3}$  do  $c = 500 \text{ mmol dm}^{-3}$  koje se upotrebljavaju kao pokretna faza za prvu kromatografsku dimenziju ne ometa osjetno ionizaciju analita). Prikaz dvodimenzijske kromatografije engleskog akronima MudPIT (multidimensional protein identification technology) namijenjene razdvajanju i analizi kompleksnih proteinskih smjesa dan je na slici 1.

Analiza proteina spektrometrijom masa obuhvaća analizu intaktnog proteina i analizu peptida nastalih digestijom proteina prije MS analize. Ovisno o konstrukciji instrumenta, poglavito vrsti analizatora, analize proteina mogu se provoditi na dva načina: odozgor nadolje i odozdol nagore



Slika 1 – Dvodimenzijska kromatografija engleskog akronima MudPIT (multidimensional protein identification technology) namijenjena razdvajanju i analizi kompleksne proteinske smjese

Fig. 1 – Two-dimensional chromatography MudPIT (Multidimensional Protein Identification Technology) developed for separation and analysis of complex protein mixtures

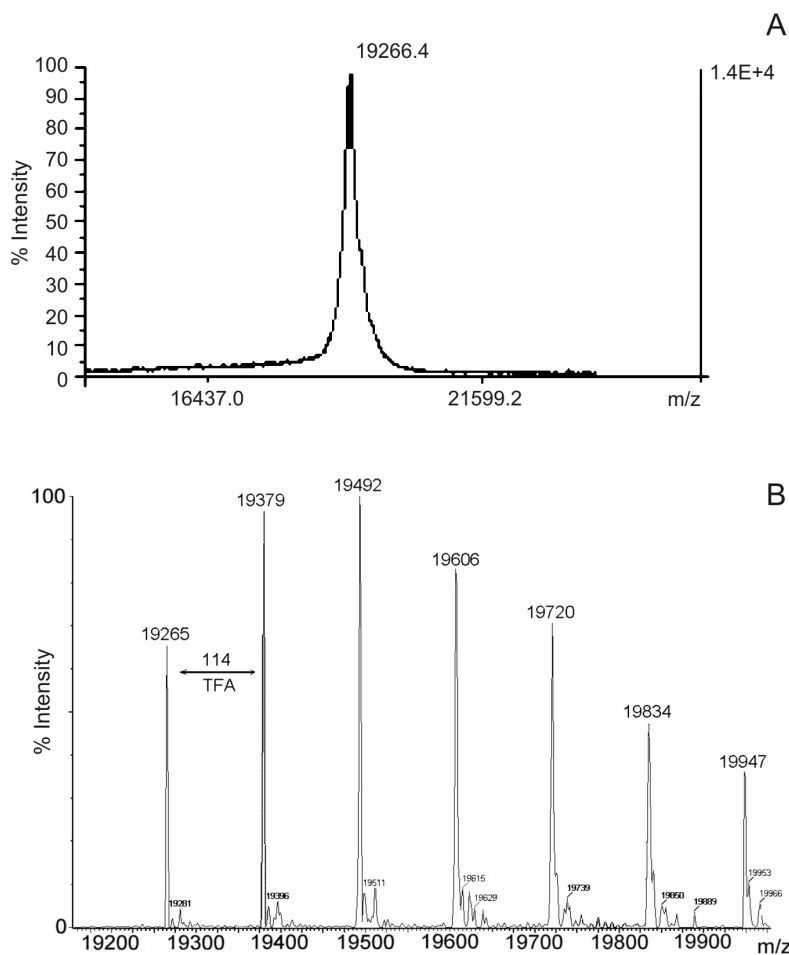
(*engl.* top-down i bottom-up).<sup>26,27</sup> Analiza proteina načinom odozgor nadolje vezana je uz potpunu analizu primarnog slijeda aminokiselina i posttranslacijskih modifikacija intaktnog proteina unutar samog instrumenta (npr. određivanje aminokiselinskog slijeda proteina ili karakterizacija glikoproteina, analiza mjesta fosforilacija, analiza disulfidnih mostova). Takav pristup zahtijeva točno mjerenje masa i počesto je opterećen gubitkom informacija o strukturi proteina nakon provedene analize. Instrumenti na kojima se mogu provesti analize načinom odozgor nadolje su FT ICR i Orbitrap.<sup>26,27</sup> Rašireniji način analize proteina načinom odozdol nagore uključuje enzimsku ili kemijsku odgradnju (cijepanje) proteina u gelu ili u otopini prikladnom proteinazom, glikozidazom, hidrolizirajućim reagensom ili nekim drugim enzimom ili kombinirano. Nakon odgradnje proteina nastali proteinski ili peptidni fragmenti se analiziraju

raju nekom od MS tehnika, a od dobivenih podataka gradi se ukupna struktura proteina i njegovih mogućih posttranslacijskih modifikacija.<sup>28</sup>

### Analiza intaktnog proteina

Molekulska masa intaktnog proteina može se odrediti tehnikama ESI ili MALDI MS. U slučaju ionizacije elektroaspršenjem više višestruko nabijenih iona nakon programskog postupka dekonvolucije spektara masa daju jedan signal koji odgovara prosječnoj molekularnoj masi proteina. Raspodjela i broj višestruko nabijenih iona ovisi o otapalu, pH, ionskoj jakosti otapala itd. Dodatkom trifluoroctene, mravlje, octene ili neke druge lako hlapljive kiseline postiže se povećana protonacija,<sup>29</sup> odnosno dodatkom 2-metoksi-etanola, dietilamina ili etilen-glikola povećana deprotonacija, čime se utječe na raspodjelu višestruko nabijenih iona.<sup>30</sup> Tehnika MALDI-TOF odlikuje se iznimnom osjetljivošću, velikom brzinom analiza (moguće je analizirati oko 100 uzoraka u 10 min) i neosjetljivošću na otapala. Ako se

uzorci analiziraju u linearnom načinu snimanja, moguće je detektirati prosječne mase i preko 200 000 Da, dok je kod snimanja uz reflektrom (ionsko zrcalo) gornja granica masa oko 20 000 Da. ESI i MALDI MS spektri rekombinantnog ljudskog interferona  $\alpha$ -2b (*rHuINF*  $\alpha$ -2b) prikazani su slikom 2.<sup>31</sup> U ESI spektru masa uočavaju se signali niza višestruko nabijenih iona *rHuINF*  $\alpha$ -2b te adukata tih iona s trifluorocetnom kiselinom. Dekonvolucijski spektar sadrži signale koji odgovaraju *rHuINF*  $\alpha$ -2b i njegovim aduktima s jednim ili više iona trifluorocetne kiseline. U spektru MALDI nalaze se samo signali jednostruko i dvostruko nabijenih iona *rHuINF*  $\alpha$ -2b. Objema tehnikama određena je masa proteina koja se dobro slaže s izračunatom masom od 19 265,16 Da. Na isti način analiziraju se i smjese proteina. Točno određivanje mase ovisi o kalibraciji instrumenta. Spektrometar masa može se kalibrirati referentnim spojem prije unošenja uzorka (vanjska kalibracija) ili se referentni spoj unosi u isto vrijeme kad i uzorak (unutarnja kalibracija), pri čemu kalibrant treba pokrivati široko područje masa (npr. polietilen-glikol ili klusteri smjese natrijevog i cezijevo jodida). Veća točnost postiže se unutarnjom kalibracijom. Točnost određivanja masa nalazi se u rasponu od  $w < 1 \cdot 10^{-6}$  (ESI- i MALDI-FTICR) do  $w < 5 \cdot 10^{-6}$  (ESI- i MALDI-TOF). Rutinskim analizama "malih molekula" i peptida na instrumentima koji kao analizator imaju jednostruki ili trostruki kvadrupol ili ionsku stupicu uz pravilno optimiranu vanjsku kalibraciju određuju se mase s pogreškom  $w < 50 \cdot 10^{-6}$ .<sup>3,32,33</sup>



Slika 2 – MALDI-TOF spektar dobiven analizom 0,01 mg interferona  $\alpha$ -2b uz linearni način snimanja. Rezolucija signala kod mase 19 266 Da računana na polovici visine iznosi 241 (A). Dekonvoluirani spektar interferona  $\alpha$ -2b (0,01 mg). Izmjerena masa interferona  $\alpha$ -2b je 19 266 Da (izračunana prosječna masa iznosi 19 265 Da). Signali s povećanjem mase 114 Da nastali su nakon stvaranja adukata interferona  $\alpha$ -2b i TFA (B)

Fig. 2 – MALDI-TOF spectrum of 0.01 mg interferon  $\alpha$ -2b in linear mode. Resolution in linear mode at mass value 19 266 Da calculated with 50 % centroid is 241 (A). Deconvoluted spectrum of interferon  $\alpha$ -2b (0.01 mg). Observed mass value of interferon  $\alpha$ -2b is 19 266 Da (calculated average mass value is 19 265 Da). Other peaks in deconvoluted spectrum with mass increment 114 Da represent interferon  $\alpha$ -2b – TFA adducts (B)

### Određivanje primarne strukture proteina

#### Identifikacija proteina i posttranslacijskih modifikacija

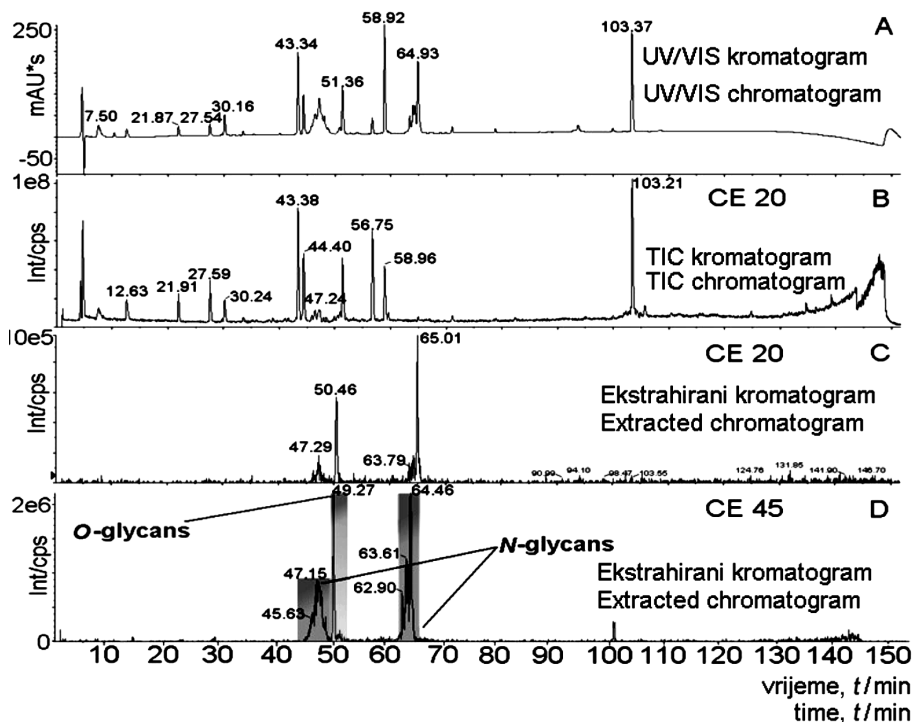
Najbrža metoda za identifikaciju proteina, koji su u većini slučajeva prethodno odjeljeni 2-D gel-elektroforezom pa pocijepani enzimom ili pocijepani enzimom pa kromatografski odvojeni je tzv. metoda otiska prsta (*engl.* peptide mass fingerprint, PMF).<sup>34,35</sup> Cijepanje proteina se osim enzimskim putem može postići i kemijskom odgradnjom, a nastala smjesa peptida analizira se spektrometrijom masa. Osnovni princip metode temelji se na uspoređivanju eksperimentalno određenih masa peptida s teoretski izračunatim masama pohranjenim u bazama podataka (The Protein Data Bank,<sup>36</sup> SwissProt<sup>37</sup>). Kod pretraživanja baza podataka uvijek treba biti svjestan krajnjeg očekivanog rezultata. Tako je krajem 2006. u bazi podataka SwissProt bilo 250 256 proteinskih sljedova razvrstanih po imenima gena, porijeklu organizma ili nekom drugom kriteriju.<sup>37</sup> Primjerice, pretraživanje baze podataka za proteine izolirane iz *Mammillaria gracillis* (slabo istraženi proteom)<sup>38</sup> neće uroditi smislenim ishodom, dok je vjerojatnost nalaženja proteina iz *Escherichia coli* (izrazito dobro istraženi proteom) povezana isključivo s kvalitetom

analitičkog postupka primijenjenog u izolaciji i analizi.<sup>39</sup> Odabir baze podataka također može biti ključan za ishod analize. Baza podataka The Protein Data Bank sadrži za razliku od baze podataka SwissProt sedam puta manji broj unesenih proteinskih sljedova, ali podaci iz te baze podataka sadrže tercijarne i kvarterne strukturalne elemente proteina (X-Ray 34 407, NMR 5996 i elektronska mikroskopija 141).<sup>22,40</sup>

Razvijeni su različiti algoritmi za identifikaciju proteina na temelju masa peptida i dostupni su putem interneta.<sup>3,41</sup> Neki od tih algoritama rade pretraživanje nakon pretvorbe spektralnih podataka u tekstualni numerički oblik (lista svih monoizotopnih masa pronađenih u određenom spektru iznad zadane granice detekcije), a nekima je dostatan sirovi podatak (*engl.* raw data), odnosno spektar masa u originalnom digitalnom zapisu za provedbu pretraživanja baze ili više baza podataka. Prvi su vremenski zahtjevniji i većinom su dostupni putem interneta (npr. ProteinProspector),<sup>42</sup> te nisu namijenjeni visokofrekventnim analizama (*engl.* high-throughput analysis), dok su drugi komercijalni i vezani uz pojedine proizvođače spektrometara masa (npr. Mascot<sup>43</sup> ili GlobalLynx<sup>44</sup>). Cijepanjem proteina specifičnom proteinazom nastaje jedinstveni skup peptida, pa je dovoljno samo nekoliko točno određenih masa peptida za identifikaciju proteina.<sup>45</sup> Ukoliko pretraživanjem baza podataka nije postignuto zadovoljavajuće slaganje, potrebno je uzorak proteina pocijepati i s drugim proteinazama, uzeti u obzir popunjenost baze podataka skupom ciljanih proteina, primijeniti dodatne analize koje će utvrditi potencijalne posttranslacijske modifikacije ili načiniti *de novo* sekvenciranje. Potencijalne posttranslacijske modifikacije, kao što su fosforilacija ili glikozilacija, mogu se ustanoviti ciljanim analizama spektrometrije masa uz prilagođen postupak pripreme uzoraka. Kod fosfoproteina osim uobičajenog načina analize proteina opisanog u ovom radu treba primijeniti metaloafinitetnu kromatografiju (*engl.* immobilized metal affinity chromatography, IMAC) sa svrhom aglomeriranja fosfoproteina. Nakon postupka obogaćivanja smjese proteina fosfoproteinima i proteolize, kod same analize spektrometrom masa, fosfopeptidi se mogu detektirati snimanjem svih iona prekursora i iona produkata čija se masa razlikuje 98 Da (neutralni gubitak fosforne kiseline) te konsekvativnim MS<sup>3</sup> sekvenciranjem iona produkata.<sup>46</sup> Dobivene podatke možemo analizirati prema postojećim bazama podataka specijaliziranim za pronalaženje fosfoproteina<sup>47</sup> imajući na umu da se potencijalna mjesta fosforilacije serin, treonin i tirozin, nalaze u specifičnom amino-kiselinskom okruženju (tzv. motiv).

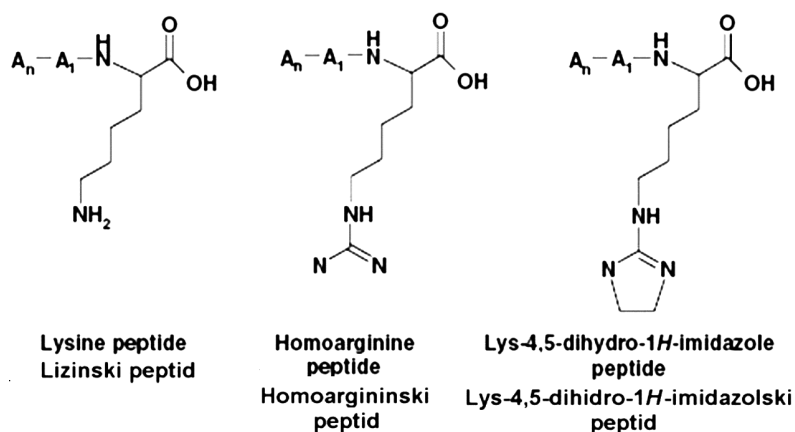
Za analizu glikoproteina također je poželjna upotreba fragmentacijske MS analize uz prethodnu pripremu uzorka koja uključuje i proteolizu i

odgradnju šećera. Šećeri se u kromatogramu ukupne ionske struje (*engl.* total ion current, TIC) mogu raspoznati ekstrakcijom pojedinih masa iz kromatograma (slika 3). Ekstrakcijom dijagnostičkih iona pri povišenoj kolizijskoj energiji (npr. 40 eV):  $m/z$  366 Hex-Hex<sup>+</sup> NAc<sup>+</sup>,  $m/z$  292 NeuAc<sup>+</sup> ili NeuGc<sup>+</sup>,  $m/z$  204 HexNAc<sup>+</sup>,  $m/z$  163 Hex<sup>+</sup> i  $m/z$  657 NeuAc-Hex-HexNAc<sup>+</sup> ili NeuGc-Hex-HexNAc<sup>+</sup>, može se utvrditi postojanje glikoproteina u uzorku.<sup>28</sup> Odgradnja *N*- ili *O*-glikana može se postići djelovanjem enzima (endoglikozidaze) ili pri povišenoj temperaturi u blago bazičnim ili izrazito kiselim uvjetima.<sup>48</sup> Nakon odgradnje ugljikohidrati se moraju izdvojiti iz peptidne ili proteinske smjese prije MS-analize jer blizina proteina ili soli u maloj količini onemogućuje analizu šećera koji nemaju povoljan afinitet za protonacijom ili deprotonacijom. Razdvajanje i pročišćavanje ugljikohidrata prije MS-analize se najučinkovitije provodi adsorpcijom šećera na aktivni ugljen, ispiranjem vodom te desorpcijom smjesom vode, acetonitrila i TFA,  $w = 0,1 \cdot 10^{-2}$ .<sup>49</sup> Analiza proteinskog djela glikoproteina može uslijediti nakon proteolize, koju slijedi spektrometrijska analiza i pretraživanja baza podataka specijaliziranih za glikoproteine algoritima koji su prilagođeni kompleksnoj strukturi ugljikohidrata.<sup>50,51</sup> Broj unosa glikoproteina u bazama Glycosuite (komercijalna baza)<sup>52</sup> ili The Protein Data Bank<sup>36</sup> (nekomercijalna baza) manji je od 10000, što je u usporedbi s proteinskim bazama podataka i poznavajući moguću kompleksnost glikoproteina, relativno malen broj. Pretraživanje se može olakšati poznavajući način vezanja i deglikozilacije *N*- i *O*-glikana.<sup>51,53</sup> Kod sisavaca i većine biljaka *N*-glikani se najčešće vežu na asparagin s motivom Asn-Xxx-(Ser, Thr), a *O*-glikani na serin ili treonin. Nakon deglikozilacije *N*-glikoproteina, asparagin prelazi u aspartat povećavajući mjerenu masu proteina ili peptida za 1 Da.



Slika 3 – UV/VIS (A) i TIC (B) kromatogrami te kromatogrami ekstrahiranih masa dijagnostičkih iona za detekciju glikana snimljeni pri kolizijskoj energiji 20 eV (C) i 45 eV (D)

Fig. 3 – UV/VIS (A) and TIC (B) chromatograms followed by extracted ion chromatograms of glycan diagnostic ions obtained at collision energies 20 eV (C) and 45 eV (D)



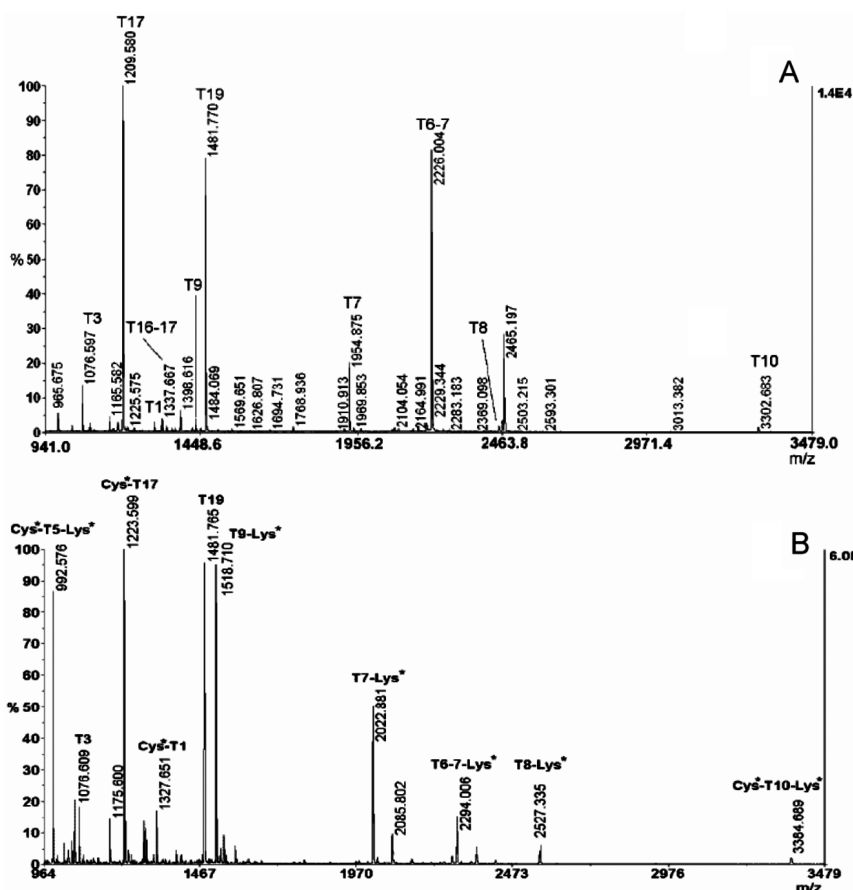
Slika 4 – Strukture C-terminalnih lizinskih peptida: prije derivatizacije i nakon derivatizacije 2-metilizouream i 2-metoksi-4,5-dihidro-1H-imidazolom kojom nastaje homoarginin i Lys-4,5-dihidro-1H-imidazol

Fig. 4 – Structures of C-terminal lysine-containing peptide before derivatization, C-terminal homoarginine- and Lys-4,5-dihydro-1H-imidazole-containing peptides after derivatization with *O*-methylisourea and 2-methoxy-4,5-dihydro-1H-imidazole, respectively

Tablica 2 – Specifično cijepanje proteina djelovanjem proteinaza

Table 2 – Specific cleavage of proteins by proteases

Proteinaza Protease	Specifično mjesto cijepanja Cleavage site
Tripsin	Iza lizina i arginina ako iza njih nema prolina
Trypsin	C-terminal side of arginine and lysine
Glu-C	Ispred asparginske i glutaminske kiseline ako iza njih nema glutaminske kiseline ili prolina
Glu-C	C-terminal side of glutaminic acid
Lys-C	Iza lizina ako iza njega nema prolina
Lys-C	C-terminal side of lysine
Asp-N	Ispred asparginske kiseline
Asp-N	N-terminal side of aspartic acid



Slika 5 – MALDI spektr masa rHuNF-2b triptičkog digesta prije (A) i nakon (B) derivatizacije 2-metoksi-4,5-dihidro-1H-imidazolom (1 pmol na MALDI pločici). Fragmenti T1 i T17 sadrže svaki po jedan cistein te im je nakon derivatizacije zabilježen porast  $m/z$  14,017. Fragmenti T5 i T10 sadrže lizin i cistein sa zabilježenim porastom  $m/z$  82,041

Fig. 5 – MALDI mass spectrum of rHuNF-2b tryptic digest before (A) and after (B) 2-methoxy-4,5-dihydro-1H-imidazole derivatization (1 pmol on the MALDI plate). Fragments T1 and T17 contain one cysteine and have observed mass increment of  $m/z$  14.017. Fragments T5 and T10 contain lysine and cysteine with observed mass increment  $m/z$  82.041

Proteinaze koje se najčešće primjenjuju su tripsin, Asp-N, Glu-C i Lys-C. Načini specifičnih cijepanja proteina proteinazama prikazani su u tablici 2.<sup>54,55</sup> Proteinaze su odabrane prema najvećoj vjerojatnosti pronalazanja fragmenata u nekoj od spektrometrijskih baza podataka.

Poznato je da detekcija peptida spektrometrijom masa ovisi o intrinzičkim svojstvima kao što su slijed aminokiselina, bazičnost C-krajnje aminokiseline, veličina i hidrofobnost peptida i mogućnost nastajanja stabilne sekundarne strukture. Učinak supresije signala opažen je u analizi smjese peptida dobivene cijepanjem tripsinom.<sup>56</sup> Zbog veće bazičnosti arginina u odnosu na lizin, fragmenti koji na C-kraju sadrže arginin prilikom ionizacije tehnikom MALDI mogu djelomično ili u potpunosti prigušiti signale lizinskih fragmenata, što utječe na određivanje peptidnog slijeda u potpunosti, a time i na identifikaciju proteina. Kako bi se povećao prinos iona s lizinom na C-kraju peptida, rabe se kemijske modifikacije lizina (slika 4). Reakcijama gvanidinacije s *O*-metilizouream<sup>57,58</sup> (2-metilizourea) lizin se prevodi u homoarginin, povećava mu se bazičnost, a time i intenzitet signala odgovarajućih peptida u spektru masa. Povećanje djelotvornosti ionizacije peptida koji sadrže lizin postiže se i pomoću 2-metoksi-4,5-dihidro-1H-imidazola (slike 4 i 5).<sup>59</sup> Reakcije derivatiziranja odvijaju se u otopini, a u novije vrijeme i na koloni (Zip-Tip). Reakcijama na čvrstoj fazi skraćuje se vrijeme potrebno za postupak derivatizacije, a ujedno se provodi i postupak pročišća-

vanja proteina od različitih onečišćenja uključujući i korištenih reagenasa za derivatizaciju.<sup>60,61</sup> Nedavno je opisana metoda gvanidinacije peptida izravno s pločice MALDI, prikladna za derivatizaciju vrlo malih količina analita.<sup>62</sup>

Derivatizacija i derivatizacijski reagensi služe i u svrhu relativne kvantifikacije pojedinih proteina. Različito izotopno obilježen derivatizacijski reagens koji specifično reagira s proteinom (N-kraj, lizin, cistein itd.)<sup>63</sup> reagira s jednakim obujmom više proteinskih otopina u kojima želimo uspojedno ustanoviti količinu pojedinog proteina prije proteolize (op. a. kod nekih reagensa derivatizacija se vrši nakon proteolize). Nakon proteolitičkog cijepanja smjesa peptida izravno se analizira tehnikom MALDI-TOF ili se razdvaja nekom od kromatografskih ili drugih separacijskih metoda te analizira spektrometrijom masa. S obzirom na veliko preklapanje signala u spektru masa kompleksne proteinske smjese, separacija na proteinske ili peptidne komponente je time poželjnija, čim je smjesa proteina kompleksnija. Relativni omjer intenziteta signala peptida istih aminokiselinskih sljedova, ali obilježenih reagensima koji se razlikuju u broju izotopa (npr. deuterij, <sup>13</sup>C, <sup>15</sup>N) mogu ukazati na različite količine određenog proteina unutar stanice u danom trenutku u odnosu na neku drugu stanicu, odnosno razvojni stupanj ili stres kojem je stanica bila podvrgnuta (npr. oksidativni stres).<sup>64,65</sup>

#### Identifikacija proteina fragmentiranjem peptida

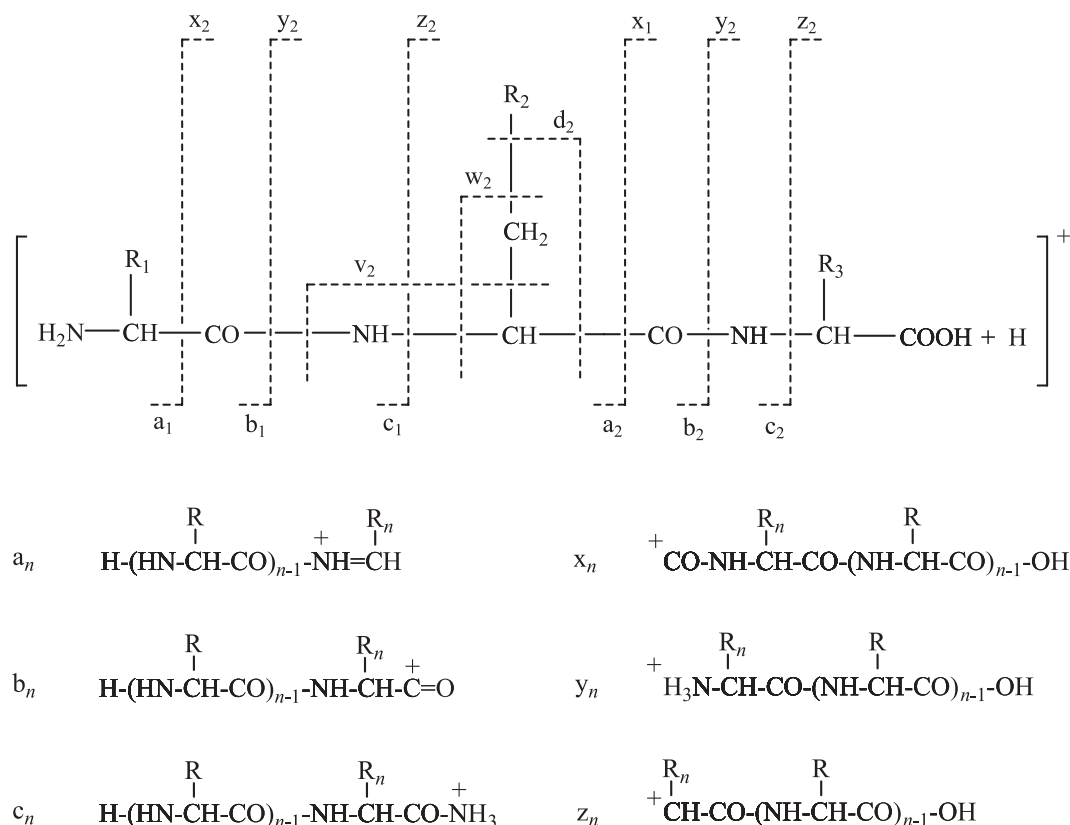
Fragmente proteina nastale cijepanjem proteinazama i određene spektrometrijom masa moguće je dalje fragmentirati tehnikama tandemne spektrometrije masa, čime se dobiva uvid u primarnu strukturu proteina. Tandemna spektrometrija masa može se odvijati u vremenu ili u prostoru, pa način fragmentacije ovisi o uvjetima ionizacije i vrsti instrumenta. U slučaju ionizacije elektroraspršenjem i uporabom spektrometra tipa Q-TOF ili QQQ odabranom ionu određene vrijednosti  $m/z$  dovodi se dodatna energija u posebnoj kolizijskoj ćeliji kako bi se izazvala fragmentacija. Postoje različite tehnike prijenosa energije, ali se najčešće rabi kolizijom inducirana disocijacija (*engl.* collision-induced dissociation, CID). MALDI-TOF tehnikom uz ion analita tijekom ionizacijskog procesa nastaje i određen broj metastabilnih iona, produkata rasapa iona prekursora.<sup>67</sup> Fragmentiranje iona prekursora može se dogoditi u ionskom izvoru (brza fragmentacija, *engl.* prompt fragmentation) ili po izlazu iona iz ionskog izvora u dijelu spektrometra masa koji prethodi TOF-u (poslijeionizacijska fragmentacija, *engl.* post source decay, PSD). Poslijeionizacijskom fragmentacijom nastaju fragmentni ioni iste brzine kao i ion prekursor, pa se uz linearan način snimanja ne opažaju u spektru masa. Moguće ih je razlikovati snimanjem uz reflektrom. Naime, zbog manje kinetičke energije ioni fragmenta ne ulaze jednako duboko u elektrostatsko polje reflektroma kao ioni prekursora, prije izlaze iz reflektroma i bilježe se na manjim vrijednostima  $m/z$  u odnosu na ion prekursora.<sup>67</sup> U tipičnom PSD-instrumentu potpuni spektar iona produkata postiže se snimanjem u nekoliko koraka. Reflektrom može analizirati usko područje energije, pa se potencijal reflektroma postupno mijenja. Snimanjem niza segmenata, od kojih je prvi fokusiran na ion prekursora, a svi ostali ka nižim masama iona produkata, dobiva se ukupni spektar masa. Kalibracija segmenata provodi se eksterno po istom segmentnom slijedu po kojem je analiziran analit.

Kalibrant mora imati molekulsku masu jednaku ili veću od analita i poželjno je da po svakom segmentu daje minimalno dvije kalibracijske točke. Nakon PSD-a u spektrima masa najčešće izostaju izotopni signali, a točnost određivanja masa je do 0,3 Da. U posljednjih pet godina tehnika MALDI-TOF analiza je unaprijeđena dodatkom još jednog analizatora masa, pa se tako ta tehnika naziva MALDI-TOF/TOF. Tandemna spektrometrija masa se kod analizatora TOF/TOF provodi selekcijom iona u prvom TOF-analizatoru, cijepanjem iona prekursora u kolizijskoj ćeliji, koja se nalazi između prvog i drugog analizatora te snimanjem nastalih fragmenata drugim TOF-analizatorom. Brzina rada, točnost mjerenih masa, kvaliteta MS/MS spektara i osjetljivost instrumenata MALDI-TOF/TOF uveliko nadmašuje performanse klasičnog instrumenta MALDI-TOF.<sup>68</sup>

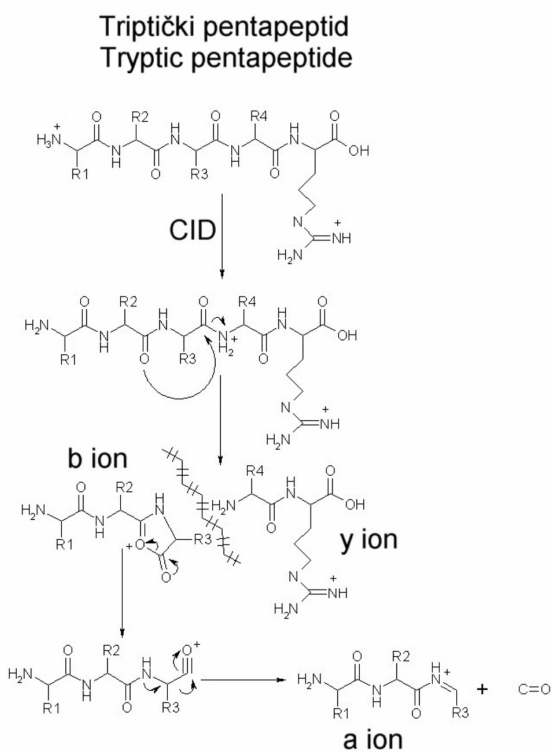
Ioni peptida nastali fragmentiranjem mogu se podijeliti u dvije osnovne skupine: 1) ioni koji su nastali cijepanjem jedne ili više veza okosnice peptidnog lanca i 2) ioni koji su nastali cijepanjem okosnice peptidnog lanca i cijepanjem bočnog ogranka.<sup>3</sup> Nomenklaturu nastalih fragmentnih iona predložili su *Roepstorff* i *Fohlman*<sup>69</sup> a kasnije ju je modificirao *Biemann*.<sup>70</sup> Okosnica peptidnog lanca može se cijepati na tri tipa veze C–C, C–N ili N–C, pri čemu nastaje šest vrsta fragmentnih iona. Fragmenti označeni kao  $a_n$ ,  $b_n$  i  $c_n$  imaju pozitivan naboj na N-kraju, dok fragmenti označeni s  $x_n$ ,  $y_n$  i  $z_n$  imaju pozitivan naboj na C-kraju peptida (slika 6). Fragmenti  $c_n$  i  $y_n$  uključuju prijenos dvaju protona, jednog odgovornog za protoniranje i drugog koji potječe s druge strane pocijepanog peptida. Indeks  $n$  označava broj aminokiselina u fragmentnom ionu. Razlika u masi između dva konsektivna iona iste serije omogućuje određivanje aminokiseline odnosno peptidnog aminokiselinskog slijeda. Izuzeci su leucin i izoleucin kao izomeri, odnosno glutamin i lizin kao mogući izobari (ovise o rezoluciji i točnosti mjerenja masa instrumenta). Ukoliko do fragmentiranja dolazi pri nižim energijama, u spektru masa obično se opažaju fragmentni ioni serije  $b$  i  $y$  (mehanizam nastajanja  $b$ -,  $y$ - i  $a$ -iona na primjeru cijepanja iona triptičkog pentapeptida prikazan je na slici 7).<sup>71</sup> Takvi fragmentni ioni mogu gubiti male molekule kao npr. vodu ili amonijak iz funkcionalnih skupina bočnih ogranka.<sup>3,72</sup>

Cijepanjem barem dvije veze peptidnog lanca nastaju dva tipa iona: imonijev ion i unutarnji fragment (*engl.* internal fragment). Unutarnji fragmenti obično sadrže tri do četiri aminokiseline i ne uključuju ishodni N- ili C-kraj. Označavaju se slovima koji iskazuju aminokiselinski slijed fragmentnog iona. Slabo su zastupljeni osim u slučaju peptida što sadrže prolin koji je zbog svoje bazičnosti više podložan protoniranju, a time i fragmentiranju, dok s druge strane prolinski prsten sterički ometa nastajanje C-krajnjih iona.<sup>73</sup> Imonijevi ioni ( $H_2N=CHR$ )<sup>+</sup> detektiraju se pri malim vrijednostima  $m/z$  i označavaju se slovom odgovarajuće aminokiseline.<sup>3</sup> Nastaju kao posljedica unutarnje fragmentacije kombinacijom  $a$ - i  $y$ -načina cijepanja iona.

U visokoenergetskim MS/MS spektrima uočena su tri dodatna tipa fragmenta  $d$ ,  $v$  i  $w$  koji nastaju cijepanjem okosnice peptidnog lanca i bočnog ogranka (slika 6). Ukoliko dolazi do cijepanja veze između  $\beta$ - i  $\gamma$ -C atoma bočnog ogranka, a naboj ostaje na N-kraju, nastaju  $d$ -ioni. Ako naboj ostaje na C-kraju, nastaju ioni tipa  $w$ . Fragmentni ioni  $d$  i  $w$  korisni su za razlikovanje izomera leucina i izoleucina. Nisu



Slika 6 – Fragmenti proteinskih iona dobiveni tehnikama CID ili PSD  
 Fig. 6 – Fragmentation paths of proteins in CID or PSD

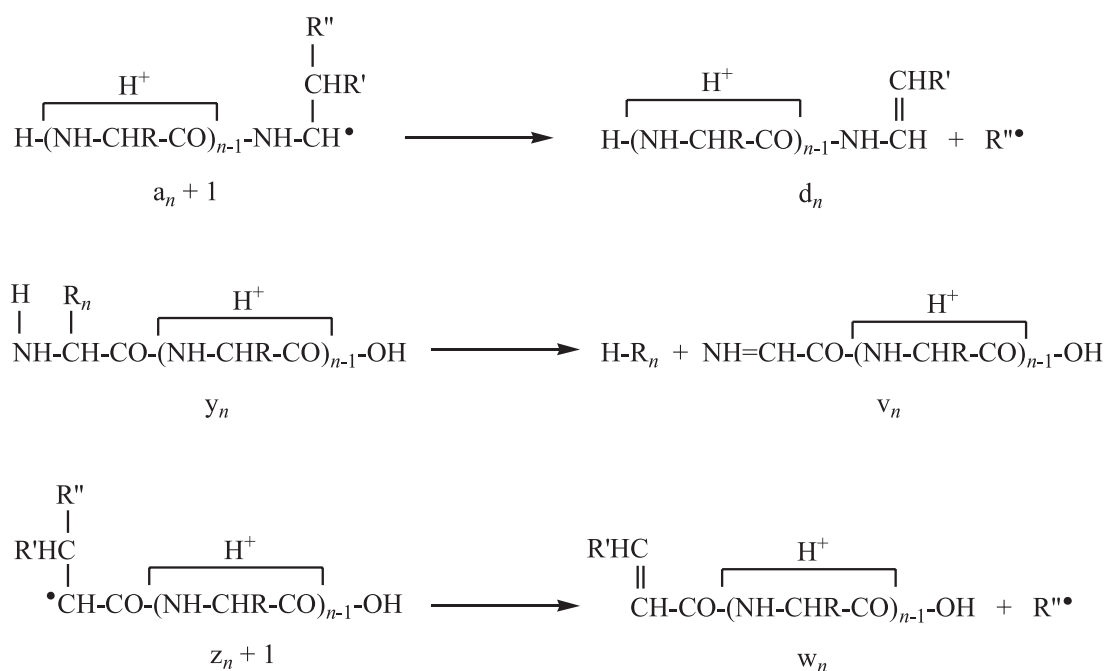


Slika 7 – Fragmentacija triptičkog dvostruko nabijenog pentapeptida na ione b i y kolizijom induciranom disocijacijom u spektrometru masa  
 Fig. 7 – Fragmentation of a doubly charged tryptic pentapeptide to b- and y-ions by collision-induced dissociation inside the mass spectrometer

uočeni kod aminokiselina koje sadrže aromatsku skupinu (His, Phe, Tyr, Trp). Ukoliko dolazi do potpunog gubitka bočnog lanca, a naboje ostaje na C-kraju, govorimo o fragmentu v. Takav tip fragmentnog iona znatnog je intenziteta za aminokiseline koje ne daju ione tipa w. Do sada nije uočen ekvivalentni ion s nabojem na N-kraju peptidnog lanca. Mehanizam nastajanja i strukture fragmentnih iona serije d, v, i w prikazani su slikom 8. Opisani fragmentni ioni nastaju i pri fragmentiranju višestruko protoniranih peptida. Svi poznati ioni nastali fragmentacijskim tehnikama MS/MS ili PSD prikazani su u tablici 3.

Na izgled spektra masa dobivenog uz CID ili PSD utječe položaj i delokalizacija naboja. Primjerice, u spektrima masa malih peptida koji ne sadrže bazične aminokiseline podjednako su zastupljeni ioni serije b i y. Povećanjem duljine peptidnog lanca povećava se prinos iona serije b. Bazične aminokiseline (Arg, Lys, His i Pro) utječu na proces fragmentiranja tako da induciraju nastajanje iona s C- ili N-krajem, ovisno u čijoj se blizini nalaze. Mehanizmi fragmentiranja protoniranih peptida sustavno su obrađeni u preglednom članku B. Paizs i S. Suhai.<sup>74</sup>

Iako MS/MS ili PSD spektri s mnoštvom signala fragmentnih iona sadrže veoma mnogo informacija o primarnoj strukturi proteina, iznimno ih je teško interpretirati. Da bi se olakšalo razlikovanje iona s N- odnosno C-krajem, primjenjuje se izotopno obilježavanje C- ili N-kraja. Primjerice, izotop kisika <sup>18</sup>O uvođa se tijekom digestije proteina u smjesi neobilježene i obilježene vode, čime se ne utječe na proces digestije i ne gubi na osjetljivosti. Dobiveni fragmenti sa C-krajem sadrže smjesu izotopa <sup>16</sup>O i <sup>18</sup>O, pa se u spektru



Slika 8 – Nastajanje iona serije d, v i w

Fig. 8 – Formation of d-, v- and w-ions

Tablica 3 – Fragmentni ioni nastali tehnikama MS/MS ili PSD  
 Table 3 – Fragment ions produced by MS/MS or PSD

Oznaka iona Denotation	Fragment Fragment	Oznaka iona Denotation	Fragment Fragment
a	N + M – CO	d	a i djelomični bočni ogranak a and partly side chain
a*	a–NH <sub>3</sub>	v	y i potpuni bočni ogranak y and side chain
a°	a–H <sub>2</sub> O	w	z i djelomični bočni ogranak z and partly side chain
a <sup>++</sup>	(a+H)/2	x	C + M + CO
b	N + M	y	C + M + H <sub>2</sub>
b*	b–NH <sub>3</sub>	y*	y–NH <sub>3</sub>
b°	b–H <sub>2</sub> O	y°	y–H <sub>2</sub> O
b <sup>++</sup>	(b+H)/2	y <sup>++</sup>	(y+H)/2
c	N + M + NH <sub>3</sub>	z	C + M – NH

N – grupa na N-kraju (N-terminal group)

M – bočni ogranak (amino acid residue)

C – grupa na C-kraju (C-terminal group)

masa razlikuju od neobilježenih iona s N-krajem peptidnog lanca.<sup>57</sup> *In vivo* obilježavanje staničnih proteina s deuteriranim lizinom (Lys-*d*<sub>4</sub>) predložili su S. Gu i suradnici.<sup>75</sup> Razlika u masi od 4 Da između obilježenog i neobilježenog lizina olakšava identifikaciju peptida koji sadrže lizin nakon digestije proteina. Gotovo istodobno S.-E. Ong razvija meto-

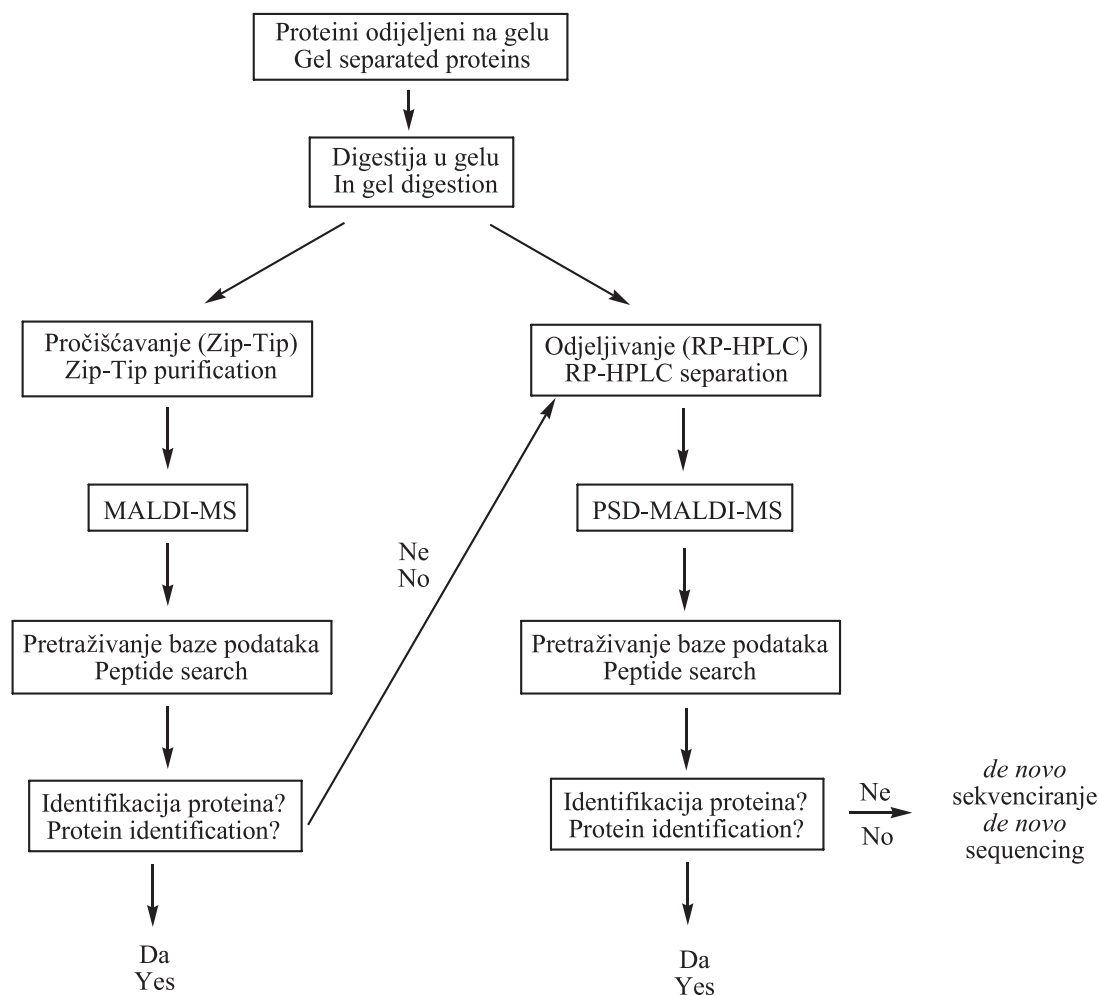
du uzgoja stanica na hranjivim podlogama koje sadrže izotopno obilježene aminokiseline u svrhu prethodno objašnjene relativne kvantifikacije proteina i naziva je SILAC (*engl.* stable isotope labeling by amino acids in cell culture).<sup>76</sup> Za razliku od kemijskih modifikacija peptida djelotvornost obilježavanja proteina unutar jedne stanične linije je 100 %. Nakon uzgoja stanica i ugradnje pojedine izotopom obilježene aminokiseline u pojedini protein, uspoređuju se i kvantificiraju spektri masa nakon proteolize "teških" (izotopno obilježenih) i "lakih" (izotopno neobilježenih) proteina.<sup>77</sup>

Kako bi se pojednostavio izgled spektra, odnosno preferirano dobili signali jedne serije iona, npr. ioni serije y, primjenjuje se kemijski aktivirana fragmentacija peptida koja omogućuje i *de novo* sekvenciranje peptida. Primjer postupka identifikacije proteina spektrometrijom masa shematski je prikazan slikom 9.

#### De novo sekvenciranje

Identifikacija proteina metodama PMF i PSD temelji se na uspoređivanju eksperimentalno dobivenih podataka s teorijski izračunatim masama u bazama podataka. Genomski slijed većine organizama je i dalje nepoznat, pa prema tome u bazama nisu zastupljeni podaci o odgovarajućim proteinima. Osim toga, ako i u bazama postoje podaci o poznatim proteinima različite modifikacije kao npr. posttranslacijske modifikacije, mogu onemogućiti identifikaciju dijela ili cijelog proteinskog aminokiselinskog slijeda. Zbog toga potpuno određivanje primarne strukture proteina zahtijeva određivanje aminokiselinskog slijeda s minimalnom uporabom baza podataka, odnosno *de novo* sekvenciranje peptida i proteina. Takvo određivanje temelji se na tandemnoj spektrometriji masa, MS/MS ili na PSD eksperimentima. Kako bi se olakšala interpretacija složenih spektara, peptidi se kemijski derivatiziraju prikladnim rea-



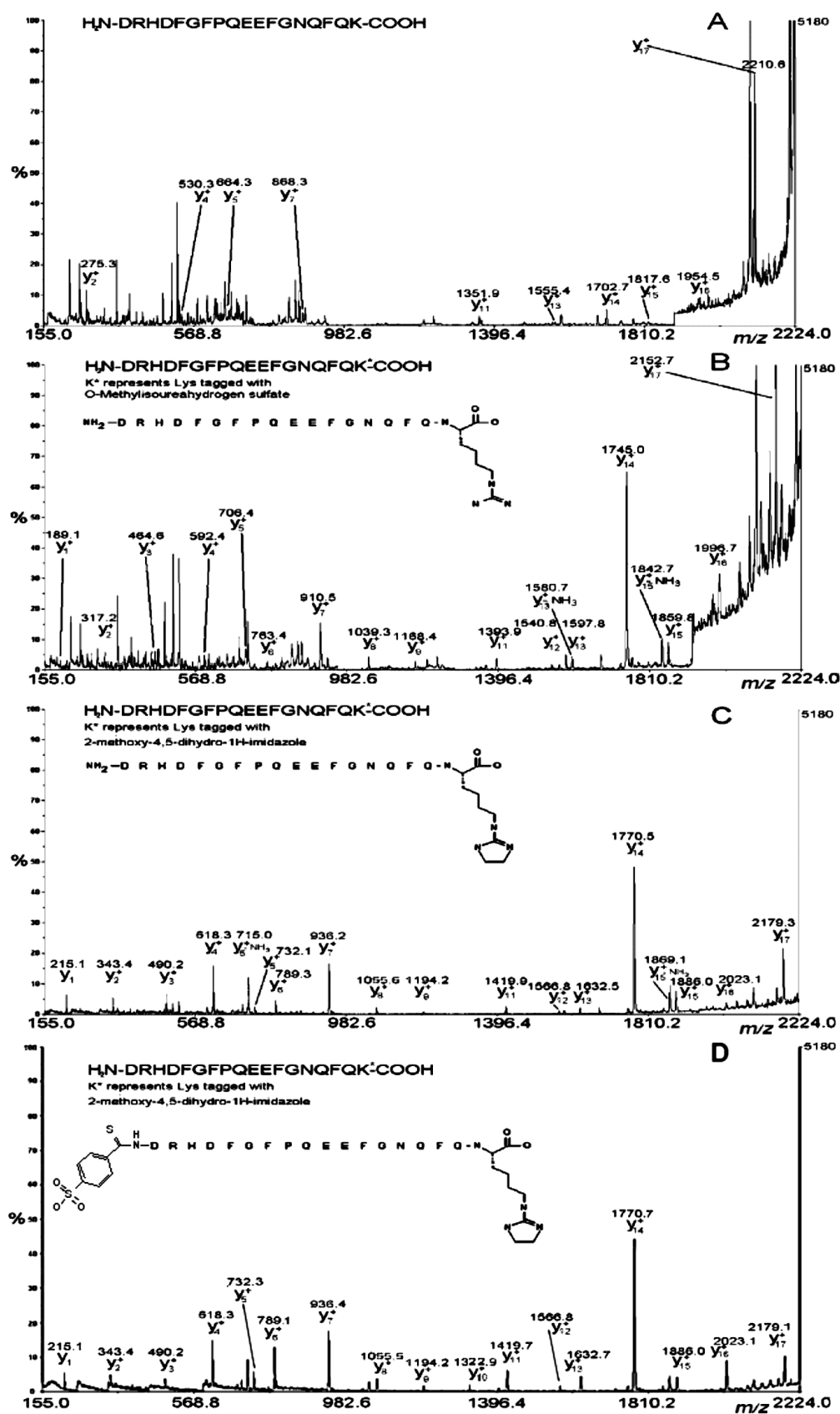


Slika 9 – Postupak identifikacije proteina odijeljenih na gelu spektrometrijom masa

Fig. 9 – Mass spectrometric identification route of gel-separated proteins

gensima koji gotovo isključivo ili pretežno daju jednu seriju fragmentnih iona. Iz razlike u masama između konsekvativnih signala određuje se slijed aminokiselina. Keoughova ideja o sulfoniranju peptida, odnosno uvođenju kiselinske skupine na N-kraj peptida pokazala se iznimno uspješnom.<sup>78–80</sup> Tako derivatiziran peptid koji na sebi nosi istodobno pozitivan i negativan naboj može se prikazati formulom  $^{-}O_3S-C_6H_5-A_1-A_2-A_3-A_4-A_5-X^{+}$ , gdje A označava bilo koju aminokiselinu, a X lizin ili arginin. Ako se radi o lizinskom peptidu,  $\epsilon$ -amino-skupina lizina se kemijski modificira kako bi se povećao prinos tih iona u plinovitoj fazi i ujedno omogućila selektivna sulfonacija N-kraja. Budući da je najbazičnija skupina peptida već protonirana, dodatni proton potreban za ionizaciju molekule nasumično odabire mjesto protoniranja. Na mjestu protoniranja dolazi do pucaanja peptidne veze i stvaranja iona produkata serije b i y.

Međutim, fragmenti koji sadrže  $SO_3^{-}$  neće biti vidljivi pri snimanju pozitivnih iona tako da se u spektru masa uočavaju jedino signali iona serije y. Najpoznatiji reagensi za sulfoniranje su 2-sulfobenzojeva kiselina, sulfo-NHS-estri (N-hidroksisukcinimid, NHS) i 4-sulfopenil-izotiocijanat koji u vremenski nezahtjevnoj reakciji (do 30 min) modificiraju peptid i čine ga pogodnim za sekvenciranje. Na slici 10 uspoređeni su PSD MALDI spektri nativnog, gvanidiranog i sulfoniranog peptida.<sup>81,82</sup> Složeni spektar nativnog peptida sadrži signale iona serije a i b, serije y i produkte unutarnje pregradnje. Povećanjem bazičnosti peptida povećava se i unutarnja energija potrebna za fragmentaciju, odnosno gvanidacijom se ne postiže poboljšano fragmentiranje. Međutim, u PSD spektru sulfoniranog homoargininskog peptida gotovo su jedino signali iona serije y, te je moguće *de novo* određivanje aminokiselinskog slijeda peptida.



Slika 10 – Poslijeionizacijski MALDI spektri masa rHuINF  $\alpha$ -2b iona triptičkog fragmenta T6-7: nederivatizirani,  $m/z$  2226,0 (A); s Lys\* derivatiziranim O-metiloureaom,  $m/z$  2268,0 (B); s Lys\* derivatiziranim 2-metoksi-4,5-dihidro-1H-imidazolom,  $m/z$  2294,0 (C); i s Lys\* derivatiziranim 2-metoksi-4,5-dihidro-1H-imidazolom i N-krajem derivatiziranim s 4-sulfofenil-izotiocijanatom,  $m/z$  2509,0 (D)

Fig. 10 – Post-source decay MALDI mass spectra of rHuINF  $\alpha$ -2b tryptic fragments: T6-7 ion,  $m/z$  2226.0 (A); T6-7-Lys\* tagged with O-methylisourea hydrogen sulfate,  $m/z$  2268.0 (B); T6-7-Lys\* tagged with 2-methoxy-4,5-dihydro-1H-imidazole,  $m/z$  2294.0 (C); and T6-7-Lys\* tagged with 2-methoxy-4,5-dihydro-1H-imidazole and sulfonated at N-terminus with 4-sulfophenylisothiocyanate,  $m/z$  2509.0 (D)

## ZAHVALA

Ovaj rad napravljen je u okviru znanstvenih projekata 119-1191342-1083 i 119-1191342-2960.

## Literatura:

## References:

1. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **246** (1989) 64.
2. M. Karas, F. Hillenkamp, *Anal. Chem.* **60** (1988) 2299.
3. E. De Hoffman, V. Stroobant, *Mass Spectrometry Principle and Applications*, Wiley, Chichester, 2002, str. 239-275.
4. T. J. Hill, D. Lafitte, P. J. Derrick, *Electrospray ionization mass spectrometry*, u S. E. Harding and B. Z. Chowdhry (uredn.), *Protein-Ligand Interactions: structure and spectroscopy*, Oxford University Press, Oxford, 2001, str. 311-346.
5. A. J. R. Heck, R. H.H. van den Heuvel, *Mass Spectrom. Rev.* **23** (2004) 368.
6. K. F. Geoghegan, M. A. Kelly, *Mass Spectrom. Rev.* **24** (2005) 347.
7. P. Kebarle, L. Tang, *Anal. Chem.* **65** (1993) 972A.
8. S. J. Gaskell, *J. Mass Spectrom.* **32** (1997) 677.
9. R. B. Cole, *J. Mass Spectrom.* **35** (2000) 763.
10. N. Galić, *Kem. Ind.* **53** (2004) 117.
11. R. Tenobi, R. Knochenmuss, *Mass Spectrom. Rev.* **17** (1998) 337.
12. M. Karas, M. Glückmann, J. Schäfer, *J. Mass Spectrom.* **35** (2000) 1.
13. M. Kussmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M. Rossel-Larsen, L. Jakobse, J. Gobom, E. Mirgorodskaya, A. Kroll-Kristen, L. Palm, P. Roepstorff, *J. Mass Spectrom.* **32** (1997) 593.
14. M. Rožman, *Arh. Hig. Rada Toksikol.* **54** (2003) 19.
15. S. Kazazić, S. Pečur, D. Srzić, *Kem. Ind.* **48** (1999) 181.
16. R. Willoughby, E. Sheehan, S. Mitrovich, *A Global view of LC/MS*, Global View Publishing, Pittsburg, Pennsylvania 2002, str. 93-109.
17. Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman, R. G. Cooks, *J. Mass Spectrom.* **40** (2005) 430.
18. S. Cristoni, L. R. Bernardi, *Mass Spectrom. Rev.* **22** (2003) 369.
19. C. L. Chepanoske, B. E. Richardson, M. Von Rechenberg, J. M. Peltier, *Rapid Commun. Mass Spectrom.* **19** (2005) 9.
20. J. Peng, S. P. Gygi, *J. Mass Spectrom.* **36** (2001) 1083.
21. A. J. Link, *Trends Biotechnol.* **20** (2002) S8.
22. A. J. Link, J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. Garvik, M. Yates JR 3<sup>rd</sup>, *Nat. Biotechnol.* **17** (1999) 676.
23. M. P. Washburn, D. Wolters, M. Yates JR 3<sup>rd</sup>, *J. Proteome Res.* **1** (2002) 242.
24. D. L. Tabb, W. H. McDonald, *Brief Funct. Genomic Proteomic* **5** (2006) 98.
25. T. Wehr, *LC GC* **24** (2006) 1004.
26. B. Bogdanov, R. D. Smith, *Mass Spectrom. Rev.* **24** (2005) 168.
27. G. Khalsa-Moyers, W. H. McDonald, *Brief Funct. Genomic Proteomic* **5** (2006) 98.
28. M. Cindrić, L. Bindila, T. Cepo, J. Peter-Katalinic, *J. Proteome Res.* **5** (2006) 3066.
29. J. Qu, W. Chen, G. Luo, Y. Wang, S. Xiao, Z. Ling, G. Chen, *Analyst.* **127** (2002) 66.
30. A. T. Iavarone, J. C. Jurchen, E. R. Williams, *J. Am. Soc. Mass Spectrom.* **11** (2000) 976.
31. M. Cindrić, M. Vuletić, *J. Sep. Sci.* **26** (2003) 1263.
32. K. Webb, T. Bristow, M. Sargent, B. Stein, *Methodology for accurate mass measurement of small molecules*, LGC Limited, Teddington UK, 2004, str. 1-15.
33. B. Domon, R. Aebersold, *Science* **312** (2006) 212.
34. J. Reinders, U. Lewandrowski, J. Moebius, Y. Wagner, A. Sickmann, *Proteomics* **4** (2004) 3686.
35. J. R. Yates, *J. Mass Spectrom.* **33** (1998) 1.
36. <http://www.rcsb.org/pdb>
37. <http://www.expasy.org>
38. B. Balen, M. Krsnik-Rasol, A. D. Zamfir, J. Milosevic, S. Y. Vakhrushev, J. Peter-Katalinic, *J. Proteome Res.* **5** (2006) 1658.
39. A. R. Blackler, A. A. Klammer, M. J. MacCoss, C. C. Wu, *Anal. Chem.* **78** (2006) 1337.
40. N. M. Luscombe, D. Greenbaum, M. Gerstein, *Methods Inf. Med.* **40** (2001) 346.
41. [http://psidev.sourceforge.net/ms/mzdata\\_implementers.html](http://psidev.sourceforge.net/ms/mzdata_implementers.html)
42. <http://prospector.ucsf.edu>
43. D. N. Perkins, D. J. C. Pappin, D. M. Creasy, J. S. Cottrell, *Electrophoresis* **74** (2002) 5593.
44. H. Wang, Shawn G. Clouthier, V. Galchev, D. E. Misek, U. Duffner, C.-Ki Min, R. Zhao, J. Tra, G. S. Omenn, J. L. M. Ferrara, S. M. Hanash, *Mol. Cell. Proteomics* **4** (2005) 618.
45. B. Kuster, M. Schirle, P. Mallick, R. Aebersold, *Nat. Rev. Mol. Cell. Biol.* **6** (2005) 577.
46. A. Gruhler, J. V. Olsen, S. Mohammed, P. Mortensen, N. J. Faergeman, M. Mann, O. N. Jensen, *Mol. Cell. Proteomics* **4** (2005) 310.
47. S. M. Wurgler-Murphy, D. M. King, P. J. Knelly, *Proteomics* **4** (2004) 1562.
48. F. G. Hanisch, M. Jovanovic, J. Peter-Katalinic, *Anal. Biochem.* **290** (2001) 47.
49. N. H. Packer, M. A. Lawson, D. R. Jardine, J. W. Redmond, *Glycoconjugate J.* **15** (1998) 737.
50. N. Farriol-Mathis, J. S. Garavelli, B. Boeckmann, S. Duvaud, El. Gasteiger, A. Gateaul, A. L. Veuthey, A. Bairoch, *Proteomics* **4** (2004) 1537.
51. A. Bohne-Lang, C. W. von der Lieth, *Nucleic Acids Res.* **33** (2005) W214.
52. <https://www.GlycoSuite.com>
53. T. H. Thanka Christlet, K. Veluraja, *Biophys J.* **80** (2001) 952.
54. S. J. Hubbard, *Biochim. Biophys. Acta* **1382** (1998) 191.
55. S. J. Hubbard, R. J. Beynon, J. M. Thornton, *Protein Eng.* **11** (1998) 349.
56. E. Krause, H. Wenschuh, P. R. Jungblut, *Anal. Chem.* **71** (1999) 4160.
57. K. G. Standing, *Curr. Opin. Struct. Biol.* **13** (2003) 595.
58. R. L. Beardsley, J. P. Reilly, *Anal. Chem.* **74** (2002) 1884.
59. E. C. Peters, D. M. Horn, D. C. Tully, A. Brock, *Rapid Commun. Mass Spectrom.* **15** (2001) 2387.
60. P. Chen, S. Nie, W. Mi, X. C. Wang, S. P. Liang, *Rapid Commun. Mass Spectrom.* **18** (2004) 191.
61. M. Cindrić, T. Čepo, A. Škrilin, M. Vuletić, L. Bindila, *Rapid Commun. Mass Spectrom.* **20** (2006) 694.
62. S. L. Cockrill, K. L. Foster, J. Wildsmith, A. R. Goodrich, J. G. Dapron, T. C. Hassell, W. K. Kappel, G. B. I. Scott, *BioTechniques* **38** (2005) 1.
63. S. E. Ong, M. Mann, *Nat. Chem. Biol.* **5** (2005) 252.
64. M. Sethuraman, M. E. McComb, T. Heibeck, C. E. Costello, R. A. Cohen, *Mol. Cell Proteomics* **3** (2004) 273.
65. B. Comuzzi, M. D. Sadar, *Cellscience* **3** (2006), 61.
66. P. Chaurand, F. Luetzenkirchen, B. Spengler, *J. Am. Soc. Mass Spectrom.* **10** (1999) 91.

67. B. Spengler, *J. Mass Spectrom.* **32** (1997) 1019.
68. K. F. Medzihradzky, J. M. Campbell, M. A. Baldwin, A. M. Fallick, P. Juhasz, M. L. Vestal, A. L. Burlingame, *Anal. Chem.* **72** (2000) 552.
69. P. Roepstorff, J. Fohlman, *J. Biomed. Mass spectrom.* **11** (19) 601.
70. K. Biemann, *Methods Enzymol.* **193** (1990) 886.
71. C. S. Lane, *Cell. Mol. Life Sci.* **62** (2005) 848.
72. R. Aebersold, D. R. Goodlett, *Chem. Rev.* **101** (2001) 269.
73. V. H. Wysocki, G. Tsaprailis, L. L. Smith, L. A. Breci, *J. Mass Spectrom.* **35** (2000) 1399.
74. B. Paizs, S. Suhai, *Mass Spectrom. Rev.* **24** (2005) 508.
75. S. Gu, S. Pan, E. M. Bradbury, X. Chen, *Anal. Chem.* **74** (2002) 5774.
76. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **1** (2002) 376.
77. C. P. Vaughn, D. K. Crockett, M. S. Lim, K. S. J. Elenitoba-Johnson, *Mol. Diagn.* **8** (2006) 513.
78. K. Gevaert, J. Vandekerckhove, *Electrophoresis* **21** (2000) 1145.
79. T. Keough, R. S. Youngquist, M. P. Lacey, *Anal. Chem.* **75** (2003) 157A.
80. T. Keough, M. P. Lacey, R. S. Youngquist, *Rapid Commun. Mass Spectrom.* **16** (2002) 1003.
81. T. Keough, M. P. Lacey, R. J. Strife, *Rapid Commun. Mass Spectrom.* **15** (2001) 2227.
82. T. Keough, M. P. Lacey, R. S. Youngquist, *Rapid Commun. Mass Spectrom.* **14** (2000) 2348.
- FT ICR – Analizator ionsko ciklotronske rezonancije uz Fourierovu transformaciju  
– Fourier transform ion cyclotron resonance
- IMAC – Metaloafinitetna kromatografija  
– Immobilized metal affinity chromatography
- LC – Tekućinska kromatografija  
– Liquid chromatography
- MALDI – Matricom pomognuta ionizacija uz desorpciju laserskim zračenjem  
– Matrix assisted laser desorption ionization
- MS – Spektrometrija masa  
– Mass spectrometry
- MS/MS – Tandemna spektrometrija masa  
– Tandem mass spectrometry
- MudPIT – Identifikacija proteina dvodimenzijском kromatografijom  
– Multidimensional protein identification technology
- PMF – Metoda otiska prsta  
– Peptide mass fingerprint
- PSD – Poslijeionizacijska fragmentacija  
– Post source decay
- Q-TOF – Hibrid kvadrupola i analizatora koji mjeri vrijeme leta  
– Hybrid quadrupole-time of flight
- QQQ – Trostruki kvadrupol  
– Triple quadrupole
- rHuINF – Rekombinantni humani interferon  
– Recombinant human interferon
- SCX – Jaki kationski izmjenjivač  
– Strong cation exchange
- SILAC – Obilježavanje proteina u kulturi stanica izotopno obilježenim aminokiselinama  
– Stable isotope labeling by amino acids in cell culture
- TFA – Trifluoroctena kiselina  
– Trifluoroacetic acid
- TOF – Analizator koji mjeri vrijeme leta  
– Time-of-flight

### Popis kratice Abbreviations

- CID – Kolizijom inducirana disocijacija  
– Collision-induced dissociation
- ESI – Elektroraspršenje  
– Electrospray

## SUMMARY

## Protein Analysis by Mass Spectrometry

N. Galić<sup>a</sup> and M. Cindrić<sup>b</sup>

Soft ionization techniques, electrospray (ESI) and matrix-assisted laser desorption/ionization (MALDI) make the analysis of biomolecules by mass spectrometry (MS) possible. MS is used for determination of the molecular weight of peptides and protein, sequence analysis, characterization of protein-ligand interactions etc. The detection limit, resolution and mass accuracy depend on instrument used (Table 1). Impurities (buffers, salts, detergents) can reduce the ion intensities or even totally suppress them, so a separation method (chromatography, 2D-gel electrophoresis) must be used for purification of the sample.

Molecular mass of intact protein can be determined by ESI or MALDI MS. Multiply charged ions are produced by ESI MS, while singly charged ions are predominant in MALDI spectra (Fig. 2).

Sequence analysis of proteins by MS can be performed using peptide mass fingerprint. In this method, proteins are separated by 2-D gel electrophoresis and digested with specific protease (Table 2) or digested and then separated by two-dimensional chromatography (Fig. 1). The obtained peptide mixtures are analyzed by MS or MALDI-TOF technique. The masses determined by MS are compared with calculated masses from database entries. Different algorithms have been developed for protein identification. Example of posttranslational modifications (*N*- and *O*-glycosylation) and protein sequence complex analysis after dual digestion (endoproteinase digestion followed by endoglycosidase digestion) is shown in Fig. 3.

It is known that detection of peptides by MS is influenced by intrinsic properties like amino acid composition, the basicity of the C-terminal amino acid, hydrophobicity, etc. Arginine-containing peptides dominate in MS spectra of tryptic digest, so the chemical derivatization of lysine terminal residue by *O*-methyloisourea or 2-methoxy-4,5-1*H*-imidazole was suggested (Fig. 4).

The peptide mass fingerprint method can be improved further by peptide fragmentation using tandem mass spectrometry (collision-induced dissociation, CID or post-source decay, PSD). Different types of fragments and main fragmentation paths are shown in Figs. 6 and 8. Fragmentation pathway of a doubly charged tryptic pentapeptide to *b*- and *y*-ions by collision-induced dissociation inside the mass spectrometer is described more in details in Fig. 7. All types of fragment ions are summarized in table 3. Since the any of the peptide bonds can be broken in several ways, the MS/MS spectra are complex, and quite difficult to interpret. Chemical derivatization is used to obtain only or predominantly one type of fragment ions. Sulfonation of *N*-terminal amino group enhance PSD sequencing, producing mainly *y*-type fragment ions. The mass difference of two consecutive *y*-ions corresponds to an amino acid mass, so the peptide sequence can be obtained with minimal or no assistance from genomic data, e. g. *de novo* protein sequencing is possible. Fig. 9 represents the strategy for the protein identification by mass spectrometry. Various chemical modifications on the peptide fragmentation patterns are shown in Fig. 10.

<sup>a</sup> Laboratory of Analytical Chemistry, Faculty of Science,  
University of Zagreb, Horvatovac 102a, 10 000 Zagreb, Croatia

Received February 22, 2007

Accepted June 21, 2007

<sup>b</sup> Mario Cindrić, Pliva – Research & Development,  
Prilaz baruna Filipovića 29, 10 000 Zagreb, Croatia