

VYUŽITÍ NESPECIFICKÝCH PEPTIDŮ PRO PROTEOMICKOU IDENTIFIKACI NÍZKOMOLEKULÁRNÍCH PROTEINŮ Z JEČMENE MALDI-TOF HMOTNOSTNÍ SPEKTROMETRIÍ

PAVEL ŘEHULKA^a, GÜNTER ALLMAIER^b
a JOSEF CHMELÍK^{a,c}

^aÚstav analytické chemie AV ČR, Veveří 97, 611 42 Brno,
^bTechnická univerzita, Getreidemarkt 9/164, A-1060, Vídeň, Rakousko, ^cVýzkumné centrum pro studium obsahových látek ječmene a chmele, Ústav analytické chemie AV ČR, Veveří 97, 611 42 Brno
rehulka@iach.cz, guenter.allmaier@tuwien.ac.at,
chmelik@iach.cz

Došlo 12.9.05, přijato 11.12.05.

Klíčová slova: nespecifické peptidy, proteiny, proteomika, ječmen, MALDI-TOF/TOF

Úvod

Proteomická identifikace nízkomolekulárních proteinů pomocí tradičního přístupu je komplikována několika skutečnostmi. Rozlišení těchto proteinů separovaných gelovou elektroforézou je obecně malé, což má za následek společnou migraci více složek proteinové směsi, které mají podobnou molekulovou hmotnost. Tyto proteiny se pak ve výsledném jednorozměrném (1-D) gelu vyskytují na stejném místě. Dalším problémem identifikace malých proteinů kombinací enzymového štěpení a hmotnostní spektrometrie (MS) je skutečnost, že tyto proteiny zpravidla poskytují pouze několik specifických peptidů – obzvláště v případě, je-li štěpení provedeno přímo v gelu a při malých množstvích vzorku¹. Výsledné hmotnostní spektrum směsi obsahuje nejen signály peptidů z identifikovaného proteinu, dalších proteinů přítomných ve vzorku, ale také autolyzátů štěpícího enzymu, peptidů keratinů, nespecifických peptidů a jiných příměsí (např. matricových klastrů při analýze metodou MALDI-TOF MS, cit.²). Za této situace analyzované proteiny zpravidla nelze identifikovat pomocí často používané metody „peptide mass fingerprinting“ (PMF), která využívá měření relativní molekulové hmotnosti peptidů a porovnání experimentálních dat s údaji v databázích proteinů^{3,4}.

Vzhledem ke složitosti peptidové směsi v enzymovém digestu nestačí malý počet specifických peptidů k jednoznačné identifikaci proteinu. Mimo tyto specifické peptidy (odpovídající očekávanému štěpení daným enzymem) však rovněž i některé další složky směsi pocházejí

z hledaného proteinu a mohly vzniknout při přípravě vzorku buď působením chemikálií nebo neobvyklou aktivitou proteolytického enzymu.

Specifické peptidy uvolněné působením trypsinu mají na *N*-konci jakoukoli aminokyselinu s výjimkou prolinu (nejedná-li se o *N*-koncový peptid) a na *C*-konci lysin nebo arginin (s výjimkou jakékoli aminokyseliny na *C*-konci proteinu). Aminokyselinou předcházející *N*-konci jakékoli tryptického peptidu (s výjimkou *N*-konce proteinu) v sekvenci proteinu je buď lysin nebo arginin. Ostatní peptidy pocházející ze stejného proteinu nazýváme nespecifické, které se nehodí k identifikaci pomocí PMF, ale mají uplatnění v metodách využívajících fragmentace peptidů pro stanovení sekvence, která je podkladem pro databázovou identifikaci. To znamená, že nespecifické peptidy, které při PMF představují vážný problém, lze úspěšně využít pro identifikaci proteinů cestou fragmentačních technik. Kombinace kapalínové chromatografie s tandemovou hmotnostní spektrometrií (MS/MS) na bázi ionizace elektrosprejem⁵ nebo MALDI-TOF MS doplněná technikou „post-source decay (PSD)“⁶ byly úspěšně využity pro identifikaci proteinů pomocí nespecifických peptidů^{5,6}.

Podrobná analýza proteomu ječmene je důležitá pro jeho efektivní využití v zemědělství, potravinářském průmyslu a zejména v pivovarnictví. Získané poznatky pomohou pochopit jeho biologické a technologické vlastnosti (odolnost vůči suchu, schopnost enzymového odbourávání škrobu, kvalitu pěny atd.).

První identifikace ječných proteinů pomocí 1-D gelové elektroforézy a MS publikovali Chmelík a spol.⁷ Pro zjednodušení směsi proteinů extrahovaných z ječných obiliek byl využit Osbornův systém rozpouštědel⁸, který tyto proteiny rozděluje do čtyř skupin: vodou extrahovatelné albuminy, globuliny extrahovatelné roztokem NaCl, hordeiny rozpustné ve vodných roztocích alkoholů a gluteliny extrahovatelné roztokem NaOH. Tento způsob extrakce zjednodušil směsi proteinů do té míry, že pro jejich další separaci nebylo nutné použít dvourozměrnou (2-D) gelovou elektroforézu. Vhodnou alternativou k elektroforéze jsou chromatografické metody, z nichž byla úspěšně využita pro separaci proteinů např. gelová chromatografie⁹, která obdobně jako výše uvedené elektroforetické metody separuje proteiny podle jejich velikosti. 2-D elektroforéza byla ale nezastupitelná, pokud proběhla extrakce proteinů jediným pufrům¹⁰. Větší rozlišení 2-D elektroforézy umožňuje rozdělit i velmi podobné proteiny, což v případě, že je rozdílná migrace způsobena posttranslačními modifikacemi, vede k tomu, že jeden a ten samý protein je identifikován na různých místech gelu. Rozlišení jednotlivých forem vyžaduje nalezení a charakterizaci modifikovaných peptidů, což je mnohem obtížnější než identifikace proteinu a v podstatě neproveditelné pomocí PMF (cit.¹¹). V případě selhání PMF je nutné využití fragmentace peptidů tandemovou hmotnostní spektrometrií. Je-li k dispozici vhodně vybavený hmotnostní spektrometr např. MALDI-TOF/TOF, lze MS/MS experiment provést se stejným vzorkem, který byl původně použit pro PMF analýzu¹².

Jednou z vhodných technik fragmentace peptidů je PSD. Tato technika je dostatečně citlivá a ve spojení s „curved-field“ reflektorem (CFR)¹³, který fokusuje současně všechny fragmentové ionty, i rychlá. Její nevýhodou je tvorba různých interních fragmentů, což částečně komplikuje interpretaci hmotnostního spektra, ve kterém se nejvíce vyskytují a-, b- a y-ionty, čímž se PSD podobá nízkoenergetické kolizně indukované disociaci. V případě tryptických peptidů je nevýhodou lokalizace mobilního protonu na bazických aminokyselinách (Lys a Arg), což snižuje účinnost fragmentace a vede k nízkému pokrytí sekvence peptidů¹⁴. K získání vhodných peptidů lze použít řady méně specifických enzymů např. chymotrypsinu, jehož výhodou je, že je aktivní za stejných podmínek jako trypsin¹⁵.

Cílem této práce bylo identifikovat nízkomolekulární proteiny extrahované z obilek ječmene MS/MS analýzou s využitím nespecifických peptidů.

Experimentální část

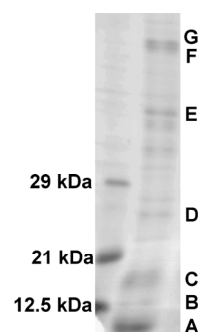
Jako vzorek pro tuto studii byla vybrána odrůda ječmene (*Hordeum vulgare* L.) Monaco. Rozemletá ječná zrna (1 g) byla extrahována 10 ml deionizované vody po dobu 20 min za stálého třepání. Směs byla centrifugována 15 min při 7000 g. Supernatant (vodný extrakt) byl smíchán se vzorkovým pufrům pro gelovou elektroforézu (tj. 50 mM Tris-HCl, pH 6,8, 4% SDS, 12% glycerol, 2% β -merkaptoethanol, 0,01% bromfenolová modř) v objemovém poměru 1:1. Tato směs byla povařena 5 min na vodní lázni a alikvot (20 μ l) byl aplikován na diskontinuální 1-D polyakrylamidový gel (6% koncentrační gel, 20% separační gel, velikost 150 mm \times 150 mm \times 1 mm). Vizualizace proteinů byla provedena pomocí Coomassie Brilliant Blue R-250. Fixace proběhla ve směsi 45,4% methanol / 4,6% octová kyselina / 50% deionizovaná voda (1 h), barvení ve směsi 45,4% methanol / 4,6% octová kyselina / 49,9% deionizovaná voda / 0,1% Coomassie Brilliant Blue R-250 (1 h) a odbarvení pozadí gelu bylo dosaženo ve směsi 5% methanol / 7,5% octová kyselina / 87,5% deionizovaná voda (24 h). Po omytí gelu v deionizované vodě (2 \times ; 10 min) byly vyřezány obarvené proužky gelu. Proteiny v nich obsažené byly přímo v gelu redukovány, alkylovány a enzymově štěpeny pomocí trypsinu¹⁶. Ječné proteiny v gelu byly štěpeny se specifickým hovězím trypsinem (Trypsin sequencing grade, Roche, Mannheim, SRN), kdežto v případě analýzy extraktu β -amylasy z ječmene (Sigma-Aldrich, Steinheim, SRN) byl použit trypsin s vysokou nespecifickou štěpnou aktivitou (Lachema, Brno, ČR).

MALDI-TOF hmotnostní spektra byla měřena pomocí Kompact MALDI SEQ a AXIMA CFR přístrojů (Shimadzu Biotech Kratos Analytical, Manchester, UK). Oba přístroje jsou vybaveny dusíkovým laserem ($\lambda=337$ nm, 3 ns délka pulsu). Urychlovací napětí bylo 20 kV a data byla naměřena za použití „time-delayed“

extrakce v reflektromovém módu. Oba přístroje jsou vybaveny reflektorem CFR umožňujícím současnou fokusaci všech fragmentů v PSD analýze. Šířka iontové brány (při analýze s AXIMA CFR přístrojem) byla nastavena na ± 10 Th pro výběr prekurzorového iontu. Pro získání celkového spektra bylo v MS módu akumulováno 100 spekter a v PSD módu 1000 spekter. Všechna MALDI spektra byla vyhlazena pomocí algoritmu Savitzky-Golay dodaného výrobcem a kalibrována interně (PMF experimenty) nebo externě (PSD měření). MALDI-TOF/TOF MS měření byla provedena na přístroji 4700 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA) vybaveném Nd:YAG laserem ($\lambda=355$ nm, 500 ps šířka pulsu). Při MS/MS analýze byl jako kolizní plyn použit vzduch.

Vzorek byl nanášen na matici technikou thin-layer. Nasycený roztok kyseliny 4-hydroxy- α -kyanoskořicové (Sigma-Aldrich) v acetonu byl smíchán s roztokem nitrocelulosity (10 mg ml⁻¹, aceton:isopropanol = 1:1, v/v) v objemovém poměru 4:1 a 0,5 μ l této směsi bylo nanášeno na MALDI terčík. Lyofilizované extrakty byly rozpuštěny v 20 μ l 1% kyseliny trifluoroctové, přečištěny pomocí ZipTip C₁₈ špiček (Millipore, Framingham, MA, USA) a eluovány přímo (asi 1 μ l eluentu) na suchou vrstvou matrice.

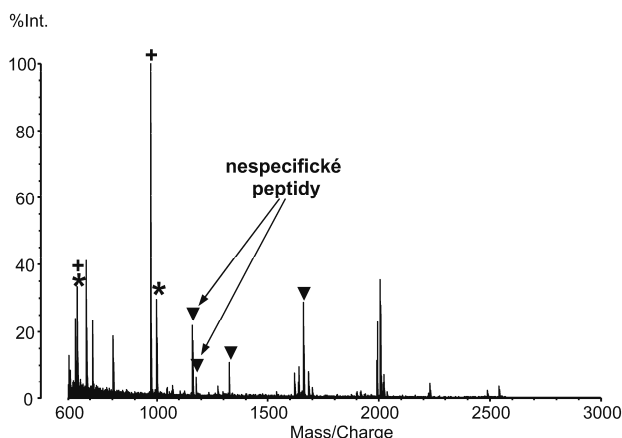
Naměřená data byla interpretována programy v internetovém souboru Protein Prospector (<http://prospector.ucsf.edu/>). Program MS-Fit byl použit pro vyhodnocení PMF experimentů, programy MS-Tag a MS-Pattern byly použity pro vyhodnocování PSD experimentů. K vyhledávání posloužily databáze SwissProt a NCBI GenBank.



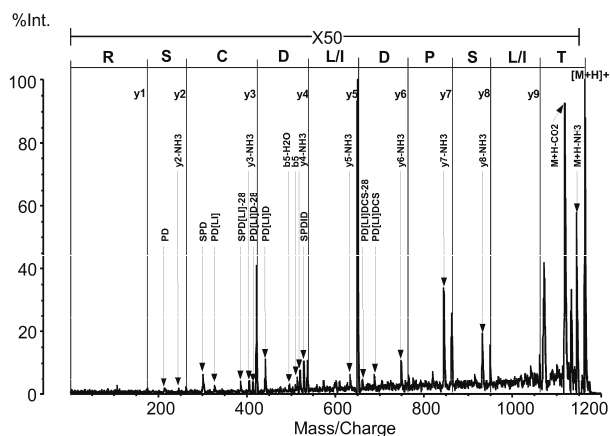
Obr. 1. 1-D SDS-PAGE elektroferogram směsi proteinů ječmene rozpustných ve vodě; označené proužky obsahují následující identifikované proteiny: A – subtilisin-chymotrypsin inhibitor CI-1A, subtilisin-chymotrypsin inhibitor CI-1B, prekurzor nespecifického lipid-transfer proteinu 1; B – trypsin/amylasa inhibitor pUP13; C – alfa-amylasa/trypsin inhibitor CMd prekurzor, alfa-amylasa inhibitor BMAI-1 prekurzor; D – cytosolová triosofosfát isomerasa (TIM); E – glyceraldehyd-3-fosfátdehydrogenasa, cytosolový protein Z (Z4) (hlavní endospermový albumin); F – beta-glukosidasa BGQ60 prekurzor; G – beta-amylasa

Výsledky a diskuse

Extrahované proteiny z obilky ječmene byly separovány 1-D gelovou elektroforézou (obr. 1). Zatímco vybrané proteiny o vyšší molekulové hmotnosti byly identifikovány metodou PMF, v případě nízkomolekulárních proteinů z proužku A tato metoda nebyla úspěšná. Hmotnostní spektrum peptidů získaných trypsinovým štěpením proteinů obsažených v proužku A je znázorněno na obr. 2. Několik peptidů z této směsi bylo fragmentováno tandemovou hmotnostní spektrometrií (MS/MS). Interpretace PSD



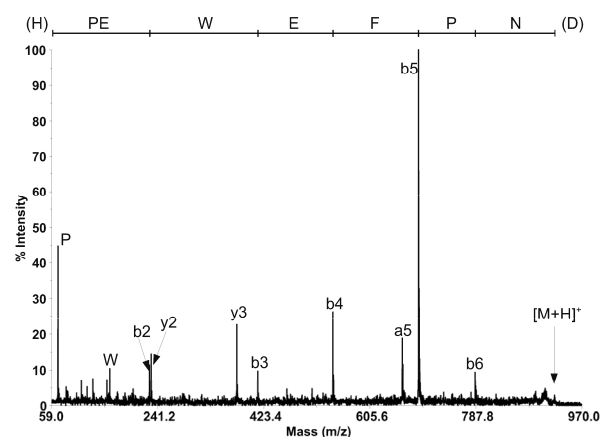
Obr. 2. MALDI-TOF hmotnostní spektrum in-gel digestu z proužku A v pozitivním reflektorném módu; + označuje peptidy ze subtilisin-chymotrypsin inhibitoru CI-1A, * označuje peptidy ze subtilisin-chymotrypsin inhibitoru CI-1B a ▼ – označuje obecné peptidy z prekursoru nespecifického lipid-transfer proteinu 1. Z nich peptidy označené šipkou jsou nespecifické peptidy z prekursoru nespecifického lipid-transfer proteinu 1



Obr. 3. PSD spektrum nespecifického peptidu (rodičovský iont $[M+H]^+$ o $m/z = 1163,54$) z prekursoru nespecifického lipid-transfer proteinu 1; primární struktura identifikovaného peptidu je zobrazena v horní části obrázku od C-konce směrem k N-konci (koresponduje s y-ionty přítomnými ve spektru)

spekter iontů s m/z 973,59 a 1001,66 Th poskytla krátké úseky peptidových sekvencí, které byly identifikovány pomocí MS-Pattern jako součásti dvou proteinů: subtilisin-chymotrypsin inhibitoru CI-1A (8882 Da, teoretický pI 5,24) a subtilisin-chymotrypsin inhibitoru CI-1B (8963 Da, teoretický pI 5,33). Z molekulových hmotností a isoelektrických bodů je zřejmé, že tyto dva proteiny by nebylo možné efektivně separovat ani při použití konvenčních gelů pro 2-D gelovou elektroforézu.

PSD analýza peptidu o m/z 1163,54 Th ukázala dostatečnou fragmentaci s kompletní sérií y-iontů, doplněnou některými b-ionty a množstvím interních peptidových fragmentů (obr. 3). To umožnilo jednoznačné určení sekvence tohoto peptidu (s výjimkou rozlišení leucinu a isoleucinu). Program MS-Pattern ale nenalezl žádný protein, v němž by se vyskytoval tryptický peptid této sekvence. Protože při analýze trypsinového digestu jiného extraktu z ječmene se nám podařilo identifikovat β -amylasu na základě fragmentace jejího nespecifického peptidu pomocí MALDI-TOF/TOF (viz obr. 4), byla specifita enzymu v programu MS-Pattern změněna z nastavení „Trypsin“ na „No enzyme“. Výsledkem byla jednoznačná identifikace prekursoru nespecifického lipid-transfer proteinu 1 (12,3 kDa). Stejný protein byl identifikován i při nastavení specifity enzymu na „Slymotrypsin FYWKR“, což odpovídá pomyslnému spojení enzymových aktivit trypsinu a chymotrypsinu. Pomocí teoretického (in silico) štěpení identifikovaného proteinu směsí obou enzymů bylo zjištěno, že hmotnostní spektrum experimentálního digestu obsahuje ještě tři další peptidy, jejichž molekulové hmotnosti odpovídají peptidům identifikovaného proteinu (na obr. 2 jsou označeny ▼). Dva z nich odpovídají teoretickým produktům trypsinového štěpení, třetí odpovídá molekulové hmotnosti fragmentovaného nespecifického peptidu s cysteinem modifikovaným nezreagovaným akrylamidem z použitého gelu.



Obr. 4. MALDI-TOF/TOF fragmentační spektrum nespecifického peptidu (rodičovský iont $[M+H]^+$ o $m/z = 918,388$) z β -amylasy; primární struktura identifikovaného peptidu je zobrazena v horní části obrázku od N-konce směrem k C-konci (koresponduje s b-ionty přítomnými ve spektru); P a W označují imoniové ionty prolinu a tryptofanu

MARAQVLLMA AALVLMITAA PRAAVALNCG QVDSKMKPCL TVVQGGPGPS
GECCNGVRDL HMQAQSSEDR QTVCNCLKGI ARGIHNLNLN NAASTPSKCN
VNVVPTTISPD IDCSTIY

Obr. 5. Primární struktura prekurzoru nespecifického lipid-transfer proteinu 1 s vyznačeným signálním peptidem (přerušovaně podtržený), identifikovanými tryptickými peptidy (podtrženými jednoduchou čarou) a identifikovaným nespecifickým peptidem (dvojitě podtržený)

Molekulová hmotnost identifikovaného prekurzoru nespecifického lipid-transfer proteinu 1 (12,3 kDa), která byla získána z databáze SwissProt, neodpovídá nalezené poloze na gelu (přibližně 9 kDa). Z databáze vyplývá, že příslušná aminokyselinová sekvence obsahuje signální peptid v pozicích 1–26 (celá primární struktura tohoto proteinu je uvedena na obr. 5). Hmotnostní spektrum na obr. 2 neobsahuje žádné hodnoty m/z, které by odpovídaly signálnímu peptidu, což potvrzuje předpoklad, že nativní protein postrádá signální peptid. Molekulová hmotnost proteinu bez signálního peptidu (9,7 kDa) odpovídá pozorované poloze na gelu. Na obr. 4 jsou rovněž označeny identifikované peptidy. Pokrytí primární struktury (bez signálního peptidu) dvěma tryptickými peptidy je 31 %, zahrnutím nespecifického peptidu se pokrytí zvýší na 42 %.

Získané výsledky jasně ukazují význam MS/MS experimentů pro identifikaci proteinů a to zejména pokud jsou v nich zahrnuty nespecifické peptidy, které nelze využít u metody PMF¹⁷. Od roku 2006 se nejvýznamnější proteomické časopisy rozhodly nepublikovat až na výjimky identifikace proteinů založené pouze na PMF. Analýza nespecifických peptidů je velmi významná zejména v případě nízkomolekulárních proteinů, jejichž enzymovým štěpením vzniká relativně málo specifických peptidů, což výrazně snižuje úspěšnost PMF. Efektivní fragmentace nespecifických peptidů rovněž zvyšuje pokrytí analyzované primární struktury, což je velmi důležité v případě studia biologických vlastností proteinů. Zatímco použití specificky štěpících enzymů¹⁸ je esenciálním předpokladem pro identifikaci proteinů pomocí PMF metody, v případě využití tandemové hmotnostní spektrometrie (např. techniky MALDI-TOF/TOF) přináší použití méně specifických enzymů další výhody při identifikaci a strukturní analýze proteinů¹⁹.

Tato práce byla podpořena prostředky Centra pro studium obsahových látek ječmene a chmele, IM6215648902 (MŠMT).

LITERATURA

1. Mattow J., Jungblut P. R., Müller E. C., Kaufmann S. H. E.: *Proteomics* 1, 494 (2001).
2. Karty J. A., Ireland M. M. E., Brun Y. V., Reilly J. P.: *J. Chromatogr.*, B 782, 363 (2002).
3. Joubert-Caron R., Le Caër J. P., Montandon F., Poirier F., Pontet M., Imam N., Feuillard J., Bladier D., Rossier J., Caron M.: *Electrophoresis* 21, 2566 (2000).
4. Lim H., Eng J., Yates J. R., Tollaksen S. L., Giometti C. S., Holden J. F., Adams M. W. W., Reich C. I., Olsen G. J., Hays L. G.: *J. Am. Soc. Mass Spectrom.* 14, 957 (2003).
5. König S., Zeller M., Peter-Katalinic J., Roth J., Sorg C., Vogl T.: *J. Am. Soc. Mass Spectrom.* 12, 1180 (2001).
6. Řehulka P., Chmelík J., Allmaier G.: *Rapid Commun. Mass Spectrom.* 19, 79 (2005).
7. Chmelík J., Řehulka P., Mayrhofer C., Allmaier G.: *Kvasný Průmysl* 47, 159 (2001).
8. Chmelík J., Řehulka P., Střelcová M., Kubáň V., Mayrhofer C., Allmaier G.: *Rostl. Výroba* 48, 261 (2002).
9. Šalplachta J., Řehulka P., Chmelík J.: *J. Mass Spectrom.* 39, 1395 (2004).
10. Finnie C., Melchior S., Roepstorff P., Svensson B.: *Plant Physiol.* 129, 1308 (2002).
11. Henzel W. J., Billeci T. M., Stults J. T., Wong S. C., Grimley C., Watanabe C.: *Proc. Natl. Acad. Sci. U. S. A.* 90, 5011 (1993).
12. Medzihradsky K. F., Campbell J. M., Baldwin M. A., Falick A. M., Juhasz P., Vestal M. L., Burlingame A. L.: *Anal. Chem.* 72, 552 (2000).
13. Cornish T. J., Cotter R. J.: *Rapid Commun. Mass Spectrom.* 7, 1037 (1993).
14. Kinter M., Sherman N. E.: *Protein Sequencing and Identification Using Tandem Mass Spectrometry*, str. 79. John Wiley & Sons, New York 2000.
15. Šalplachta J., Marchetti M., Chmelík J., Allmaier G.: *Rapid Commun. Mass Spectrom.* 19, 2725 (2005).
16. Jensen O. N., Wilm M., Shevchenko A., Mann M., v knize: *2-D Proteome Analysis Protocols* (Link A. J., ed.), kap. 52. Humana Press, Totowa 1999.
17. Chmelík J.: *Chem. Listy* 99, 883 (2005).
18. Stosová T., Havliš J., Lenobel R., Šebela M.: *Chem. Listy* 99, 896 (2005).
19. MacCoss M. J., McDonald W. H., Saraf A., Sadygov R., Clark J. M., Tasto J. J., Gould K. L., Wolters D., Washburn M., Weiss A., Clark J. I., Yates J. R.: *Proc. Natl. Acad. Sci. U. S. A.* 99, 7900 (2002).

P. Řehulka^a, G. Allmaier^b, and J. Chmelík^c
(^a *Institute of Analytical Chemistry, Academy of Sciences of the Czech Republic, Brno, Czech Republic,* ^b *Vienna University of Technology, Vienna, Austria,* ^c *Research Centre for Study of Extractives from Barley and Hop, Institute of Analytical Chemistry, Academy of Sciences of the Czech Republic, Brno, Czech Republic*): **The Use of Non-specific Peptides for Proteomic Identification of Low-Molecular-Weight Proteins from Barley by MALDI-TOF Mass Spectrometry**

Nonspecific peptides were analyzed for proteomic identification of low-molecular-weight barley proteins either by MALDI-TOF mass spectrometry involving post-source decay analysis or using a MALDI-TOF/TOF instrument. Identification of low-molecular-weight barley proteins was successfully accomplished by MS/MS analysis of both tryptic and nonspecific peptides in the cases where peptide mass fingerprinting failed. Utilization of nonspecific peptides also increased protein sequence coverage in database searching.

**ANALYTICKÁ LABORATOŘ NABÍZÍ K PRODEJI GAMMA RAY SPECTROMETR
miniSPEC GR-130 (PŘÍSTROJ NA MĚŘENÍ RADIACE),
CENA DOHODOU.**

Bližší informace u vedoucí laboratoře - tel. 555 530 361,
e-mail: eliska.vybiralova@macco.cz