

PROTEOMICKÝ PRŮVODCE

JOSEF CHMELÍK

Ústav analytické chemie AV ČR, Veveří 97, 611 42 Brno,
Česká republika
chmelik@iach.cz

Došlo 4.9.05, přijato 16.11.05.

Klíčová slova: proteomika, definice, přístupy, metody, cíle

Obsah

1. Úvod
2. Přístupy k proteomickému studiu
3. Metody proteomického studia
4. Cíle proteomiky
5. Obsah proteomického čísla Chemických listů

1. Úvod

Proteomika se během posledního desetiletí stala jedním z nejdynamičtějších oborů v přírodních vědách. Za to vděčí rozvoji molekulární biologie, genomiky, bioinformatiky a spektroskopických a separačních metod, což přispělo k zdokonalení analýzy nukleových kyselin (např. rychlé sekvenování), separace a charakterizace bílkovin (např. dvourozměrná gelová elektroforéza a chromatografie, nové ionizační metody hmotnostní spektrometrie). Využití výpočetní techniky pro zpracování dat získaných hmotnostní spektrometrií umožnilo rychlou identifikaci bílkovin.

Existuje celá řada definic proteomiky. Za jednu z nejlepších považují tuto: proteomika je obor, který se zabývá systematickou analýzou bílkovin z hlediska jejich identity, množství a funkcí. Jiné definice popisují spíše jednotlivé oblasti studia bílkovin (např. identifikaci, vztahy mezi strukturou a funkcí) nebo způsob získávání údajů (např. rychlá identifikace bílkovin zejména pro diagnostické účely).

Prudce rostoucí počet publikací znesnadňuje začínajícím proteomikům získat ucelenou představu. Toto číslo Chemických listů je věnováno především chemickým a analytickým metodám vhodným k identifikaci bílkovin (tedy tzv. analytické proteomice), bioinformatice, některým vztahům mezi genomem a proteomem, vybraným skupinám bílkovin a několika příkladům identifikace bílkovin a je určeno zejména studentům a dalším zájemcům, kteří nemají přístup k zahraniční literatuře. Tento článek proto upozorňuje na další práce věnované proteomické problematice publikované v českých časopisech. Protože dosud publikované proteomické příspěvky v české literatu-

ře v důsledku různých přístupů jednotlivých autorů působí poněkud roztržtěně, pokusil jsem se v tomto článku shrnout a utřídit některé častěji používané termíny a metody s cílem usnadnit čtenářům pochopení dalších prací v tomto čísle, které jsou shrnuty v závěrečné části tohoto příspěvku.

2. Přístupy k proteomickému studiu

Existuje celá řada rozdílných přístupů k proteomickému studiu jak z hlediska použité metodologie, tak účelu a cílů. V následující části budou zmíněny jenom vybrané častěji používané přístupy:

Analytická proteomika je založena na kombinaci separace bílkovin ze složitých biologických směsí, jejich charakterizaci hmotnostní spektrometrií a na bioinformatickém zpracování získaných údajů. Hlavním cílem je identifikace bílkovin včetně stanovení molekulové hmotnosti, sekvence (pořadí) aminokyselin a určení posttranslačních modifikací. Dalšími cíli jsou např. charakterizace nekovalentních interakcí bílkovin s jinými látkami a kvantitativní stanovení heterogenity.

Strukturní proteomika se zabývá studiem struktury bílkovin s použitím krystalografie, NMR, MS a řady dalších technik s cílem pochopit strukturní chování bílkovin (tvorbu a stabilitu nativní konformace) a využít tyto poznatky v proteinovém inženýrství při modifikacích bílkovin pro specifické účely (léčiva, průmyslové enzymy atd.).

Funkční proteomika se zabývá nejen funkcí bílkovin, ale zejména studiem komplexních životních procesů, jejichž pochopení má zásadní význam pro studium vývoje organismu či mechanismu a léčbu nemocí.

Diferenční (srovnávací) proteomika je založena na analýze složitých směsí bílkovin, která sleduje změny složení bílkovin při různých stavech organismu (např. normální versus patologický stav) s cílem nalézt a identifikovat rozdílně se vyskytující bílkoviny.

High-throughput proteomika je zaměřena na rychlé získávání velkého množství údajů o bílkovinách, což je důležité zejména pro screeningové účely ve zdravotnictví, zemědělství, kontrole potravin atd.

High-coverage proteomika se zabývá získáváním údajů o sekvenci aminokyselin včetně posttranslačních modifikací bílkovin s cílem co nejvyššího pokrytí primární struktury (tj. stanovením pořadí co nejvyššího počtu aminokyselinových zbytků v bílkovině), což je důležité při studiu role bílkovin v životních procesech.

Bottom-up proteomika představuje „klasický“ postup identifikace bílkovin, kdy je daná bílkovina nejprve izolována ze směsi, potom enzymově rozštěpena na směs peptidů, které jsou charakterizovány hmotnostní spektrometrií (molekulová hmotnost, případně sekvence). K tomu jsou většinou využívány specifické peptidy, tj. peptidy, které odpovídají specifitě použitého enzymu (zpravidla trypsinu).

Shotgun proteomika je obecný postup pro identifikace bílkovin, kdy je neseparovaná směs bílkovin nejprve enzymově rozštěpena na směs peptidů, které jsou následně separovány vhodnou metodou (nejčastěji HPLC) a potom sekvenovány tandemovou hmotnostní spektrometrií (často jsou využívány i nespecifické enzymy např. proteinasa K). Jedna z metod se nazývá MudPIT (multidimensional protein identification technology, vícerozměrná technologie identifikace bílkovin).

Top-down proteomika je postup identifikace bílkovin, kdy je daná bílkovina nejprve izolována ze směsi a potom charakterizována (nejvýhodněji je celá molekula fragmentována v hmotnostním spektrometru). Tento přístup přináší rychlé informace o molekulové hmotnosti, sekvenci aminokyselin a posttranslačních modifikacích.

3. Metody proteomického studia

Příprava vzorku pro identifikaci je velmi důležitá. Není příliš rozumné začít homogenizací celého organismu a z této komplexní směsi se snažit charakterizovat celý proteom. I kdyby se to podařilo, což v současnosti se nedaří plně ani pro mikroorganismy, nezískáme více informací než ty, které známe již z genomu. Proto je nutné si definovat na začátku každého projektu jeho cíle a na tomto základě vzorek dostatečně zjednodušit (vybrat jen určitou část organismu nebo typ bílkovin) a tomu přizpůsobit izolační postup (výběr extrakčních činidel, separační metody atd.).

Tradiční metodou analytické proteomiky je dvourozměrná gelová elektroforéza na polyakrylamidovém gelu, využívaná nejen k separaci bílkovin, ale i k získávání dvourozměrných map bílkovin, které lze přímo využít k proteomické charakterizaci organismů, tkání, buněk, metabolických procesů atd.¹⁻³. Další tradiční metodou je Edmannova degradace bílkovin využívaná k jejich sekvenaci, která je nyní nahrazována hmotnostní spektrometrií⁴. Dvourozměrná kapalinová chromatografie spíše doplňuje než nahrazuje dvourozměrnou gelovou elektroforézu⁵. Její největší předností je, že separované bílkoviny jsou v roztoku, a lze je tedy snadno použít ke stanovení jejich molekulové hmotnosti, ze které lze rychle zjistit, zda je bílkovina posttranslačně modifikovaná. Tyto bílkoviny lze použít pro další studium. Existuje řada dalších metod, které umožňují izolovat bílkoviny pro proteomické účely. Část z nich je založena na principu kapilární elektroforézy a isoelektrické fokusace v kapalině⁶⁻⁸, další využívají field-flow frakcionaci⁹ pro izolaci bílkovin¹⁰ nebo biologických částic¹¹. Nejdůležitější jsou stále metody chromatografické. Kromě dvourozměrné kapalinové chromatografie se využívají v proteomice i další techniky (např. gelová chromatografie¹²), ale stále důležitější úlohu hrají afinitní techniky¹³⁻¹⁵.

Dynamický rozvoj analytické proteomiky byl umožněn zejména díky pokrokům hmotnostní spektrometrie, a to hlavně využitím nových ionizačních technik MALDI a ionizace elektrospřejem (ESI), za které byla v roce 2002 udělena Nobelova cena K. Tanakovi a J. Fennovi. Existuje několik hmotnostně spektrometrických způsobů identifi-

ce bílkovin: stanovení molekulové hmotnosti intaktních bílkovin, PMF (peptide mass fingerprinting, peptidové mapování), fragmentační analýza peptidů a fragmentační analýza intaktních bílkovin.

Stanovení molekulové hmotnosti intaktní bílkoviny sice nelze považovat za dostatečnou identifikaci, na druhou stranu poskytuje okamžitou informaci o možných posttranslačních modifikacích.

PMF využívá stanovení molekulové hmotnosti peptidů vzniklých specifickým štěpením bílkovin určitým enzymem, stanovené molekulové hmotnosti peptidů jsou porovnány s teoretickými molekulovými hmotnostmi peptidů vzniklých specifickým štěpením stejným enzymem všech bílkovin v databázích. Výhodou je, že tento postup vyžaduje velmi malé množství bílkovin a že ho lze snadno automatizovat. Nevýhodou je pak nejednoznačnost výsledků při analýze směsi bílkovin a to, že jde pouze o nepřímou identifikaci, která selhává u mutací a posttranslačních modifikací.

Fragmentační analýza využívá stanovení molekulové hmotnosti fragmentových iontů vzniklých v hmotnostním spektrometru z peptidů ke stanovení jejich aminokyselinové sekvence. Ta je potom využita pro identifikaci v databázích. Původně byla tato technika náročnější a pomalejší než PMF, protože byla vyvinuta pro elektrospřejovou tandemovou hmotnostní spektrometrii (ESI MS/MS), ale rychlý rozvoj tandemových technik MALDI (např. MALDI TOF/TOF, MALDI Q-TOF) umožnil použití stejného vzorku pro oba přístupy. Navíc fragmentační analýza peptidů umožňuje i analýzu nespecifických peptidů¹⁶ nebo stanovení sekvence aminokyselin u bílkovin, které nejsou v databázích, tzv. *de novo* sekvenací. Fragmentační analýza intaktních bílkovin s použitím MS přináší v případech, kdy je daná bílkovina nejprve izolována ze směsi a potom fragmentována tandemovou hmotnostní spektrometrií, informace o molekulové hmotnosti, sekvenci aminokyselin a posttranslačních modifikacích¹⁸.

Údaje získané z MS/MS analýzy peptidů lze interpretovat několika postupy. Dříve se používala ruční interpretace kolizních spekter, kdy je získaná sekvenční informace použita k prohledávání databáze bílkovin vyhledávacím programem. Zpracování obrovského množství dat získaného hmotnostní spektrometrií si vyžádalo možnost přímého využití experimentálních MS/MS dat pro vyhledávací programy. Další rozvoj výpočetní techniky a bioinformatiky nejen pro rychlou identifikaci bílkovin, ale i pro zjišťování jejich biologické role, je nezbytnou podmínkou splnění cílů proteomického výzkumu.

4. Cíle proteomiky

Proteomika, která prochází v posledních letech rychlým vývojem, zasahuje do všech oblastí vědy i aplikací, kde nějakou roli hrají bílkoviny. Hlavním cílem proteomiky je popis a vysvětlení role bílkovin v procesech, které probíhají v organismech, což zahrnuje identifikaci bílkovin (včetně stanovení molekulové hmotnosti, sekvence aminokyselin, kvantitativní stanovení heterogenity, určení vazeb

S–S a posttranslačních modifikací), interakce bílkovin s jinými látkami, kvalitativní i kvantitativní stanovení změn složení bílkovin během biologických procesů atd. Nejdůležitější jsou aplikace proteomiky v medicíně (studium příčin a mechanismů chorob, nalezení markerů patologických stavů), farmacii (hledání a vývoj léčiv), potravinářství (kvalita a bezpečnost potravin, nové výrobní procesy), průmyslu (technické enzymy), ochraně životního prostředí (likvidace škodlivin enzymy), zemědělství (zvyšování výnosů, pěstování nových odolnějších odrůd s nižším obsahem škodlivých látek) atd.

Proteomika generuje stále větší množství dat o bílkovinech v makromolekulárních komplexech, buňkách, organelách, tkáních i celých organismech a tempo získávání dat se neustále zrychluje. Současně se zlepšují a vyvíjejí nové proteomické techniky, a to jak analytické, biologické, tak zejména bioinformatické, neboť produkované množství dat přesahuje možnosti nepočítačového zpracování.

5. Obsah proteomického čísla Chemických listů

Prudký rozvoj proteomiky nebyl dosud příliš zachycen v české odborné literatuře³, proto se výbor proteomické sekce České společnosti pro biochemii a molekulární biologii spolu s redakcí Chemických listů dohodl na přípravě proteomického čísla, které by poskytlo českým zájemcům pokud možno ucelený přehled této problematiky. Na počátku jsem oslovil asi 30 pracovníků, kteří se u nás proteomikou zabývají, aby napsali příspěvek na určité téma. Asi dvě třetiny oslovených vyhovělo mé žádosti a první část společného úsilí nyní čtenářům předkládáme. Druhá část příspěvků bude v Chemických listech publikována v roce 2006, kde se mimo jiné objeví referáty o glykoproteinech, měření proteomických dat a studiu trojrozměrné struktury bílkovin hmotnostní spektrometrií.

Úvodní články první části shrnují historii, vývoj a obecnější přístupy proteomiky^{19,20}. Následující příspěvky popisují moderní chemické proteomické postupy, využití proteolytických enzymů²¹, integrovaných řešení²² a mikrofluidiky²³ v proteomice. Další příspěvek se zabývá důležitou skupinou bílkovin fosfoproteiny²⁴. Tři firemní příspěvky popisují moderní hmotnostně spektrometrickou instrumentaci^{25–27}.

Aplikace jsou jednak zaměřeny na využití dvojrozměrné kapalinové chromatografie pro izolaci bílkovin²⁸, jednak na využití proteomiky v molekulární toxikologii²⁹, na charakterizaci proteomu bakteriofága 812 (cit.³⁰) a na proteomickou identifikaci glutenových bílkovin³¹.

Tato práce byla podpořena prostředky výzkumnému záměru MŠMT č. Z40310501.

LITERATURA

1. Bouchal P., Kučera I.: Chem. Listy 97, 29 (2003).
2. Váňa P., Šmarda J.: Chem. Listy 98, 1130 (2004).
3. Collinsová M., Jiráček J.: Chem. Listy 98, 1112 (2004).
4. Cantin G. T., Yates J. R.: J. Chromatogr. A 1053, 22

- (2004).
5. Vollmer M., Horth P., Vad C., Nagele E.: LC GC Europe 17, 14 (2004).
6. Chmelík J., Janča J.: Chem. Listy 83, 321 (1989).
7. Chmelík J., Thormann W.: J. Chromatogr. 600, 305 (1992).
8. Thormann W., Caslavská J., Molteni S., Chmelík J.: J. Chromatogr. 589, 321 (1992).
9. Chmelík J.: Chem. Listy 93, 670 (1999).
10. Kang D., Moon M.H.: Anal. Chem. 77, 4207 (2005).
11. Urbánková E., Vacek A., Chmelík J.: J. Chromatogr., B 687, 449 (1996).
12. Šalplachta J., Řehulka P., Chmelík J.: J. Mass Spectrom. 39, 1395 (2004).
13. Madera M., Mechref Y., Novotny M. V.: Anal. Chem. 77, 4081 (2005).
14. Opitck G. J., Scheffler J. E.: Expert Rev. Proteomics 1, 57 (2004).
15. Baumann M., Meri S.: Expert Rev. Proteomics 1, 207 (2004).
16. Zdráhal Z., Plocek J., Konečný P., Chmelík J.: Chem. Listy 91, 811 (1997).
17. Řehulka P., Chmelík J., Allmaier G.: Rapid Commun. Mass Spectrom. 19, 79 (2005).
18. Kelleher N. L.: Anal. Chem. 76, 196A (2004).
19. Kovářová H.: Chem. Listy 99, 886 (2005).
20. Weiser J., Holub M., Nezbedová Š., Bezoušková S.: Chem. Listy 99, 890 (2005).
21. Štosová T., Havliš J., Lenobel R., Šebela M.: Chem. Listy 99, 896 (2005).
22. Herick K.: Chem. Listy 99, 906 (2005).
23. Grym J., Foret F.: Chem. Listy 99, 915 (2005).
24. Halada P.: Chem. Listy 99, 922 (2005).
25. Godula M.: Chem. Listy 99, 930 (2005).
26. Verner P.: Chem. Listy 99, 937 (2005).
27. Boháč M., Ingendoh A., Fuchser J., Witt M.: Chem. Listy 99, 943 (2005).
28. Skalníková H., Kovářová H., Moos J., Filová V., Halada P.: Chem. Listy 99, 952 (2005).
29. Fella K., Glückmann M., Kruft V., Kamer P.-J., Kröger M.: Chem. Listy 99, 957 (2005).
30. Zdráhal Z., Eyer L., Konečná H., Preisler J.: Chem. Listy 99, 962 (2005).
31. Šalplachta J., Allmaier G., Chmelík J.: Chem. Listy 99, 967 (2005).

J. Chmelík (*Institute of Analytical Chemistry, Academy of Sciences of the Czech Republic, Veveří 97, 611 42 Brno, Czech Republic*) **Proteomic Guide**

Proteomics is a fast developing field aimed at protein identification. Its development was made possible by improvement of separation and spectroscopic techniques, and development of bioinformatics necessary for data evaluation. This review summarizes some basic approaches, aims and methods (including MS, electrophoresis, chromatography, focusing field-flow fractionation, etc.). The review opens a special proteomics issue of the journal.