

PROTEOMIKA V POSTGENOMOVÉ DOBĚ

HANA KOVÁŘOVÁ

Ústav živočišné fyziologie a genetiky AV ČR, Rumburská 89, 27721 Liběchov
kovarova@iapg.cas.cz

Došlo 3.9.05, přijato 11.11.05.

Klíčová slova: proteom, genom, dvojrozměrná gelová elektroforéza, hmotnostní spektrometrie, posttranslační modifikace

Obsah

1. Úvod
2. Rozvoj proteomiky
 - 2.1. Cesta od genomu k proteomu
 - 2.2. Dělení proteinů
 - 2.3. Identifikace proteinů hmotnostní spektrometrií
3. Současná tvář proteomiky
 - 3.1. Elektromigrační technika
 - 3.2. Chromatografické techniky
 - 3.3. Identifikace a kvantifikace proteinů
4. Aplikace proteomiky
 - 4.1. Proteinové profilování
 - 4.2. Posttranslační modifikace
 - 4.3. Proteinové interakce
5. Kam směřuje proteomika

Úvod

Termín „genom“ označuje veškerou výbavu nukleových kyselin, které zajišťují buněčnou dědičnost a kódují strukturu i funkční projev buňky. Analogicky „proteom“ zahrnuje všechny bílkoviny (proteiny), které tvoří převažující součást buněčných struktur a vykonávají prakticky všechny buněčné funkce: převádějí a zpracovávají biologickou informaci, uskutečňují výměnu látek a energie (metabolismus).

Slovo „proteom“ (proteome; *protein complement of the genome*) použil poprvé australský doktorand Mark Wilkins v r. 1994 na konferenci v Sieně. Vznikl nový obor zvaný proteomika, jehož náplní je globální identifikace proteinů v určité zdravé nebo nemocné buňce či tkáni a studium jejich funkce či interakce¹. Velmi brzy se ukázalo, že proteomika bude v porovnání s genomikou, studiem na úrovni nukleových kyselin, daleko složitější. Příčinou je proměnlivost určitého proteomu během buněčného růstu a diferenciaci, ale i vliv funkčních produktů jiných buněč-

ných typů či vnitřního prostředí organismu, které se mění v závislosti na stárnutí, při adaptaci na změny vnějšího prostředí a v průběhu nemoci. Proteiny jsou stále pozměňovány. Navazují se na ně fosforečné, cukerné nebo lipidové skupiny, spojují se s jinými proteiny do větších molekulárních celků. Děje se tak v průběhu jejich syntézy i po dobu jejich setrvání v organismu. Množství jednotlivých proteinů v jediné eukaryotické buňce se odhaduje v rozmezí od 1 do 10⁶ molekul a dosahuje až 10¹⁰ v séru². Není tedy divu, že základní skladba lidského genomu byla rozluštěna během jediného roku na robotických zařízeních a DNA sekvenátorech americké firmy Celera Genomics Craiga Ventera, zatímco pro sestavení proteomu jsou nezbytné citlivé metody dělení proteinů, jejich automatické zpracování a náročné určování proteinů hmotnostní spektrometrií.

Proteomika je klíčem k systematickému studiu biologických dějů. Vývoj směřuje ke stále citlivějším metodám. Cílem je identifikovat a charakterizovat proteiny v nejnižších možných koncentracích.

2. Rozvoj proteomiky

2.1. Cesta od genomu k proteomu

Technologický rozvoj analytické proteinové chemie kolem roku 1990 byl původně zaměřen na zvyšování citlivosti metod používaných k identifikaci proteinů dělených na gelech. Nejčastěji bylo používáno určování sledu aminokyselin od *N*-konce neporušených proteinů nebo enzymově štěpených proteinových fragmentů (Edmanovo sekvenování), ale tento způsob byl pomalý a málo citlivý³. V těchto letech byly už rozluštěny genomy několika mikroorganismů a to vytvořilo předpoklady k identifikaci proteinů pomocí genových databází. Efektivnost tohoto přístupu byla však omezena schopností určit částečnou sekvenční informaci z daného proteinu a srovnat ji s daty uloženými v genových databázích.

2.2. Dělení proteinů

K dělení proteinů se stále používá dvojrozměrná gelová elektroforéza (2DE). Poprvé byla popsána v roce 1975 nezávisle třemi různými autory⁴⁻⁶. Předností této metody je rozdělení jednotlivých proteinů ze směsi získané z buňky, tkáni či tělních tekutin na gelech podle dvou nezávislých biochemických charakteristik. Nejprve jsou proteiny v prvním směru rozděleny podle svých nábojů (do isoelektrických bodů, pI) a poté v kolmém směru podle svých molekulových hmotností. Výsledkem je tzv. proteinová mapa, v níž každý protein zaujímá charakteristickou pozici. Proteiny v gelu jsou vizualizovány barvením různými způsoby, přičemž nejvyšší citlivosti dosahuje stříbření⁷. Technika 2DE jako první umožnila a podpořila současnou

kvantitativní analýzu velkého počtu proteinů.

Metoda 2DE přispěla již v 80. letech k sestavování proteinových databází. V této době byly také navrženy statistické metody (shlukovací algoritmy či metoda hlavních komponent) hodnocení proteinových map. Identifikaci proteinů z gelů však umožnilo až použití nových analytických přístrojů, které byly vyvinuty na počátku 90. let.

2.3. Identifikace proteinů hmotnostní spektrometrií

V hmotnostní spektrometrii (MS) je měřena hmotnost molekuly (přesněji poměr hmotnosti k náboji), která musí být nejprve ionizována a přenesena do vysokého vakua přístroje. Zatímco analýza malých molekul byla úspěšná, látky s vyšší molekulovou hmotností se rozkládaly. Vývoj dvou metod, ionizace elektrosprejem⁸ (ESI) a laserová desorpce za přítomnosti matrice⁹ (MALDI, matrix-assisted laser desorption/ionization) ke konci 80. letech druhého tisíciletí vyřešil tento problém a otevřel cestu k výrobě nových typů přístrojů. Iontové zdroje MALDI jsou nejčastěji spojeny s analyzátozem hmotnosti TOF (time of flight), zatímco ESI je obvykle spojen s iontovou pastí (IT) nebo kvadrupolovým (Q) radiofrekvenčním typem analyzátoru. Vedle MALDI-TOF jsou také dostupné hmotnostní spektrometry využívající ESI iontové zdroje a hybridní Q-TOF analyzátoři¹⁰.

Četné experimenty potvrdily, že stanovení hmotnosti proteinů, pokud nedosahuje extrémní přesnosti, nestačí k jejich určení. To umožní teprve hmotnosti peptidů, které vzniknou štěpením proteinu na specifických místech. Identifikace pak spočívá v srovnání spektra experimentálních peptidů s teoretickými spektry vypočítanými ze sekvenčních databází. V roce 1993 bylo publikováno pět nezávislých sdělení o vyhledávacích databázových algoritmech^{11–15}. Tak vznikly metody známé jako peptidové hmotnostní mapování či otiskování (peptide mass mapping/fingerprinting).

3. Současná tvář proteomiky

3.1. Elektromigrační technika: 2DE a MS

Dvojměrná polyakrylamidová elektroforéza zůstává základem proteomiky¹⁶. Použití imobilizovaných pH gradientů do velké míry vyřešilo problém reprodukovatelnosti prvního rozměru a usnadnilo dělení bazických proteinů. Zdokonaluje se účinnost solubilizace membránových proteinů, i když přetrvávají problémy s hydrofóbními proteiny nebo proteiny o extrémní velikosti či pI. Objevují se i další postupy jako techniky diferenčního fluorescenčního značení proteinů (DIGE) či fluorescenční detekce separovaných proteinů, která výrazně rozšiřuje oblast linearity při kvantifikaci proteinů. Počítačové vyhodnocování velkého počtu proteinových map, nezbytné pro úplné objasnění různých biologických procesů, se vylepšuje a automatizuje díky novému programovému vybavení, přesto stále vyžaduje určité manuální zásahy. Hmotnostní spektrometrie MALDI-TOF se stala metodou první volby při identifikaci proteinů dělených 2DE.

3.2. Chromatografické techniky: LC-MS

Omezení související s použitím klasické elektromigrační techniky vedla k rychlému rozvoji pokročilejší generace proteomických technik, zejména kombinace kapalinové chromatografie (LC) a tandemové hmotnostní spektrometrie (MS/MS), a značení stabilním izotopem¹⁰. Tyto techniky jsou založeny na podmínce, že měření peptidů vzniklých proteolytickým štěpením komplexní směsi proteinů může být použito jak pro identifikaci proteinů, tak pro jejich relativní kvantifikaci. Při použití LC-MS postupu jsou peptidy děleny kapalinovou chromatografií s reverzní fází (RP-HPLC) a poté jsou nastříknuty na hmotnostní spektrometr ESI. Softwarové vybavení přístroje může vybrat jednotlivé prekurzorové ionty pro tandemovou MS/MS analýzu. Ty jsou pak fragmentovány za vzniku dceřinných iontů, které jsou dále změřeny a zobrazeny ve výsledném spektru MS/MS a udávají aminokyselinovou sekvenci prekurzoru. Vzhledem k vysoké kapacitě kolon RP-HPLC a vysoké analytické schopnosti ESI-MS je možné analyzovat velmi komplexní směsi peptidů, což bylo poprvé demonstrováno na příkladu peptidů prezentovaných hlavním histokompatibilním komplexem imunitního systému¹⁷. Je zřejmé, že analyzátor hmotnostního spektrometru je jeho rozhodující součástí a schopnost generovat hmotnostní spektra udávající informace o iontech odvozených ze studovaných peptidů/proteinů je klíčovým prvkem v kontextu současné proteomiky. Vedle již zmíněných analyzátorů typu IT, TOF a Q, je nutné zmínit i analyzátor typu FT-ICR (Fourier transform ion cyclotron resonance)¹⁸, který vychytává ionty podobně jako iontová past, avšak za podmínek vysokého vakua v silném magnetickém poli. Jeho předností je vysoká přesnost měření hmotnosti a rozlišení. V posledních letech se úspěšně prosazují nové konfigurace iontových zdrojů a hmotnostních analyzátorů, např. MALDI-TOF/TOF, kde je mezi dva analyzátoři TOF vložena kolizní cela. MALDI-generované prekurzorové ionty jsou v prvním analyzátoru TOF analyzovány podle poměru jejich hmotnosti a náboje (m/z) a vybrané ionty pak procházejí kolizní celou, kde jsou dále fragmentovány a hmotnosti iontových fragmentů jsou měřeny ve druhém analyzátoru TOF (cit.¹⁹). Obdobně může být vložena kolizní cela i mezi kvadrupolový iontový filtr a analyzátor TOF (Q-TOF). Tyto hybridní přístroje poskytují informačně bohatá hmotnostní spektra a v současné proteomice zaujímají dominantní místo při identifikaci i *de novo* sekvenaci proteinů.

3.3. Identifikace a kvantifikace proteinů

Ve snaze přidat LS-MS přístupům kvantitativní aspekt studia proteomů byly navrženy různé postupy, které využívají skutečnosti, že páry chemicky identických analytů s různým izotopickým složením mohou být rozlišeny v hmotnostním spektrometru díky rozdílu v jejich hmotnosti a že poměr signálních intenzit těchto párových analytů udává jejich relativní množství. Stabilní izotopové značky mohou být do peptidů zavedeny metabolicky pomocí značených aminokyselin, enzymově prostřednictvím přenosu ¹⁸O z vody do peptidů²⁰, nebo chemickou reakcí vyu-

živající např. ICAT (isotope-coded affinity tags)²¹. V současné době jsou popsány chemické postupy pro specifické izotopové značení skupin SH, aminoskupin, *N*-vázaných glykosacharidů aj. Značení určitých skupin či struktur je závislé na navržení vhodných chemických reakcí. V posledním období bylo popsáno tzv. značení SILAC (stable isotope labeling with amino acid in cell culture)²², kdy je do proteinů jednoho z porovnávaných vzorků inkorporován ¹³C-arginin. Nezanedbatelnou výhodou je, že takto může být označeno převážně množství peptidů aniž by bylo nezbytné použít chemické reakce.

4. Aplikace proteomiky

4.1. Proteinové profilování

Proteiny buněk, tkání či tělních tekutin mohou být v současné době charakterizovány pomocí výše popsaných metod. Jedno možné měřítko technického pokroku je počet identifikovaných proteinů. Pro komplexní vzorek tato čísla v současné době dosahují tisíce. Je jistě lákavé systematicky analyzovat proteiny exprimované v buňkách a tkáních, získat úplné proteinové mapy a uspořádat příslušné proteinové databáze. Nicméně, žádnou kompletní mapu jediného proteomu se dosud nepodařilo získat a tento fakt dokumentuje potřebu dalšího technologického vývoje v proteomice. Na rozdíl od jednoduchého vyzdvihování analytického charakteru použitých metod je však logické zdůraznit u proteomických studií potřebu řešit určitou biologickou otázku. Mnohé studie jsou proto zaměřeny na sledování změn proteinových profilů v závislosti na funkčním stavu buňky či tkáně. Získaná data jsou však spíše semikvantitativní založená na relativní kvantifikaci a představují seznam kandidátních proteinů typických pro sledovaný biologický jev. Aby byl pohled ucelenější, je nutné kombinovat i mnohé proteomické techniky, které často poskytují komplementární výsledky.

4.2. Posttranslační modifikace

Mnohé proteiny jsou posttranslačně modifikovány a počet možných modifikací aminokyselin dosahuje 200. Tyto posttranslační modifikace (PTM) jsou většinou reverzibilní a mají významnou regulační úlohu²³. Vzhledem k tomu, že prakticky všechny PTM jsou spojeny se změnou molekulové hmotnosti proteinu, jsou metody hmotnostní spektrometrie ideálním prostředkem k určení typu modifikace, ale i aminokyseliny, na které se modifikace nachází. Proteomová analýza PTM je však mnohem obtížnější než jednoduchá identifikace proteinu, a to z důvodů nízké stechiometrie PTM, lability a reverzibility vazby, a potřeby izolovat modifikovaný peptid obsahující PTM (cit.²⁴).

4.3. Proteinové interakce

Interakce protein-protein jsou důležité nejen z hlediska regulace funkce proteinů, ale i jejich vnitrobuněčné lokalizace. Mnohé biologicky relevantní interakce probíhají s nízkou afinitou, jsou přechodné a závislé na

specifickém buněčném mikroprostředí. Proteomický přístup ke studiu proteinových komplexů vyžaduje jejich specifické obohacení či purifikaci s následnou analýzou navázaných proteinů. Získání dostatečného množství proteinového komplexu s požadovanou čistotou však představuje nesnadný úkol. Specifické strategie musí být formulovány pro každý sledovaný proteinový komplex a jemné separační techniky musí být zvoleny tak, aby nedocházelo k narušení proteinových interakcí. Mezi nejčastěji používané izolační postupy patří afinitní purifikace interagujících proteinů a identifikační část experimentů kombinuje elektromigrační i chromatografické separace s hmotnostní spektrometrií²⁵. Proteinové interakce mohou být určovány také pomocí proteinových čipů (arrays)²⁴. Na malých destičkách jsou navázané specifické proteinové sondy a po kultivaci s buněčným lyzátem je interagující cílový protein prokázán pomocí fluorescence či jinou optickou metodou, popřípadě hmotnostní spektrometrickou technikou SELDI (surface enhanced laser desorption ionization).

5. Kam směřuje proteomika

Proteomika je dynamickým oborem, který se zabývá systematickým studiem proteomů neboli souborů buněčných a tkáňových bílkovin s cílem poskytnout podrobné poznatky o funkci, struktuře a kontrole biologických systémů. Většina velkých farmaceutických firem otevřela vlastní proteomová centra a mnoho biotechnologických společností má proteomiku jako hlavní program. Veřejný výzkum lidského proteomu je řízen mezinárodní organizací HUPO (Human proteome organization – www.hupo.org) a nedávno založená EuPA (European proteomics association) se zaměřuje i na výzkum proteomů mnoha dalších živočišných druhů a podporuje aktivity národních proteomických společností v Evropě. Také Proteomická sekce České společnosti pro biochemii a molekulární biologii je členem EuPA a aktivně vstupuje do evropského ale i celosvětového proteomického výzkumu.

Přestože má proteomika multidisciplinární charakter, nejvýznamnější přínosy se očekávají v oblasti biomedicíny. Poznání proteinů, které se vyskytují v lidském organismu v nepředstavitelně mizivých množstvích miliontin mikrogramů nebo po nesmírně krátkou dobu, umožní poznat složité regulace a hierarchické úrovně lidského organismu a položí solidní i když složité základy pro klinickou medicínu zítřka.

LITERATURA

1. Wilkins M. R., Sanchez J.-C., Gooley A. A., Apper R. D., Humphery-Smith I., Hochstrasser D. F., Williams K. L.: *Biotech. Gen. Eng. Reviews* 13, 19 (1995).
2. Patterson S. D., Aebersold R. H.: *Nature Genetics Suppl.* 33, 311 (2003).
3. Aebersold R. H., Leavitt J., Saavedra R. A., Hood L.

- E., Kent S. B.: Proc. Natl. Acad. Sci. U.S.A. 84, 6970 (1987).
4. O'Farrell P. H.: J. Biol. Chem. 250, 4007 (1975).
 5. Scheele G. A.: J. Biol. Chem. 250, 5275 (1975).
 6. Klose J.: Humangenetik 26, 231 (1975).
 7. Hochstrasser D. F., Merrill C. R.: Appl. Theor. Electrophor. 1, 35 (1988).
 8. Fenn J. B., Mann M., Meng C. K., Wong S. F., Whitehouse C. M.: Science 246, 64 (1989).
 9. Karas M., Hillenkamp F.: Anal. Chem. 60, 2299 (1988).
 10. Aebersold R. H., Mann M.: Nature 422, 198 (2003).
 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. Mann M., Hojrup P., Roepstorff P.: Biol. Mass Spectrom. 22, 338 (1993).
 13. Pappin D. J. C., Hojrup P., Bleasby A. J.: Curr. Biol. 3, 327 (1993).
 14. James P., Quadroni M., Carafoli E., Gonnet G.: Biochem. Biophys. Res. Commun. 195, 58 (1993).
 15. Yates J. R. III, Speicher S., Griffin P. R., Hunkapiller T.: Anal. Chem. 214, 397 (1993).
 16. Griffin T. J., Aebersold R. H.: J. Biol. Chem. 276, 45497 (2001).
 17. Hunt D. F., Henderson R. A., Shabanowitz J., Saka-guchi K., Michel H., Sevilir N., Cox A. L., Appella E., Engelhard V. H.: Science 255, 1263 (1992).
 18. Martin S. E., Shabanowitz J., Hunt D. F., Marto J. A.: Anal. Chem. 72, 4266 (2000).
 19. Medzihradzky K. F., Campbell J. M., Baldwin M. A., Falick A. M., Juhasz P., Vestal M. L., Burlingame A. L.: Anal. Chem. 72, 552 (2000).
 20. Yao X., Freas A., Ramirez J., Demirev P. A., Fenselau C.: Anal. Chem. 73, 2836 (2001).
 21. Gygi S. P., Rist B., Gerber S. A., Turecek F., Gelb M. H., Aebersold R. H.: Nature Biotechnol. 17, 994 (1999).
 22. Ong S. E., Blagoev B., Kratchmarova I., Kristensen D. B., Steen H., Pandey A., Mann M.: Mol. Cell. Proteomics 1, 376 (2002).
 23. Krishna R., Wold F., v knize: *Protein structure – A practical Approach* (Creighton T. E., ed.), 2. vydání, str. 91. Oxford University Press, New York 1997.
 24. Simpson R. J.: *Proteins and Proteomics*. Cold Spring Harbor Laboratory Press 2003.
 25. Huber L. A.: Nature Rev. 4, 74 (2003).

H. Kovářová (*Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Liběchov, Czech Republic*): **Proteomics in Postgenome Era**

Proteomics studies protein properties to provide an integrated view of cellular processes. Proteome, which is the protein complement of a genome, provides a platform for proteomics. Proteomics, a burgeoning field at present, is aimed at profiling and characterizing gene expression at the protein level, their post-translational modifications and interactions. The review is devoted to the most important events in the short history of proteomics and to recent trends and developments in the field.