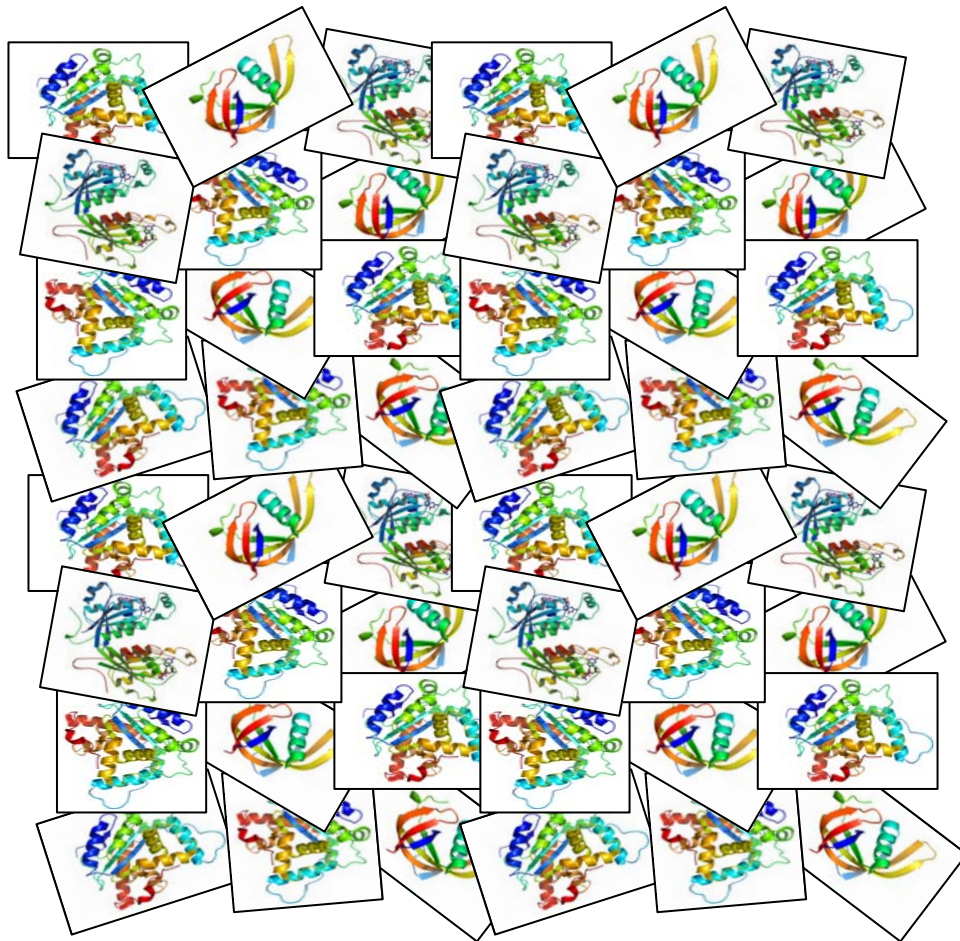


# *Proteomika – historie a současnost*



*Jiří Petrák*

**Praha 2012**

Tento text vznikl jako součást habilitační práce autora.

## **Obsah**

<b>Proteom a proteomika</b>	5
<b>Historický úvod - tři technické pilíře proteomiky</b>	5
Elektroforéza	5
Chromatografie	8
Hmotnostní spektrometrie	9
<b>Od biochemie proteinů k moderní proteomice</b>	11
<b>Kvantitativní hmotnostní spektrometrie</b>	15
Využití stabilních izotopů	15
Metody kvantifikace bez značení	16
<b>Analýza kvalitativních změn, proteinové komplexy a proteinové čipy</b>	17
<b>Maxima, limity a problémy současné proteomiky</b>	19
<b>Závěr</b>	21
<b>Literatura</b>	22

## Proteom a proteomika

Termín **proteom** spatřil světlo světa poměrně nedávno, poprvé ho použil Marc R. Wilkins v roce 1994 pro označení proteinového komplementu genomu (1). Proteom lze definovat jako kompletní sadu proteinů nacházejících se v konkrétním okamžiku v daném organismu, tkáni či buňce. Širší definice proteomu však zahrnuje nejen jednotlivé proteiny, ale také jejich varianty, veškeré jejich posttranslační modifikace, jejich metabolický obrat, subcelulární lokalizaci a interakce mezi jednotlivými bílkovinami.

**Proteomika** je souborem metod, technik, přístupů a konceptů, které usilují o kvantitativní a kvalitativní popis a porovnání jednotlivých proteomů. Je přístupem globálním (large scale), čímž se odlišuje od klasických metod molekulární biologie či biochemie proteinů, cílených zpravidla jen na jeden či několik málo proteinů. Proteomika se snaží identifikovat veškeré přítomné proteiny, odpovědět na otázku jak jsou modifikovány, kdy a v jaké míře jsou exprimovány, kde se nacházejí a jak spolu interagují. **Expresní proteomika**, která představuje hlavní směr proteomického výzkumu a je také těžištěm této práce, usiluje o kvantitativní popis změn v proteomu vyvolaných nějakým fyziologickým či patologickým procesem, nebo vnější změnou či cíleným zásahem.

Z technického nebo metodického hlediska je proteomika vybudována na třech technických pilířích. Jsou to **elektroforéza, chromatografie a hmotnostní spektrometrie**, které umožňují separaci, identifikaci a kvantifikaci stovek či tisíců proteinů a peptidů v jediném experimentu. Právě o těchto třech základních nástrojích, o jejich vývoji i limitech, které definují hranice expresní proteomiky, budeme postupně hovořit na následujících stránkách. Budeme diskutovat možnosti, problémy a omezení současné proteomiky a zmíníme se o tom, co proteomika biologii a medicíně nabízí. Čtyři vybrané studie nám v závěru práce poslouží jako konkrétní ilustrace různých aplikací a problémů současné expresní proteomiky v biologii a medicíně.

### Stručný historický úvod – tři technické pilíře proteomiky

V úvodním textu se vypravíme, trochu odlehčenou formou, k historickým kořenům tří základních nástrojů proteomiky – elektroforézy, chromatografie a hmotnostní spektrometrie. Exkurzi pojmem chronologicky, nehlouběji v čase se vrátíme v případě elektroforézy – na samotný úsvit éry moderních separačních metod.

#### Elektroforéza

Svítání začíná na východě. Ano, na východě – v Moskvě, na březích stejnojmenné řeky, odkud si v roce 1807 profesor Ferdinand Friedrich Reuss, toho času vedoucí katedry chemie na Moskevské univerzitě, přinesl vzorky jílu a s pomocí volta článek sestaveného z 92 provrtaných stříbrných rublů proložených zinkovými plíšky popsal to, co dnes známe jako elektroosmotický tok (přesun molekul vody v elektrickém poli k negativnímu pólu) a rovněž přesun částic jílu v elektrickém poli (elektroforéza) (2). Reuss tak položil základy elektroforetické separace, která se stala jednou ze základních metod separace biomolekul, a představuje jeden z

tří metodických pilířů proteomiky. Stalo se tak o několik desetiletí dříve, než švédský učenec Jöns Jacob Berzelius rozšířil náš slovník o termín protein.

Zásadního pokroku se elektromigrační metody dočkaly až o více než 100 let později, ve třicátých letech dvacátého století, díky práci Arneho Tiselie, který v roztoku mezi elektrodami dokázal separovat sérové proteiny a opticky sledovat pohyblivé rozhraní jednotlivých molekul (3, 4). Za svůj přínos v oboru obdržel Tiselius v roce 1948 Nobelovu cenu za chemii. Na Tiseliovy elektroforetické metody navázali mnozí, z nichž zmiňme především Linuse Paulinga, který v roce 1949 demonstroval odlišnou elektroforetickou migraci hemoglobinu u pacientů se srpkovitou anémií a zdravých jedinců (5).

Negativní vlivu gravitace na separaci nabitých částic ve volném roztoku však rozlišení Tiseliovy elektroforézy zásadně omezoval. Řešení problému bylo nalezeno převedením vodivého pufru do pevného média. Prvním použitým podpůrným médiem byl papír (6). Papírová elektroforéza získala ve čtyřicátých a padesátých letech na značné oblibě a byla využívána k dělení především menších molekul, nukleotidů, aminokyselin. Papírová elektroforéza hrála například klíčovou roli v experimentech Vernona Ingrama, který následoval Paulinga ve studiích srpkovité anémie. Ingramm zkombinoval papírovou elektroforézu s papírovou chromatografií a získal tak dvojrozměrnou „mapu“ peptidů hemoglobinů. Mezi 26 peptidů v této 2D mapě byl odhalen jediný peptid v hemoglobinu ze srpkovitých erytrocytů, který se svými vlastnostmi lišil od peptidů "zdravého" hemoglobinu (7). S pomocí nově vyvinuté metody sekvenování proteinů Pehra Emany byla v nalezeném peptidu identifikována i jediná aminokyselinová záměna glutaminu za valin, která je za srpkovitou anémií zodpovědná, a byly tak položeny základy moderní molekulární genetiky (8). Linus Pauling a Vernon Ingram jsou tak právem považováni za otce molekulární medicíny.

Kromě papíru se jako podpůrné medium pro elektroforetické separace v padesátých letech ujal agarový gel (9). Agarová elektroforéza umožnila v roce 1953 například vývoj zásadní imunologické metody – křížové (imuno) elektroforézy, která dala vznik mnoha dnešním imunologickým metodám (10). Nepříznivý elektroosmotický efekt záporně nabitého agaru na elektroforetickou migraci vyústil později v zavedení čisté agarózy jako elektroforetického média, široce využívaného dodnes především pro dělení nukleových kyselin (11).

Současně s agarem se však rozvíjelo i využití jiných podpůrných medií. V roce 1955 demonstroval Oliver Smithies separační schopnost škrobového gelu, když detekoval sérové haptoglobiny (12). Elektroforetická separace biomolekul byla dále zdokonalena zavedením vícesložkových diskontinuálních pufrů biochemikem českého původu Miroslavem Poulikem (13).

Léta šedesátá přinesla několik zásadních pokroků, především izoelektrickou fokusaci a polyakrylamidové gely. Již v roce 1961 upozornil Harry Svensson na možnost separace amfoterních molekul v gradientu pH, byl to však až jeho žák Olof Vesterberg, kdo na přelomu 60 a 70 let vypracoval syntézu nosičových amfolytů pro gradient pH a položil tak základy izoelektrické fokusace (IEF) proteinů jako standardní separační techniky (14, 15). Když v roce 1962 Samuel Raymond představil polyakrylamidovou gelovou elektroforézu, byla cesta k účinným jedno-, či více-rozměrným elektroforetickým separacím komplexních směsí proteinů otevřena (16). Stačilo už pouze odhalit skutečnost, že množství negativně nabitého detergentu SDS je funkcí jeho molekulové hmotnosti a že tuto vlastnost lze využít k elektroforetickému dělení proteinů dle molekulové hmotnosti (17,19) a zkombinovat polyakrylamidovou elektroforézu s diskontinuálním

pufrovým systémem (19). Vznikla tak účinná elektroforetická metoda rozdělující proteiny podle jejich molekulové hmotnosti, SDS-PAGE (SDS-polyakrylamidová gelová elektroforéza), jak jí známe dnes.

Dvě metody elektroforetické separace založené na dvou odlišných vlastnostech bílkovin tak byly na světě – izoelektrická fokusace (IEF) dělicí bílkoviny na základě jejich náboje či přesněji izoelektrického bodu (pI), a SDS-PAGE dělicí na základě molekulové hmotnosti. Zkombinovat tyto metody do dvojrozměrné separace a zásadně tak zvýšit rozlišení se podařilo v roce 1975 několika týmům najednou. Nejprecizněji však metodu zpracoval a za otce dvojrozměrné elektroforézy (2-DE) založené na pI a MW je považován Patrick O'Farrel, který na jediném gelu dokázal rozdělit 1100 proteinových skvrn lyzátu *E. coli*. (20).

O'Farrell tak otevřel cestu k dvojrozměrnému rozdělování, vizualizaci a kvantifikaci stovek až tisíců bílkovinných skvrn na jediném gelu. Byla to cesta, která s pomocí několika dalších zdokonalení umožnila, alespoň teoreticky, uvažovat o kompletním kvantitativním i kvalitativním popisu úplné proteinové sady buňky, tkáně či organismu. V následujícím odstavci se zaměřím na ta zmíněná, zdánlivě drobná, ale přesto velmi významná vylepšení dvojrozměrné elektroforézy (2-DE) bílkovin.

Jedním z důvodů detekce vysokého počtu skvrn byla skutečnost, že O'Farrel použil radioaktivně značené vzorky a radiografii. Jeho konkurenti byli odkázáni na barvení bílkovin v gelu pomocí relativně málo citlivého pigmentu Coomassie Brilliant Blue, používaného k detekci bílkovin již od roku 1965 (21). Na dostatečně citlivou, ale neradioaktivní metodu detekce bílkovin v gelech si museli vědci počkat až do roku 1979, kdy Switzer, Merril a Shifrin zavedli redukční barvení solemi stříbra a zvýšili tak detekční limit z mikrogramů na pouhé desítky nanogramů proteinu (22).

Jedním z nedostatků 2-DE v 70. a 80. letech byla nízká reprodukovatelnost, způsobená především nejednotným postupem syntézy amfolytů a migrací gradientu pH v elektrickém poli v průběhu izoelektrické fokusace. Klasický systém IEF navržený Svenssonem a Vesterbergem (14, 15) využíval totiž k vytvoření gradientu pH roztoky amfolytů, které ovšem v elektrickém poli migrovaly, a gradient pH se tak posouval a destabilizoval. Tento problém byl vyřešen v roce 1982 Bejellqvistem, který derivatizoval kyselými a bazickými skupinami (karboxy- a aminoskupinami) akrylamid a vyvinul tak imobilizované pH gradienty (IPG) (23). Metoda pro reprodukovatelnou 2-DE analýzu komplexních proteinových vzorků včetně citlivé detekce tak byla k dispozici začátkem 80. let a otevřela cestu k robustní analýze kompletní sady proteinů. Až o dvanáct let později bude poprvé použit termín „proteom“. Ale k tomu se ještě dostaneme.

Dnes je dvojrozměrná elektroforéza využívána prakticky ve stejné podobě jako tomu bylo v polovině osmdesátých let, po zavedení imobilizovaných gradientů. Podstatného vývoje se však dočkala vlastní zařízení k provozování 2-DE, algoritmy a software na analýzu obrazu z 2-DE gelů a především nové způsoby detekce proteinů v gelech. Z metod detekce stojí za zmínku fluorescenční metody barvení proteinů v gelu s detekčním limitem pod 1 ng proteinu (24) a především revoluční multiplexní metoda DiGE (Difference gel electrophoresis) založená na značení vzorků před elektroforézou jedním ze tří fluoroforů s odlišnými spektrálními charakteristikami (25). Vzorky pak mohou být smíchány a rozděleny pomocí 2-DE, gel totiž při skenování při různých vlnových délkách poskytne 3 různé obrazy. Metoda DiGE kromě vysoké citlivosti detekce také redukuje technickou variabilitu 2-DE experimentu a snižuje minimální počet gelů potřebných k vyhodnocení. Detekční limit DiGE je zhruba 0,1 ng proteinu, v nejcitlivějším, takzvaném saturačním uspořádání, vzroste citlivost metody ještě zhruba desetkrát (26).

Dvojměrná elektroforéza umožnila první pokusy o zmapování proteomu a odstartovala v devadesátých letech nástup proteomiky. Své specifické místo v proteomice si, i přes razantní nástup tzv. shot-gun technologií založených na chromatografické separaci, uchovala dodnes.

### Chromatografie

Podobně jako v případě elektroforézy se musíme obrátit směrem na východ, k carskému Rusku. Nebo přesněji řečeno na severovýchod. V roce 1903 patřila pod carskou moc také Varšavská univerzita, na jejíž půdě prezentoval, v první jarní den, botanik Michail Cvet své poznatky o rozdělování roztoku rostlinných pigmentů na křídovém adsorbentu (27). Protože pracoval s barevnými chlorofyly a karotenoidy, nazval tento jev barvopisem – chromatografií.

Cvetovy pionýrské publikace zůstaly bez zásadní odezvy a následovníky získaly až ve třicátých letech v osobách Richarda Kuhna a Edgara Lederera, kteří chromatograficky dělili různé přírodní látky v preparativním měřítku (28, 29). Skutečný boom chromatografie však představovaly až práce Archera Martina a Richarda Syngeho v letech čtyřicátých. Ti přeměnili Cvetův model, mimo jiné použitím silikagelu jako stacionární fáze, v metodu označovanou jako rozdělovací chromatografie (partition chromatography) (30). Chromatografie Syngeho a Martina byla již použitelná k dělení aminokyselin a peptidů.

Metoda separace přírodních látek na kolonách se silikagem trpěla kvůli vlastnostem silikagelu a manuálnímu plnění kolon nízkou reprodukovatelností. To vedlo Consdena, Martina a Syngeho ke vývoji metody, kde roli silikagelu jako stacionární fáze zastoupil list papíru (31). Světlo světa tak spatřila papírová chromatografie. Za zásadní přínos pro chemii byli v roce 1952 Martin a Synge odměněni Nobelovou cenou za chemii. Sami autoři demonstrovali výhody papírové chromatografie při separaci aminokyselin a oligopeptidů při sekvenaci peptidu gramicidinu S (32). O něco později ji v dvojměrném uspořádání využili pro dělení peptidů také Frederick Sanger (při sekvenaci inzulinu) a Vernon Ingram (při analýze hemoglobinu srpkovité anémie) (33, 7).

S nízkou reprodukovatelností separace na silikagelových kolonách byli nespokojeni také sovětsí farmakologové Nikolaj Izmailov a Maria Shraiber, když místo plnění kolon zkoušeli silikagel nanášet v tenké vrstvě na skleněné desky (34). Jejich práce, která dala vzniknout tenkovrstevné chromatografii, pochází již z roku 1938, metoda se však rozšířila až po roce 1956, kdy na ní navázal německý chemik Egon Stahl (35). Tenkovrstevná chromatografie si svůj význam, především pro separaci menších molekul, uchovala dodnes.

Zavedení malých sférických částic stacionární fáze v šedesátých letech a zároveň možnost chemické modifikace jejich povrchu vedly k zásadnímu nárůstu v rozlišení kapalinových chromatografií, ale také k nutnosti jejich provozování za vysokých tlaků. Svět vědy a průmyslu tak byl obohacen o HPLC (high performance (high pressure) liquid chromatography). Termín byl zaveden již v roce 1966 Csabou Horváthem (36), ujal se ovšem až později. Léta sedmdesátá přinesla porézní částice o velikosti několika mikrometrů (32) a chemicky modifikované stacionární fáze, především již dříve objevenou tzv. reverzní (obrácenou) fází (Reversed phase, RP) (38).

Právě vysokotlaká kapalinová chromatografie, provozovaná na kolonách s mikrolitrovými až nanolitrovými průtoky (nano-, mikro-HPLC), je dnes centrální metodou separace peptidů v proteomice. Je k tomu předurčena nejen vysokou rozlišovací schopností, ale také reprodukovatelností a flexibilitou a přímou

propojitelností s hmotnostními spektrometry. Další její výhodou je možnost vícerozměrného uspořádání, tedy postupné separace založené na dvou různých vlastnostech peptidů. K účinnému 2D rozdělování komplexních peptidových směsí v proteomice se nejčastěji kombinuje chromatografie na iontoměničích (různý náboj peptidů) se separací v reverzní fázi (RP) (různá hydrofobicita peptidů) (39, 40).

Přestože chromatografie dominuje v separaci peptidů, lze ji samozřejmě využívat i k dělení intaktních proteinů podle velikosti metodou gelové filtrace, nebo například při nativních separacích proteinových komplexů (41). Kromě klasické adsorpční nebo permeační chromatografie využívá proteomika poměrně často také chromatografii afinitní, založenou na biospecifické reverzibilní interakci zakotveného ligandu s proteinem či peptidem. Afinitní chromatografie slouží většinou k izolaci specifických proteinů, nebo skupin proteinů z komplexních směsí. Metoda byla poprvé použita v roce 1968 Pedrem Cuatrecasasem pro izolaci enzymů (42). Ligandem pro afinitní chromatografii může být jakákoliv specificky a dostatečně silně interagující molekula – například protilátka, metabolit, interagující protein nebo peptid a podobně. Nejčastějšími aplikacemi v proteomice je izolace biotinylovaných proteinů či peptidů avidinem, izolace fosfopeptidů na metal-afinitních kolonách, glykopeptidů na pektinech, nebo odstraňování majoritních složek vzorku pomocí specifických protilátek (43, 44).

Chromatografické separace peptidů a proteinů, především však mikrokapilární a nanokapilární HPLC dělení peptidů v reverzní fázi, představují základní metody moderní proteomiky. Vhodná chromatografická zařízení jsou proto neodmyslitelnou součástí výbavy dnešních proteomických laboratoří.

### Hmotnostní spektrometrie (Mass spectrometry, MS)

Na začátku této kapitoly se vrátíme do posledních let devatenáctého století. Vypravíme se pro změnu na severozápad, do Cambridge.

Za otce hmotnostní spektrometrie je právem považován fyzik J. J. Thomson (1886-1940), díky jeho zásadním pracím o vychylování katodového záření v elektrickém poli (45) a objevu a popisu elektronu. Thomson s pomocí Francise W. Astona sestrojil první hmotnostní spektrometr schopný měřit hmotnost nabitých atomů. Zařízení využívalo plynové výbojky jako zdroj iontů, které následně procházely elektrickým a magnetickým polem a byly vychylovány a detekovány na fotografické desce. Ashton s pomocí tohoto hmotnostního spektrometru zaznamenal izotopy některých prvků (46).

V průběhu prvních třiceti let 20. století došlo k postupnému vylepšení rozlišení hmotnostních spektrometrů, ale zásadní vliv na rozvoj metody měla až druhá světová válka svým požadavkem na separaci a charakterizaci izotopů uranu. Zde nelze nezmínit přínos Alfréda O. C. Niera, který s pomocí obřího hmotnostního spektrometru s magnetickým sektorem izoloval nanogramová množství  $^{235}\text{U}$  pro projekt Manhattan (47).

V polovině čtyřicátých let už byly hmotnostní spektrometry komerčně dostupné a dobývaly svou pozici esenciálních nástrojů fyziky a chemie. Jednalo se především o hmotnostní spektrometry s magnetickým sektorem nebo dvojitou fokusací (double-focussing), kde zdrojem iontů analytu bylo nejčastěji jeho bombardování elektrony. Hmotnostní spektrometry však v té době byly především nástroji kvantitativní analýzy. Léta padesátá představoval zásadní pokrok v instrumentaci. Koncept zcela nového analyzátoru typu TOF (měření doby letu, time-of-flight) byl zformován již v roce 1946 Williamem Stephensem (48), fyzické podoby

se však dočkal až v roce 1955 (49,50). Kvadrupólový analyzátor a kvadrupólová iontová past byly navrženy v roce 1955 (51).

Díky mnoha technickým vylepšením dokázal koncem padesátých let J. H. Beynon (52) identifikovat pomocí hmotnostní spektrometrie známé organické sloučeniny. O něco později Klaus Biemann určil dosud neznámou strukturu složité organické sloučeniny (53), stanovil pravidla pro fragmentaci peptidů a navrhl první metodu sekvenování peptidů pomocí MS, která se stala základem dnešních postupů k identifikaci proteinů (54). Analýza malých organických molekul pomocí MS se stala do 80. let rutinou. Velké molekuly, jako nukleové kyseliny a proteiny, však představovaly zásadní výzvu. Bylo třeba nalézt účinný, ale přitom „jemný“ způsob ionizace velkých molekul, který by nezpůsoboval jejich zásadní fragmentaci. To se podařilo koncem 80. let. John Fenn využil k ionizaci šetrnou metodu označovanou jako ESI (electrospray ionization) a demonstroval její účinnost na spektrech dimerů a monomerů albuminu (55, 56). Koichi Tanaka využil odlišný princip ionizace s pomocí laseru a změnil tímto způsobem spektra intaktního lysozymu a chymotrypsinu (57). Franz Hillenkamp a Michael Karas použili k ionizaci proteinů rovněž metodu laserové ionizace, ale v přítomnosti organické matrice a jsou tak považováni za otce metody MALDI (matrix-assisted laser desorption ionization) (58).

Vývoj „měkkých“ ionizačních technik ESI a MALDI měl zcela zásadní vliv na možnosti využití MS v biologii a byl kritickým objevem, který umožnil identifikaci proteinů i jejich kvantitativní analýzu. Další dveře ke studiu proteomu tak byly otevřené. S možností ionizovat nejen malé, ale i velké molekuly došlo k velmi rychlému vývoji na poli hmotnostních analyzátorů. Kvadrupólová iontová past se stala prvním z komerčně dostupných analyzátorů, který umožňoval identifikaci proteinů díky možnosti provozovat tandemovou hmotnostní spektrometrii peptidů vzniklých enzymatickým štěpením bílkovin (MS/MS). Následovaly hybridní přístroje kombinující kvadrupólový analyzátor s analyzátozem typu time-of-flight (q-TOF) (59). V roce 1993 série klíčových prací demonstrovala využití hmotnostní spektrometrie pro identifikaci proteinů nikoliv sekvenováním *de novo*, ale na základě porovnávání MS a MS/MS dat s teoretickými hmotnostmi peptidů a jejich fragmentů vypočtených na základě známých sekvencí kódujících genů (60-65).

Za nástup nejmodernějších hmotnostních analyzátorů typu FT-ICR (fourierovsly transformovaná iontová cyklotronová rezonance) v posledním desetiletí vděčíme především pionýrským pracím z let sedmdesátých (66). Z novodobých pokroků v konstrukci hmotnostních spektrometrů stojí jistě za zmínku analyzátor typu Orbitrap, vyvinutý Alexandrem Makarovem (67). Nejnovější hybridní hmotnostní spektrometry, které kombinují orbitrap s jinými typy analyzátorů umožňují velmi rychlé proteomické analýzy s identifikací několika tisíc proteinů v jediném experimentu (68, 69).

Hmotnostní spektrometrie, dříve výhradní nástroj fyziky, se v posledních deseti letech stala esenciálním nástrojem biologického výzkumu, především jako metoda identifikace proteinů a metabolitů. Centra biologického a medicínského výzkumu bez hmotnostních spektrometrů jsou ve světě dnes již těžko představitelná.

Za poznámku „pod čarou“ stojí fakt, že historie vývoje hmotnostní spektrometrie je ověněna několika Nobelovými cenami. Otec hmotnostní spektrometrie J. J. Thomson ji získal již v roce 1906, F.W. Aston v roce 1922, Wolfgang Paul s Hansem Dehmlem v roce 1989 a John Fenn s Koichi Tanakou v roce 2002.



## Od biochemie proteinů k moderní proteomice

Úvodním exkurzem do historie jsme stručně popsali vývoj tří základních nástrojů proteomiky. Pojďme se teď podívat na ranná období studia komplexních proteinových směsí a na vznik proteomiky jako souboru nástrojů, přístupů a konceptů biologického studia.

V polovině osmdesátých let byla již k dispozici dvojrozměrná elektroforéza se schopností reprodukovatelné separace, a díky citlivým metodám barvení i detekce, až několika tisíc proteinových skvrn. V té době však nebyly k dispozici dostatečně citlivé metody identifikace nanogramového množství neznámých proteinů obsažených v jednotlivých skvrnách.\* Snahy výzkumníků se proto ubíraly především směrem k vytvoření proteinové databáze založené na 2-DE obrazu (například Human Protein Index (70) a později 2-DE databáze J. E. Celise (71)) s předpokladem, že takto získané buněčné „mapy“ bude možné katalogizovat, porovnávat mezi sebou a získávat tak důležité biologické informace. To však bylo komplikováno a odsouzeno k neúspěchu nepředpokládanou rozsáhlostí a variabilitou proteomu, relativně nízkou inter-laboratorní reprodukovatelností 2-DE separací a dalšími překážkami.

V předchozím odstavci jsme zmínili, že v osmdesátých letech nebyly k dispozici dostatečně citlivé metody identifikace neznámých proteinů z gelů. Zde musíme pro upřesnění učinit krátkou odbočku a zmínit se o chemických metodách sekvenování proteinů, které byly vyvinuty již v letech čtyřicátých a padesátých. Přestože tyto postupy nemají v dnešní proteomice prakticky žádnou roli, genialita a nezdolná energie jejich tvůrců si to zcela jistě zaslouhují.

Anglický biochemik Frederick Sanger vypracoval již v roce 1945 metodu postupného sekvenování peptidů (72) založenou na derivatizaci N-koncových aminokyselin dinitrofenolem (DNP), hydrolýze peptidu, extrakci značené aminokyseliny a její charakterizaci.

Touto náročnou a zdlouhavou metodou stanovili v roce 1947 Consden, Gordon, Martin a Syngé aminokyselinovou sekvenci pentapeptidu gramicidinu S (32). Sám tvůrce DNP metody Sanger použil tento přístup začátkem padesátých let k dosažení velmi ambiciózního cíle – k určení kompletní aminokyselinové sekvence inzulínu. K separaci peptidů používal sofistikovaných postupů včetně elektroforézy a dvojrozměrné papírové chromatografie a po několika letech úsilí, v roce 1953, kompletní sekvenci inzulínu stanovil (33, 73-75).

Mezitím švédský biochemik Pehr Victor Edmann dokončil vývoji odlišného sekvenačního postupu založeného na značení N-koncové aminokyseliny fenylothiokyanátem, jejím selektivním odštěpením a stanovením (76). Edmannovo N-koncové odbourávání přineslo zásadní technologický posun, neboť poskytovalo sekvenční informaci o úsecích až 30 aminokyselin, bylo rychlejší a vyžadovalo podstatně menšího množství výchozího proteinu.

*\* Malá, ale dobře viditelná proteinová skvrna na 2-DE gelu barvená koloidní Coomassie Brilliant Blue představuje cca 1 pikomol průměrně velkého proteinu. Podobně velká skvrna po barvení stříbrem obsahuje pouze stovky až desítky femtomolů proteinu. Přenos proteinu z gelu na membránu či extrakce peptidů z gelu jsou navíc zatíženy značnými ztrátami a reálné množství získaného materiálu je tak zpravidla v řádech femtomolů.*

Metodu již bylo možné použít například na proteiny přenesené z gelu na membránu, citlivost však byla stále limitující. Metoda byla navíc omezena na proteiny s nemodifikovanými N-konci. Přestože se mnozí výzkumníci s větším či menším úspěchem pokoušeli v osmdesátých letech o N-koncové sekvenování proteinů z blotů po 2-DE (77), jednalo se o cestu trnitou. V nejlepším případě tento postup poskytoval informaci o sekvenci několika aminokyselin. Jejich znalost mohla posloužit k navržení sond k detekci genu, nebo mRNA, k jeho sekvenování, případně k zaklonování a následnému studiu. Situace se však zásadně změnila v devadesátých letech s dekodováním genomů, s genovými databázemi a s nástupem hmotnostní spektrometrie, jako nástroje přesného stanovení hmotností peptidů a jejich fragmentů.

V předchozím textu jsme se narazili na sekvenování genomů. Jakou roli však hraje genomika, nebo přesněji řečeno znalost kódujících sekvencí genomů? Dle převládajícího genocentrického pohledu je protein definován a identifikován na základě jeho kódujícího genu. (O problematice geno-centrického versus proteo-centrického pohledu se více rozepisují na str 26). Identifikace proteinů pomocí hmotnostní spektrometrie je zpravidla založena na porovnávání hmotnosti peptidů, či jejich fragmentů, s hmotnostmi teoretických peptidů kódovaných geny přítomnými v genových databázích. Identifikace bílkovin v proteomických experimentech je tedy těžko představitelná bez předchozího přečtení kódujících oblastí patřičného genomu, nebo alespoň genomu nějakého blízkce příbuzného organismu. Bez znalosti genomu můžeme sice pomocí hmotnostní spektrometrie získat sekvenováním *de novo* menší či větší část aminokyselinové sekvence proteinu, ale jednoznačně identifikovat ho můžeme jen tehdy, známe-li kódující gen.

Sekvenování genomů, nebo přesněji kódujících oblastí genomů, je tedy nutnou podmínkou identifikace proteinů pomocí hmotnostní spektrometrie a tedy i podmínkou existence proteomiky, tak jak ji dnes chápeme a používáme. Proto si dovoluji, především pro uvedení do historických souvislostí a doplnění celkového obrazu, alespoň velmi stručně zmínit některé zásadní milníky historie sekvenování DNA a celých genomů.

V roce 1977 publikovali Allan Maxam a Walter Gilbert svou metodu sekvenace DNA založenou na chemické degradaci (78) a ve stejnou dobu nezávisle na nich Frederick Sanger představil sekvenování založené na terminaci DNA řetězce (79). Sanger současně publikoval také první kompletní genom, sekvenci DNA bakteriofága  $\phi$ X174. (80). V roce 1991 byl v laboratořích Craiga Ventera spuštěn projekt mapování lidského genomu. První buněčný genom (bakterie *Haemophilus Influenzae*) se podařilo přečíst v roce 1995 (81). O rok později je dekodován genom kvasinky *Saccharomyces cerevisiae* (82) a v roce 1998 přečtena sekvence DNA prvního vícebuněčného organismu – haďátka *Caenorhabditis elegans* (83). V roce 2001 byla publikována předběžná verze lidského genomu (84), předběžná verze genomu myši byla zveřejněna v roce 2002 (85) o dva roky později bylo dokončeno sekvenování potkana (86).

V současnosti je podle databáze [genomesonline.org](http://genomesonline.org) kompletně osekvenováno 150 archebakterií, 2802 bakterií a 168 eukaryot. Vzhledem k neobyčejnému zrychlení a zlevnění technologií lze předpokládat, že počet plně sekvenovaných organismů, v případě člověka i jedinců, poroste v blízké budoucnosti velmi rychle.

Vraťme se však nyní zpátky k proteomice. Na sklonku osmdesátých a v první polovině devadesátých let se začala otevírat možnost identifikace proteinů nebo stanovení jejich sekvence jinak než *de novo*. Databáze kódujících úseků DNA a později i kompletních genomů začaly být k dispozici, a k rychlé identifikaci proteinů na základě znalosti kódujících genů tak chyběla jen dostatečně citlivá metoda k získání částečné informace o aminokyselinové sekvenci proteinu. Ta byla nalezena v hmotnostní spektrometrii.

Vývoj metod „měkké“ ionizace proteinů a peptidů (ESI a MALDI) odstartoval zcela novou éru studia bílkovin. Metody hmotnostní spektrometrie dokázaly se slušnou přesností stanovit molekulovou hmotnost celého proteinu, ta však byla k jednoznačné identifikaci bílkoviny nedostatečná. Ukázalo se však, že molekulové hmotnosti více peptidů, vzniklých štěpením izolované bílkoviny specifickou endoproteázou (například trypsinem), poskytují specifičtější informaci a mohou být využity k porovnávání s genovými a proteinovými databázemi a identifikaci proteinu tak poskytnout. V roce 1993 bylo publikováno několik průlomových studií, které tento přístup využily (s pomocí hmotnostních spektrometrů MALDI-TOF) a položily základ identifikace proteinů, například z 2-DE gelů. Tento postup označujeme jako **peptidové mapování nebo peptidový fingerprinting** (peptide mass fingerprinting, PMF) (60-63). Změřené hmotnosti peptidů jsou porovnávány s hmotnostmi všech teoretických všech (tryptických) peptidů, vypočítaných na základě kódujících sekvencí DNA přítomných v databázi.

Prakticky ve stejnou dobu, tentokrát s využitím ionizace ESI ve spojení s iontovou pastí nebo trojitým kvadrupólem bylo k identifikaci proteinu využito **tandemové hmotnostní spektrometrie (MS/MS) a fragmentace peptidů**. V tomto případě je nejprve stanovena molekulová hmotnost peptidu (první MS), ten je v hmotnostním spektrometru dále fragmentován (nejlépe v peptidické vazbě) a zároveň je stanovena hmotnost jednotlivých fragmentů (druhá MS). Jejich hmotnosti jsou pak opět porovnávány s databázemi a peptid a tedy i protein je tak identifikován. Právě algoritmy, které umožnily identifikaci proteinů v databázích na základě MS/MS dat, byly po peptidovém mapování druhým zásadním pokrokem v identifikaci proteinů z minimálních množství metodami hmotností spektrometrie (64, 65). Protože informace o hmotnostech fragmentů peptidu odráží aminokyselinovou sekvenci peptidu, což je z hlediska identifikace proteinu informací vysoce specifickou, je metoda MS/MS vhodná i pro identifikaci proteinů ze složité směsi peptidů, kde je peptidový fingerprinting nepoužitelný. Vzhledem k tomu, že takto získaná informace vypovídá o aminokyselinové sekvenci peptidu, je tato metoda někdy nepřesně označována jako mikrosekvenování peptidů.

Polovina devadesátých let, jak již bylo řečeno, byla ve znamení sestavení prvních genomů, plnicích se genových databází, a nově nabytých možností identifikace proteinů pomocí hmotnostní spektrometrie. V této atmosféře byl v roce 1994 na prvním 2-DE setkání v Sieně v Itálii poprvé použit termín „**proteom**“ pro proteinový komplement genomu (1). Formálně se tak zrodila proteomika jako obor usilující o kvalitativní a kvantitativní popis proteomu.

Dvojrozměrná elektroforéza byla v devadesátých letech v podstatě jediným nástrojem proteomiky. V druhé polovině devadesátých let, kdy se identifikace proteinů z 2-DE gelů pomocí MS staly zcela běžnými, se však ukázalo, že ani rozlišení několika tisíc proteinových skvrn není pro komplexní vzorky dostatečné a skvrny

často obsahují více než jeden protein (87). Tato skutečnost samozřejmě omezuje spolehlivost relativní kvantifikace proteinu založené na integrované optické densitě patřičné skvrny, jež je založena na předpokladu jediného proteinu ve skvrně. Když k tomu přičteme technickou i časovou náročnost, nízkou reprodukovatelnost a některá další omezení 2-DE, je zjevné, že bylo nutné hledat zcela nové přístupy. Když ne elektroforéza intaktních proteinů, tak co tedy?

Základy zcela novému přístupu, vyhýbajícímu se 2-DE, položila skupina D. F. Hunta v první polovině devadesátých let. Ta na příkladu antigenních peptidů uvolněných z MHC molekul demonstrovala obrovský potenciál kapalinové chromatografie spojené s tandemovou hmotnostní spektrometrií (LC-MS/MS) pro separaci a identifikaci složitých směsí peptidů (88, 89). Byla tak započata cesta k zásadní proměně proteomiky a o své místo na slunci se začaly hlásit tzv. „shot-gun“ metody (shot-gun, brokovnice).

Podstatou shot-gun přístupů je naštěpení komplexního proteinového vzorku na peptidy specifickou endoproteázou a separace peptidů (zpravidla) pomocí kapalinové chromatografie s následnou identifikací proteinů v databázích na základě MS/MS dat. Potenciál shot-gun metod dělit a identifikovat stovky, případně tisíce proteinů v jediném vzorku byl demonstrován v několika následujících letech (90-92 a další) na různých biologických vzorcích. Samozřejmě i shot-gun přístupy měly a stále mají, podobně jako 2-DE, svá omezení. Zvážíme-li, že v běžném savčím buněčném lyzátu po jeho naštěpení trypsinem existuje nejméně několik set tisíc odlišných peptidů, je jasné, že tak složitý vzorek vyžaduje buď **vícetupňovou separaci peptidů** nebo nějaký způsob **redukce počtu analyzovaných peptidů**.

Prvním ze způsobů řešení jsou dvojrozměrné chromatografické separace, předznamenané pionýrskou prací skupiny J. R. Yatese III z roku 2001 (93). Ta využívá k separaci tryptických peptidů z homogenátu kvasinky mikrokapilární chromatografii s iontoměničím v prvním rozměru, následovaným separací v reverzní fázi. Uvedenou technologií se podařilo identifikovat téměř 1500 bílkovin *Saccharomyces cerevisiae*. Tento postup zdomácněl pod označením „multidimensional protein identification technology“, zkráceně MudPIT; v různých technických obměnách představuje dodnes hlavní metodu shot-gun technologií. Své místo ve spektru shot-gun přístupů však mají i metody hybridní, využívající místo dvou chromatografických dělení například kombinaci izoelektrické fokusace peptidů s chromatografií (94), nebo kombinující SDS-PAGE s chromatografií v reverzní fázi (95).

Druhým způsobem vypořádání se s vysokou komplexitou vzorku je její snížení. Jedním z možných způsobů redukce počtu peptidů je izolace peptidů určitého typu na základě jejich specifických chemických vlastností. Lze například specificky izolovat peptidy obsahující –SH skupinu (96), peptidy fosforylované (97) nebo glykosylované (98). Oba přístupy, tedy vícerozměrné separace i redukce komplexity, jsou dnes hojně využívány a kombinovány a představují metodický základ moderní proteomiky.

## Kvantitativní hmotnostní spektrometrie

### Využití stabilních izotopů

Dvojměrná elektroforéza poskytuje na základě rozdílné optické hustoty proteinových skvrn informaci semikvantitativní, tedy informaci o rozdílech v expresi jednotlivých proteinů v porovnávaných vzorcích. Shotgun přístupy však kvantitativní data samy o sobě neposkytují. Intenzita MS signálu peptidu závisí, kromě množství ionizovaných peptidů, na mnoha dalších proměnných, mimo jiné na dalších složkách vzorku. Hmotnostní spektrometrie tak sama o sobě není metodou kvantitativní. Platí však, že intenzitu signálu dvou peptidů identického chemického složení, které se však z důvodu odlišného izotopického složení liší hmotností, lze v rámci jednoho spektra kvantitativně porovnávat (99).

Právě na tomto poznatku je založeno několik zásadních metod identifikace peptidů pomocí MS zároveň s jejich kvantifikací. Historicky první takovou metodou vyvinula skupina R. Aebersold již v roce 1999 (100). Metoda se vžila pod názvem „isotope-coded affinity tags“ (i-CAT). Je založena na chemickém značení proteinů jednou ze dvou specifických značek, které jsou chemicky identické, ale liší se zastoupením izotopů vodíku (H versus D). Značky zároveň obsahují biotin pro afinitní izolaci značených peptidů. Dva vzorky určené k porovnání jsou po označení na –SH skupinách smíchány, rozštěpeny trypsinem a značené peptidy jsou izolovány s pomocí biotinové značky. Izolované peptidy jsou pak podrobeny chromatografické separaci a identifikaci. Poměr intenzity MS signálu obou verzí (H a D) každého peptidu poskytuje relativní kvantitativní informaci o expresi proteinu ze kterého daný peptid pochází. Identita proteinu je stanovena na základě MS/MS spekter jednotlivých peptidů.

Na podobném principu, tedy vnesení hmotnostního rozdílu při zachování chemické identity peptidů, je založeno několik dalších kvantitativních metod. Velkou popularitu získala metoda označovaná zkratkou iTRAQ (isobaric tags for relative and absolute quantitation) založená rovněž na chemickém značení čtyřmi (v novějších obměnách dokonce až osmi) chemicky identickými, ale izotopicky odlišnými značkami (101, 102). Značení v tomto případě probíhá až na úrovni peptidů. Identifikace i kvantifikace značených peptidů pak probíhá na základě MS/MS spekter, kdy kromě fragmentace peptidu dochází k odštěpení značek a jejich relativní intenzita signálu vypovídá o množství peptidu v porovnávaných vzorcích.

Kromě iCAT a iTRAQ metod bychom měli zmínit také značení inkorporací  $^{18}\text{O}$  a  $^{16}\text{O}$  do peptidu při štěpení trypsinem v  $\text{H}_2^{18}\text{O}$  nebo  $\text{H}_2^{16}\text{O}$  a také značení volných aminoskupin dimethylací s použitím vodíku a deuteria (103, 104).

Kromě chemického značení proteinů či peptidů se zhruba ve stejnou dobu podařilo využít k relativní kvantifikaci proteinů i značení metabolického, tedy přirozené inkorporace izotopické značky do organismu žijícího v prostředí s izotopem. První takovou prací byla analýza kvasinek rostoucích v běžném ( $^{14}\text{N}$ ) médiu a médiu obsahujícím  $^{15}\text{N}$  (105). Pro savčí buňky byla později metoda vylepšena,  $^{15}\text{N}$  byl v médiu nahrazen izotopicky značenou esenciální aminokyselinou (106). Pro tuto velmi citlivou metodu zatíženou minimálním množstvím artefaktů se vžila zkratka SILAC (Stable isotope labeling by amino acids in cell culture) (107).

Metabolická inkorporace stabilního izotopu umožňuje semikvantitativní proteomickou analýzu mikroorganismů či buněk rostoucích v médiu. V posledních letech však byly částečně metabolicky označeny také vícebuněčné

zástupci rostlinné říše, ale také drosophila, hadíátko, pstruh, slepice, myš a potkan (108). Krmením myší syntetickou dietou se značenou aminokyselinou po dobu tří generací se v roce 2008 podařilo týmu Mathiasa Manna získat „SILAC myši“ kompletně proznačené  $^{13}\text{C}_6$  lysinem (109). SILAC myši jsou dnes již dostupné komerčně a tento přístup se i přes vysokou cenu rychle ujal, protože představuje citlivý a nejmenším množstvím artefaktů zatížený nástroj ke studiu globálních změn exprese bílkovin v eukaryotních organismech.

### Metody kvantifikace bez značení (Label-free methods)

Přestože metody kvantifikace pomocí MS založené na značení stabilními izotopy poskytují uspokojivé výsledky, nejsou prosty omezení a komplikací. Značením vzorku roste cena experimentu i jeho časová náročnost a použitím více kroků se zvyšuje možnost vzniku experimentálních artefaktů. Proto jsou již několik let zkoumány možnosti relativní kvantifikace proteinů ve směsi bez jejich značení.

V jednom z předcházejících odstavců jsme se zmínili o tom, že intenzita peptidového signálu ve směsi závisí na mnoha proměnných, mimo jiné na přítomnosti jiných peptidů soutěžících při ionizaci o náboj. To znamená, že hmotnostní analýza komplexních směsí není sama o sobě metodou kvantitativní. Je však nezpochybnitelné, že mezi množstvím peptidu a intenzitou MS signálu je přímý vztah. Je tedy možné derivovat semikvantitativní data z párového shot-gun experimentu s neznačenými proteiny?

V posledních letech se ukazuje, že pokud jsou dva vzorky A a B zpracovány zcela identickým způsobem a peptidy separovány (například dvojrozměrnou kapalinovou chromatografií) při zachování naprosto stejných separačních podmínek, potom lze relativní množství jednotlivých proteinů ve vzorcích A a B odvodit z počtu nebo intenzity MS nebo MS/MS spekter jednotlivých peptidů. Právě tohoto pozorování využívají dvě základní metody „label-free“ přístupů. Odvozují informaci o relativním množství proteinu na základě:

- 1) porovnávání intenzit MS odpovídajících peptidových signálů daného proteinu mezi vzorky A a B
- 2) porovnávání počtu MS/MS spekter získaných ve vzorcích A a B pro jednotlivé proteiny.

První přístup bývá označován jako kvantifikace dle integrované intenzity peptidových píků (**integrated measurement of chromatographic peak areas**) druhý jako kvantifikace počtem spekter (**spectral counting**) (110, 111). Spolehlivost obou metod je závislá na statistické analýze dat z více peptidů, podobně jako u metod s využitím stabilních izotopů.

Práce přímo porovnávající kvantifikace stejného vzorku label-free metodou s izotopovým značením potvrzují značnou korelaci kvantitativních dat těmito metodami získaných (112, 113). Lze tedy předpovídat, že label-free metody nabudou v blízké budoucnosti zásadního významu.

## **Analýza kvalitativních změn, proteinové komplexy a speciální metody.**

S výjimkou některých specifických a úzce zaměřených technik či instrumentací jsme si v historickém úvodu zmínili vše podstatné pro uvedení do současné popisné a expresní proteomiky. Neměli bychom však zapomínat, že kromě úrovně exprese dokáže proteomika popisovat i kvalitativní vlastnosti proteinů a jejich změny, a usiluje rovněž o analýzu funkčních proteinových komplexů. Protože těžiště této práce spočívá v expresní proteomice, tedy v kvantitativní analýze proteinové exprese, zmíníme se o těchto ostatních metodách jen velmi stručně.

### Analýza posttranslačních modifikací

Posttranslační modifikace mění fyzikálně chemické vlastnosti analyzovaných proteinů a ovlivňují tak, až na výjimky, separaci proteinů a peptidů a/nebo jejich hmotnostní spektra. Na základě těchto změn je založena většina metod pro globální analýzu posttranslačních modifikací.

Nejdéle studovanými posttranslačními modifikacemi jsou pravděpodobně fosforylace a glykosylace. Fosforylace posouvá izoelektrický bod proteinu směrem k nižšímu pI a různé fosforylované formy téhož proteinu tak lze často pozorovat na 2-DE jako horizontální řády skvrn - „vláčky“ (spot trains) porovnatelné molekulové hmotnosti. Na základě MS analýz peptidů z těchto skvrn lze, na základě změněné hmotnosti peptidu nebo jeho fragmentu, stanovit fosforylovaný peptid nebo dokonce určit konkrétní fosforylovanou aminokyselinu. Alternativně lze ke značení fosforylovaných proteinů použít také radioaktivní  $^{32}\text{P}$ , fosfo-specifické protilátky nebo také fosfo-specifické pigmenty.

Posttranslační proteolýza, tedy zkracování proteinu od jeho N- nebo C-konce vede ke změně jeho molekulové hmotnosti a při větší změně tedy může být rovněž pozorována jako změna migrace na 2-DE ve vertikálním směru.

V současnosti je však globální analýza posttranslačních modifikací převážně doménou shot-gun přístupů. Vlastní modifikace jsou zde identifikovány opět na základě změny hmotnosti peptidu nebo jeho fragmentu. Cílené analýze fosfopeptidů nebo glykopeptidů (nebo jinak modifikovaných peptidů) v komplexním vzorku většinou předchází nějaká metoda jejich obohacení. Nejčastěji je to jejich afinitní izolace, založená na odlišných fyzikálně chemických vlastnostech modifikovaných peptidů (např. izolace glykopeptidů pomocí lektinů, izolace fosfopeptidů pomocí specifických protilátek nebo metal-afinitní chromatografií) nebo na chemické derivatizaci posttranslačních modifikací (například biotinylace oxidovaných glykosylací).

O specifických metodách analýzy posttranslačních modifikací by bylo možné napsat mnoho stran, ale vzhledem k tomu, že těžiště této práce spočívá jinde, odkáží laskavé čtenáře na některý z mnoha přehledných článků, které se proteomikou posttranslačních modifikací zabývají (114-116).

### Analýza proteinových komplexů

Separace a izolace intaktních proteinových komplexů, případně nekovalentních komplexů proteinů s jinými molekulami, vyžaduje práci v nedenaturujících podmínkách. Tento základní imperativ tak samozřejmě definuje metody, které lze k separaci proteinových komplexů použít.

Klasickou metodou pro dělení nativních proteinů nebo proteinových komplexů je některá z mnoha variant nativní elektroforézy (117). Z hlediska rozlišení a potenciálního vícerozměrného uspořádání, musíme

zmínit především metodu Blue native elektroforézy (Blue native electrophoresis), ve které umožňuje migraci bílkovin (včetně velkých hydrofobních komplexů) vazba pigmentu Coomassie Blue (118). Blue native elektroforézu lze kombinovat s klasickou SDS-PAGE do dvojrozměrné elektroforézy a rozdělit tak separované komplexy na jednotlivé podjednotky (119). Blue native elektroforéza byla využita například při mapování proteinových komplexů v buněčném jádře nebo membránových proteinových komplexů v tenkém střevě (120, 121).

Nativní elektroforézu lze samozřejmě kombinovat do dvojrozměrných separací také s nativními chromatografiemi. Tento postup byl využit například k izolaci proteinových komplexů nekovalentně vázících železo (41).

Kromě tradičních proteomických přístupů založených na separačních metodách a jejich kombinacích, lze do škatulky „proteomika“ zahrnout také impresivní práce Anne C. Gavinové založené na globální afinitní purifikaci bílkovinných komplexů a identifikaci jejich jednotlivých složek (122, 123).

#### Proteinové čipy (protein arrays)

Jedním ze specifických a dosud minoritních směrů v proteomice jsou i tzv. proteinové čipy. Jedná se o obdobu DNA nebo RNA čipů, tedy o membrány či destičky (mikro nebo makro) se stovkami či tisíci bílkovin zakotvených v definovaných pozicích. Na proteinovém čipu by, alespoň teoreticky, mělo být možné zkoumat interakci proteinů nebo jiných látek se zakotvenými proteiny (124). Variantou proteinových čipů jsou tzv. reverse arrays (obrácené čipy) na kterých jsou zakotveny protilátky a lze tak studovat přítomnost specifických antigenů například v séru pacientů (125).

Zdali se podaří využít proteinových čipů zahrnujících stovky nebo dokonce tisíce bílkovin v základním výzkumu či dokonce v klinické diagnostice, ukáže čas.





## Maxima, limity a problémy současné proteomiky.

Zásadní technologický pokrok v posledních pěti letech, ztělesněný novými rychlými a přesnými hmotnostními spektrometry, umožnil přiblížit se k identifikaci proteinových produktů kódovaných všemi geny *Saccharomyces cerevisiae*. Skupina Mathiasa Manna identifikovala více než 4000 kvasničných bílkovin, přičemž počet ověřených otevřených čtecích rámců v kvasničném genomu je zhruba 4600 (126). Nejnovější práce na lidských buňkách identifikovaly více než 10 000 proteinů v jednom experimentu (68, 69), což je číslo, které se blíží počtu transkripčně aktivních genů v typické savčí buňce. Na druhou stranu je však třeba upozornit, že pro třetinu ze zhruba 20 400 lidských genů dosud nebyl žádný proteinový produkt nalezen (127).

Chceme-li posoudit limity či omezení jakékoliv techniky, musíme znát kvantitativní parametry zkoumaného objektu. Naskytá se tedy zcela zásadní otázka: **Jak velký je kvasničný, myší nebo lidský proteom?** Je-li proteom definován jako „proteinový komplement genomu“ nebo „soubor všech proteinů produkovaných v té či oné buňce, tkáni či jedinci“, musíme nutně vycházet z definice proteinu. Ať se to zdá překvapivé, současná biologie nemá zcela jasnou a obecně akceptovanou definici proteinu. Historická definice „jeden gen-jeden protein“ je překonána, jeden gen může kódovat více mRNA, ze kterých může vznikat značné množství strukturně i funkčně odlišných polypeptidů, jež mohou být dále diverzifikovány posttranslačně. Z jednoho genu tak vzniká (potenciálně velké) množství variantních polypeptidů, které se mohou lišit strukturou, lokalizací a funkcí. Liší-li se bílkovina strukturou, funkcí a lokalizací, můžeme pak ještě hovořit o jednom proteinu? Je tedy protein definován kódujícím genem? Strukturou? Funkcí?

Z převládajícího **geno-centrického** pohledu jsou všechny produkty téhož genu jen variantami téhož proteinu. Počet proteinů tak odpovídá počtu exprimovaných genů. Z pohledu **proteo-centrického** je však každá taková varianta individuální molekulou – proteinovou entitou, bytostí. Typickým příkladem „variantních proteinů“ jsou řetězce imunoglobulinů nebo modifikované varianty histonů. Různé proteinové molekuly s odlišnou strukturou, lokalizací a funkcí kódují i dobře známé geny jako například ENO1 (enoláza 1) (128) či GAPDH (129) nebo HSP27. Právě HSP27 známý jako gen pro heat shock protein 27 (HSPB1) dává vzniknout minimálně 30 takovým variantám odlišitelným na 2-DE (130). Pro tyto „varianty“, „formy“ či „proteinové produkty“, pro jejichž zařazení či nomenklaturu není dosud jasné pravidlo, byl navržen termín „protein species“ (131).

Kompletní lidský proteom má tedy z převládajícího geno-centrického pohledu pouze zhruba 20 400 proteinů, z hlediska proteo-centrického  $n$  krát 20 400, přičemž  $n$  je dosud neznámé číslo.

Jak je z předchozího textu patrné, má současná proteomika značné problémy koncepční, větší pozornost však věnuje problémům či omezením technickým. Neobyčejná komplexnost proteomu (i z geno-centrického pohledu) a vysoký dynamický rozsah koncentrací jednotlivých proteinů jsou stále hlavní technickou překážkou detekce všech proteinů nebo proteinových variant, které jednotlivé proteomy zahrnují.

Podívejme se teď stručně na technická omezení obou základních proteomických přístupů. Začneme s dvojrozměrnou elektroforézou: Vyjdeme-li z kombinace maximální možné nanášky vzorku, kterou lze rozdělit na běžném IPG-IEF gelu (500  $\mu$ g až 1 mg proteinu), a z maximální citlivosti kompatibilní detekce proteinů po SDS-PAGE, můžeme snadno odvodit, že bez předcházející vícestupňové frakcionace lze na 2-DE

gelu zobrazit pouze proteiny, které se v typické savčí buňce vyskytují nejméně v řádu desítek tisíc kopií. Tedy jen ty nejhojnější. Kromě toho je známo, že 2-DE elektroforéza není vhodná pro velmi velké a velmi malé proteiny (nad 150 kDa a pod 15 kDa) a rovněž pro proteiny s extrémním izoelektrickým bodem. Dalším, a to zcela zásadním omezením, je to, že 2-DE elektroforéza neumožňuje analýzu transmembránových proteinů z důvodu jejich vysoké hydrofobicity (132). Častý překryv proteinových skvrn, který brání kvantitativní analýze, představuje další komplikaci (133)

S trochou skepse lze konstatovat, že 2-DE elektroforéza umožňuje analýzu několika set nejhojnějších hydrofilních bílkovin, které představují jen malý zlomek celého, ať již bakteriálního nebo eukaryotického proteomu. V neprospěch 2-DE hovoří také nemožnost automatizace. Dvojměrná elektroforéza na druhou stranu uchovává informaci o posttranslačních modifikacích a jednotlivých variantách daného proteinu lišících se velikostí a/nebo nábojem. Další výhodou jsou menší nároky na MS analýzu z hlediska potřebného strojového času i typu hmotnostního spektrometru.

Kromě zmíněných omezení stojí v souvislosti s 2-DE za zmínku i dosud neprostudovaný fenomén, pro který se užívá termín „dělá vu“. Ukazuje se totiž, že bez ohledu na analyzovanou tkáň, nebo typ experimentu či biologický druh, ze kterého vzorek pochází, figurují na seznamu identifikovaných diferencially exprimovaných proteinů po 2-DE velmi často stejné bílkoviny (134). Dosud není jasné, zdali se jedná o biologický jev nebo o artefakt 2-DE analýzy. O fenoménu proteomického dělá vu se podrobněji rozepisují na straně 45.

Shot-gun metody umožňují, a to z hlediska počtu identifikovatelných proteinů, hlubší vhled do proteomu. Práce, které s pomocí jedno- nebo vícerozměrné separace peptidů v on-line kombinaci s rychlými a přesnými hmotnostními spektrometry identifikují a relativně kvantifikují 1500-5000 proteinů, jsou dnes již zcela běžné. Je však počet identifikovaných proteinů tím nejdůležitějším kritériem?

Proteiny jsou v shot-gun experimentech identifikovány na základě MS/MS jednotlivých peptidů přítomných ve směsi. Sekvenční informace obsažená v MS/MS spektrech je porovnávána s teoretickými hodnotami hmotností fragmentů všech peptidů přítomných v databázích. Protein, nebo přesněji řečeno gen, který protein kóduje, je pak nalezen na základě shody jednoho nebo více spekter s databází. Pouze větší či menší část sekvence proteinu je přítom identifikovanými peptidy pokryta. Slabinou tohoto přístupu je to, že nedokáže rozlišit různé varianty či „protein species“ kódované jedním genem. Nelze rozlišit verze proteinů vzniklé alternativním sestřihem (pokud je zachován čtecí rámec) či posttranslačními úpravami. Informace o úrovni exprese tak vypovídají o množství jednotlivých identifikovaných peptidů, ale nikoliv o konkrétní variantě proteinu složeného z těchto peptidů.

Zmínili jsme, že transmembránové proteiny není možné efektivně analyzovat 2-DE. Problém s analýzou transmembránových hydrofobních proteinů se do jisté míry týká i shot-gun přístupů. Přestože tyto postupy nepracují s intaktními proteiny, jejich uspořádání upřednostňuje hydrofilní peptidy. Pokud má transmembránový protein dostatečně velké mimomembránové domény, které poskytnou dostatek peptidů, může být v shot-gun experimentech identifikován. Na selektivní izolaci hydrofilních povrchových domén transmembránových proteinů plazmatické membrány jsou využívány účinné specifické postupy založené na značení povrchových domén na intaktních buňkách s následnou afinitní izolací (135). Problémem však zůstávají proteiny s malými, nebo trypsinu nepřístupnými extra-membránovými doménami. K jejich analýze nezbyvá než zvolit specifický postup selektivně zaměřený na hydrofobní transmembránové domény (136, 137)

Nejmodernější shot-gun metody umožňují detekci a kvantifikaci až několika tisíc bílkovin v jednom experimentu, a blíží se tak počtu všech genů exprimovaných v daný okamžik v dané tkáni. Z geno-centrického pohledu se tak blížíme popisu kompletního genomu, ale z pohledu proteo-centrického pohledu se dosud pohybujeme jen „po povrchu“, pokud ignorujeme jednotlivé „protein species“. Přes značná omezení 2-DE i shot-gun přístupů poskytují tyto metody velmi cenným vhled do buněčné a tkáňové fyziologie a patofyziologie a doplňují tak velmi vhodně metody fyziologie a buněčné a molekulární biologie.

## **Závěr**

Proteomika jako specifický soubor nástrojů, technik a přístupů ke studiu živé přírody se v posledním desetiletí etablovala v základním výzkumu v mnoha biologických oborech, ale i v aplikovaných zemědělských a potravinářských studiích a rovněž ve výzkumu medicínském. Od roku 1995 každoročně přibývá publikací, které proteomiku využívají nebo rozvíjejí. Jen za posledních 10 let se podle databáze Medline počet článků zveřejněných v jednotlivých letech zvýšil téměř devítinásobně (ze 747 v roce 2001 na 6246 v roce 2011).

Spolu s počtem proteomických laboratoří a publikovaných článků narůstá, především v posledních pěti letech, také počet kritických hlasů. Ty upozorňují především na různá technická omezení a limity současných metod, ale také na nedostatky koncepční. Kritické hlasy „zevnitř“, tedy z komunity odborníků-proteomiků však vypovídají především o tom, že proteomika již překonala stádium nekritické dětské fascinace sama sebou a stává se zralou metodou, která si uvědomuje své možnosti a kriticky zkoumá svá omezení a nedostatky.

Kromě technických limitů daných parametry separačních metod a hmotnostní spektrometrie je třeba mít na paměti, že bez jasné definice proteinu nelze definovat proteom. Proteomika přelomu milénia založená převážně na 2-DE separaci intaktních proteinů byla schopná odlišit jednotlivé varianty proteinů („protein species“ nebo „proteoforms“) kódované jedním genem. Moderní shot-gun metody, které v posledních letech nahradily klasickou 2-DE analýzu poskytují z hlediska počtu identifikovaných bílkovin hlubší vhled do proteomu, ale různé proteinové varianty kódované jedním genem rozlišit neumí (pokud se nejedná o změnu čtecího rámce). Cesta k řešení tohoto problému zřejmě povede zpátky k práci s intaktními proteinovými molekulami. Metody pro separaci i MS analýzu celých bílkovinných molekul jsou již k dispozici a tento směr budoucího vývoje proteomiky naznačují i některé nejnovější publikace (138).

**Proteomika je, stejně jako ostatní nástroje zkoumání živé přírody, metodou nedokonalou. Přesto však může poskytovat velmi cenné informace. Jejich kvalita však záleží především na vhodném uspořádání experimentu a na kritickém přístupu experimentátora.**

## Literatura

1. Wilkins MR, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotech Gen Eng Rev.* 1995;13:19.
2. Reuss FF. Notice sur un nouvel effet de lélectricite galvanique. *Mémoires de la Société Impériale des Naturalistes de l'Université Imperiale de Moscou.* 1809:327.
3. Tiselius A. The moving-boundary method of studying the electrophoresis of proteins. *Nova Acta Regiae Societatis Scientiarum Upsaliensis* 1930; Ser. IV,7(4).
4. Tiselius A. A new apparatus for electrophoretic analysis of colloidal mixtures. *Transactions of the Faraday Society* 1937; 33:524.
5. Pauling L, et al. Sickle cell anemia a molecular disease. *Science* 1949;110(2865): 543.
6. von Klobusitzky D, König P. Biochemische studien über die gifte der schlangengattung Bothrops. VI. *Arch Exp Pathol Pharmacol.* 1939;192:271.
7. Ingram VM. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature* 1956;178(4537):792.
8. Ingram VM. Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. *Nature* 1957;180(4581):326.
9. Gordon AH, et al. Electrophoresis of proteins in agar jelly. *Nature* 1949; 164 (4168): 498.
10. Grabar P, Williams CA. Method permitting the combined study of the electrophoretic and the immunochemical properties of protein mixtures; application to blood serum. *Biochim Biophys Acta* 1953;10:193.
11. Hjerten S. Agarose as an anticonvection agent in zone electrophoresis. *Biochim Biophys Acta* 1961;53:514.
12. Smithies O. Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *Biochem J.* 1955;61(4):629.
13. Poulik MD. Starch gel electrophoresis in a discontinuous system of buffers. *Nature* 1957;180:1477.
14. Svensson H. Isoelectric fractionation, analysis, and characterization of ampholytes in natural pH gradients. I. The differential equation of solute concentrations at a steady state and its solution for simple cases. *Acta Chem Scand.* 1961;15:325.
15. Vesterberg O. Synthesis and isoelectric fractionation of carrier ampholytes. *Acta Chem Scand.* 1969;23:2653.
16. Raymond S, et al. Acrylamide gel as an electrophoresis medium. *Nature* 1962;195:697.
17. Shapiro AL, et al. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys Res Commun.* 1967;28(5):815.
18. Pitt-Rivers R, Impiombato FS. The binding of sodium dodecyl sulphate to various proteins. *Biochem J.* 1968;109(5):825.
19. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227(5259):680.
20. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.* 1975;250(10):4007.

21. Meyer TS, Lamberts BL. Use of coomassie brilliant blue R250 for the electrophoresis of microgram quantities of parotid saliva proteins on acrylamide-gel strips. *Biochim Biophys Acta* 1965;107(1):144.
22. Switzer RC 3rd, et al. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal Biochem.* 1979;98(1):231.
23. Bjellqvist B, et al. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* 1982;6(4):317.
24. Weiss W, et al. Protein detection and quantitation technologies for gel-based proteome analysis. *Methods Mol Biol.* 2009;564:59.
25. Unlü M, et al. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997;18(11):2071.
26. Shaw J, et al. Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. *Proteomics* 2003;3(7):1181.
27. Tswett MS. O novoy kategorii adsorbtsionnykh yavleniy i o primenenii ikh k biokhimiicheskomu analizu. *Trudy Varhavskago Obshestva Estestvoispytatelei, Otdelenie Biologii* 1905;14(6):20.
28. Kuhn R, et al. Zur Kenntnis der Xanthophylle *Z Physiol Chem.* 1931;197:141.
29. Kuhn R., Lederer E. Über die Farbstoffe des Hummers (*Astacus gammarus* L.) und ihre Stammsubstanz, das Astacin. *Berichte der deutschen chemischen Gesellschaft* 1933;66:488.
30. Martin AJ, Synge RL. A new form of chromatogram employing two liquid phases: A theory of chromatography. 2. Application to the micro-determination of the higher monoamino-acids in proteins. *Biochem J.* 1941;35(12):1358.
31. Consden R, et al. Qualitative analysis of proteins: a partition chromatographic method using paper. *Biochem J.* 1944;38(3):224.
32. Consden R, et al. Gramicidin S: the sequence of the amino-acid residues. *Biochem J.* 1947;41(4):596-602.
33. Sanger F, Tuppy H. The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates. *Biochem J.* 1951;49(4):463.
34. Schreiber MS. 75 Years of Chromatography-an Historical Dialogue. Elsevier Scientific Publishing Company, Amsterdam-Oxford New York (1979) 413.
35. Stahl E. Thin-layer chromatography; methods, influencing factors and an example of its use. *Pharmazie* 1956;11(10):633.
36. Horvath CG, et al. Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers. *Anal Chem.* 1967;39(12):1422.
37. Majors RE. High-performance liquid chromatography on small particle silica gel. *Anal Chem.* 1972;44(11):1722.
38. Howard GA, Martin AJ. The separation of the C12-C18 fatty acids by reversed-phase partition chromatography. *Biochem J.* 1950;46(5):532.
39. Sandra K, et al. Highly efficient peptide separations in proteomics Part 1. Unidimensional high performance liquid chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2008;866(1-2):48.
40. Sandra K, et al. Highly efficient peptide separations in proteomics. Part 2: bi- and multidimensional liquid-based separation techniques. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2009;877(11-12):1019.

41. Babusiak M, et al. Identification of heme binding protein complexes in murine erythroleukemic cells: study by a novel two-dimensional native separation - liquid chromatography and electrophoresis. *Proteomics* 2005;5(2):340.
42. Cuatrecasas P, et al. Selective enzyme purification by affinity chromatography. *Proc Natl Acad Sci U S A.* 1968;61(2):636.
43. Azarkan M, et al. Affinity chromatography: a useful tool in proteomics studies. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2007;849(1-2):81.
44. Urh M, et al. Affinity chromatography: general methods. *Methods Enzymol.* 2009;463:417.
45. Thomson JJ. Cathode Rays. *Phil. Mag.* 1897;44:293.
46. Aston FW. The Mass-Spectra of Chemical Elements. *Phil. Mag.* 1919;37I:707.
47. Nier AO. Some reminiscences of mass spectrometry and the Manhattan Project. *J Chem Edu.* 1989;66(5):385.
48. Stephens W. Pulsed Mass Spectrometer with Time Dispersion. *Bull. Am. Phys. Soc.* 1946;21(2):22.
49. Katzenstein HS, Friedland WE. New Time-Of-Flight Mass Spectrometer. *The Rev of Sci Instrum.* 1955;26(4):324.
50. Wiley WC, McLaren IH. Time-Of-Flight Mass Spectrometer with Improved Resolution. *The Rev of Sci Instrum.* 1955;26(12):1150.
51. Paul W, Steinwedel H. Ein neues Massenspektrometer ohne Magnetfeld. *Z. Naturforschg.* 1953;8:448
52. Beynon JH. The use of the mass spectrometer for the identification of organic compounds. *Mikrochim Acta* 1956;44(1-3):437.
53. Bieman K., The determination of the carbon cytoskeleton of sarpagine by mass spectrometry. *Tetrahedron Lett.* 1960;15: 9.
54. Biemann K, et al. Determination of the amino acid sequence in oligopeptides by computer interpretation of their high-resolution mass spectra. *J. Am. Chem. Soc.* 1966;88(23):5598.
55. Yamashita M, Fenn JB. Electrospray Ion Source. Another Variation on the free-jet theme. *J Phys Chem.* 1984;88(20):4451.
56. Fenn JB, et al. Whitehouse C.M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* 1989; 246(4926):64.
57. Tanaka K, et al. Protein and polymer analysis up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 1988;2(8):151.
58. Karas M, et al. Influence of the Wavelength in High Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal Chem.* 1985;57:2935.
59. Glish GL, Goeringer DE. Tandem Quadrupole/Time-of-Flight Instrument for Mass Spectrometry/Mass Spectrometry. *Anal Chem.* 1984;56:2291.
60. Henzel WJ, et al. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci U S A.* 1993;90(11):5011.
61. Mann M, et al. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol Mass Spectrom.* 1993;22(6):338.

62. James P, et al. Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun.* 1993;195(1):58.
63. Yates JR 3rd, et al. Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem.* 1993;214(2):397.
64. Eng JK, et al. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom.* 1994;5:976.
65. Mann, M. Sequence database searching by mass spectrometric data. In *Microcharacterization of Proteins* (eds. Kellner, R., Lottspeich, F. & Meyer, H.E.) 223-245 (VCH, Weinheim, 1994).
66. Comisarow MB, Marshall AG. Fourier transform ion cyclotron resonance [FT-ICR] spectroscopy. *Chem Phys Lett.* 1974; 25(2):282.
67. Makarov A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal Chem.* 2000; 72(6):1156.
68. Beck M, et al. The quantitative proteome of a human cell line. *Mol Syst Biol.* 2011;7:549 (in press).
69. Nagaraj N, et al. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol.* 2011;7:548 (in press).
70. Anderson NG, Anderson L. The Human Protein Index. *Clin Chem.* 1982;28:739.
71. Celis JE, et al. Comprehensive, human cellular protein databases and their implication for the study of genome organization and function. *FEBS Lett.* 1989;244(2):247.
72. Sanger F. The free amino groups of insulin. *Biochem J.* 1945;39(5):507.
73. Sanger F, Tuppy H. The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates. *Biochem J.* 1951;49(4):481.
74. Sanger F, Thompson EOP. The amino-acid sequence in the glycyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates. *Biochem J.* 1953;53(3):353.
75. Sanger F, Thompson EOP. The amino-acid sequence in the glycyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates. *Biochem J.* 1953;53(3):366.
76. Edman P. A method for the determination of amino acid sequence in peptides. *Arch Biochem.* 1949;22(3):475.
77. Bauw G, et al. Protein-electroblotting and -microsequencing strategies in generating protein data bases from two-dimensional gels. *Proc Natl Acad Sci U S A.* 1989;86(20):7701.
78. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA* 1977;74(2):560.
79. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977;74(12):5463.
80. Sanger F, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 1977;265(5596):687.
81. Fleischmann RD, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 1995;269(5223):496.
82. Galibert F, et al. Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome X. *EMBO J.* 1996;15(9):2031.
83. *C. elegans* Sequencing Consortium, et al. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 1998;282(5396):2012.

84. Lander ES, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860.
85. Lindblad-Toh K, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002;420(6915):520.
86. Gibbs RA, et al. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* 2004;428(6982):493.
87. Gygi SP, et al. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A.* 2000;97(17):9390.
88. Hunt DF, et al. Characterization Of Peptides Bound To The Class-I Mhc Molecule Hla-A2.1 By Mass Spectrometry. *Science* 1992;255(5049):1261.
89. Appella E, et al. Analysis of the structure of naturally processed peptides bound by class I and class II major histocompatibility complex molecules. *EXS.* 1995;73:105.
90. Link AJ, et al. Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol.* 1999;17(7):676.
91. Mintz PJ, et al. Purification and biochemical characterization of interchromatin granule clusters. *EMBO J.* 1999;18(15):4308.
92. Spahr CS, et al. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics* 2001;1(1):93.
93. Washburn MP, et al. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol.* 2001;19(3):242.
94. Cargile BJ, et al. Immobilized pH gradient isoelectric focusing as a first-dimension separation in shotgun proteomics. *J Biomol Tech.* 2005;16(3):181.
95. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc.* 2006;1(6):2650-60.
96. Spahr CS, et al. Simplification of complex peptide mixtures for proteomic analysis: reversible biotinylation of cysteinyl peptides. *Electrophoresis* 2000;21(9):1635.
97. Ficarro SB, et al. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol.* 2002;20(3):301.
98. Hayes BK, et al. Specific isolation of O-linked N-acetylglucosamine glycopeptides from complex mixtures. *Anal Biochem.* 1995;228(1):115.
99. De Leenheer AP, Thienpont LM. Application of isotope dilution-mass spectrometry in clinical chemistry, pharmacokinetics, and toxicology. *Mass Spectrom Rev.* 1992;11:249.
100. Gygi SP, et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol.* 1999;17(10):994.
101. Unwin RD, et al. Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. *Mol Cell Proteomics* 2005;4(7):924.
102. Choe L, et al. 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* 2007;7(20):3651.
103. Mirgorodskaya OA, et al. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun Mass Spectrom.* 2000;14(14):1226.



104. Hsu JL, Huang SY, Chow NH, Chen SH. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem.* 2003 Dec 15;75(24):6843.
105. Oda Y, et al. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci U S A.* 1999;96(12):6591.
106. Veenstra TD, et al. Proteome analysis using selective incorporation of isotopically labeled amino acids. *J Am Soc Mass Spectrom.* 2000;11(1):78.
107. Ong SE, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;1(5):376.
108. Gouw JW, et al. Quantitative proteomics by metabolic labeling of model organisms. *Mol Cell Proteomics* 2010;9(1):11.
109. Krüger M, et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* 2008;134(2):353.
110. Neilson KA, et al. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 2011;11(4):535.
111. Zhu W, et al. Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol.* 2010;2010:840518.
112. Patel VJ, et al. A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res.* 2009;8(7):3752.
113. Collier TS, et al. Comparison of stable-isotope labeling with amino acids in cell culture and spectral counting for relative quantification of protein expression. *Rapid Commun Mass Spectrom.* 2011;25(17):2524.
114. Johnson H, Eyers CE. Analysis of post-translational modifications by LC-MS/MS. *Methods Mol Biol.* 2010;658:93.
115. Witze ES, et al. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* 2007;4(10):798.
116. Kiernan UA. Quantitation of target proteins and post-translational modifications in affinity-based proteomics approaches. *Expert Rev Proteomics* 2007;4(3):421.
117. Wittig I, Schägger H. Native electrophoretic techniques to identify protein-protein interactions. *Proteomics* 2009;9(23):5214-23.
118. Schägger H and Jagow G. Blue native electrophoresis for isolation of membrane protein complexes. *Anal Biochem.* 1991; 19:223-231.
119. Krause F. Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (Membrane) protein complexes and supercomplexes. *Electrophoresis* 2006 Jul;27(13):2759-81.
120. Novakova Z, et al. Separation of nuclear protein complexes by blue native polyacrylamide gel electrophoresis. *Electrophoresis* 2006;27(7):1277-87.
121. Babusiak M, et al. Native proteomic analysis of protein complexes in murine intestinal brush border membranes. *Proteomics* 2007;7(1):121-9.
122. Gavin AC, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 2002;415(6868):141-7.

123. Kühner S, et al. Proteome organization in a genome-reduced bacterium. *Science* 2009;326(5957):1235-40.
124. Stoevesandt O, et al. Protein microarrays: high-throughput tools for proteomics. *Expert Rev Proteomics* 2009;6(2):145-57.
125. Sanchez-Carbayo M. Antibody microarrays as tools for biomarker discovery. *Methods Mol Biol.* 2011;785:159-82.
126. de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 2008;455(7217):1251.
127. Legrain P, et al. The human proteome project: current state and future direction. *Mol Cell Proteomics* 2011;10(7): (in press)
128. Pancholi V. Multifunctional alpha-enolase: its role in diseases. *Cell Mol Life Sci.* 2001;58(7):902.
129. Hara MR, Snyder SH. Nitric oxide-GAPDH-Siah: a novel cell death cascade. *Cell Mol Neurobiol.* 2006;26(4-6):527.
130. Jungblut PR, et al. The speciation of the proteome. *Chem Cent J.* 2008;2:16.
131. Hoehenwarter W, et al. The necessity of functional proteomics: protein species and molecular function elucidation exemplified by in vivo alpha A crystallin N-terminal truncation. *Amino Acids.* 2006;31(3):317.
132. Wilkins MR, et al. Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number. *Electrophoresis* 1998;19(8-9):1501.
133. Gygi SP, et al. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A.* 2000;97(17):9390.
134. Petrak J, et al. Déjà vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins. *Proteomics* 2008;8(9):1744.
135. Hofmann A, et al. Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells. *Blood* 2010;116(13):e26-34.
136. Blackler AR, et al. A shotgun proteomic method for the identification of membrane-embedded proteins and peptides. *J Proteome Res.* 2008;7(7):3028-34.
137. Blackler AR, Speers AE, Wu CC. Chromatographic benefits of elevated temperature for the proteomic analysis of membrane proteins. *Proteomics* 2008;8(19):3956-64.
138. Tran JC, et al. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* 2011;480(7376):254-8.

