

Proteomické metody a hmotnostní spektrometrie

Proteomika je vědní disciplína zkoumající proteom, podobně jako např. genomika zkoumá genom. Zatímco genom je relativně stálý, proteom je velmi dynamický. Různé proteiny přítomné v různých buňkách určují, jak buňka vypadá a jaké funkce vykonává. Znalost proteomu nám umožňuje sledovat biologické děje na molekulární úrovni a porozumět změnám, ke kterým dochází v reakci na různé podněty nebo během stárnutí či onemocnění. Vzhledem k obrovské dynamice proteomu a jeho komplexitě není v současné době možné zmapovat všechny formy proteinů přítomné ve studovaném biologickém objektu. Přesto, za využití moderních proteomických technik a jejich kombinací, můžeme do proteomu stále detailněji nahlížet.

Před 30 lety, v r. 1995, australský molekulární biolog Marc R. Wilkins zavedl termín proteom, jímž označoval veškeré proteiny (bílkoviny) produkované v buňce nebo tkáni na základě genetické informace uložené v genomu. Již tehdy spolu s kolegy upozorňovali na významné rozdíly v definování proteomu ve srovnání s genomem. Genom je, pokud neuvažujeme mutace, ve většině buněk lidského těla stejný – s výjimkou např. pohlavních buněk, které mají jen poloviční počet chromozomů, nebo červených krvinek (erytrocytů), které nemají jádro a mitochondrie, a tedy ani jadernou či mitochondriální DNA. Oproti tomu proteom je vysoce dynamický a liší se významně v čase, např. v reakci na nějaký podnět, jako je infekce, hladovění, stárnutí apod., i prostoru – např. neurony mozkové kůry mají jiný proteom než neurony mozečku. Pokud mluvíme o proteomu, myslíme tím tedy soubor všech proteinů přítomných v konkrétním čase v konkrétním prostoru – např. v buněčné organelle, buňce, tkáni nebo v celém organismu. Proteiny představují

skutečný fenotyp, proto je studium proteomu (proteomika) zásadní pro porozumění fyziologickým i patologickým stavům.

Mnoho tváří proteomu

U člověka známe přibližně 20 700 genů, které kódují proteiny. Počet různých druhů proteinů je však výrazně vyšší než počet genů. Jak je to možné? Při prepisu z DNA do RNA (transkripci) dochází u savců u řady genů k alternativnímu sestřihu vznikající molekuly mediátorové RNA (mRNA), kdy z každého genu jsou průměrně čtyři různé transkripční varianty. Alternativní sestřih mRNA tak dává vzniknout více než 80 tisícům různých proteinových izoform. Proteiny kódované stejným genem, ale jinou transkripční variantou, izoformy proteinu, mohou mít zcela odlišné funkce. Příkladem je gen *ENO1*, který může kódovat cytoplasmatický enzym dráhy glykolýzy s délkou 434 aminokyselin, ale také jaderný DNA-vazebný protein s délkou 341 aminokyselin.

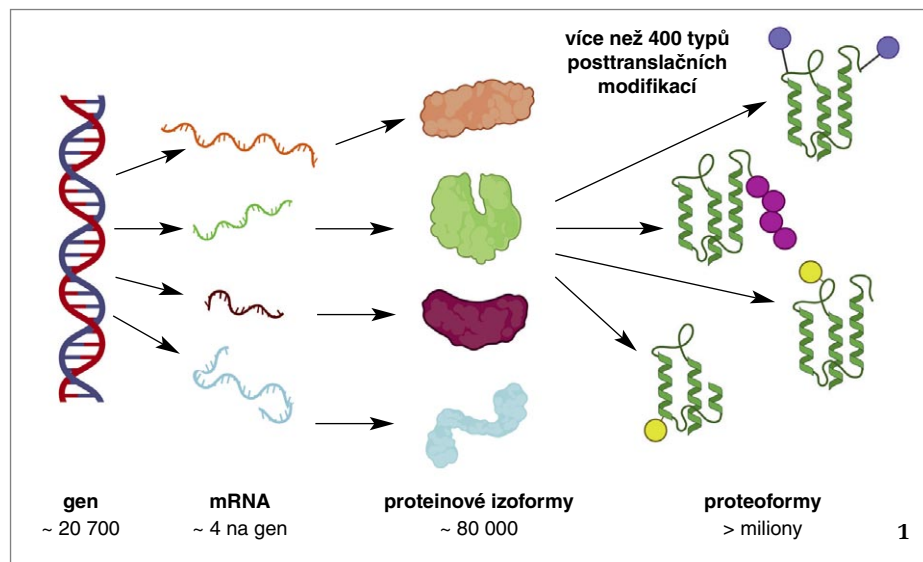
Regulace buněčných drah a procesů vyžadují možnost rychlé změny aktivity

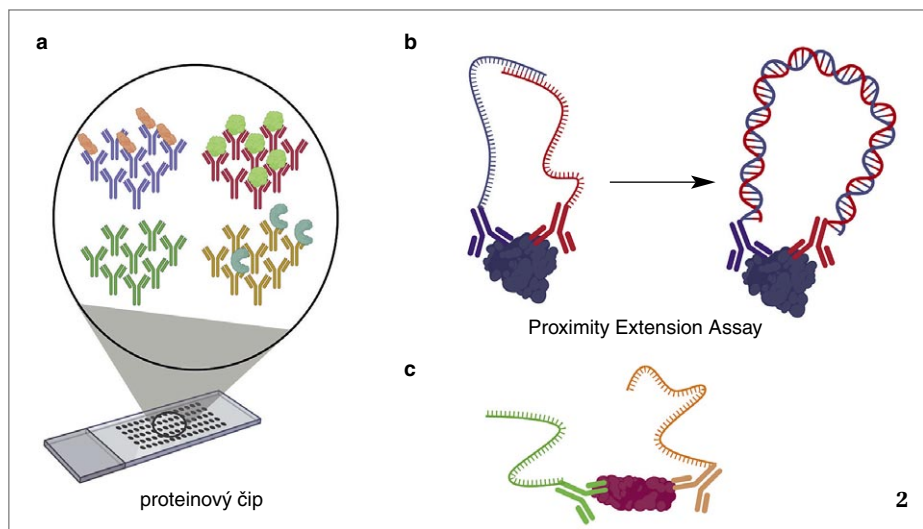
1 Od genu po proteoformu. Lidský genom obsahuje přibližně 20 700 genů, které kódují proteiny. Každý gen má průměrně čtyři transkripční varianty dané alternativním sestřihem, což umožňuje vznik více než 80 tisíc proteinových izoform. Proteiny poté procházejí posttranslačními modifikacemi (úpravami, blíže v textu), jichž existuje více než 400 různých typů a vedou ke vzniku milionů unikátních proteoformů. Proteoformy: modré kolečko – glykosylace, fialové – ubiquitinace, žluté – fosforylace, zkrácení proteinu – proteolýza

2 Protilátkové proteomické metody. Proteinový čip (a) – protilátky jsou imobilizovány na sklíčku a specifickou vazbou rozpoznávají antigeny (proteiny ve vzorku). Proximity Extension Assay (PEA, b) – analyt (protein) je rozpoznáván pomocí dvou protilátek, které na sobě mají jednovláknovou DNA s komplementárním úsekem k vlákně DNA na druhé protilátce. Po vazbě na protein se dostanou do vzájemné blízkosti, dojde k hybridizaci DNA a syntéze dvouvláknové DNA pomocí polymerázové řetězové reakce (PCR). PEA (c) – pokud se na protein navážou protilátky nespecificky, nebudou jednovláknové DNA vzájemně komplementární a k hybridizaci nedojde. Vytvořeno v programu BioRender, orig. J. Červenka (obr. 1 a 2, odkazy uvedeny na webové stránce Živý)

(funkce) nebo lokalizace proteinů. Protože syntéza nových proteinů nebo odbourání již přítomných jsou časově a energeticky náročné (u savců je přibližně 30 % veškeré energie spotřebováno na syntézu proteinů), mají buňky jiný mechanismus pro rychlou regulaci proteinů – posttranslační modifikace (blíže v článku Ze života bílkovin aneb Na tvaru záleží, Živa 2020, 6: 282–285). Posttranslační úpravy umožňují velmi rychlou a přesnou regulaci aktivity a lokalizace proteinů v buňkách i mimo ně. Při těchto úpravách dochází ke změně chemického složení molekuly proteinu. Typickými příklady posttranslačních modifikací jsou odstranění iniciačního metioninu na N-konci proteinu, fosforylace nebo defosforylace proteinu (navázání nebo odstranění fosfatové skupiny, které se využívá např. pro regulaci signálních drah) či glykosylace (navázání molekuly nebo více molekul sacharidu na vybrané aminokyseliny). Stejná izoforma proteinu může mít v různém čase nebo v různých buňkách jiné posttranslační modifikace – může mít různé typy úprav a ty mohou být na různých aminokyselinách i v jiné intenzitě. Různé varianty stejného proteinu, kódovaného tedy jedním genem, se označují jako proteoformy nebo proteinové druhy.

V současnosti bylo popsáno u savců více než 400 různých typů posttranslačních modifikací, což dále zásadně zvyšuje komplexitu proteomu. Jestliže u člověka předpokládáme více než 80 tisíc proteinových izoform, a pokud započítáme i posttranslační varianty proteinů, dostáváme se na úroveň milionů až miliard proteoformů v lidském proteomu (obr. 1). Pro kompletní zmapování všech proteoformů v dané buň-





ce i tkáni nám však zatím chybí analytické nástroje. Přesto je právě studium proteoformů jedním z cílů moderní proteomiky.

Kromě komplexity proteomu způsobené rozmanitostí jednotlivých proteoformů komplikuje studium proteomu také obrovský rozptýl v zastoupení jednotlivých proteinů, tzv. dynamický rozsah koncentrací proteinů. Zatímco některé proteiny, typicky cytoskeletu nebo enzymy základních metabolických drah, jsou v buňkách zastoupeny ve velkém množství kopií, jiných, např. některých proteinů signálních drah nebo transkripčních faktorů, je v buňkách velmi málo. Ještě více se rozdíl projevuje při studiu krevní plazmy, kdy rozdíl koncentrace mezi vysoce zastoupenými a málo zastoupenými proteiny dosahuje i více než 10 řádů (10^{10}). Zatímco interleukin-6 (signální molekula regulující zánět) je v 1 ml krevní plazmy zdravého člověka přítomen v řádu jednotek pikogramů (pg), albumin (hlavní protein krevní plazmy) je v řádu desítek mg (viz také Živa 2013, 2: 50–53, obr. 2). Analyzovat celkový proteom takto složitě směsi vyžaduje pokročilé separační a analytické metody. Žádná současná technika ale neumožňuje rozlišit a přesně kvantifikovat proteom v celém jeho dynamickém rozsahu koncentrací proteinů v rámci jediné analýzy.

Metody studia proteomu

V proteomice se k identifikaci a kvantifikaci proteinů využívají zejména protilátkové metody a hmotnostní spektrometrie (anglicky Mass Spectrometry, MS). Každý z těchto přístupů má své výhody i svá omezení, a proto se pro studium proteomu často používají oba. Některé moderní metody, např. hmotnostní cytometrie, kombinují obě techniky v rámci jedné analýzy. Protilátkové metody obvykle analyzují celé (intaktní) proteiny, metody založené na hmotnostní spektrometrii mohou také analyzovat intaktní proteiny, ale častěji jsou předem enzymaticky štěpeny na menší peptidy, které jsou pro analýzu vhodnější.

● Protilátkové metody

Jsou nejstaršími a zároveň nejrozšířenějšími technikami studia konkrétních předem vybraných proteinů. Metod s využitím protilátek je celá řada, my se zde zaměříme na multiplexní postupy, které umožňují sledovat více proteinů zároveň a zmapovat

tak významnější část proteomu v jednom experimentu.

Protilátky vysoce specificky rozpoznávají svůj antigen (v angličtině protein of interest – náš protein zájmu) a dokážou se na něj navázat i v přítomnosti mnoha jiných proteinů, které mohou být kvantitou i ve velkém nadbytku, a za využití neprotilátkových technik by náš protein zájmu mohly maskovat. Protilátkové metody tak lze s výhodou použít i ke sledování proteinů přítomných v daném vzorku v malém počtu kopií, např. zmíněného interleukinu-6 a dalších signálních molekul.

Typickým příkladem protilátkové metody je ELISA (Enzyme-Linked Immunosorbent Assay), která se široce uplatňuje v klinických a výzkumných laboratořích (viz článek Biochemické metody využívané v lékařství 1., Živa 2025, 2: 73–76). ELISA je však metodou biochemickou, protože umožňuje analyzovat pouze jeden protein. Pro studium proteomu je nutné analyzovat současně, v jednom experimentu nejméně stovky, lépe však tisíce proteinů ve vzorku. Toho je dosaženo tím, že primární protilátky jsou imobilizovány na destičku (např. mikroskopické sklo) v předem určených pozicích. Proteiny ve studovaném vzorku jsou fluorescenčně označeny a přidány na sklíčko, kde protilátky vychytají příslušné antigeny (proteiny). Kvantifikace probíhá na základě intenzity signálu fluorescence jednotlivých pozic protilátek na sklíčko. Tuto techniku označujeme jako proteinový čip (protein microarray) a je obdobou DNA čipu (DNA microarray) využívaného třeba pro sledování genové exprese (obr. 2a).

Moderní protilátkovou metodou je technika Proximity Extension Assay (PEA) založená na dvou specifických protilátkách rozeznávajících dvě různá místa (epitopy) sledovaného proteinu. Každá z protilátek má na sobě navázanou jednovláknovou DNA, která je komplementární k vlákně DNA na druhé protilátce. Po vazbě obou protilátek na antigen se obě vlákna DNA dostanou do těsné blízkosti a hybridizují spolu (obr. 2c). Hybridizovaná dvouvláknová DNA funguje jako primer pro navažující polymerázovou řetězovou reakci (Polymerase Chain Reaction, PCR), při níž dojde k syntéze celé dvouvláknové DNA (obr. 2b). Každá dvojice protilátek rozpoznávající jeden protein má specifickou

sekvenci přiřazené DNA, podle které je protein identifikován, a množství syntetizované konkrétní dvouvláknové DNA je přímo úměrné množství daného proteinu ve vzorku. Identifikace a kvantifikace proteinu se následně provádí sekvenací nebo kvantitativní PCR – díky amplifikaci signálu v tomto posledním kroku je metoda vysoce citlivá i pro málo zastoupené proteiny. PEA metoda byla nedávno použita k analýze 3 000 proteinů v krevní plazmě 54 tisíc pacientů (Sun a kol. 2023) a v letošním roce byla ve Velké Británii zahájena rozsáhlá studie 5 400 proteinů v 600 tisících vzorců krevní plazmy. První výsledky ukázaly, že změny hladin některých proteinů by mohly sloužit jako biomarkery pro posuzování rizika rozvoje určitých onemocnění. Např. u nádorů plic nebo některých leukemií bylo možné zaznamenat změny hladin vybraných proteinů v krvi již 7 let před stanovením diagnózy.

Alternativou k protilátkovým metodám je využití aptamerů namísto protilátek. Aptamery jsou syntetické jednovláknové DNA molekuly s modifikovanými nukleotidy, které se podobně jako protilátky vážou na cílové proteiny. Specifita vazby aptameru na hledaný protein je zajištěna sekvencí a chemickými modifikacemi nukleotidů v aptameru. Identifikaci proteinů provádíme na základě sekvenace aptameru, která je pro každý protein unikátní. Aptamery jsou fluorescenčně značené a kvantifikace se provádí pomocí DNA čipu. V současnosti jsou k dispozici aptamerové metody pro analýzu až 11 tisíc proteinů.

Hlavní výhodou protilátkových (a aptamerových) metod je jejich vysoká citlivost – díky specifické vazbě na hledaný protein umožňují ve vzorku nalézt i málo zastoupené proteiny, jako např. růstové faktory, chemokiny a cytokiny. Tyto metody se dají využít i k rozpoznání konkrétních proteoformů, např. proteinů fosforylovaných v určitém místě, což nám může poskytnout nepřímou informaci i o aktivitě cílového proteinu. Princip a provedení protilátkových metod jsou v porovnání s metodami hmotnostní spektrometrie srozumitelné i bez předchozích zkušeností s pokročilými předchemickými technikami. Navíc využívají technologie a postupy analýzy dat, které jsou dlouhodobě zavedené a používané pro genomiku a transkriptomiku. Na druhou stranu, protilátkové (a aptamerové) metody umožňují detekci a kvantifikaci pouze předem vybraných proteinů a proteoformů – pokud není protein rozpoznáván protilátkou/aptamerem, nemůže být studován. Zásadní nevýhodou těchto metod je však problematická specifita protilátek (aptamerů), které nemusejí rozpoznávat pouze konkrétní protein, případně mohou interagovat mezi sebou a narušit tak specifitu detekce.

● Metody separace proteomu

pro následnou hmotnostní spektrometrii
Jak jsme si již ukázali, proteom je vysoce komplexní a dynamická směs proteinů. Pro analýzu proteomu pomocí hmotnostní spektrometrie je proto nutné tuto směs zjednodušit – rozdělit (separovat). Dvěma základními přístupy separace proteinů jsou dvojrozměrná gelová elektroforéza a kapalinová chromatografie.

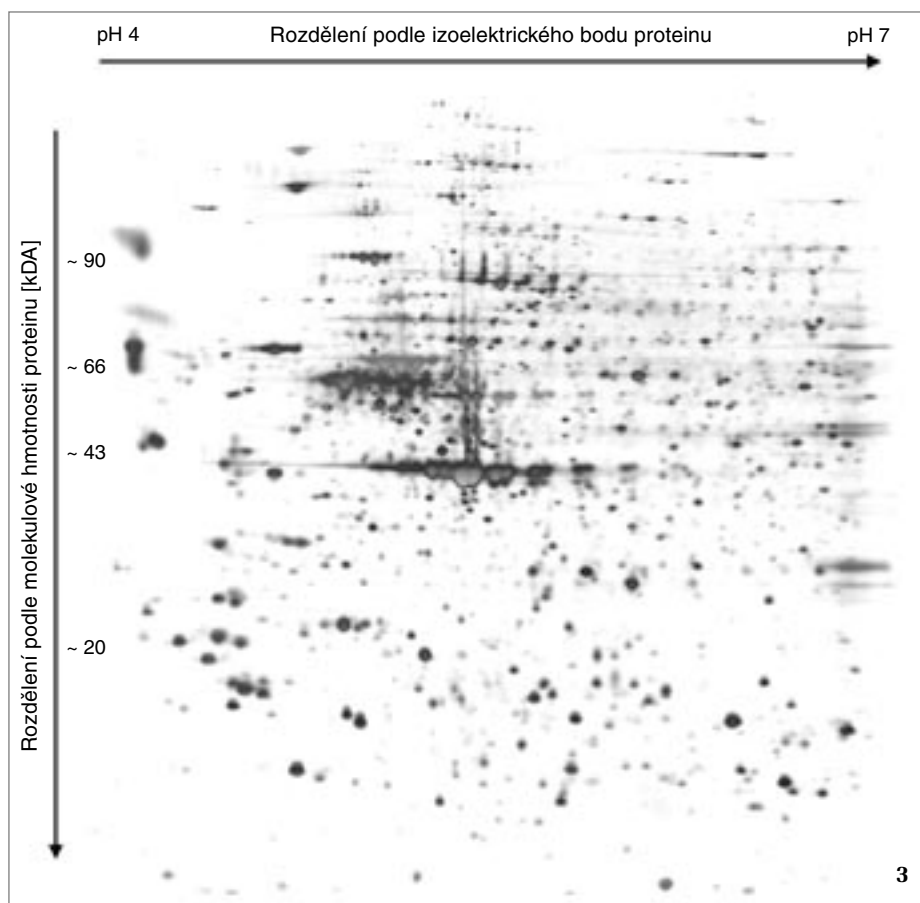
K separaci proteinů na základě izoelektrického bodu (pI) a molekulové hmotnosti slouží dvojrozměrná gelová elektroforéza (2-DE). Proteiny se nejprve rozdělí podle svého pI pomocí izoelektrické fokusace v úzkém proužku gelu s gradientem pH. Jakmile protein doputuje do oblasti, kde pH gelu odpovídá pI proteinu, celkový náboj proteinu bude roven nule a pohyb proteinu v elektrickém poli v gelu se zastaví. Proužek gelu s fokusovanými proteiny se následně umístí na polyakrylamidový gel a elektroforézou s dodecylsulfátem sodným (SDS-PAGE) se proteiny rozdělí na základě své velikosti, resp. molekulové hmotnosti. Před nebo po separaci jsou obarveny, což umožní určit jejich polohy a intenzitu proteinové skvrny v gelu (obr. 3). Obrazovou analýzou pomocí speciálních programů lze poté nalézt proteinové skvrny s odlišnou intenzitou mezi vzorky, ty z gelu vyříznout a daný protein identifikovat hmotnostní spektrometrií. Metoda 2-DE poskytuje cennou informaci o izoelektrickém bodu proteinů, velikosti a případně i posttranslačních modifikacích (ovlivňují separaci v obou rozměrech). Nevýhodou je složité a časově náročné provedení, nízká reprodukovatelnost a poměrně malý počet kvantifikovatelných proteinů – obvykle okolo 1 000 proteinů zastoupených ve vzorcích v největším množství.

Nejrozšířenější separační metodou současných proteomických experimentů je vysokoúčinná kapalinová chromatografie (High-Performance Liquid Chromatography – HPLC) s reverzní fází, při které dochází k interakcím separovaných molekul se dvěma fázemi – pevnou fází ve formě kuliček s hydrofobním povrchem (nejčastěji se používá povrch tvořený nasycenými osmnáctuhlíkovými řetězci) a mobilní fází, kapalinou, která kolem kuliček protéká a unáší peptidy do hmotnostního spektrometru (viz obr. 2 na str. CLXXXVI kuléru).

Po nanesení vzorku se peptidy rozpuštěné ve vodě navážou na hydrofobní povrch kuliček. Postupně je přiváděna mobilní fáze se zvyšujícím se podílem organického rozpouštědla a uvolňuje nám nejprve hydrofilnější peptidy z kuliček do roztoku, a tedy následně do hmotnostního spektrometru. Hydrofobnější peptidy drží na kuličkách pevněji a do mobilní fáze se uvolňují až při vyšších koncentracích organického rozpouštědla. HPLC s reverzní fází umožňuje rozdělit i vysoce komplexní směsi peptidů a zároveň peptidy o stejné hydrofobitě zakonzcentruje, čímž umožní jejich detekci a kvantifikaci pomocí hmotnostní spektrometrie. Nevýhodou HPLC je potřeba specifických přístrojů a nutnost optimalizace protokolu pro detekci některých peptidů (proteinů).

● Hmotnostní spektrometrie

Zřejmě nejrozšířenějším proteomickým analytickým přístupem současnosti je hmotnostní spektrometrie, obvykle v podobě kapalinového chromatografu přímo napojeného na tandemový hmotnostní spektrometr (LC-MS/MS). Hmotnostní spektrometr si lze představit jako vysoce přesné váhy, které měří hmotu nabitých částic (iontů) s přesností na tisíce daltonu (je-



3 Separace proteinů dvojrozměrnou gelovou elektroforézou. Proteiny z lidských fibroblastů byly v prvním rozměru rozděleny na základě izoelektrického bodu (pI) v gradientu pH 4 až 7 (izoelektrická fokusace). Rozdělení ve druhém rozměru bylo provedeno pomocí elektroforézy s dodecylsulfátem sodným (SDS-PAGE) na základě jejich molekulové hmotnosti (jednotky kDa – kilodaltony). Proteiny byly následně barveny stříbřením. Orig. H. Kupcová Skalníková

den dalton odpovídá hmotnosti přibližně jednoho protonu nebo neutronu, což je zhruba $1,66 \times 10^{-24}$ g). Pomocí MS tak lze velmi přesně určit molekulovou hmotnost studovaných analytů. Při tandemové MS, kdy jsou dva hmotnostní analyzátoři napojeny za sebou (viz obr. 1 na str. CLXXXVI kuléru), lze též srážkami s molekulami plynného dusíku či jinými technikami přerušit některé chemické vazby ve studované molekule a získat tak specifické fragmenty. V oblasti proteomiky se dají peptidy fragmentovat v místě peptidových vazeb, tím nám LC-MS/MS poskytne i částečnou informaci o aminokyselinové sekvenci daného peptidu, což napomáhá jeho spolehlivému určení. Moderní LC-MS/MS systémy umožňují během 30 minut „zvážit“ desetitisíce peptidů a zaznamenat jejich intenzity signálů, čímž získáme informace o přítomnosti a kvantitě až 10 tisíc proteinů ve studovaném vzorku. Zároveň využívají velmi malá množství vstupního materiálu (řádově stovky ng proteinů), ale mohou analyzovat i jednotlivé buňky. Mimo kvantifikace peptidů, resp. proteinů, umožňuje LC-MS/MS analýzy proteoforem, včetně určení typů posttranslačních modifikací

a jejich lokalizace. Pro analýzy posttranslačních modifikací je často nutné proteiny/peptidy s danou modifikací vyextrahovat a zakonzcentrovat, aby mohly být v hmotnostním spektrometru detekovány. Např. fosforylované proteiny/peptidy se dají ze složité směsi ve vzorku vychytat pomocí protilátek rozpoznávajících fosforylovaný tyrozin nebo pomocí některých iontů kovů (např. oxidu titaničitého). Počet proteinů/peptidů, které lze identifikovat a kvantifikovat ve vzorku, silně závisí na typu vzorku a použitých separačních a jiných technikách.

Výhodami MS metod je vysoká specifita identifikace a kvantifikace proteinů, možnost necílených i cílených analýz (není nutné předem vybrat proteiny, které chceme studovat) a velké pokrytí proteomu. Hlavní nevýhodou je detekce pouze nabitých peptidů (nelze studovat všechny peptidy/proteiny), dále nutnost specifického a finančně nákladného vybavení a složitější analýza dat než v případě protilátkových metod.

V navazujícím příspěvku v kulérově části tohoto čísla Živy (na str. CLXXXVI až CLXXXVII) naleznete podrobnější popis a grafické znázornění fungování hmotnostního spektrometru a příklady využití MS v medicíně i jiných oblastech.

Tato práce byla podpořena projektem Národní ústav pro výzkum rakoviny (číslo LX22NPO5102) financovaným Evropskou unií – Next Generation EU.

Použitou a doporučenou literaturu uvádíme na webové stránce Živy.