

HMOTNOSTNÍ SPEKTROMETRIE VE STRUKTURNÍ BIOLOGII: URČOVÁNÍ VYŠŠÍ STRUKTURY PROTEINŮ A PROTEINOVÝCH KOMPLEXŮ

TOMÁŠ VAISAR

*UW Medicine Diabetes Institute, University of Washington, Seattle, 850 Republican St, Seattle, WA 98109, USA
tvaisar@uw.edu*

Došlo 15.1.20, přijato 24.2.20.

Klíčová slova: strukturální biologie, hmotnostní spektrometrie, zesílení proteinů, H/D výměna, proteinové komplexy, afinitní hmotnostní spektrometrie, nativní hmotnostní spektrometrie

Obsah

1. Úvod
2. Metody chemického zesílení
3. Hmotnostní spektrometrie s výměnou vodíku za deuterium
4. Nativní hmotnostní spektrometrie
5. Afinitní hmotnostní spektrometrie

1. Úvod

Biologická aktivita proteinů je převážně podmíněna jejich terciární a kvartérní strukturou. Ve své nativní formě se mnoho proteinů navíc nachází v různých oligomerních stavech nebo tvoří proteinové komplexy. Tyto komplexy se mohou skládat nejen z několika málo proteinů, ale celá řada esenciálních komplexů (například ribosom) se skládá z desítek až stovek proteinových jednotek. Charakterizace jak vyšší struktury jednotlivých proteinů, tak i složení a struktury proteinových komplexů, je kritická pro porozumění mechanismům procesů v buňkách a živých organismech. Klasickými metodami pro strukturálně biologické studie jsou rentgenová difrakce a nukleární magnetická rezonance (NMR). Obě tyto metody poskytují strukturální informace s vysokým rozlišením. Obě metody však zároveň mají řadu omezení, která z velké části pramení ze základní vlastnosti vyšší struktury proteinů, kterou je flexibilita proteinové páteře umožňující dynamické změny konformace. Podle termodynamické hypotézy navržené Anfinsenem představuje nativní struktura uspořádání proteinu a rozpouštědla (vodného prostředí obklopujícího protein v přirozeném stavu) s nejnižší volnou energií¹. Tato struktura je ale často oddělena od energeticky nejbližších alternativních struktur pouze nízkými bariérami, a tak je struktura proteinu v nativním stavu většinou velmi dynamická a je dána spíše sbírkou celé řady struktur než jednou přesně

určenou strukturou. Tento dynamický charakter proteinových struktur představuje podstatnou překážku pro určení struktury pomocí rentgenové krystalografie nebo NMR, neboť v prvním případě je nutné protein či komplex pro analýzu zkrystalizovat a v druhém případě zabraňuje změření signálu z flexibilních částí proteinu. V prvním případě je tak změřená struktura tou nejstabilnější za podmínek krystalizace, v druhém případě pak je výsledná struktura „průměrem“ struktur, které protein zaujme za dobu měření, která se pohybuje v řádu minut. Navíc jsou možnosti NMR značně omezené velikostí studovaného proteinu.

Od zavedení měkkých ionizačních metod, především elektrosprejové ionizace, hmotnostní spektrometrie (MS) přispívala k charakterizaci proteinů především schopností identifikovat proteiny a určit jejich molekulovou hmotnost s přesností daleko přesahující ostatní metody. Těchto schopností hmotnostní spektrometrie se ve stále větší míře využívá i při strukturálně-biologických studiích. Hmotnostně spektrometrické metody pro strukturální biologii neposkytují kompletní strukturální informaci jako rentgenová difrakce. Také jejich strukturální rozlišovací schopnost je nízká ve srovnání s rentgenovou difrakcí a NMR (desítky Å oproti jednotkám Å u rentgenové a NMR spektroskopie). Informace získané z MS analýzy jsou tak používány především v kombinaci s komplementárními strukturálními daty získanými jinými technikami a jako vstupní restriční kritéria pro molekulární modelování struktury proteinů^{2,3}. MS metody mají však i řadu výhod – kromě identifikačních schopností se k nim řadí citlivost, v podstatě neomezená velikost studovaného proteinu či proteinového komplexu, přesnost měření a použitelnost pro řešení široké řady proteinových systémů od izolovaných proteinů a komplexů až po komplexní směsi, jako jsou celé buňky nebo plazma.

Mezi MS metody používané pro strukturální biologii se řadí: chemické zesílení (chemical cross-linking MS, CX-MS, někdy také XL-MS), MS s výměnou deuteria (hydrogen-deuterium exchange MS, HDX-MS), nativní hmotnostní spektrometrie proteinových komplexů a hmotnostní spektrometrie s použitím protilátek (affinity-purification-capture MS, AP-MS)⁴. Protože je každá z těchto metod založena na velmi odlišných principech, poskytují specifickou informaci o struktuře. Metody nativní MS a AP-MS jsou především používány k určení identity proteinů či podjednotek proteinových komplexů, metoda CX-MS poskytuje informaci o prostorovém uspořádání proteinu nebo vzájemné orientaci podjednotek proteinových komplexů. Metoda HDX-MS může poskytnout informaci o strukturální flexibilitě celého proteinu, i jeho jednotlivých částí. Může také sloužit k charakterizaci proteinových komplexů. Kromě zmíněných metod byly popsány i metody kombinující hmotnostní spektrometrii s chemickými metodami založenými na modifikaci po-

vrchu proteinu (tzv. reaktivní otisk), například oxidací (oxidativní otisk) nebo modifikací povrchu proteinu činidlem s reaktivitou specifickou pro vybrané skupiny postranních řetězců aminokyselin. Tyto metody jsou založené na předpokladu, že s daným činidlem reagují pouze funkční skupiny proteinu, které jsou na povrchu, a jsou tedy dostupné pro použité činidlo, zatímco skupiny, které jsou v důsledku vyšší struktury proteinu či komplexu uvnitř struktury, jsou stericky bráněné a nereagují^{5,6}. Podobně lze využít také omezené enzymové degradace za nativních podmínek.

Protože úspěšnost aplikace hmotnostní spektrometrie při řešení struktur proteinů a jejich komplexů silně závisí na volbě vhodných metod, jsou často používány v kombinaci. Ve většině případů jsou informace získané z hmotnostně spektrometrických experimentů komplementární s výsledky dalších technik (NMR, rentgenová difrakce, elektronová mikroskopie atp.) a ve spojení s nimi pak přispívají k řešení struktury proteinů. Často jsou také MS výsledky (např. informace o interagujících aminokyselinových zbytcích a jejich přibližných vzdálenostech z chemického zesílení) využívány při počítačovém molekulárním modelování k výběru biologicky významných vstupních struktur pro počítačovou optimalizaci.

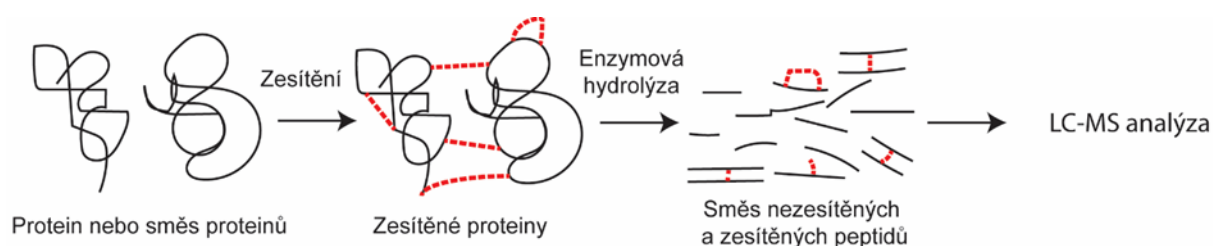
2. Metody chemického zesílení

Metod chemického zesílení se převážně používá k určení interagujících proteinů v proteinových komplexech, nebo ke strukturní charakterizaci proteinu, při které chemické zesílení poskytuje informaci o interagujících doménách proteinu a sterické restriktce pro molekulární modelování. Základním principem metod založených na chemickém zesílení proteinů je kovalentní propojení funkčních skupin v postranních řetězcích dvou aminokyselinových zbytků, které jsou si v nativní konformaci proteinu prostorově blízké (avšak od sebe vzdálené v primární sekvenci proteinu) nebo náleží dvěma různým spolu interagujícím proteinům⁷. Pro identifikaci propojených (zesílených) aminokyselin se po reakci se síťovací činidlem (cross-linkerem) používá enzymové degradace a standardních metod využívaných v proteomice s jistými

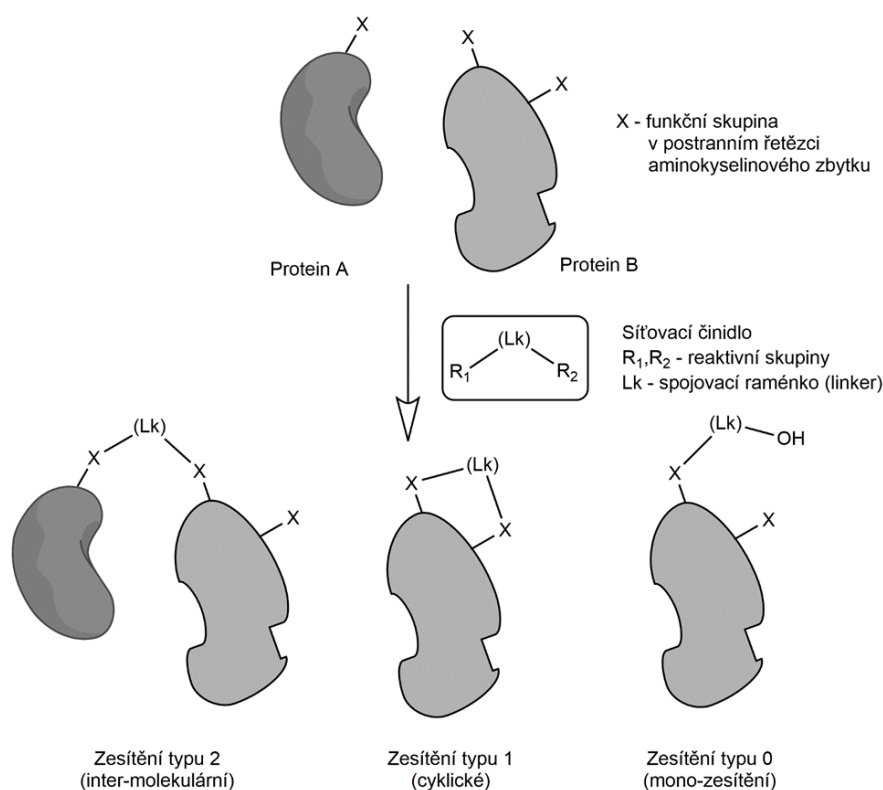
úpravami, které budou diskutovány později (obr. 1).

Síťovací činidla se vyznačují tím, že mají dvě reaktivní skupiny spojené inertním řetězcem, spojkou (linkerem) určující maximální vzdálenost aminokyselinových zbytků, které daný „cross-linker“ může spojit^{8–10}. Při reakci „cross-linkeru“ s proteinem nedochází pouze k žádoucím reakcím propojujícím dvě prostorově vzdálené funkční skupiny jednoho proteinu či dvou interagujících proteinů (internebo intra-proteinové cross-linky mezi dvěma peptidy nazývané také zesílení typu 2), ale i k několika nežádoucím reakcím⁷. Síťovací činidlo může spojit funkční skupiny dvou aminokyselin nacházejících se blízko sebe v primární sekvenci proteinu, které se po následné enzymové hydrolyze ocitnou v rámci jednoho peptidu. Tím se vytvoří tzv. cyklické linky nebo zesílení typu 1. Síťovací činidlo také může reagovat s proteinem pouze jednou ze svých reaktivních skupin, zatímco k hydrolyze druhé skupiny dochází dříve, než může reagovat s reaktivní skupinou druhého aminokyselinového zbytku. Výsledkem jsou tzv. mono-linky, také nazývané zesílení typu 0. Zesílení typu 0 a typu 1 nejsou většinou pro studium struktury proteinu užitečná (obr. 2).

Jednou z důležitých charakteristik síťovacího činidla je dobrá rozpustnost za podmínek, v nichž je studovaný protein či komplex v nativní konformaci, nejčastěji tedy v prostředí vodního pufru. Dalším předpokladem je vysoká reaktivita síťovacího činidla s funkčními skupinami proteinu, aby byla upřednostněna pokud možno kvantitativní reakce síťovacího činidla s funkčními skupinami aminokyselinových zbytků oproti hydrolyze reaktivní skupiny síťovacího činidla. V současné době existuje celá řada síťovacích činidel, které se dají rozdělit podle několika kritérií. Jednak podle reaktivity (např. vůči aminokyselinám, thiolovým skupinám, karboxylovým skupinám nebo jejich kombinacím), podle délky spojovacího řetězce, raménka, a také podle toho, zda spojovací raménko obsahuje další specifické funkční skupiny (chemicky štěpitelné skupiny, isotopické značky nebo afinitní značky apod.). Existuje široká škála síťovacích činidel, podrobný souhrn je dostupný v cit.^{8,11}. Nejčastěji se používají síťovací činidla s dvěma stejnými reaktivními skupinami, tzv. homobifunkční síťovací činidla (obr. 3), reagující s aminokyselinami postranního řetězce lysinu, jednak z důvodu



Obr. 1. Schématické shrnutí měření chemického zesílení proteinů s hmotnostně spektrometrickou analýzou (LC-MS, kapalinová chromatografie s hmotnostní spektrometrií)

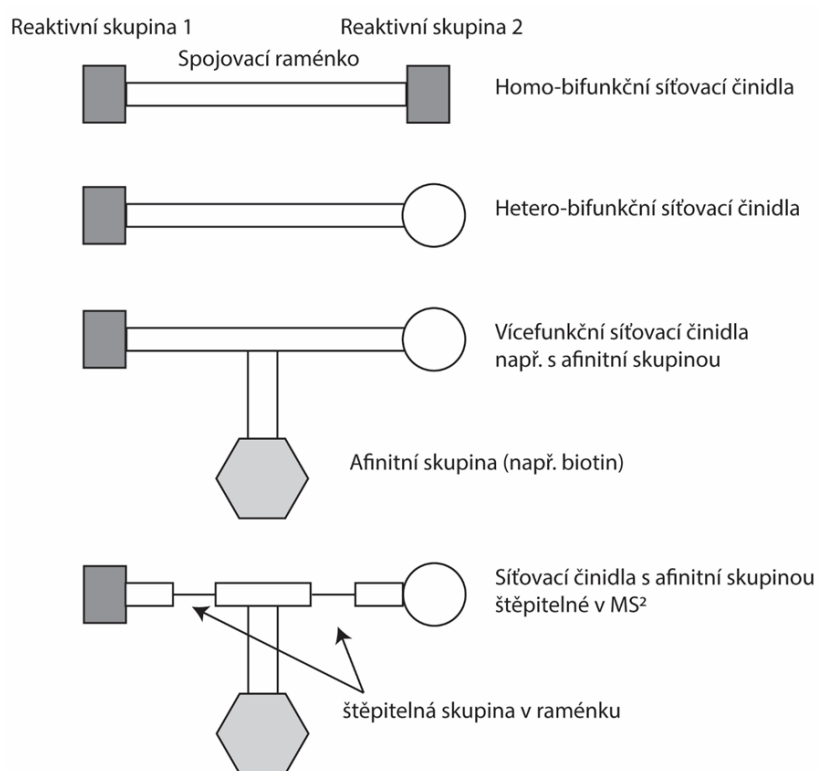


Obr. 2. Schématické znázornění produktů chemického zesíťení proteinů

frekvence výskytu této aminokyseliny v proteinech (v průměru 6 %) a také pro jejich vysokou reaktivitu a specificitu. Je možná i reakce s N-koncovou aminoskupinou proteinu, avšak tato skupina je často modifikovaná a také její pK_a je mnohem nižší než pK_a lysinové aminoskupiny. Reaktivní skupiny v těchto síťovacích činidlech jsou převážně estery *N*-hydroxysukcinimidu nebo sulfosukcinimidu, což jsou tzv. aktivní estery, které velmi ochotně reagují s aminoskupinou, ale jsou také snadno hydrolyzovatelné, a to s rychlostí, která je řádově srovnatelná s rychlostí reakce s aminoskupinami. Přestože je díky této vlastnosti obtížné s těmito síťovacími činidly dosáhnout vysokých výtěžků zesíťení, disukcinimidsuberát (DSS, raménko obsahující šest uhlíkových atomů mezi reaktivními skupinami) a disukcinimidglutarát (DSG, raménko obsahující tři uhlíkové atomy) a především jejich sulfo-analogy (např. sulfo-analog DSS, bis(sulfosukcinimidsuberát), BS3) se zvýšenou rozpustností ve vodě, jsou nejrozšířenější síťovací činidla (obr. 4A). Alternativou k lysinovým síťovacím činidlům jsou síťovací činidla na bázi maleinimidu, které reagují s thiolovými skupinami cysteinu. Avšak relativně nízký výskyt cysteinu v proteinech (kolem 2 %) výrazně limituje využitelnost těchto síťovacích činidel. Dalšími reaktivními skupinami v proteinech jsou karboxylové skupiny kyselin glutamové

a asparagové nebo guanidylová skupina argininu, ale činidla pro tyto skupiny postrádají specificitu nebo jejich reakce poskytují nehomogenní či nestabilní produkty⁸. Výjimku tvoří dihydrazidová činidla, které je možno použít k zesíťení dvou karboxylových skupin po jejich aktivaci karbodiimidu (např. EDC, 1-ethyl-3-(3-dimethylamino-propyl)-karbodiimid)¹². Vhodná volba dihydrazidového činidla pak umožňuje tvorbu zesíťení s různou délkou¹³.

Kromě homo-bifunkčních síťovacích činidel byly vyvinuty i síťovací činidla s dvěma různými reaktivními skupinami (např. jednou reagující s aminoskupinou lysinu a druhou reagující s thiolovou skupinou cysteinu) (obr. 3). Existují i síťovací činidla s fotoaktivovanými reakčními skupinami (obr. 4B)^{9,10}. Tyto síťovací činidla poskytují většinou komplikované směsi produktů, protože mohou propojit velkou řadu aminokyselinových zbytků jak intramolekulárně, tak intermolekulárně a představují tak složitý problém pro identifikaci a analýzu dat. Dalším specifickým druhem síťovacího činidla je formaldehyd, který, přestože má jen aldehydovou skupinu, je schopen zesíťení proteinů. Tato reakce probíhá přes tvorbu Schiffovy báze na postranním řetězci lysinu, která následně reaguje s nukleofilními skupinami, např. postranním řetězcem tryptofanu^{8,14}.



Obr. 3. Obecné znázornění hlavních typů síťovacích činidel používaných k chemickému zesílení proteinů

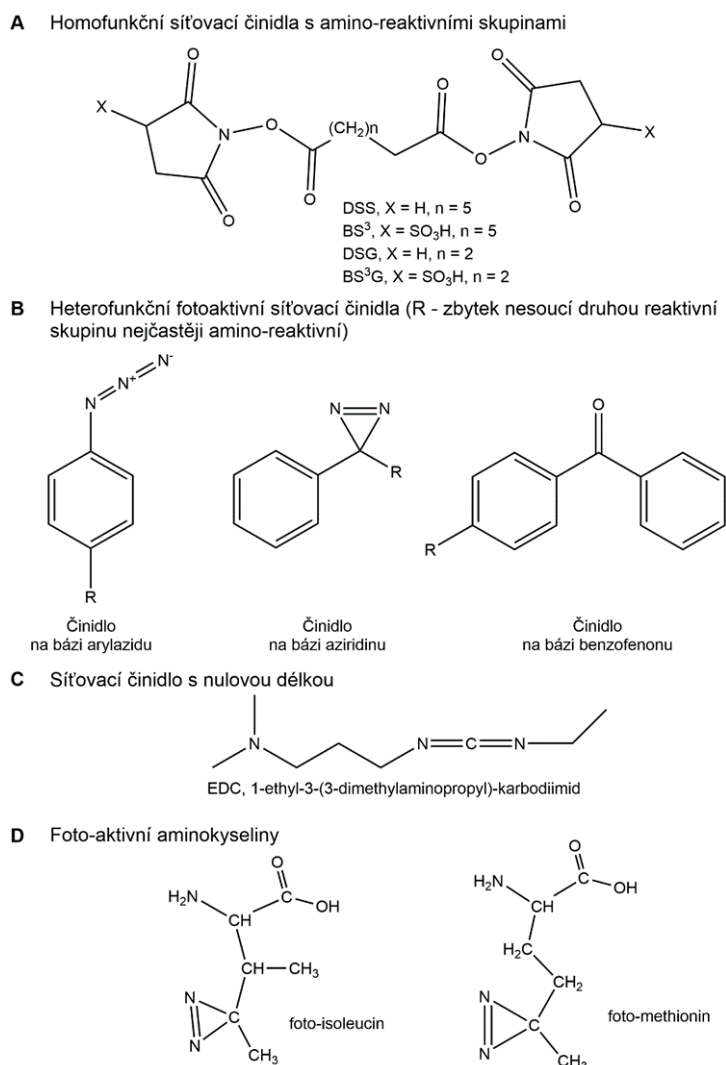
Specifickou skupinou síťovacích činidel jsou tzv. síťovací činidla s „nulovou délkou“, mezi které patří karbodiimidy (např. již zmiňované EDC) (obr. 4C)¹². Tyto „cross-linkery“ se účastní pouze zahájení tvorby zesílení, protože tvoří aktivní estery s karboxylovými funkčními skupinami kyseliny asparagové a glutamové, které následně reagují s proximálními aminoskupinami lysinových postranních řetězců a vytvoří tak přímou amidovou vazbu s „nulovou vzdáleností“ mezi spojenými aminokyselinovými zbytky. Protože reakce vyžaduje, aby dané funkční skupiny byly v těsné blízkosti, jsou tyto „cross-linkery“ ideální pro studium struktury jednotlivých proteinů nebo pro studium blízkých interakcí v proteinových komplexech.

Kromě využití síťovacích činidel *in vitro*, je možné zesílení proteinů provádět i *in vivo*, v živých systémech, především buňkách. Používají se jednak některá síťovací činidla popsaná výše¹⁵, ale byly použity i specifické strategie využívající přirozených biosyntetických procesů buněk a specifických sekvenčních motivů, např. tetra-Cys (selektivně reaguje s diarzenovými činidly) nebo azidy (reagují s fosfiny či mohou podstoupit „click“ reakci, při které je jako katalyzátor použita měď). Dále je možné provádět zesílení cílené na specifická místa v sekvenci proteinu zavedením foto-aktivního aminokyselinového analogu s aziridinovou funkční skupinou (např. foto-Met,

foto-Leu nebo foto-Ile) do sekvence proteinu⁸. Expresce proteinu s takovýmto aminokyselinovým zbytkem ve studovaných buňkách pak umožňuje tvorbu kovalentních „cross-linků“ přímo v živých buňkách po ozáření UV světlem.

Analýza produktů po zesílení proteinů má velmi specifické nároky, především protože výtěžky síťovacích reakcí jsou velmi nízké (v převážné většině případů pod 10 %). Ve směsi peptidů po enzymové degradaci je tedy pouze malá část žádoucích inter-peptidových (a inter-proteinových) „cross-linků“, které jsou překryté množstvím nemodifikovaných peptidů a peptidů s nežádoucími „cross-linky“ typu 0 a 1 (viz výše). Ke zvýšení pravděpodobnosti identifikace zesílených peptidů je používáno několik strategií. Separace na iontoměničích využívá skutečnosti, že po štěpení trypsinem mají zesílené peptidy vyšší kladný náboj, protože obsahují dva peptidy, a tudíž také dvě bazické C-koncové aminokyseliny (Lys nebo Arg) a dvě N-koncové aminoskupiny, a jsou tak pevněji vázány na katexu a mohou být vymyty odděleně od většiny ostatních tryptických peptidů.

Rada metod pak využívá nejrůznějších funkčních modifikací spojovacího raménka mezi reaktivními skupinami síťovacího činidla⁷. První druh představují síťovací činidla s izotopovou značkou. Tyto síťovací činidla v sobě



Obr. 4. Chemické struktury vybraných typů „cross-linkerů“. A) Nejčastěji používané homo-bifunkční síťovací činidla, B) heterofunkční foto-aktivované síťovací činidla, C) síťovací činidla s nulovou vzdáleností, D) příklady foto-aktivních aminokyselin používaných pro sekvenčně-specifické modifikace proteinů a studie lokálních interakcí v živých buňkách

mají zabudovanou isotopovou značku, a to nejčastěji ve formě čtyř až osmi atomů deuteria. Při použití ekvimolární směsi „lehkého“ a „těžkého“ síťovacího činidla jsou zesítené peptidy ve spektru snadno rozeznatelné podle specifické isotopové obálky a je možné nastavit sběr dat tak, aby hmotnostní spektrometr přednostně sbíral tandemová spektra zesítených peptidů. Alternativním způsobem zavedení isotopové značky je provedení tryptického štěpení v pufru obsahujícím H₂¹⁸O. Zesítené peptidy se pak vyznačují posunem o 8 Da ve srovnání s posunem o 4 Da typickým pro tryptické peptidy, které nejsou zesítené. Limitací této metody je možná zpětná výměna, která ztěžuje interpretaci.

Použití afinitních značek, časté v kvantitativní proteomice, našlo uplatnění i pro zesítení proteinů^{7,8}. Obdobně jako v případě síťovacích činidel s isotopovou značkou,

i zde je afinitní značka zabudována do spojovacího raménka reaktivních skupin. Bylo vyvinuto několik typů afinitních značek, které se používají ve spojení s řadou druhů síťovacích činidel. Nejčastěji používanou afinitní značkou je biotin. S použitím afinitních kolonek s avidinovými nebo streptavidinovými kuličkami je možné dosáhnout značného zakoncentrování zesítených peptidů v analyzovaném roztoku. Nevýhodou je značný nárůst molekulové hmotnosti síťovacího činidla s biotinovou afinitní značkou, a tím i možnost sterického bránění při reakci síťovacího činidla s proteinem. Alternativou k biotinu jsou síťovací činidla nesoucí azidovou afinitní značku, která umožňuje specifickou vazbu zesítených peptidů na kuličkách s cyklooktynem pomocí tzv. „click“ reakce. Síťovací činidla PIR (protein interaction reporters) kombinují biotinovou afinitní značku pro izolaci modifikovaných peptidů

s funkční skupinou, která se snadno a specificky štěpí při kolizní aktivaci, a poskytuje tak diagnostický fragmentový ion spolu se sekundárními fragmentovými ionty odpovídajícími zesítěným peptidům (obr. 3)¹⁶. Sekvence těchto peptidů je pak identifikována z MS³ spekter (spekter z třístupňové hmotnostní spektrometrie), přičemž MS³ měření je iniciováno automaticky po detekci diagnostického iontu síťovacího činidla v MS² spektru (spekter z dvoustupňové hmotnostní spektrometrie). Síťovací činidla obsahující nitroskupinu při fragmentaci poskytují diagnostickou neutrální ztrátu NO₂. Přestože všechny tyto síťovací činidla s modifikovaným spojovacím raménkem mezi reaktivními skupinami slibují výhody při analýze, jejich podstatnou limitací je velikost těchto modifikujících skupin, a tím i omezení minimální vzdálenosti interagujících domén proteinu či proteinů v komplexu, které je možno zesíťovat. Omezením také může být rozpustnost daného síťovacího činidla ve vodném prostředí, která je nezbytná pro zachování proteinů v nativní konformaci.

Úspěšný pokus vždy zahrnuje několik kroků optimalizace podmínek zesíťování, aby bylo dosaženo co nejvyššího výtěžku zesíťovaných produktů. Kritické parametry zahrnují např. výběr vhodného pufru, v němž musí být zachována nativní struktura studovaného proteinu či proteinového komplexu, ale zároveň s ním síťovací činidlo nesmí reagovat. Například pufr Tris je nevhodný pro síťovací činidla reagující s aminoskupinou lysinu nebo síťovací činidla nultého řádu, které síťují karboxylové skupiny s aminoskupinami lysinu. Dále je síťovací reakce efektivnější při vyšších koncentracích proteinu. Ačkoli jsou síťovací reakce z praktických důvodů často prováděny i při nižších koncentracích, neefektivněji probíhají při koncentracích proteinu (1 mg ml⁻¹ a vyšších). Důležitý parametr je také rozpustnost daného síťovacího činidla, který musí být dodán v koncentraci s dostatečným molárním přebytkem (v 2–10násobném přebytku) k celkové molární koncentraci všech reaktivních skupin (pro dané síťovací činidlo) v proteinu, nikoliv pouze molární koncentraci proteinu. Většina činidel je ve vodných roztocích málo rozpustná a jsou proto připravována v organických rozpouštědlech, nejčastěji DMSO (dimethylsulfoxid). V závislosti na studovaném proteinu či velikosti komplexu tak tento přebytek může vyžadovat milimolární koncentrace síťovacího činidla. Důležitým parametrem je také délka spojovacího raménka mezi reaktivními skupinami síťovacího činidla, která určuje vzdálenost aminokyselin, které mohou být zesíťované. I relativně krátké spojovací raménko (např. 11,4 Å pro činidlo DSS) může umožnit spojení aminokyselinových zbytků vzdálených až 30 Å (měřeno mezi alfa-uhlíky aminokyselinových zbytků). Tato vzdálenost je dána délkou a flexibilitou postranních řetězců reagujících aminokyselinových zbytků. Tak síťovací činidlo DSS se šesti CH₂ skupinami spojovacího raménka umožňuje zesíťování aminokyselin ve vzdálenosti 7–30 Å s průměrnou vzdáleností 15 Å (cit.¹⁷). I když vývoj síťovacích činidel přináší stále nové účinnější a specifitější činidla, je při plánování pokusu potřeba mít na paměti, že v praxi je experimentálně detegována pouze malá část teoreticky možných zesíťování. Je řada důvodů, proč nejsou některá zesíťování

detegována: nízký výtěžek reakce, nevhodné chromatografické chování zesíťovaného peptidu, jeho nízká ionizace, nedostatečná fragmentace, která neumožní identifikaci zesíťovaných peptidů nebo nevhodná délka zesíťovaného produktu (jeden či oba peptidy jsou velmi dlouhé nebo velmi krátké v důsledku rozložení Lys a Arg zbytků v sekvenci proteinu). Zesíťování aminokyselinových zbytků není dáno jen reaktivitou síťovacího činidla, ale je ovlivněno i dostupností reagující skupiny na proteinu, která je snížena u skupin, které tvoří nekovalentní interakce, např. vodíkové můstky¹⁸.

Hmotnostně-spektrometrická analýza zesíťovaných proteinů je typicky prováděna metodami používanými pro proteomickou analýzu „zdola“. Po enzymovém štěpení zesíťované směsi jsou tryptické peptidy analyzovány nejčastěji po chromatografické separaci v hmotnostním spektrometru s ionizací elektrosprejem datově závislým měřením (data-dependent acquisition, DDA). Interpretace spekter je pak prováděna počítačově porovnáváním experimentálních spekter fragmentů s fragmenty předpovězenými teoreticky *in silico* z proteinových sekvencí uložených v databázi. Na rozdíl od nezesíťovaných tryptických peptidů obsahují MS² spektra fragmentové ionty pocházející z obou zesíťovaných peptidů a jejich interpretace je tak mnohem komplikovanější. I když teoreticky je možné identifikaci provádět na přístrojích s nízkým rozlišením, je žádoucí analyzovat zesíťované peptidy na přístrojích s vysokým rozlišením, neboť přesné určení hmotnosti a sekvence peptidů významně přispívá ke správné identifikaci. Fragmentace metodami zachytu elektronu mohou být výhodně použity pro tuto analýzu, neboť zesíťované peptidy nesou většinou vyšší náboj než typické tryptické peptidy. Vlastní identifikace zesíťovaných peptidů a odpovídajících proteinů představuje velmi náročný problém⁷. Zesíťované peptidy ve směsi všech tryptických peptidů z enzymové degradace zesíťovaného vzorku představují jen malou frakci (viz výše) a identifikace je dále komplikována obrovským počtem teoreticky možných produktů zesíťování. Počet možných kombinací peptidových zesíťovaných párů je dán vztahem:

$$\binom{n+k-1}{k}$$

(n = počet reaktivních aminokyselinových zbytků, k = počet reaktivních skupin síťovacího činidla) a přibližně může být odhadnut pro síťovací činidlo s dvěma reaktivními skupinami jako $n^2/2$. Například pro lidský albumin (609 aminokyselin, 35 Lys zbytků) je 595 možných kombinací zesíťování a pro lidský proteom (uvažovány jsou pouze teoreticky možné tryptické peptidy s 0–2 neúplnými enzymovými štěpeními s délkou 5–45 aminokyselinových zbytků, ale bez dalších modifikací jako např. oxidace methioninu) je možných $7,5 \cdot 10^{11}$ zesíťovaných peptidů. Ve srovnání s normální proteomickou analýzou „zdola“ (analýzou tryptických peptidů) se komplexita pro lidský proteom zvýší přibližně 600 000krát a výpočetní problém pro identifikaci srovnáváním experimentálních spekter s teoretickými spektry tak přesahuje reálné možnosti běžně dostupných počítačů.

V praxi jsou tak studie zesíťování proteinů limitovány na

relativně jednoduché systémy izolovaných proteinů nebo proteinových komplexů, kde je identita jednotlivých komponent známa a problém je zjednodušen na identifikaci zesíťovaných peptidů z těchto proteinů, i když byly publikovány i studie využívající metodu zesíťování v celých organismech (shrnuto v cit.¹³). Prvním přístupem ke zjednodušení automatické identifikace je tedy omezení analýzy na malý počet proteinů, jejichž struktura je studována. Další zjednodušení pak přináší interpretací algoritmy vyvinuté specificky pro analýzu zesíťovaných proteinů. Jeden z přístupů je založen na předpokladu, že fragmentace zesíťovaných peptidů poskytne fragmentové ionty, které odpovídají fragmentovým iontům stejných peptidů, ve kterých zesíťovaný aminokyselinový zbytek nese modifikaci odpovídající síťovacímu činidlu. Při tomto přístupu je v prvním kroku generována nová databáze linearizovaných sekvencí všech permutací možných zesíťování. V druhém kroku jsou pak použity standardní databázové vyhledávače (Mascot, Sequest, Phenix, OMSSA)¹⁹. Jiný přístup vychází z předpokladu, že zesíťované peptidy mohou být považovány vždy za jeden peptid s velkou modifikující funkční skupinou (která zahrnuje síťovací činidlo a druhý zesíťovaný peptid). Tento postup tak identifikuje dva zesíťované peptidy individuálně ve dvou krocích^{20,21}. V obou případech je nutné identifikace kontrolovat a potvrdit manuální interpretací odpovídajících spekter. Přestože tyto přístupy identifikaci zesíťovaných peptidů výrazně usnadňují, proces je díky nutnosti manuální verifikace velmi zdoluhavý. Alternativním přístupem je využití isotopové značky v síťovacím činidle při interpretaci²². Při této metodě je interpretace prováděna ve dvou krocích, v prvním je provedena předběžná identifikace peptidů pomocí sekvenčních značek (založená na tom, že už sekvence fragmentových iontů odvozená od 5 a více aminokyselin může být charakteristická pro daný peptid). V druhém kroku jsou pak tyto sekvence zkombinovány a zesíťované peptidy a odpovídající proteiny identifikovány. Isotopová značka identifikuje zesíťované peptidy v MS spektrech. Srovnání MS² spekter prekurzorů s lehkou a těžkou značkou umožňuje určit, které fragmentové ionty nesou zbytek síťovacího činidla. Tato informace výrazně usnadňuje identifikaci zesíťovaných peptidů. Přestože vyhledávače (search-engines) pro automatickou interpretaci MS² spekter zesíťovaných peptidů a proteinů jsou schopné poskytnout správnou interpretaci řady spekter, výsledky obsahují velké procento nesprávných identifikací (falešně pozitivních). Jedním ze zdrojů těchto chyb jsou situace, ve kterých je jeden peptid identifikován správně, ale identifikace druhého je nesprávná. Toto často nastává v případech, kdy je druhý peptid velmi krátký a splňuje kritérium molekulové hmotnosti zesíťovaného peptidu. V takových případech spektra poskytují jen málo charakteristických fragmentů pro tento peptid a v analyzátoch s omezeným rozsahem v nízké oblasti hmotností (iontové pasti) jsou b₂, y₁ či y₂ fragmenty ztraceny. Metody zesíťování proteinů a využití síťovacích činidel jsou proto stále intenzivně vyvíjeny a vylepšovány²³.

3. Hmotnostní spektrometrie s výměnou vodíku za deuterium (HDX MS)

Protony ve funkčních skupinách postranních řetězců aminokyselin (-OH, -SH, -NH) a protony na dusíku přítomném v amidových vazbách proteinových řetězců jsou labilní a snadno vyměnitelné s protony molekul vody obklopujících molekuly proteinu. Výměna vodíku za deuterium při použití deuterované vody (D₂O) jako rozpouštědla je tradičně využíváno ve spojení s NMR (cit.²⁴). Vzhledem k tomu, že výměna vodíku za deuterium změni hmotnost studovaného proteinu, kterou je možné změřit pomocí hmotnostní spektrometrie, je HDX ve spojení s MS metodou vhodnou pro analýzu dynamiky a struktury proteinů. Metoda HDX je založena na předpokladu, že protony v amidových vazbách, které se nacházejí na povrchu proteinu a neúčastní se intramolekulárních interakcí, jsou snadno a rychle vyměnitelné. Naproti tomu protony, zapojené do vodíkových vazeb důležitých pro strukturu proteinu, stejně jako protony ukryté uvnitř sekundární či terciální struktury proteinu nebo proteinového komplexu, vykazují pomalejší kinetiku výměny. Rychlost jejich výměny odpovídá síle interakcí daného protonu a rigiditě proteinové struktury, která proton chrání. Protony vázané na heteroatomech ve funkčních skupinách postranních řetězců aminokyselinových zbytků jsou obecně labilní a podléhají rychlé výměně, a proto je obtížné kinetiku jejich výměny měřit. Poslední skupinou vyměnitelných protonů jsou protony vázané na alfa-uhlíku aminokyselinových zbytků. Ty však vykazují extrémně pomalou kinetiku výměny, kterou nelze pomocí hmotnostní spektrometrie měřit. Dá se tedy říci, že HDX metoda měří dostupnost rozpouštědla k jednotlivým amidovým protonům. Ve srovnání s NMR umožňuje hmotnostní spektrometrie studie HDX s mnohem větší citlivostí a bez omezení velikosti proteinu. Tato metoda tak nachází časté uplatnění při strukturálních studiích, ale i při studiích interakcí proteinů s nejrůznějšími ligandy (proteiny, peptidy, malými molekulami, apod.), mapování epitopů, srovnávání proteinů s jejich mutanty, atd.^{25–27}

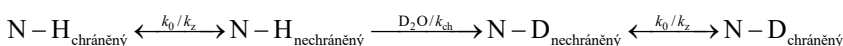
Vlastní HDX-MS měření se provádí ve třech krocích. První krok zahrnuje výměnu vodíku a deuteria, která se provádí v několika časových intervalech, aby byl zachycen postupný průběh výměny a mohla být změřena jejich kinetika. Po iniciaci výměny je reakce zastavena v daných časových intervalech a vzorek je připraven pro hmotnostně spektrometrickou analýzu. Ve většině případů se tak děje enzymovou degradací na peptidy, jak bude popsáno později. Druhý krok zahrnuje hmotnostně spektrometrickou analýzu a třetí krok analýzu a interpretaci dat.

Výměna vodíku za deuterium je obvykle prováděna přidáním pufru obsahujícím D₂O do roztoku proteinu na přesně stanovený časový úsek. Je ale možné provést inkubaci v opačném pořadí. V takovém případě je nejprve protein plně deuterován (inkubací v D₂O pufru), t.j. všechny dostupné vodíky jsou vyměněny za deuterium, a následně je provedena časově závislá reverzní výměna zpět na protonovanou formu (přidáním H₂O pufru)²⁷. Kompletní výměna protonů za deuteria však není snadná a reverzní vý-

měna dále předpokládá, že se protein navrátí do původní, fyziologicky relevantní struktury (předpoklad, který nemusí být vždy splněn). V důsledku těchto limitací se reverzní výměny používá jen výjimečně.

HDX měření je možné provádět dvěma způsoby – kontinuální výměnou nebo pulzní výměnou^{27,28}. Při kontinuální výměně je protein v nativním stavu vystaven D₂O a rozsah výměny je měřen v závislosti na čase v rozmezí minut až hodin. Oblasti proteinu, které jsou flexibilnější a které stráví více času v otevřené konformaci, jsou také přístupnější výměně protonů za deuteria. Oproti tomu oblasti proteinu, které jsou rigidní, jsou méně náchylné k výměně. Podle rovnice (1) zachycuje měření HDX metodou kontinuální výměny spíše dynamiku struktury proteinu než vlastní strukturu. Měření HDX pulzní výměnou vychází z denaturovaného proteinu, který je rychle převeden do nativního prostředí, čímž je odstartován proces nativního skládání. Následně je protein v přesně definovaných intervalech vystaven krátkému pulzu D₂O (řádově se jedná o milisekundy). Série měření je provedena v postupně vzrůstajících časových intervalech od iniciace skládání proteinu. Reakční doba s D₂O je konstantní a reakce je ukončena okyselením roztoku na pH 2,5, při kterém je rychlost výměny velmi pomalá. Závislosti kinetiky H/D výměny, která probíhá s maximální rychlostí při pD 8–10, se používá při pulzních měřeních také k urychlení výměny a dosažení většího rozsahu výměny v průběhu milisekundového pulzu, čímž se vylepší přesnost kvantifikace výměny. HDX měřená pulzní výměnou tak umožňuje zachytit přechodné stavy skládání proteinu do nativní konformace. Při těchto konformačních změnách jsou některé domény proteinu kompletně otevřené, zatímco jiné jsou provázané vodíkovými můstky. Pečlivá volba podmínek měření pak umožní kompletní výměnu amidových protonů otevřených oblastí proteinu a minimální výměnu protonů účastnících se vodíkových vazeb. Například při provedení měření s pulzem 5 ms při pD 9 může dojít k výměně více než 99 % H⁺ ($k_{\text{HDX}} \approx k_{\text{ch}} \approx 10^3 \text{ s}^{-1}$), zatímco u chráněných amidových skupin (faktor chránění $k_{\text{HDX}}/k_{\text{ch}}$ 10^2 – 10^6) je za stejných podmínek výměna nižší než 5 %. Takto provedené měření snadno umožňuje identifikovat oblasti proteinu, které jsou esenciální pro konformační změny jeho struktury.

K správnému provedení HDX měření je potřeba rozumět parametrům, které jsou pro kinetiku výměny důležité. Kinetika výměny je řízena dvěma rychlostními konstantami – chemickou, která odpovídá výměně protonu kompletně vystaveného rozpouštědla (k_{ch}), a rychlostní (k_0 , konstanta otevření), která zachycuje kinetiku, se kterou proton, v nativní konformaci chráněný interakcemi a strukturou proteinu ($H_{\text{chráněný}}$), přechází do stavu, ve kterém je kompletně vystavený okolním molekulám vody ($H_{\text{nechráněný}}$).²⁷ Po výměně za deuterium se pak protein vrací do původní konformace (rychlostní konstanta zavření, k_2). Souhrnně je kinetika výměny zachycena rovnicí (1)²⁷.



kde N je dusíkový atom amidové vazby.

Chemické prostředí každého amidového protonu je specifické, dané charakterem sousedících aminokyselin i interakcemi s ostatními částmi proteinu. V důsledku toho má každý proton specifické konstanty k_{ch} , k_0 a k_2 . Reakce vyjádřené rovnicí (1) mají dvě limitní situace: (a) $k_{\text{ch}} \gg k_2$ – k výměně dochází při prvním „otevření“ struktury a vystavení protonu okolnímu rozpouštědlu a kinetika (EX1) je řízena rychlostí „otevření“ struktury ($k_{\text{HDX}} \approx k_0$), (b) $k_2 \gg k_{\text{ch}}$ (kinetika EX2), při které je celková rychlost daná kombinací rychlosti chemické výměny a rychlosti otevření struktury ($k_{\text{HDX}} \approx k_0 k_{\text{ch}}$). Za podmínek EX2 protein a především lokální prostředí vyměňovaného protonu podstoupí řadu konformačních změn, a to rychleji než je tento proton vyměněn za deuterium. V nativní konformaci proteinu jsou přítomné i amidové protony, které jsou volně přístupné rozpouštědlu a rychlost jejich výměny je tak dána chemickou konstantou k_{ch} . Tyto protony jsou po iniciaci výměny rychle vyměněny a představují první rychlou fázi výměny. Chemická rychlostní konstanta je silně závislá na pH a pD s maximem při pH 8–10. Naopak při pH 2,5 je rychlost výměny minimální²⁹. Kinetiku HDX je také možno modulovat nastavením teploty, za které je výměna prováděna, neboť snížení teploty má za následek zpomalení kinetiky výměny.

Hmotnostně spektrometrická analýza při HDX měřeních je prováděna v zásadě dvěma způsoby – přímou analýzou celého proteinu představující „globální“ přístup obdobný proteomice „shora“ (top-down), nebo po enzymové degradaci, „lokální“, jako při proteomice „zdola“ (bottom-up)^{27,30}. První přístup je vhodný pouze pro specifické aplikace, např. studium interakce s ligandem, protože poskytuje pouze globální informaci o celkové H/D výměně v celém proteinu v závislosti na změně podmínek, a tedy o termodynamice procesu. Druhý způsob analýzy naopak umožňuje studovat H/D výměnu s detailní informací až na úroveň jednotlivých aminokyselinových zbytků z kinetiky H/D výměny v jednotlivých peptidech změřené z MS i MS/MS spekter. Při obou metodách H/D výměny, kontinuální i pulzní^{27,28}, je reakce ukončena snížením pD na 2,5. Protože však může stále docházet ke zpětné výměně, je nezbytné současně snížit teplotu vzorku na 0 °C a analyzovat vzorky co nejrychleji nebo je zamrazit pomocí tekutého dusíku a uchovávat při –80 °C. K proteolytickému štěpení jsou pak používány enzymy, které jsou aktivní v oblasti pH < 3 a provádí se za teploty blízké 0 °C, aby k výměně nedocházelo ani u nově vytvořených peptidů. Nejčastěji se používá pepsin nebo další vybrané enzymy³¹, které navíc poskytují enzymové štěpení proteinu ve velmi krátké době (v řádu minut). Protože pepsin i další kyselá proteasy nemají přesně definovanou specifitu (na rozdíl třeba od trypsinu), je před vlastní kvantifikací H/D výměny nezbytné nejprve jednotlivé peptidy jednoznačně iden-

tifikovat. Relativní rozsah H/D výměny v závislosti na čase se vypočítá podle vzorce:

$$\text{rozsah deuterace } (t) = \frac{m(t) - m_0}{m_{100} - m_0} \quad (2)$$

kde $m(t)$ je střední/průměrná hodnota hmotnosti peptidu v čase t , a m_0 a m_{100} jsou hmotnosti odpovídající peptidu s nulovou a s kompletní H/D výměnou. Kompletně vyměněný vzorek je připraven například několikanásobnou dlouhou inkubací proteinu v D₂O pufru za podmínek maximální výměny (denaturovaný protein, neutrální až mírně zásadité pH, zvýšená teplota). Celková kinetika H/D výměny proteinu s počtem amidových skupin N je složena z N individuálních k_{HDX} hodnot a jako takovou je jí téměř nemožné rozložit na jednotlivé příspěvky. Pro zjednodušení je modelována ze dvou komponent – příspěvku amidových vazeb podléhajících okamžité (velmi rychlé) výměně A_0 , a příspěvku amidových protonů podléhajících pomalejší výměně, která je ve většině případů modelována dvěma členy (3):

$$\text{rozsah H/D výměny } (t) = A_0 + A_1 (1 - \exp[-k_1 t]) + A_2 (1 - \exp[-k_2 t]) \quad (3)$$

kde A_1 a A_2 jsou frakce celkové populace vyměnitelných protonů, kterým jsou připsány průměrné rychlostní konstanty k_1 a k_2 . Ve většině případů jsou dva členy dostatečné k popsání pozorované výměny. Jen velmi zřídka dochází k situacím, kdy je k modelování třeba rovnic s více členy. Naopak v některých případech je kinetiku výměny možno popsat pouze jedním členem.

HDX v kombinaci s enzymovou degradací pepsinem a kapalinovou chromatografií s hmotnostní spektrometrií (LC-MS) poskytuje informaci o dynamice proteinu se strukturním rozlišením, které je limitované délkou peptidu vzniklého enzymovým štěpením. Zvýšeného rozlišení lze dosáhnout použitím dvou proteas s odlišnou specificitou. Strukturního rozlišení až na úroveň jednotlivých aminokyselin je potenciálně možné dosáhnout z tandemových fragmentačních spekter³². Úspěšnost tohoto přístupu je však omezená tím, že při fragmentaci klasickými aktivačními metodami (např. fragmentace kolizní aktivací) dochází k intra-molekulárním přesmykům labilních protonů, a tak ke ztrátě polohové specificity H/D výměny. Této randomizaci je možné předejít použitím aktivačních metod, jako jsou aktivace záchytem nebo přenosem elektronu (electron-capture dissociation, ECD, electron-transfer dissociation, ETD). Aplikace těchto metod pro HDX je bohužel omezená, protože pepsin, nejpoužívanější enzym pro HDX, poskytuje především krátké peptidy, které tvoří převážně ionty s nízkým nábojem (2+), které nejsou vhodné pro ECD/ETD fragmentace. Použití proteas, které poskytují delší peptidy, může překonat tyto limitace.

Studium proteinových komplexů metodou HDX přináší další specifické problémy, které je nutno vzít v úvahu. Především se jedná o stabilitu komplexu (vyjádřenou disociační konstantou K_d) a o její vliv na kinetiku výměny. Protože v závislosti na K_d komplexu je vždy v roztoku přítomna směs komponent komplexu disociovaná a váza-

ná, je v typickém měření zvláště studována kinetika výměny jednotlivých komponent komplexu a celého komplexu. Metody studia proteinových komplexů pomocí HDX jsou v detailu diskutovány v přehledném článku Kochert a spol.³³.

4. Nativní hmotnostní spektrometrie

Řada biologických procesů je zprostředkována nekovalentními komplexy proteinů. Metody pro analýzu nekovalentních proteinových komplexů založené na hmotnostně spektrometrické analýze poskytují informaci na několika úrovních³⁴. Metoda chemického zesílení poskytuje informaci o složení komplexu a přímých kontaktech interagujících proteinů, zatímco metody HDX dávají informaci o dynamice komplexu. Nativní hmotnostní spektrometrie pak umožňuje přímé měření velikosti komplexu a jeho stability³⁴. Základním předpokladem nativní MS je schopnost převést proteinový komplex z roztoku do plynné fáze v neporušené formě. Prvním předpokladem je existence proteinu v nativní formě v roztoku. Vzhledem ke kompatibilitě pufru s MS (požadavek na těkavost pufru) byl často používán octan amonný. Protože je ale jeho pK_a nízké (pK_a 4,75), není octan amonný jako pufr vhodný za neutrálního pH a jeho použití může vést k artefaktům díky snížení pH roztoku a následné změně proteinové struktury³⁵. Převod iontů proteinu do plynné fáze je umožněn v první fázi měkkými ionizačními technikami, především elektrosprejovou ionizací následovanou převodem iontů z atmosférického tlaku do vakua. K udržení stability je tlak v druhé fázi diferenčního pumpování hmotnostního spektrometru vyšší ve srovnání se standardními přístroji, a to pro lepší fokusaci a desolvaci velkých, pomalu letících iontů. Analýza nativních proteinů dále vyžaduje analyzátor s velkým rozsahem m/z umožňující detekci iontů s nízkým nábojovým stavem. Pro analýzu jsou tak nejčastěji používány hybridní analyzátor kombinující kvadrupólový analyzátor (se sníženou frekvencí umožňující průchod těchto iontů) s analyzátozem času letu (Q-TOF)^{36–38}. Ačkoliv hmotnostní spektrometry s iontovou cyklotronovou rezonancí a Fourierovou transformací (FT-ICR) jsou také používány především pro možnost tandemové hmotnostní spektrometrie, tyto analyzátor mají ve srovnání s Q-TOF analyzátorů omezený rozsah m/z a také větší nárok na množství vzorku. Na rozdíl od analýzy pro charakterizaci proteinů proteomikou „shora“, při které je protein obvykle rozpuštěn ve vodném roztoku organického rozpouštědla (nejčastěji acetonitrilu nebo methanolu) okyseleném na pH 2–3 kyselinou mravenčí nebo octovou, jsou pro nativní hmotnostní spektrometrii proteiny či proteinové komplexy rozpuštěné v těkavém vodném pufru, nejčastěji hydrogenuhličitanu amonném při fyziologickém pH. Typickým rysem spekter nativních proteinů a komplexů je poměrně úzká distribuce iontů odpovídajících násobně nabitým stavům proteinu. Ve srovnání s denaturovanými proteiny elektrosprejovaným z kyselého rozpouštědla, které jsou rozbalené a mohou tak typicky nést řadu protonů a jsou tedy mnohonásobně nabitě, proteinové komplexy

v nativní formě jsou kompaktní a nesou tedy mnohem nižší náboj. Příkladem může být ureasa z *H. pylori*, která se za denaturujících podmínek vyskytuje jako kombinace alfa a beta podjednotek nesoucích v průměru 30 a 50 nábojů (M_r 26,6 a 61,7 kDa), zatímco za nativních podmínek existuje jako komplex čtyř alfa a čtyř beta podjednotek a ve spektru pak poskytuje sérii iontů s úzkou distribucí nabitých stavů a průměrným nábojem 80, přestože molekulová váha komplexu (1063 kDa) je skoro 50krát větší než molekulová hmotnost jednotlivých podjednotek³⁹.

Přestože je v současné době jasné, že v průběhu přenosu iontů nativních proteinů a komplexů z kapalné do plynné fáze může dojít k částečné změně vyšší struktury v důsledku ztráty molekul vody stabilizujících terciární a kvartérní strukturu, nativní hmotnostní spektrometrie v mnoha případech prokázala svoji užitečnost pro strukturní biologii. Nativní hmotnostní spektrometrie byla např. úspěšně použita k charakterizaci proteasomu, centrálního enzymového komplexu degradace proteinů. Celková molekulová hmotnost komplexu je 2,5 MDa a komplex se skládá z katalytického jádra, 20S proteasomu a dvou 19S proteasomů zakrývajících oba konce jádra. Proteasom je často používán pro validaci technik strukturní biologie. Za použití nativní MS byla prokázána stechiometrie bakteriálního a králičího 20S proteasomu a bylo poprvé zjištěno, že savčí proteasom vykazuje větší heterogenitu ve srovnání s bakteriálním proteasomem⁴⁰.

Důležitým nástrojem při analýze proteinových komplexů nativní hmotnostní spektrometrií je tandemová MS. Při fragmentaci kolizně indukovanou disociací (collision-induced dissociation, CID) je dominantní fragmentací iontů proteinových komplexů odštěpení monomerní podjednotky, které obvykle poskytuje ionty podjednotky s vysokým nábojem a „okleštěný“ zmenšený komplex se sníženým nábojem. Ladění fragmentační energie pak umožňuje postupnou fragmentaci a analýzu relativní afinity jednotlivých podjednotek komplexu, a nejen určení složení a stechiometrie komplexu, ale i relativní stability jednotlivých podjednotek. Kromě disociace komplexu při CID nativních komplexů také dochází k částečnému rozbalení proteinů, které je závislé na náboji fragmentovaného iontu komplexu. V některých případech může být zvýšení náboje odštěpené podjednotky při CID doprovázeno fragmentací peptidového řetězce a poskytuje tak i informaci vhodnou pro identifikaci dané podjednotky⁴¹. Zatímco CID je charakterizována primárně disociací nekovalentních vazeb a rozkladem komplexu na podjednotky, při disociaci pomocí metod založených na záchytu elektronu (ECD nebo ETD) dochází především ke štěpení peptidových vazeb, a tak tato metoda disociace nekovalentních nativních komplexů může být využita k sekvenování a identifikaci jednotlivých podjednotek. Slibnou pro charakterizaci nativních komplexů se ukazuje metoda disociace kolizí s povrchem (surface induced dissociation, SID)⁴². Protože přenos energie při kolizi iontu s povrchem probíhá velmi rychle a získaná vnitřní energie se nestačí rovnoměrně rozložit do všech vazeb a stupňů volnosti iontu, jsou produktem fragmentace iontů nativních komplexů na povrchu jak jednotlivé podjednotky, tak i části komplexu

složené z několika podjednotek, které si pravděpodobně zachovávají složenou strukturu⁴³. Tandemová hmotnostní spektrometrie tak umožňuje stanovení stechiometrie, stability a složení komplexů.

K analýze proteinových komplexů je také využíváno technik založených na mobilitě iontů v plynné fázi. Separace iontů těmito technikami je založena na kolizním průřezu iontu, a tak nepřímou charakterizuje tvar iontu^{44–46}. Tyto techniky tak umožňují studovat konformaci proteinů a v kombinaci s vhodnými změnami pufrálních podmínek i vliv nativního prostředí na stabilitu proteinů i komplexů. Za příklad může být uvedena studie konformační stability beta-2-mikroglobulinu (β_2m) pomocí iontově mobilní MS s použitím cestující napěťové vlny (travelling voltage wave MS, TWIMS). V TWIMS spektru β_2m spolu se snižováním pH dochází k rozšíření distribuce násobně nabitých stavů pozorovaných na m/z škále z převážně binární s molekulárními ionty 7+ a 8+ na širokou obálku molekulárních iontů s nábojem 7+ až 13+. Zároveň dochází i k rozdělení mobility určitých výše nabitých stavů na dva stavy indikující tvorbu nové konformace iontu⁴⁷. Nižší mobilita nově vytvořeného stavu je pak indikací většího kolizního průřezu iontu ukazujícího na otevřenou konformaci v důsledku denaturace proteinu. Podobně lze pomocí iontově mobility rozlišit například rozdílné geometrie kapsid viru hepatitidy B, přičemž dominantní kompaktní forma viru je doprovázena menší populací otevřenější formy s větším kolizním průřezem a pomalejší mobilitou⁴⁸. Speciálním typem iontově mobilní analyzátoru je analyzátor elektroforetické mobility v plynné fázi (gas phase electrophoretic mobility molecular analyzer, GEMMA), také nazývaný analyzátor diferenční iontové mobility (differential ion mobility analyzer, DMA)^{49,50}. Na rozdíl od typického analyzátoru pro iontovou mobilní spektrometrii jsou pro analýzu DMA ionty s vysokým nábojem produkované elektrospřejem nejprve redukovány na převážně neutrální a jednou nabitě a pak separovány v DMA na základě tření v proudu vzduchu za atmosférického tlaku. Separované ionty jsou detegovány jako jednotlivé částice v kondenzačním počítači částic. Přestože rozlišení DMA je ve srovnání s jinými iontově mobilními technikami malé, má DMA několik významných výhod – umožňuje analýzu částic v rozsahu od několika nm až do velikosti μm , a tedy od proteinů o hmotnosti 20–30 kDa až po komplexy s hmotností v řádu MDa a byly analyzovány i celé virové částice. DMA také umožňuje přímé měření počtu částic. Mimo jiné byla tato metoda použita ke studiu proteinových komplexů a virových částic, a dále i k přímému měření koncentrace a velikosti lipoproteinových částic, HDL a LDL-komplexů lipidů a proteinů s důležitou funkcí v homeostáze cholesterolu, a to jako biomarkerů pro diagnostiku kardiovaskulárních nemocí^{51,52}.

5. Afinitní hmotnostní spektrometrie

Afinitní izolace proteinových komplexů spojená s hmotnostní spektrometrií (affinity-purification mass spectrometry, AP-MS) není sice přímou metodou struktur-

ní analýzy, nicméně je užitečným předstupněm ostatních výše popsanych metod, a i sama o sobě přispívá ke strukturní analýze komplexů. Zatímco výše diskutované hmotnostně spektrometrické metody slouží k charakterizaci struktury proteinových komplexů, AP-MS má za primární cíl poskytnout informaci o složení proteinových komplexů za fyziologických podmínek a v širším měřítku má pak sloužit ke globálnímu mapování proteinových interakcí. Dále je možné kombinovat AP-MS s chemickým zesítěním a dalšími výše zmíněnými technikami k podrobnějšímu studiu struktury proteinového komplexu. AP-MS metoda je založena na izolaci určitého proteinu („návnady“) spolu s jeho přirozenými vazebnými partnery pomocí specifických protilátek. Izolovaná směs je pak charakterizována standardními proteomickými metodami, nejčastěji proteomikou „zdola“.

K samotné imunoafinitní izolaci se používají tři přístupy. Zprv je možno použít protilátek specifických k danému proteinu. Tento přístup na jednu stranu umožňuje studovat komplexy za podmínek, které jsou nejbližší přirozenému stavu *in vivo* (v buňkách, tkáních či tělesných tekutinách), na druhou stranu je ale omezen dostupností kvalitních protilátek. Druhý přístup je založen na genetické manipulaci buněk. Daný protein je v buňkách produkován jako konjugát s afinitní značkou (nejčastěji značka *flag*), pro kterou existují kvalitní specifické protilátky. Poslední přístup, nazývaný tandemová afinitní izolace, kombinuje dvě afinitní značky spojené enzymově štěpitelnou spojkou, které jsou konjugované k danému proteinu za účelem zvýšené specifity izolace⁵³. Pro vlastní izolaci jsou protilátky imobilizované kovalentně na pevném nosiči, nejčastěji agarosových kuličkách nebo nověji na magnetických nosičích.

Hlavním problémem AP-MS je odlišení specifických vazebných partnerů studovaného proteinu od proteinů, které nespecificky interagují s protilátkou nebo jejím nosičem. V nejjednodušším přístupu je možné pouze eliminovat proteiny známé pro svou afinitu k protilátkám či k používaným nosičům, které jsou shrnuté ve veřejné databázi Crapome⁵⁴. Avšak zahrnutí řady kontrolních podmínek je nezbytné pro správnou identifikaci proteinových komplexů pomocí AP-MS. Klasickým přístupem je použití kontrolní nespecifické protilátky. Vzorek je paralelně připraven izolací na nosiči se specifickou protilátkou a nosiči s nespecifickou protilátkou. Tento přístup může být dále kombinován s použitím izotopového značení chemickými značkami (ICAT – isotope-coded affinity tag, iTRAQ – isobaric tag for relative and absolute quantitation, TMT – tandem mass tag). Efektivním přístupem je izotopové značení přímo v buněčné kultuře přístupem SILAC (stable isotope labeling by amino acids in cell culture)⁵⁵. V tomto případě jsou připraveny dvě buněčné kultury. První z nich je kultivována na médiu s běžnými aminokyselinami a dochází v ní k produkci proteinu konjugovaného s afinitní značkou („lehká izotopová značka“). Druhá kultura, produkující protein bez afinitní značky, je kultivována v mediu s aminokyselinami značenými stabilními izotopy („těžká izotopová značka“, nejčastěji se používají lysin a arginin obsahující 8 resp. 10 atomů ¹³C a ¹⁵N). Po lýze

buněk jsou buněčné extrakty zkombinovány (ekvimolárně vzhledem ke studovanému proteinu) a dále zpracovány ve směsi. Proteiny specificky asociující se studovaným proteinem jsou pouze ty, které obsahují jen lehkou izotopovou značku. Proteiny identifikované s oběma izotopovými značkami nebo jen těžkou značkou jsou pak nespecifické kontaminanty. Tandemová afinitní izolace pak přistupuje k problému nespecifických kontaminantů zvýšením specifity prostřednictvím dvoustupňové izolace. Po první afinitní izolaci je první afinitní značka enzymově odštěpena a specificky vázané komplexy jsou dále vybrány ze směsi pomocí druhé afinitní značky. Kromě falešně pozitivních interakčních partnerů je také možné, že při AP-MS nedojde k identifikaci již známých interakčních partnerů (proteinové komplexy identifikované dříve nezávislými metodami). K této situaci (falešně negativní identifikace) může dojít, pokud podmínky AP-MS experimentu naruší relativně slabou interakci proteinů, například použitím detergentu při eluci proteinů nespecificky adsorbovaných na nosiči protilátky. Při plánování i interpretaci výsledků AP-MS experimentu je tedy na místě postupovat s opatrností a do protokolu celého pokusu zahrnout nezbytná kontrolní měření.

Hmotnostní spektrometrie s afinitní izolací byla úspěšně použita ke globální charakterizaci proteinové interakční mapy (protein-protein interaction, PPI) v *Saccharomyces cerevisiae*⁵⁶ i dalších mikroorganismech. Kromě statické charakterizace proteinových interakcí je možné použít AP-MS také k charakterizaci dynamiky proteinových komplexů⁵⁷ a pro stanovení vazebné afinity jednotlivých komponent komplexu⁵⁸.

Hmotnostně spektrometrické metody tedy nabízejí širokou škálu nástrojů pro studium struktury jak izolovaných proteinů, tak proteinových komplexů. Díky citlivosti a flexibilitě jsou důležitým doplňkem klasických metod strukturní biologie – rentgenové difrakce, NMR a v poslední době i elektronové mikroskopie. Jejich kombinací lze dosáhnout jak identifikace jednotlivých komponent komplexů (AP-MS), tak i stechiometrie (AP-MS, CX-MS), struktury (nativní MS, CX-MS) a strukturní dynamiky (HDX-MS). Použití hmotnostní spektrometrie ke strukturně biologickým studiím se neustále rozšiřuje díky vývoji nových činidel a biochemických metod a díky pokroku v hmotnostně spektrometrické instrumentaci. Strukturní studie s použitím hmotností spektrometrie tak k analýze proteinových struktur mohou poskytnout komplementární informaci nedosažitelnou jinými metodami.

Seznam zkratk

AP-MS	hmotnostní spektrometrie s afinitním záchytem a izolací (affinity-purification-capture MS)
BS3	bis(sulfosukcinimid)suberát
CX-MS	hmotnostní spektrometrie s chemickým zesítěním (chemical cross-linking MS, také XL-MS)
CID	kolizně indukovaná disociace (collision-induced dissociation)

- DDA datově závislé měření (data-dependent acquisition)
- DMA analyzátor diferenční iontové mobility (differential ion mobility analyzer)
- DMSO dimethylsulfoxid
- DSG disukcinimidglutarát
- DSS disukcinimidsuberát
- ECD disociace záchytem elektronu (electron-capture dissociation)
- EDC 1-ethyl-3-(3-dimethylaminopropyl)-karbodiimid
- ETD disociace přenosem elektronu (electron-transfer dissociation)
- FT-ICR hmotnostní spektrometr s iontovou cyklotronovou rezonancí a Fourierovou transformací
- GEMMA analyzátor elektroforetické mobility v plynné fázi (gas phase electrophoretic mobility molecular analyzer)
- HDX-MS hmotnostní spektrometrie s vodíko-deuteriová výměnou (hydrogen-deuterium exchange MS)
- ICAT afinitní značka s izotopovým tříděním (isotope-coded affinity tag)
- iTRAQ izobarické značky pro relativní a absolutní kvantifikaci (isobaric tag for relative and absolute quantitation)
- MS hmotnostní spektrometrie
- MS² dvoustupňová hmotnostní spektrometrie
- MS³ třístupňová hmotnostní spektrometrie
- PIR protein interaction reporters
- PPI proteinové interakční mapy (protein-protein interaction)
- LC-MS kapalinová chromatografie s hmotnostní spektrometrií
- Q-TOF kvadrupól s analyzátelem doby letu (quadrupole time of flight)
- SID povrchem indukovaná disociace (surface induced dissociation)
- SILAC analýza stabilních izotopových poměrů aminokyselin v buněčných kulturách (stable isotope labeling by amino acids in cell culture)
- TMT tandemová hmotnostní značka (tandem mass tag)
- TWIMS iontově mobilní MS s použitím cestující napěťové vlny (travelling voltage wave MS)
- Rinner O., Beck M., Aebersold R.: *Mol. Cell Proteomics* 9, 1634 (2010).
8. Sinz A.: *Mass Spectrom. Rev.* 25, 663 (2006).
9. Piotrowski C., Ihling C. H., Sinz A.: *Methods* 89, 121 (2015).
10. Leo G., Altucci C., Bourgoin-Voillard S., Gravagnuolo A. M., Esposito R., Marino G., Costello C. E., Vellotta R., Birolo L.: *Rapid Commun. Mass Spectrom.* 27, 1660 (2013).
11. Merkley E. D., Cort J. R., Adkins J. N.: *J. Struct. Funct. Genomics* 14, 77 (2013).
12. Novak P., Kruppa G. H.: *Eur. J. Mass Spectrom.* 14, 355 (2008).
13. Leitner A., Joachimiak L. A., Unverdorben P., Walzthoeni T., Frydman J., Forster F., Aebersold R.: *Proc. Natl. Acad. Sci. U. S. A.* 111, 9455 (2014).
14. Hoffman E. A., Frey B. L., Smith L. M., Auble D. T.: *J. Biol. Chem.* 290, 26404 (2015).
15. Chavez J. D., Bruce J. E.: *Curr. Opin. Chem. Biol.* 48, 8 (2019).
16. Tang X., Bruce J. E.: *Mol. Biosyst.* 6, 939 (2010).
17. Leitner A., Faini M., Stengel F., Aebersold R.: *Trends Biochem. Sci.* 41, 20 (2016).
18. Novak P., Kruppa G. H., Young M. M., Schoeniger J.: *J. Mass Spectrom.* 39, 322 (2004).
19. Maiolica A., Cittaro D., Borsotti D., Sennels L., Ciferri C., Tarricone C., Musacchio A., Rappsilber J.: *Mol. Cell. Proteomics* 6, 2200 (2007).
20. Singh P., Shaffer S. A., Scherl A., Holman C., Pfuetzner R. A., Larson Freeman T. J., Miller S. I., Hernandez P., Appel R. D., Goodlett D. R.: *Anal. Chem.* 80, 8799 (2008).
21. Chu F., Baker P. R., Burlingame A. L., Chalkley R. J.: *Mol. Cell. Proteomics* 9, 25 (2010).
22. Rinner O., Seebacher J., Walzthoeni T., Mueller L. N., Beck M., Schmidt A., Mueller M., Aebersold R.: *Nat. Methods* 5, 315 (2008).
23. Yilmaz S., Shiferaw G. A., Rayo J., Economou A., Martens L., Vandermarliere E.: *Mass Spectrom. Rev.* 37, 738 (2018).
24. Krishna M. M., Hoang L., Lin Y., Englander S. W.: *Methods* 34, 51 (2004).
25. Kaltashov I. A., Eyles S. J.: *Mass spectrometry in structural biology and biophysics: architecture, dynamics, and interaction of biomolecules*, 2. vyd. Wiley, Hoboken 2012.
26. Eyles S. J., Kaltashov I. A.: *Methods* 34, 88 (2004).
27. Konermann L., Pan J., Liu Y. H.: *Chem. Soc. Rev.* 40, 1224 (2011).
28. Deng Y., Zhang Z., Smith D. L.: *J. Am. Soc. Mass Spectrom.* 10, 675 (1999).
29. Smith D. L., Deng Y., Zhang Z.: *J. Mass Spectrom.* 32, 135 (1997).
30. Zhang Z., Smith D. L.: *Protein Sci.* 2, 522 (1993).
31. Kadek A., Mrazek H., Halada P., Rey M., Schriemer D. C., Man P.: *Anal. Chem.* 86, 4287 (2014).
32. Rand K. D., Zehl M., Jorgensen T. J.: *Acc. Chem. Res.* 47, 3018 (2014).
33. Kochert B. A., Jacob R. E., Wales T. E., Makriyannis

LITERATURA

- Anfinsen C. B.: *Science* 181, 223 (1973).
- Walzthoeni T., Leitner A., Stengel F., Aebersold R.: *Curr. Opin. Struct. Biol.* 23, 252 (2013).
- Heck A. J.: *Nat. Methods* 5, 927 (2008).
- Hyung S. J., Ruotolo B. T.: *Proteomics* 12, 1547 (2012).
- Liuni P., Zhu S., Wilson D. J.: *Antioxid. Redox Signaling* 21, 497 (2014).
- Konermann L., Pan Y., Stocks B. B.: *Curr. Opin. Struct. Biol.* 21, 634 (2011).
- Leitner A., Walzthoeni T., Kahraman A., Herzog F.,

- A., Engen J. R.: *Methods Mol. Biol.* 1764, 153 (2018).
34. Taverner T., Hernandez H., Sharon M., Ruotolo B. T., Matak-Vinkovic D., Devos D., Russell R. B., Robinson C. V.: *Acc. Chem. Res.* 41, 617 (2008).
 35. Konermann L.: *J. Am. Soc. Mass Spectrom.* 28, 1827 (2017).
 36. van den Heuvel R. H. a 11 spoluautorů: *Anal. Chem.* 78, 7473 (2006).
 37. Sobott F., Hernandez H., McCammon M. G., Tito M. A., Robinson C. V.: *Anal. Chem.* 74, 1402 (2002).
 38. Sobott F., McCammon M. G., Hernandez H., Robinson C. V.: *Philos. Trans. R. Soc., A* 363, 379 (2005).
 39. Pinkse M. W., Maier C. S., Kim J. I., Oh B. H., Heck A. J.: *J. Mass Spectrom.* 38, 315 (2003).
 40. Loo J. A., Berhane B., Kaddis C. S., Wooding K. M., Xie Y., Kaufman S. L., Chernushevich I. V.: *J. Am. Soc. Mass Spectrom.* 16, 998 (2005).
 41. Benesch J. L., Ruotolo B. T., Sobott F., Wildgoose J., Gilbert A., Bateman R., Robinson C. V.: *Anal. Chem.* 81, 1270 (2009).
 42. Wysocki V. H., Joyce K. E., Jones C. M., Beardsley R. L.: *J. Am. Soc. Mass Spectrom.* 19, 190 (2008).
 43. Ma X., Loo J. A., Wysocki V. H.: *Int. J. Mass Spectrom.* 377, 201 (2015).
 44. Bush M. F., Hall Z., Giles K., Hoyes J., Robinson C. V., Ruotolo B. T.: *Anal. Chem.* 82, 9557 (2010).
 45. Uetrecht C., Rose R. J., van Duijn E., Lorenzen K., Heck A. J.: *Chem. Soc. Rev.* 39, 1633 (2010).
 46. Ruotolo B. T., Benesch J. L., Sandercock A. M., Hyung S. J., Robinson C. V.: *Nat. Protoc.* 3, 1139 (2008).
 47. Smith D. P., Giles K., Bateman R. H., Radford S. E., Ashcroft A. E.: *J. Am. Soc. Mass Spectrom.* 18, 2180 (2007).
 48. Uetrecht C., Versluis C., Watts N. R., Wingfield P. T., Steven A. C., Heck A. J.: *Angew. Chem. Int., Ed. Engl.* 47, 6247 (2008).
 49. Kaddis C. S., Lomeli S. H., Yin S., Berhane B., Apostol M. I., Kickhoefer V. A., Rome L. H., Loo J. A.: *J. Am. Soc. Mass Spectrom.* 18, 1206 (2007).
 50. Kaddis C. S., Loo J. A.: *Anal. Chem.* 79, 1778 (2007).
 51. Caulfield M. P., Li S., Lee G., Blanche P. J., Salameh W. A., Benner W. H., Reitz R. E., Krauss R. M.: *Clin. Chem.* 54, 1307 (2008).
 52. Hutchins P. M. a 11 spoluautorů: *Clin. Chem.* 60, 1393 (2014).
 53. Gingras A. C., Gstaiger M., Raught B., Aebersold R.: *Nat. Rev. Mol. Cell Biol.* 8, 645 (2007).
 54. Mellacheruvu D. a 37 spoluautorů: *Nat. Methods* 10, 730 (2013).
 55. Ong S. E., Mann M.: *Nat. Protoc.* 1, 2650 (2006).
 56. Krogan N. J. a 52 spoluautorů: *Nature* 440, 637 (2006).
 57. Rinner O., Mueller L. N., Hubalek M., Muller M., Gstaiger M., Aebersold R.: *Nat. Biotechnol.* 25, 345 (2007).
 58. Sharma K., Weber C., Bairlein M., Greff Z., Keri G., Cox J., Olsen J. V., Daub H.: *Nat. Methods* 6, 741 (2009).
- T. Vaisar** (*UW Medicine Diabetes Institute, University of Washington, Seattle, USA*): **Mass Spectrometry for Structural Biology: Determination of Higher Structure of Proteins and Protein Complexes**
- This review summarizes current status of applications of mass spectrometry to structural biology and structure of proteins and protein complexes. Methods of chemical cross-linking allow investigations of protein structure and interactions through identification of amino acid residues that appear in close proximity providing information on spatial arrangement of proteins and identification of interacting components of protein complexes. Methods of hydrogen-deuterium exchange can provide information on protein dynamics, spatial restriction and indirectly identifying residues involved in molecular interacting. Native mass spectrometry is a method that directly studies proteins in their biologically relevant conformation and composition and stability of protein complexes. Affinity purification mass spectrometry is a method for specific isolation and characterization of protein complexes including their stoichiometry and in combination with other techniques discussed herein can provide structural information on structure of protein complexes. Collectively these techniques provide unique structural information that complements other methods for study of proteins structure include NMR, x-ray crystallography, CryoEM as well as molecular modelling, contributing to our understanding of protein structure and their biological activity.
- Keywords:** structural biology, mass spectrometry, protein cross-linking, H/D exchange, protein complexes, affinity mass spectrometry, native mass spectrometry