

Univerzita Pardubice
Fakulta chemicko-technologická

Příprava rekombinantních proteinů v různých expresních systémech

Bakalářská práce

Univerzita Pardubice
Fakulta chemicko-technologická
Akademický rok: 2020/2021

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Barbora Pokorná**
Osobní číslo: **C18269**
Studijní program: **B3912 Speciální chemicko-biologické obory**
Studijní obor: **Zdravotní laborant**
Téma práce: **Příprava rekombinantních proteinů v různých expresních systémech**
Zadávací katedra: **Katedra biologických a biochemických věd**

Zásady pro vypracování

V dostupné odborné literatuře mladší 10 let se seznámte se způsoby přípravy rekombinantních proteinů, expresními systémy prokaryotní, eukaryotní aj., a expresní vektory. Zpracujte rešerši na téma odlišných posttranslačních modifikací rekombinantních proteinů v odlišných expresních systémech. Zařadte fosforylace, acetylace, glykosylace.

Rozsah pracovní zprávy: **25 s.**
Rozsah grafických prací: **dle potřeby**
Forma zpracování bakalářské práce: **tištěná**

Seznam doporučené literatury:

Vedoucí bakalářské práce: **Mgr. Marcela Slováková, Ph.D.**
Katedra biologických a biochemických věd

Datum zadání bakalářské práce: **18. prosince 2020**

Termín odevzdání bakalářské práce: **2. července 2021**

L.S.

prof. Ing. Petr Kalenda, CSc.
děkan

prof. Mgr. Roman Kandár, Ph.D.
vedoucí katedry

V Pardubicích dne 26. února 2021

Prohlašuji:

Práci s názvem Příprava rekombinantních proteinů v různých expresních systémech jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury. Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne

Barbora Pokorná

Poděkování

Chtěla bych poděkovat Mgr. Marcele Slovákové, Ph.D., za trpělivost a ochotu při vedení práce.

Dále bych chtěla poděkovat své rodině a přátelům, za podporu během celého studia.

Anotace

Tato práce se zabývá přípravou rekombinantní proteinů v různých prokaryotních a eukaryotních expresních systémech a popisem těchto expresních systémů. Dále se věnuje posttranslačním modifikacím rekombinantních proteinů, hlavně fosforylaci, acetylaci a glykosylaci. Popisuje rozdíly v posttranslačních modifikacích mezi jednotlivými expresními systémy.

Klíčová slova

rekombinantní proteiny, expresní systémy, posttranslační modifikace, fosforylace, glykosylace, acetylace

Title

Preparation of recombinant proteins in various expression systems

Annotation

This work deals with the preparation of recombinant proteins in various prokaryotic and eukaryotic expression systems and the description of these expression systems. It also deals with posttranslational modifications of recombinant proteins, mainly phosphorylation, acetylation and glycosylation. It describes the differences in posttranslational modifications between expression systems.

Keywords

recombinant proteins, expression systems, posttranslational modifications, phosphorylation, glycosylation, acetylation

Obsah

Seznam zkratek	7
Úvod	11
1. Rekombinantní proteiny	12
1.1. Historie rekombinantní technologie	12
1.2. Příprava rekombinantních proteinů	13
1.3. Vektory	14
2. Expresní systémy	16
2.1. Prokaryotní expresní systémy	16
2.2. Kvasinkové expresní systémy	17
2.3. Hmyzí expresní systémy	18
2.4. Savčí expresní systémy	20
2.5. Ostatní expresní systémy	21
3. Posttranslační modifikace proteinů	22
3.1. Fosforylace	22
3.1.1. Analýza fosforylací	24
3.2. Acetylace	25
3.2.1. N-koncová acetylace	26
3.3. Glykosylace	29
3.3.1. N-glykosylace	30
3.3.2. O-glykosylace	31
3.3.3. Analýza glykosylací	31

4.	Posttranslační modifikace proteinů v expresních systémech	33
4.1.	Posttranslační modifikace v prokaryotických expresních systémech	33
4.2.	Posttranslační modifikace v kvasinkových expresních systémech	34
4.3.	Posttranslační modifikace v expresních systémech hmyzích buněk	36
5.	Závěr	38
	Literatura	39

Seznam zkratek

Ac-CoA	acetylkoenzym A
ADP	adenosindifosfát
BEVS	expresní vektorový systémem bakulovirů
BHK21	linie buněk fibroblastů křeččích mlád'at
cDNA	komplementární DNA
DNA	deoxyribonukleová kyselina
<i>E. coli</i>	<i>Escherichia coli</i>
ER	endoplazmatické retikulum
Fuc	fukóza
GA	Golgiho aparát
Gal	galaktóza
GalNAc	<i>N</i> -acetylgalaktosamin
Glc	glukóza
GLCA	kyselina glukuronová
GlcNAc	<i>N</i> -acetylglucosamin
GTP	guanosintrifosfát
HEK293	linie lidských buněk embryonální ledvinové tkáně
CHO	linie vaječnickových buněk křečička čínského
IdoA	kyselina iduronová
IMAC	iontová metaloafinitní chromatografie
KATs	lysin acetyltransferázy
KDACs	lysin deacetylázy
Man	manóza
MCS	multiple cloning site (polylinker)
MOAC	afinitní chromatografie na oxidech kovů
mRNA	messengerová ribonukleová kyselina
MS	hmotnostní spektrometrie
NAD ⁺	nikotinamid adenin dinukleotid
NATs	<i>Nt</i> -acetyl transferázy
NS0	linie buněk myších hybridomů
<i>P. pastoris</i>	<i>Pichia pastoris</i>
pArg	fosfoarginin

pAsp	fosfoasparagin
PCR	polymerázová řetězová reakce
pCys	fosfocystein
pGlu	fosfoglutamin
pHis	fosfohystidin
pLys	fosfolysin
pSer	fosfoserin
pThr	fosfothreonin
PTMs	posttranslační modifikace
pTyr	fosfotyrosin
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SA	kyselina sialová
SIRT3	Sirtuin 3
tRNA	transferová RNA
Xyl	xylóza

Úvod

Technologie rekombinantních proteinů získala v posledních letech velký význam jak ve výzkumu, tak v průmyslu. Široké uplatnění má hlavně v medicíně a farmakologickém průmyslu. Rekombinantní proteiny nabídlu alternativu k izolaci proteinů z jejich přirozených zdrojů, která často vyžadovala velké množství vstupních surovin a složitou purifikaci. Produkce proteinů v expresních systémech umožňuje vysoké výtěžky velmi čistých proteinů.

Pro produkci rekombinantních proteinů je používáno množství různých expresních systémů, jako jsou bakteriální buňky, kvasinky, hmyzí buňky, nebo buňky savčí. Tyto systémy se mezi sebou liší řadou vlastností jako rychlost produkce proteinu nebo schopnost provádět různé posttranslační modifikace. Na posttranslačních modifikacích často závisí správné sbalení a funkčnost proteinu, a jsou tedy důležité pro jeho následné použití. Tato práce se zabývá především fosforylací acetylací a glykosylací rekombinantních proteinů v jednotlivých expresních systémech a rozdílly mezi těmito systémy.

1. Rekombinantní proteiny

Rekombinantní proteiny jsou proteiny kódované rekombinantní DNA, tedy DNA, která vznikla uměle, spojením dvou nebo více úseků různých DNA z různých zdrojů. Tato genetická informace je pak vnesena do hostitelského organismu, kde je replikována, transkribována a translatována. Vzniklý protein je pak izolován a purifikován. Tímto způsobem lze získat velké množství velmi čistého proteinu, pro jehož získání z přirozeného zdroje by bylo potřeba velké množství vstupní suroviny a často složité purifikační kroky. Rekombinantní proteiny nachází široké uplatnění jak ve výzkumu struktury a funkce proteinů, tak v biofarmaceutickém průmyslu, při přípravě léků a vakcín [1].

1.1. Historie rekombinantní technologie

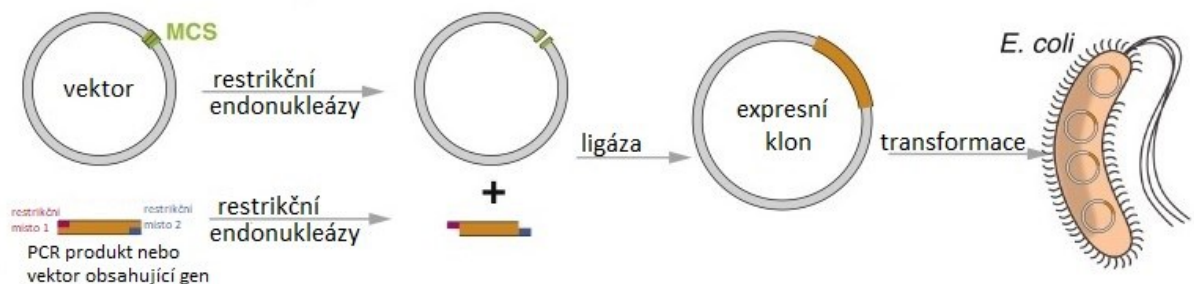
Zásadní pro molekulární biologii a biotechnologie byla práce Watsona a Cricka z roku 1953, ve které autoři popsali strukturu DNA a následně pak objevení enzymů, které umožňují modifikaci molekul DNA, jako restriční endonukleázy a ligáza. Izolace genů a jejich přenášení mezi organismy daly 70. letech vzniknout technologii rekombinantní DNA. V roce 1972 uskutečnili Jackson, Symons a Berg práci popisující tvorbu první rekombinantní molekuly DNA SV40- λ *dvgal* 120. Do molekuly DNA viru SV40 vložili pomocí endonukleázy EcoRI genetickou informaci malého plazmidu, obsahující geny bakteriofágu λ , které mu umožňují replikaci v *Escherichia coli* (*E. coli*), a tři geny *E. coli* potřebné pro metabolismus galaktózy [2]. Bergovi byla za studii a rozvoj rekombinantní technologie v roce 1980 udělena Nobelova cena za chemii. První plazmid byl editován v roce 1973, kdy byly v plazmidu kombinovány geny pro rezistenci na tetracyklin a na kanamycin. Plazmidy se tak postupně staly vektory, hlavními prostředky transportu cizorodé DNA do hostitelských buněk [3].

Roku 1977 byl v *E. coli* připraven rekombinantní somatostatin a o dva roky později pak firma Genentech připravila lidský inzulin. Savčí inzulin však v buňkách vzniká odštěpením prostředního peptidu C a aktivní hormon se skládá ze dvou řetězců propojených disulfidickými můstky. Tuto vazbu je v *E. coli* složité vytvořit, a tak byly oba řetězce exprimovány odděleně a následně chemicky spojeny. Tento produkt nahradil do té doby používaný inzulin izolovaný ze slinivky zvěřat, který často způsoboval imunologické reakce [4]. Od 80 let se na trh postupně dostávaly další rekombinantní hormony, terapeutické látky ale i první rekombinantní vakcína proti hepatitidě B [1].

1.2. Příprava rekombinantních proteinů

Základem přípravy jakéhokoliv rekombinantního proteinu je izolace kódujícího genu. Pokud je tento gen prokaryotický, stačí tento gen izolovat a pomocí polymerázové řetězové reakce (PCR) ho namnožit. Pokud je zvolený gen původu eukaryotického, obsahuje nejspíše introny. Pro expresi je tedy výhodnější izolovat mRNA a reverzní transkripcí ho převést na cDNA. V současnosti je také možné cílovou DNA vytvořit synteticky [5].

Takto připravený gen je následně vložen do vhodného vektoru. Jedna z možností je založena na restrikčních enzimech a ligaci, při které vektory obsahují polylinker (MCS), což je sekvence DNA se štěpnými místy pro restrikční enzymy. Pomocí PCR primerů je do připravené DNA přidáno restrikční místo. Vektor i DNA jsou pak štěpeny stejným restrikčním enzymem, přičemž vznikají kohezní konce a oba fragmenty jsou pak ligázou spojeny, jak je naznačeno na obrázku 1. Dalšími možnostmi je na ligaci nezávislé klonování, a rekombinační klonování [6]. Jsou dostupné také komerční systémy využívající rekombinaci, jako je Gateway (Life Technologies). Ten využívá enzymy z bakteriofága lambda, které rozeznávají specifické att místa ohraničující cílové geny a zprostředkovávají na těchto místech rekombinaci mezi heterologními sekvencemi.



Obrázek 1: Klonování DNA do vektoru s použitím restrikčních endonukleáz a ligázy. Převzato a upraveno z [7].

Následuje vnesení rekombinantního genu do hostitelské buňky určené k expresi proteinu. Tento děj se označuje jako transformace, je-li DNA vnášena do prokaryotické buňky, a jako transfekce, když je vnášena do buňky eukaryotické. Jedna z nejčastěji používaných metod je elektroporace, při které krátkým elektrickým pulzem dojde k dočasné destabilizaci membrány a DNA může proniknout do buňky [8]. Další možností je použití CaCl_2 . Při inkubaci buněk ve studeném roztoku vápenatých iontů dojde ke vzniku pórů v membráně a následným

tepelným šokem dojde ke vstupu DNA do buňky [9]. Rekombinantní DNA může být do buňky také vpravena pomocí viru.

Dále je třeba odlišit transformované či transfekované klony od těch, ve kterých přesun DNA nebyl úspěšný. K tomu se používá selekčních markerů, nejčastěji jde o geny poskytující rezistenci k antibiotikům. Dnes se stále častěji přistupuje k selekci bez použití antibiotik. Jsou používány buňky s delecí v esenciálních genech, které bez správné transformace nepřežívají [10].

Konečným krokem přípravy rekombinantních proteinů je jejich purifikace. Pokud je protein produkován do cytozolu, je nejprve potřeba lyzovat buňky. To se provádí buď enzymaticky např. lysozymem, nebo pomocí fyzikálních metod jako sonikace, zmražení nebo pomocí homogenizátoru [11]. Purifikaci usnadňuje připojení afinitních značek k proteinu a následné použití afinitní chromatografie. Nejpoužívanější je polyhistidinová značka His-tag, kterou tvoří obvykle šest histidinových zbytků. Pro izolaci a purifikaci tvoří histidin koordinační vazbu s ionty přechodných kovů. Používají se ionty Co^{2+} , Cu^{2+} a nejčastěji Ni^{2+} imobilizované na nosič v koloně, principiálně jde tedy o metaloafinitní chromatografii (IMAC) [12].

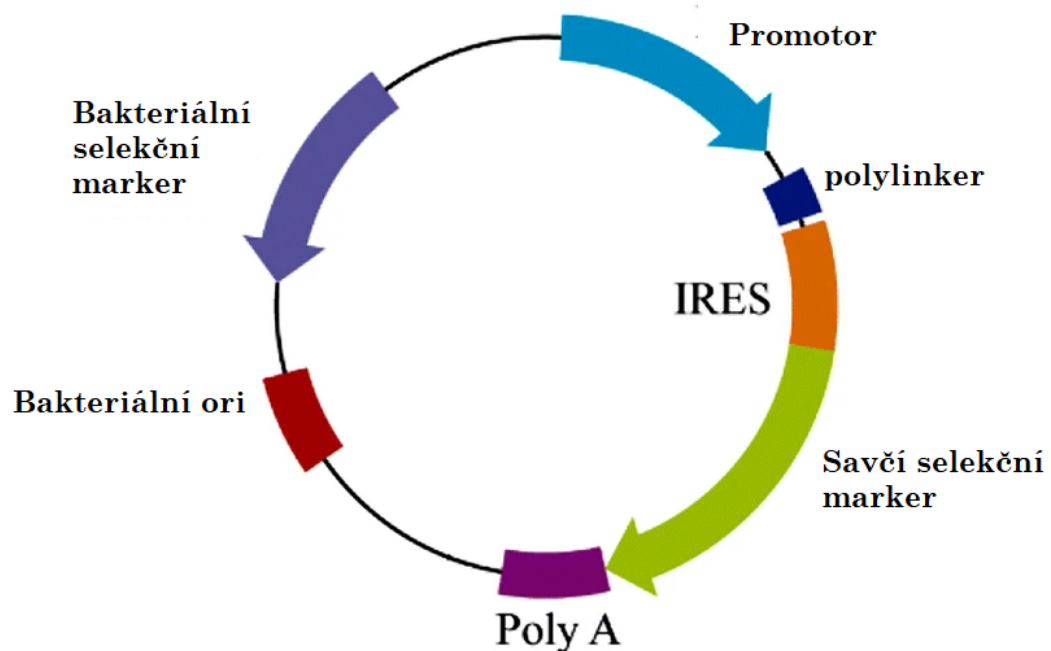
1.3. Vektory

Vektory slouží pro dopravu cizorodé DNA do hostitelské buňky, kde je replikována nebo exprimována. Vektory jsou nejčastěji plazmidy, kruhové molekuly extrachromozomální DNA, která je schopna se nezávisle množit a je přenášena do dalších buněčných generací [13]. Existuje velké množství vektorů, které se liší svými vlastnostmi a způsobem použití. Výběr vektoru záleží na exprimovaném proteinu i použitém expresním systému. Všechny vektory obsahují kombinace nezbytných sekvencí pro expresi proteinu. Patří sem počátek replikace, promotor, polylinker, selekční marker a afinitní značka.

Počátek replikace, nebo také ori, je podmínkou tvorby nových kopií vektoru a záleží na něm výsledný počet kopií vektoru v buňce. Promotor je regulační oblast, na kterou v prokaryotech nasedá polymeráza, a v eukaryotech se sem vážou transkripční faktory. Promotor a okolní regulační sekvence také regulují, s jakou měrou bude transkripce probíhat. Volba promotorů zásadně ovlivňuje výsledný výtěžek proteinu. Konstitutivní promotory jsou stále aktivní a poskytují konstantní transkripci. Zatímco promotory indukibilní se dají regulovat pomocí induktorů a represorů. Patří sem například lac promotor, součást lac operonu velmi často

používaného v expresních systémech *E. coli*. Polylinker umožňuje vložení genu cílového proteinu do vektoru. Obsahuje sekvence, které rozeznávají a štěpí restriční endonukleázy. Selekční marker slouží k rozpoznání buněk, do kterých nebyl správně vnesen vektor. Afinitní značky slouží k purifikaci získaného proteinu [14,15].

Vektory určené pro expresi v eukaryotických systémech se skládají z prokaryotické a eukaryotické části. Prokaryotická část obsahuje počátek replikace a selekční marker a slouží k zmnožení plasmidu v bakteriálních buňkách a následné selekci transfekovaných buněk. V eukaryotické části je pak promotor a selekční marker odlišný od toho v prokaryotické části. Může být přítomný i eukaryotický počátek replikace. Takovéto vektory jsou běžně komerčně dostupné, například firma Life Technologies nabízí vektory, které se pomnožují v *E. coli* a k expresi proteinů dochází v *Pichia pastoris* (*P. pastoris*). Obsahují inducibilní promotor reagující na methanol a selekční markery pro rezistenci k antibiotikům. Pro selekci v bakteriích to bývá ampicillin a pro selekci v kvasinkách pak například zeocin [14]. Takovýto vektor je zobrazený na obrázku 2.



Obrázek 2: Schéma vektoru obsahujícího prokaryotickou a eukaryotickou část. Převzato a upraveno z [16].

2. Expresní systémy

Volba expresního systému, tedy hostitelské buňky, kde bude protein produkován, je jedno z hlavních rozhodnutí při přípravě rekombinantního proteinu. Na expresním systému závisí použité technologie, molekulární nástroje i vybavení. Je třeba zohlednit, zda je protein původu prokaryotického nebo eukaryotického, jeho chemické vlastnosti a velikost a případné posttranslační modifikace (PTMs). Záleží na požadované kvalitě a funkčnosti proteinu, potřebné rychlosti produkce a výtěžku. Nelze také opominout finanční dostupnost jednotlivých systémů.

2.1. Prokaryotní expresní systémy

Bakteriální systémy v čele s *E. coli* jsou nejpoužívanějšími expresními systémy. První rekombinantní proteiny byly produkovány právě v *E. coli*, která je i dnes preferovaným organismem. *E. coli* je častá volba při snaze o expresi nových proteinů a při průmyslové produkci proteinů. To je díky dobře prozkoumanému genomu, fyziologii a metabolismu. Kultivace buněk je snadná a levná a protein je produkován rychle.

Nevýhody použití *E. coli* spočívají hlavně při produkci eukaryotních proteinů, kdy buňky nejsou schopné tvořit eukaryotní PTMs a často tak dochází k nesprávnému složení proteinů anebo tvorbě nefunkčních proteinů. Při vysoké produkci takovýchto proteinů pak dochází k tvorbě inkluzních tělísek v cytoplasmě. Tyto proteiny pak většinou nejsou biologicky aktivní a jejich sbalení *in vitro* je složité. Produkované proteiny jsou také kontaminované endotoxinem a je třeba jejich další přečištění [17,18]. Dalším problémem může být kodonová preference, kdy je u různých organismů rozdílné zastoupení kodonů kódujících stejnou aminokyselinu a tím i tRNA. Vyskytují-li se v genu proteinu kodony, které nejsou v *E. coli* časté, může docházet k chybám v translaci [19].

Pro produkci rekombinantních proteinů se nejčastěji používá buněčná linie BL21 (DE3) a linie odvozené od B21. Tyto buňky více syntetizují aminokyseliny a nemají některé proteázy, které by degradovaly produkováný protein. Také jim chybí genetická výbava pro tvorbu bičíků a mají sníženou pohyblivost. Linie BL21 (DE3) nese profág bakteriofága λ , který obsahuje gen pro T7 RNA polymerázu řízenou promotorem lacUV5 [20,21].

Dalším používanými prokaryotními systémy jsou Gram pozitivní bakterie *Bacillus* a *Streptomyces*. Produkované proteiny jsou u nich sekretovány do okolního média. *Bacillus*

subtilis poskytuje relativně vysoké výtěžky proteinu a neprodukuje endotoxin. Heterologní proteiny jsou ale často degradovány jeho extracelulárním proteolytickým systémem. *Streptomyces lividans* proti tomu má extracelulární proteolytickou aktivitu nízkou [22].

Jako expresní systémy se používají také bakterie mléčného kvašení. Patří sem několik různých druhů bakterií například *Lactococcus*, *Lactobacillus*, *Streptococcus*. Hlavně je využíván *Lactococcus lactis* díky relativně malému a prozkoumanému genomu a dostupnosti molekulárních nástrojů a vektorů. Lze ho mimo jiné použít k produkci rekombinantních alergenů, tyto bakterie mohou být využity k dopravení alergenů do trávicího traktu za vzniku imunotolerance [23].

2.2. Kvasinkové expresní systémy

Kvasinky jsou hojně využívány pro expresi proteinů jak pro průmyslové využití, tak pro medicínské účely. Poskytují totiž současně výhody jednobuněčných organismů a organismů eukaryotických. Tedy malý genom a snadnou genetickou manipulaci, rychlý růst, nízké náklady a schopnost provádět některé PTMs. Díky tomu že jsou to nepatogenní organismy, které mají dlouhou historii využití v potravinářském průmyslu, jsou často první volbou při výběru expresního systému. Nejčastěji používané jsou *Saccharomyces cerevisiae* (*S. cerevisiae*) a *P. pastoris* [24,25].

S. cerevisiae byla prvním použitým kvasinkovým systémem a díky podrobně prozkoumané fyziologii, genetice a fermentačním technikám zůstává nevíce používaným. Je to první eukaryotický organismus, jehož genom byl plně sekvenován a slouží jako modelový organismus pro studium eukaryot. K výhodám *S. cerevisiae* patří její tolerance nízkého pH, a vysokého osmotického tlaku. Překážkou pak může být nadměrná glykosylace heterologních proteinů manózou, která může vést k jejich špatnému sbalení a snížené sekreci a aktivitě. K expresi se používají jak přirozeně se vyskytující, tak geneticky upravené buněčné linie. Bylo vyvinuto množství linií s různými vlastnostmi jako je snížená proteolýza, nadměrná exprese genů pro transport proteinů nebo pro chaperony [26]. *S. cerevisiae* je významy producent proteinů pro farmakologické využití, jako jsou deriváty inzulinu, granulocytární a makrofágový kolonie stimulující faktor nebo povrchový antigen viru hepatitidy B používaný ve vakcínách proti tomuto viru [27].

Pichia pastoris se pro produkci rekombinantních proteinů stala v posledních letech velmi populární volbou. Patří do skupiny methanotropních kvasinek, je tedy schopná využívat methanol jako hlavní zdroj energie. Je zde dostupný silný a snadno regulovatelný promotor alkohol oxidáza 1, který je indukovaný methanolem a jako represor slouží glukóza. Na rozdíl od *S. cerevisiae* zde nedochází k hyperglykosylaci proteinů. Genetická podobnost mezi *S. cerevisiae* a *P. pastoris* umožňuje využívání některých postupů a genetických prvků ze *S. cerevisiae* v *P. pastoris*, jako jsou například sekreční signály, které pak dovolují vyšší výtěžky proteinu [28].

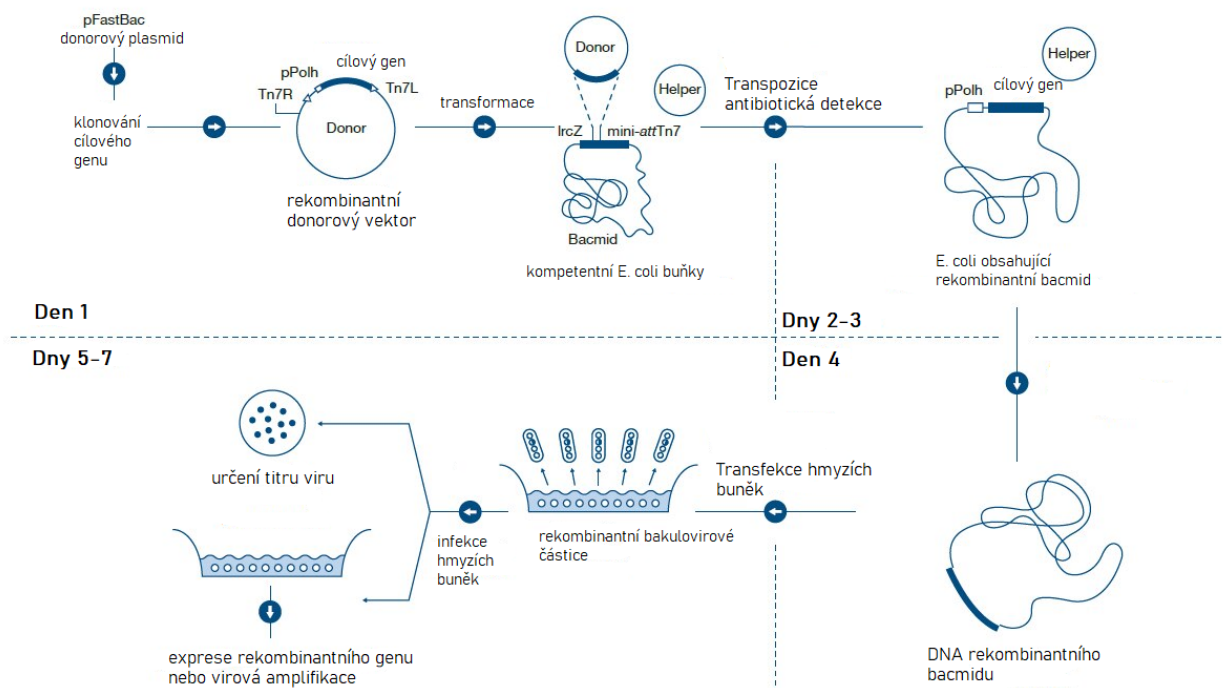
2.3. Hmyzí expresní systémy

Hmyzí buňky, hlavně z řádu lepidoptera jsou hojně využívány pro přípravu eukaryotních proteinů. Umožňují posttranslační modifikace podobné savcím a poskytují funkční, správně sbalené proteiny. Oproti savcím buňkám jsou méně náročné na čas a náklady spojené s jejich kultivací. Nejčastěji jsou používány ve spojení s expresním vektorovým systémem bakulovirů (BEVS). BEVS nabízí možnost produkce jakéhokoli proteinu v jedné buněčné linii kdy je upravován vektor, není tedy potřeba časově náročné přípravy a bezpečnostního schvalování nových buněčných linií pro nové proteiny [29].

Bakuloviry jsou viry, které specificky infikují členovce. Jsou to obalené dsDNA viry, které se replikují v jádrech hmyzích buněk. Pro přípravu rekombinantních proteinů je používán hlavně *Autographa californica*. Rekombinantní bakuloviry jsou připraveny nahrazením genu pro polyhedrin (polh) genem cizího proteinu, který je pak regulován velmi silným polh promotorem. Polyhedrin je protein který tvoří sekundární proteinové pouzdro virových částic. *In vivo* je polyhedrin nutný pro tvorbu okludovaných virů, které jsou schopny přežít ve vnějším prostředí a infikovat larvy. *In vitro* není polyhedrin nezbytný. V pozdní fázi infekce, kdy by běžně probíhala genová exprese pro tvorbu okludovaného viru, je tak produkováno velké množství rekombinantního proteinu řízené polh promotorem [30].

Rekombinantní bakuloviry mohou být konstruovány různými způsoby. Klasickou cestou je homologní rekombinace mezi virovým genomem a transferovým plazmidem, který obsahuje gen cílového proteinu. Pro větší efektivitu metody je genom bakuloviru v místě inserce linearizován [31]. Další možností je využití bacmidů, tedy bakteriálního plazmidu s vloženým genomem bakuloviru. Bacmidy jsou pak propagovány v buňkách *E. coli*, kde je do nich pomocí transportního vektoru transponován cílový gen, jak je naznačeno na obrázku 3. Tento systém

umožňuje snadnou selekci a izolaci rekombinantních bakulovirů, které jsou pak použity k infekci hmyzích buněk [32,33].



Obrázek 3: Schématické znázornění přípravy rekombinantního bacmidu a infekce hmyzích buněk bakulovirem. Převzato a upraveno z [34]

Po infekci buněk rekombinantním virem dojde ke vstupu jeho genetické informace do jádra, kde převezme kontrolu nad genovou expresí buňky. Dochází k její replikaci a syntéze virových částic a zároveň k produkci rekombinantního proteinu. Bakulovir nakonec způsobí lýzu hostitelské buňky, čímž dojde k uvolnění rekombinantního proteinu do média. Dojde ale také k uvolnění množství degradačních enzymů, což může vést k proteolýze [35].

Bakuloviry jsou nepatogenní pro člověka díky úzkému spektru hostitelů, kdy nejsou schopny replikace v obratlovcích. BEVS jsou používány k produkci komplexních, biologicky aktivních proteinů, jako jsou virové a parazitární antigeny. Slouží také k přípravě vakcín, první lidská vakcína produkovaná v hmyzích buňkách je proti lidskému papilomaviru, pro prevenci rakoviny děložního čípku pod označením Cervarix [29].

2.4. Savčí expresní systémy

V posledních letech se savčí expresní systémy dostávají do popředí hlavně pro klinické využití a strukturním studiím proteinů. Tyto systémy jsou vhodné k produkci velkých komplexních molekul s posttranslačními modifikacemi. Nicméně poskytují relativně nízké výtěžky a náklady s nimi spojené jsou vysoké [36].

Pro produkci proteinů se využívá dvou různých postupů. A to buď transientních (přechodných) buněčných linií nebo linií stabilních. Při produkci v transientních liniích se příslušný gen klonuje do savčího expresního vektoru, který je pak amplifikován v buňkách *E. coli*. Následně je transfekován do příslušných savčích buněk, kde ale nedochází k jeho zabudování do genomu buňky. Transientní příprava je méně náročná na přípravu i čas, je ale potřeba velké množství expresních plazmidů a produkce je omezena na jednu generaci buněk [37]. Oproti tomu při využití stabilní transfekce dochází k integraci cílového genu do genomu hostitelské buňky. Tato integrace může být na náhodné místo v genomu, k čemuž je využíváno např. transpozomů nebo virových vektorů, nebo může probíhat cíleně. K tomu je využíváno místně specifické rekombinace a v posledních letech jsou stále oblíbenější metody využívající systém CRISPR/Cas9 [38,39]. Nevýhodou stabilně exprimujících linií je velká časová náročnost na jejich přípravu.

Nejvíce používanými buněčnými liniemi je linie odvozená z vaječnickových buněk křečička čínského (CHO), a lidská linie odvozená embryonální ledvinové tkáně (HEK293). Obě linie dosahují v suspenzi vysoké buněčné hustoty a jsou snadno transfekovatelné.

CHO je v současné době nejvíc používaným systémem pro přípravu terapeuticky významných proteinů. Využívá se k přípravě řady monoklonálních protilátek, enzymů, hormonů jako např. luteinizační hormon, choriový gonadotropin alfa a erythropoetin, faktorů krevní srážlivost VIII a IX, a řady dalších [40]. HEK293 buňky jsou používané hlavně pro produkci proteinů k výzkumným účelům a několika terapeutických proteinů. Poskytují plně lidské posttranslační modifikace a je u nich menší šance imunogenicity [41].

K dalším používaným buněčným liniím patří linie myších hybridomů (NS0), používaná hlavně pro přípravu monoklonálních protilátek, dále fibroblasty z křeččích mláďat (BHK21) a buňky pocházející z lidských embryonálních retinálních buněk (PER.C6) [42].

2.5. Ostatní expresní systémy

Kromě výše uvedených organismů je pro přípravu proteinů využíváno dalších systémů, i když nejsou tak obvyklé. Pro průmyslové využití jsou využívány také vláknité houby, hlavně druhy *Aspergillus*, které jsou schopny poskytovat vysoké výtěžky proteinu a jsou nenáročné na kultivaci. Tvoří ale množství proteolytických enzymů, které získaný protein znehodnocují [43].

Jedním z využívaných systémů jsou transgenní rostliny, kdy je používáno rostlin různých druhů např. tabáku, rýže nebo brambor. Tyto systémy jsou lákavou alternativou k produkci proteinů ve velkém měřítku. Slouží mimo jiné k produkci řady vakcín a protilátek. K výhodám patří absence lidských patogenů, schopnost produkovat komplexní protein i relativně nízké náklady na produkci. Používá se jednak stabilní exprese, a to buď jaderná, nebo chloroplastová, kde je cílový gen začleněn do chloroplastového genomu. Chloroplastová transformace v porovnání s jadernou nabízí větší výtěžky proteinu, ale je technicky náročnější a chybí zde PTMs. Transientní exprese je zde považována za spolehlivější volbu, která nabízí rychlejší produkci a relativně vysoké výtěžky [44].

Pomalejší pokrok zaznamenalo použití transgenních živočichů, a to z důvodů technologických a náročnějším schvalovacím procesům, ale také z důvodů etických. Při použití transgenních živočichů je protein získáván z mléka, krve, případně z vaječného žloutku. Tyto systémy nabízejí stejné PTMs jako savčí buňky, ale nabízejí větší výtěžky. Mají tedy velký potenciál hlavně ve farmaceutickém průmyslu pro výrobu ve velkém měřítku [45].

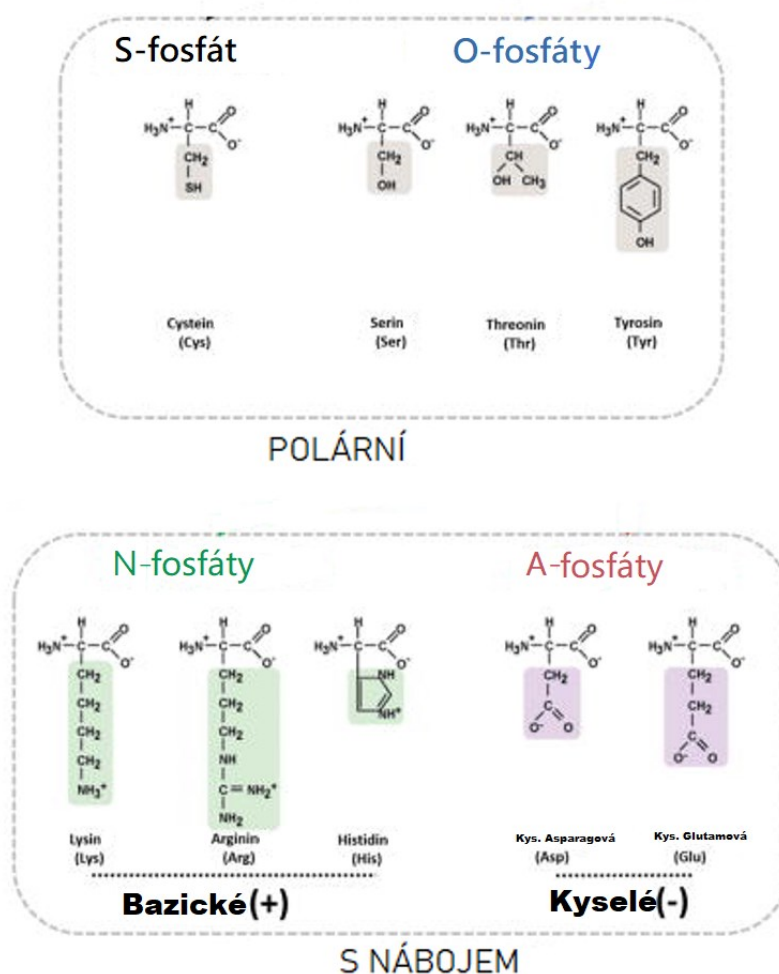
Pro tvorbu malého množství proteinu, hlavně pro výzkumné účely, je možné použít bezbuněčné expresní systémy. Používá se surového buněčného extraktu, získaného lýzou příslušných buněk, nejčastěji CHO, králíčích retikulocytů, ale může být použito i lidských lymfocytů [46]. Bezbuněčné systémy dovolují velmi rychlou přípravu proteinů bez nutnosti tvorby a selekce buněčných linií, kdy je možno aktivně sledovat proces syntézy. Je možno takto připravovat toxické nebo nestabilní proteiny, jejichž produkce v buňkách je omezena [47].

3. Posttranslační modifikace proteinů

U většiny eukaryotických proteinů dochází během translace nebo těsně po ní, ke kovalentním reverzibilním nebo ireverzibilním změnám, označovaným jako posttranslační modifikace. Tyto modifikace ovlivňují fyzikální a funkční vlastnosti proteinu, a regulují tak složení proteinu, enzymovou aktivitu, buněčnou lokalizaci a mezimolekulové interakce. PTMs zahrnují různé změny, jako připojení funkčních skupin při fosforylaci nebo methylaci, připojení delších oligosacharidových řetězců při glykosylaci nebo připojení peptidů nebo menších proteinů například při ubiquitinaci. Různorodost PTMs záleží na buněčném typu, druhu tkáně organismu, v kterém jsou proteiny syntetizovány. Většina modifikací je prováděna vlastní specifickou enzymatickou cestou. Některé modifikace probíhají hlavně u intracelulárních proteinů, zatímco další převažují u proteinů extracelulárních [48]. PTMs jsou nepostradatelné pro regulaci řady biologických funkcí jako transkripce DNA nebo degradace proteinů. Jsou také spojovány s řadou vážných chorob, jako jsou kardiovaskulární a nádorová onemocnění, diabetes nebo Alzheimerova choroba [49,50].

3.1. Fosforylace

Fosforylace je velmi častou posttranslační modifikací proteinů, která se podílí na regulaci velkého množství buněčných dějů. Jde o děj, při kterém dochází k reverzibilní adici zbytku kyseliny fosforečné na boční řetězce aminokyselin proteinu. Fosforylováno může být devět aminokyselin pomocí čtyř různých druhů vazeb za vzniku fosforylovaných aminokyselin pSer, pThr, pTyr, pHis, pLys, pArg, pAsp, pGlu a pCys (obrázek 4). V eukaryotických buňkách dochází hlavně k fosforylaci aminokyselinových zbytků serinu, threoninu, tyrosinu a histidinu, v zastoupení přibližně 60 % pSer, 25 % pThr a 15 % pTyr. Fosforylační reakce jsou řízeny enzymy ze skupiny transferáz, proteinkinázami, kterých je známo něco přes 500 a proteinfosfatázami [51]. Kinázy umožňují transfer zbytku kyseliny fosforečné z vysokoenergetických látek jako ADP nebo GTP, fosfatázy pak způsobují defosforylaci, při které hydrolyticky odštěpují navázanou fosfátovou skupinu. Souhra obou druhů enzymů pak hraje podstatnou roli v signálních drahách.



Obrázek 4: Fosforylace jednotlivých aminokyselin. Převzato a upraveno z [51].

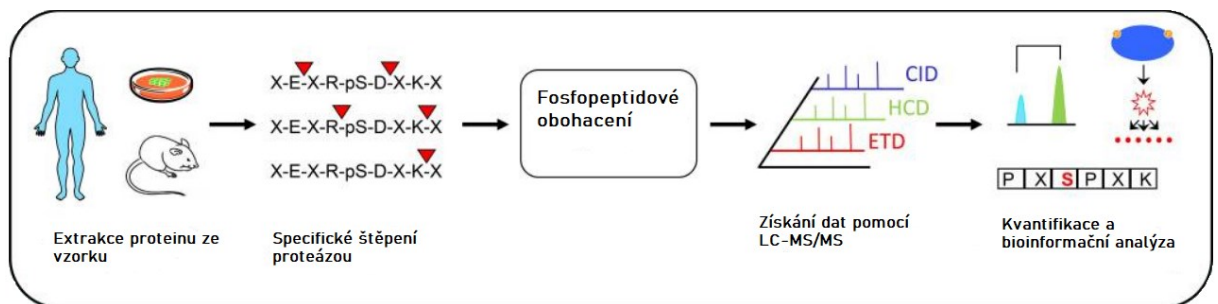
Většina proteinů je fosforylována na několika místech a jejich regulace tak může být značně komplikovaná. Může tak mít podpurný nebo opačný efekt na jednu funkci nebo může regulovat dvě funkce na sobě nezávislé. Mechanismus fosforylace jednotlivých míst může být náhodný bez zvláštních závislostí mezi sebou, nebo musí následovat v přesném pořadí. Třetí možností je struktura s jasnou hierarchií, kdy se fosforylační místo uvolní až po fosforylaci míst ostatních.

Fosforylace reguluje proteiny dvěma hlavními mechanismy, změnou struktury nebo pomocí protein-proteinových interakcí. Při strukturních změnách vytváří záporně nabitá fosfátová skupina vodíkové vazby s okolními řetězci. Nejčastěji dochází k tvorbě vodíkových vazeb s postranními řetězci argininu a lysinu. A výrazná je také interakce s dusíkem na C-konci α -

helikální struktury, kde také dochází k tvorbě vodíkových můstků. Fosforylace způsobuje lokální změny ve struktuře, které se pak mohou projevit celkovými změnami v terciální a kvarterní struktuře. Přejít mezi konformacemi s různou aktivitou nebo vazebnou specifičtostí pak vede k aktivaci, případně deaktivaci proteinu. Mnoho kontrolních buněčných mechanismů je řízeno na úrovni protein-proteinové interakce. Fosforylace hraje roli v regulaci síly těchto interakcí. Dojde-li k fosforylaci v blízkosti vazného místa, může to ovlivnit vazebnou energii komplexu. Ale také fosforylace ve větší vzdálenosti od vazného místa mohou vazbu ovlivnit přes konformační změny proteinu a regulovat tak proteinovou aktivitu [52].

3.1.1. Analýza fosforylací

Identifikací a charakteristikou fosforylovaných proteinů se zabývá fosfoproteomika. Přesto, že je fosforylace velmi častou posttranslační modifikací, které podléhá alespoň jedna třetina všech eukaryotických proteinů, její studium je náročné. Fosforylace je velmi dynamická a neprobíhá stechiometricky, v jednu chvíli se je tedy v buňce přítomno jen malé množství fosforylovaných proteinů, a fosforylované a nefosforylované formy proteinů se vyskytují zároveň. Ke studiu fosfoproteinů se nejčastěji využívá technik založených na hmotnostní spektrometrii (MS) nejčastěji tandemové MS [53]. Kvůli malému množství fosforylované formy proteinů v proteomu je jejich přímá detekce obtížná. Fosforylované proteiny jsou při hmotnostní spektrometrii také ionizovány s menší účinností než jejich nefosforylované formy. Z těchto důvodů byly vyvinuty obohacovací metody, kdy je před samotnou MS analýzou izolována část fosforylovaných proteinů. Schéma průběhu analýzy je na obrázku 5.



Obrázek 5: Schéma analýzy fosforylovaných proteinů. Přejato a upraveno[53]

Při přípravě proteinů pro analýzu je potřeba zabránit aktivitě endogenních fosfatáz pomocí jejich inhibitorů nebo tepelným rozkladem. Proteiny jsou frakcionovány pomocí SDS-PAGE

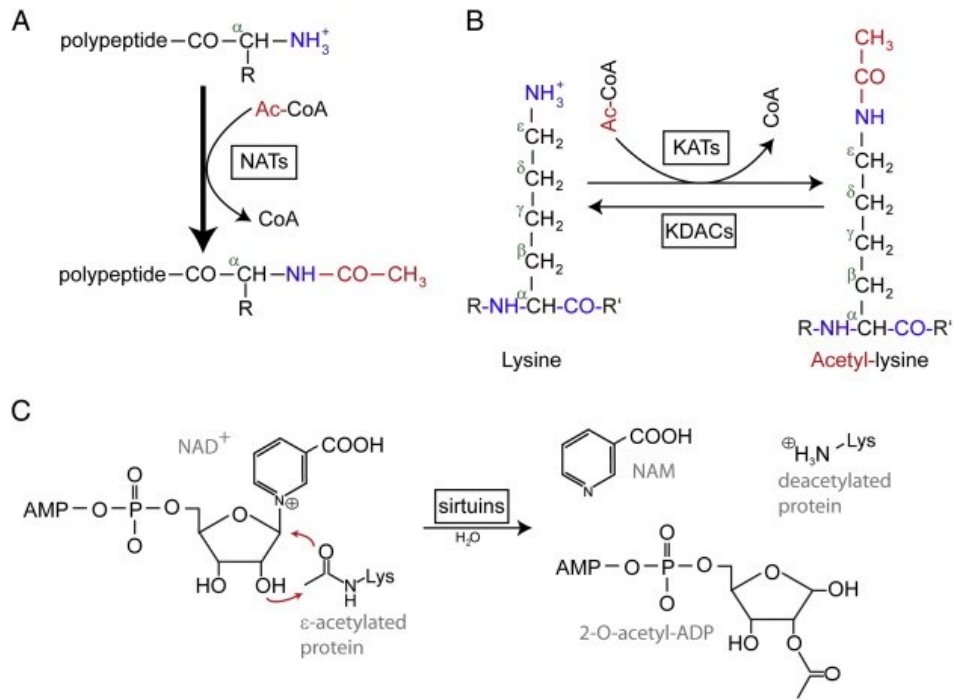
nebo pomocí izoelektrické fokusace. Následně jsou štěpeny na peptidy, nejčastěji pomocí trypsinu [54]. Lze také použít frakcionace již obohacených fosfopeptidů pomocí chromatografických metod [55]. Pro obohacování fosfopeptidů byla vyvinuta řada metod, k nejčastějším patří afinitní chromatografie a imunoprecipitace.

Afinitní chromatografie na imobilizovaných iontech kovů (IMAC) využívá elektrostatických interakcí mezi kationty kovu, nejčastěji Fe^{3+} ukotvenými na nosiči a záporně nabitou fosfátovou skupinou peptidů. Na podobném principu je založena afinitní chromatografie na oxidech kovů (MOAC), ve které je využíváno hlavně TiO_2 . Při použití IMAC dochází k obohacování hlavně několikanásobně fosforylovaných peptidů, zatímco u MAOC jsou obohacovány peptidy monofosforylované [53].

Fosfopeptidy je možné izolovat pomocí specifických protilátek. Imunoprecipitace je používána hlavně pro obohacení tyrosin-fosforylovaných proteinů. Fosforylace tyrosinu je nejméně zastoupena a tato metoda výrazně přispívá k odhalení tyrosinových fosforylačních míst. Protilátky proti fosfoserinu a fosfothreoninu nejsou dostatečně specifické a nejsou běžně používány [56].

3.2. Acetylace

Acetylace je velmi častou posttranslační modifikací jak u eukaryotických, tak u prokaryotických organismů. U eukaryot dochází k přenesení acetylové skupiny z acetylkoenzymu A na peptidový řetězec. U prokaryot může acetyl pocházet také z acetyl fosfátu [57]. Většina lidských proteinů je kotranslačně nebo posttranslačně acetylována na volném N-koncovém alfa dusíku. Tyto reakce jsou katalyzovány pomocí Nt-acetyl transferáz (NATs). Dalším možným způsobem acetylace je reverzibilní připojení acetylu na ϵ -aminoskupinu lysinu. Tyto reakce jsou regulovány pomocí lysin acetyltransferáz (KATs) a lysin deacetyláz (KDACs). Acetylace na lysinu byla poprvé objevena na histonech, kde se podílí na regulaci genové transkripce [58]. Na obrázku 6 je zobrazeno schématické znázornění acetylace a deacetylace.



Obrázek 6: A) Transfer acetylové skupiny z acetyl koenzymu A na N-koncovou skupinu polypeptidu, katalyzovaný pomocí NAT. B) Reverzibilní acetylace ε-aminoskupiny lysinu pomocí KATs, deacetylace je pak katalyzována pomocí KDACs. C) NAD⁺ dependentní deacetylace ε-acetylovaného proteinu, katalizovaná pomocí sirtuinů [58].

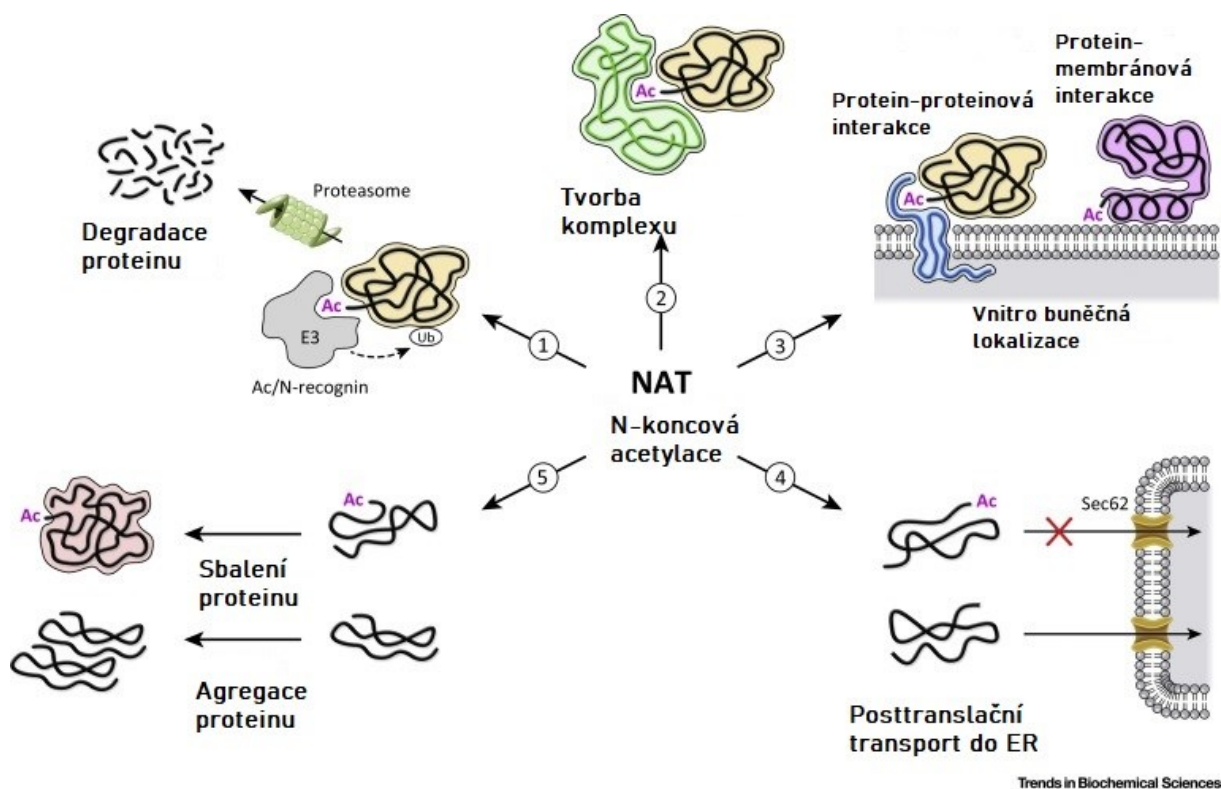
3.2.1. N-koncová acetylace

Nt-acetylace je velmi běžnou posttranslační modifikací, která se vyskytuje u většiny organismů a její frekvence stoupá s komplexitou organismu. Nt-acetylace jsou ireverzibilní. Ne u všech substrátů dojde k úplné Nt-acetylaci, v buňce tak mohou být zároveň přítomny jak acetylované, tak neacetylované proteiny. Navázáním acetylu na N-konec proteinu dochází k neutralizaci jeho kladného náboje, a tak zvýšení jeho hydrofobního charakteru a změně sterických vlastností. To ovlivňuje širokou škálu vlastností proteinů a buněčných procesů, jako je buněčná proliferace, apoptóza a stresové odpovědi. Poškození Nt-acetylace může způsobovat řadu závažných patologií, jako jsou nádorová onemocnění, X-vázaná genetická onemocnění a neurodegenerativní změny [59,60].

NAT enzymy, které tuto acetylaci katalyzují, jsou rozděleny do skupin komplexů od NatA do NatG podle jejich substrátové specifity a složení podjednotek. Zatímco skupiny NatA až NatE jsou geneticky konzervované a jsou přítomny jak u kvasinek, tak u vyšších eukaryot a vyskytují se v cytoplasmě, NatF skupina je přítomna pouze u mnohobuněčných živočichů a vyskytuje se v Golgiho aparátu. Skupina NatG se vyskytuje v chloroplastech rostlin [61,62]. Substrátová

specifita NATs je většinou určena prvními dvěma aminokyselinovými zbytky. NatA tak například acetyluje N-koncové aminokyseliny Ala, Ser, Cys, Val, Thr, a Gly, které se na konec řetězce dostanou odštěpením iniciačního methioninu pomocí Met-aminopoptidázy [63].

Vliv Nt-acetylce je velice různorodý a u jednotlivých proteinů se velmi liší. Všechny důsledky acetylce nebyly zcela objasněny a zůstávají předmětem studia. K nejdůležitějším patří polyubiquitinace a následná degradace Nt-acetylovaného proteinu v rámci dráhy Ac/N-koncového pravidla. Nt-acetylovaná skupina sloužící jako degradační signál byla poprvé popsána v roce 2010 a označena jako Ac/N-degron [64]. Nicméně Ac/N-koncové pravidlo je velmi citlivé na okolní prostředí a podmínky, kdy degron může být chráněn před ubiquitinací pomocí proteinového komplexu a řada proteinů je naopak přítomností Nt-acetylce stabilizována [65]. U některých proteinů může Nt-acetylce fungovat také jako lokalizační determinanta, kdy acetylce brání proteinům vstupu do endoplazmatického retikula [66]. Změna vlastností N konce proteinu pomocí acetylce vede také ke změnám v protein-proteinových interakcích a může vést ke zlepšení tvorby komplexů [67]. Stabilizace N konce acylací může být také důležitá pro správné sbalení proteinu, kdy při absenci NatA v kvasinkách bylo zaznamenáno množství agregovaného a špatně sbaleného proteinu a zvýšená exprese chaperonů [68]. Různé příklady vlivu Nt-acetylce jsou shrnuty na obrázku 7.



Obrázek 7: Shrnutí různých příkladů vlivu Nt-acetylce na protein. Převzato a upraveno z [69].

3.2.2. Acetylace na ϵ -aminoskupině lysinu

Acetylace na ϵ -amino skupině lysinu je reverzibilní PTM a za fosforylací je to nejčastější a nejvíce využívané modifikace pro buněčnou signalizaci a metabolismus. KDACs potřebné k deacetylaci se dělí do dvou skupin, KDACs závislé na zinečnatých iontech a sirtuiny, které využívají NAD^+ jako kosubstrát. KATs jsou závislé na Ac-CoA. NAD^+ a Ac-CoA jsou důležité substráty v buněčném metabolismu a propojují ho tak se signálními drahami [70]. V mitochondriích může díky nízkému pH a vysokým koncentracím Ac-CoA, kterého je zde přítomno 3 - 50x více, docházet ke spontánním acetylacím bez přítomnosti katalyzujícího enzymu. Deacetylace je zde řízena pomocí NAD^+ dependentního Sirtuinu 3 (SIRT3). Acetylací mitochondriálních enzymů dochází k inhibici oxidativního metabolismu. Při nedostatku SIRT3 dochází k hyperacetylaci mitochondriálních proteinů a může tak docházet k rozvoji řady onemocnění, mezi které patří diabetes, metabolický syndrom nebo nádorová onemocnění [71,72].

Acetylace byla poprvé pozorována na histonech, u kterých je její vliv i nejlépe popsán. Vysoká míra acetylace koreluje se zvýšenou transkripcí. Při acetylaci histonů dochází k destabilizaci interakce mezi DNA a histonem, kdy histon ztrácí svůj pozitivní náboj na lysinu, a tak schopnost tvořit solný můstek s negativně nabitým fosfátovým řetězcem DNA. Dochází tak k rozvolnění chromatinu a mohou se navázat transkripční faktory. Proteiny schopné rozpoznávat acetylované lysiny většinou obsahují tzv. bromodoménu. Deacetylace pak naopak míru transkripce snižuje [73].

Acetylována je i široká skupina nehistonových proteinů jako jsou transkripční faktory, cytoskeletální proteiny, chaperony, některé enzymy a receptory a mnoho dalších. Prvním rozpoznávaným nehistinovým KAT substrátem byl tumorový represor p53. Při acetylaci dochází ke zvýšení transkripční aktivity a následnému zastavení buněčného cyklu. Acetylace zde hraje výraznou roli v aktivaci p53. Při mutacích na acetylačních místech dochází k poškození schopnosti potlačovat tumorový nárůst [74].

3.2.3. Analýza acetylací

Pro kvantifikaci a charakteristiku acetylací na lysinu se používá obohacovací technika založená na imunoprecipitaci. Proteiny jsou naštěpeny, většinou pomocí trypsinu na peptidy, které jsou inkubovány s anti-Ac-lysin protilátkou ukotvenou na nosiči. Po štěpení proteinů může

proběhnout frakcionace peptidů. Peptidy jsou pak z nosiče uvolněny a následuje hmotnostní spektrální analýza [72].

U Nt-acetylovaných proteinů je analýza o něco složitější. Obohacování je zde většinou prováděno ve dvou krocích. Nejprve jsou zablokovány volné primární aminy, a to α i ϵ , pomocí stabilní izotopové skupiny $^{13}\text{C}_2\text{D}_3$, aby se odlišily *in vivo* a *in vitro* acetylace. Následně dojde k naštěpení trypsinem a nově vzniklé N-konce jsou separovány pomocí iontově výměnné chromatografie na silném katexu při nízkém pH. Zbylé peptidy jsou dvakrát za sebou frakcionovány pomocí kombinované frakcionační diagonální chromatografie (COFRADIC), kdy dojde k odseparování lysin-acetylovaných peptidů. Pak následuje tandemová MS [75].

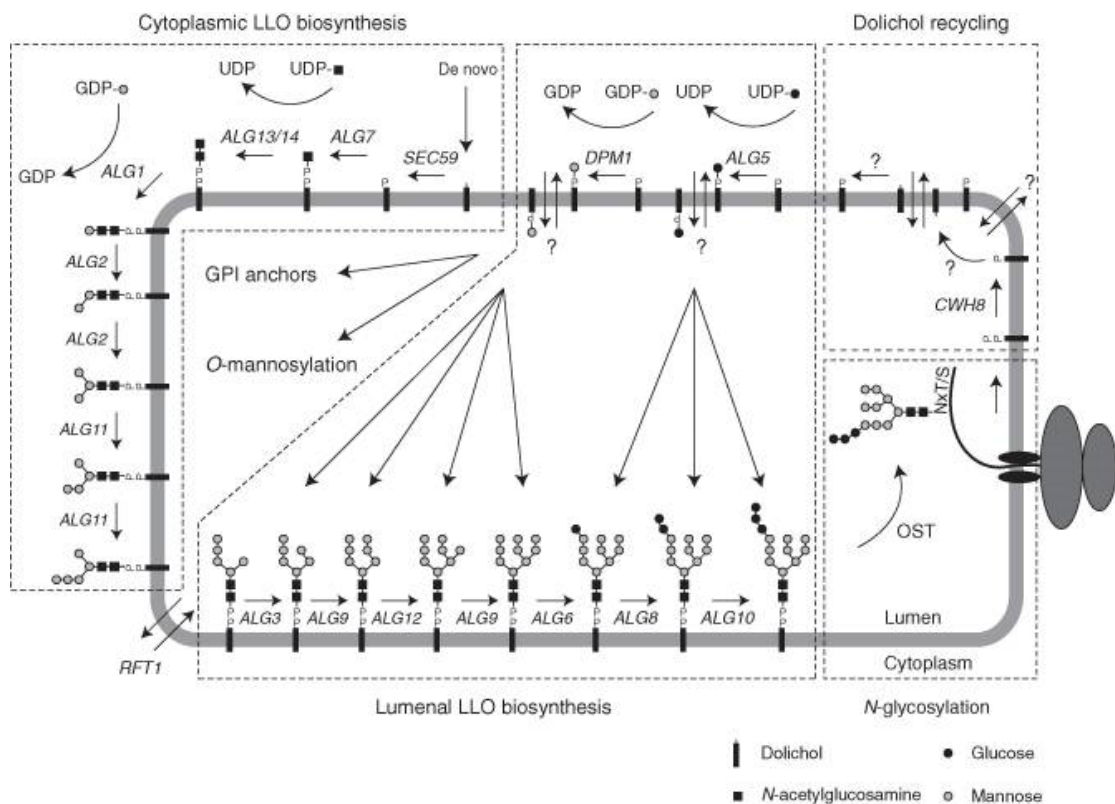
3.3. Glykosylace

Glykosylace je kotranslační a posttranslační modifikace, při které dochází ke kovalentní vazbě glykanu, tedy polysacharidového řetězce na polypeptidový řetězec. Jde o velmi komplexní modifikaci proteinu, která probíhá postupně v mnoha enzymatických krocích. Monosacharidové jednotky mohou být propojovány různými způsoby a mohou zaujímat různé konformace. Savčí glykany jsou syntetizovány složitými biosyntetickými drahami a obsahují deset druhů monosacharidových jednotek: fukózu (Fuc), galaktózu (Gal), glukózu (Glc), N-acetylgalaktosamin (GalNAc), N-acetylglucosamin (GlcNAc), kyselinu glukuronovou (GLCA), kyselinu iduronovou (IdoA), mannózu (Man), kyselinu sialovou (SA) a xylózu (Xyl). Glykany jsou k proteinům připojeny pomocí N-glykosylace, O-glykosylace, C-manosylace nebo O-GlcNAcylation. Odhaduje se, že na glykosylaci se u savců podílí přibližně 700 proteinů, z čehož asi 200 jsou glykosyltransferázy. Glykosylace mohou mít zásadní vliv na strukturní a funkční vlastnosti proteinu. Glykany mohou být přímo rozpoznávány vazebnými proteiny, což může vést ke zvýšené buněčné adhezi, buněčné signalizaci a buněčným interakcím. Glykosylace také ovlivňuje správné sbalení proteinů, a jejich cílení do buněčných organel [76]. Defekty v glykosylaci mají vliv na řadu onemocnění a jsou spojovány s nádorovými onemocněními a s cystickou fibrózou. Vrozená porucha glykosylace je vzácné genetické onemocnění, které způsobuje vážná orgánová selhání [77].

3.3.1. N-glykosylace

N-glykosylace je nejčastějším a nejstudovanějším druhem glykosylace v eukaryotních organismech. N-glykosidická vazba vzniká mezi oligosacharidem a amidovým dusíkem postranního řetězce asparaginu. Je katalyzována enzymem oligosacharyltransferázou (OST), která připojuje oligosacharid na asparagin v aminokyselinové sekvenci N-X-S/T, ve které X je jakákoli aminokyselina kromě prolinu.

Syntéza glykanu začíná na cytosolové straně membrány endoplazmatického retikula, kde dojde pomocí transferáz k postupnému navázání dvou GlcNAc a pěti Man na dolicholfosfát. Jako substrát zde slouží aktivované sacharidy v podobě UDP-GlcNAc a GDP-Man. Vzniklý $\text{Man}_5\text{GlcNAc}_2\text{-P-P-Dol}$ je pak translokován do lumen endoplazmatického retikula, kde jsou na něj připojeny další monosacharidy, za vzniku oligosacharidu $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$. Pro navázání monosacharidů uvnitř lumen jsou tyto sacharidy translokovány z cytosolu v podobě Dol-P-Man and Dol-P-Glc. Vzniklý oligosacharid je pak pomocí oligosacharyltransferázového komplexu přenesen na asparaginový zbytek proteinu [78,79]. Průběh je naznačený na obrázku 8.



Obrázek 8: Průběh N-glykosylace v ER [80].

N-glykosylace v endoplazmatickém retikulu napomáhá správnému sbalení proteinu, chybně sbalené proteiny jsou značeny a degradovány v proteazomu [79]. Glykosylované proteiny jsou pak dále přepravovány do Golgiho aparátu (GA). Předtím jsou z řetězce odštěpeny tři glukózové a jeden manózový zbytek. Glykosylovaný protein je pak v GA dále upravován a jsou tvořeny komplexní glykany. V jednotlivých kompartmentech GA dochází k odštěpování manózových jednotek a připojení fukózy na vnitřní část řetězce. Na neredukující konce řetězců jsou navázány GlcNAc, Gal a SA [77].

3.3.2. O-glykosylace

V porovnání s N-glykosylací je O-glykosylace méně zastoupená, oligosacharidové řetězce jsou více variabilní a její mechanismus je méně objasněn. O-glykosylace probíhá hlavně v Golgiho aparátu a jednotlivé sacharidy jsou připojovány postupně. Navázání glykanu probíhá přes OH skupiny serinu nebo threoninu. V eukaryotických buňkách se vyskytují dvě hlavní formy O-glykanů. Glykany mucinového typu a glykosaminoglykanové (GAG) řetězce proteoglykanů.

Syntéza O-glykanů mucinového typu je zahájena připojením GalNAc pomocí GalNAc transferázy. Na protein se pak při průchodu jednotlivými částmi GA vážou další sacharidy. Jsou přítomny na extracelulárních a sekretovaných glykoproteinech, včetně mucinů. Muciny mají ochrannou funkci na povrchu epitelu. Obsahují tandemová opakování s množstvím Pro, Ser a Thr které nabízí vhodná glykosylační místa [81].

GAG řetězce jsou navázané na serin pomocí řetězce xylóza-Gal-Gal-GlcA Na který se v Golgiho aparátu dále vážou disacharidové jednotky GlcNAc-GlcA nebo GlcNAc-kyselina iduronová a tvoří dlouhé lineární polymery. V GA se na tyto řetězce často vážou sulfáty. Patří sem například heparan sulfát, chondroitin sulfát, keratan sulfát and hyaluronan [82].

3.3.3. Analýza glykosylací

Podobně jako u analýzy předchozích posttranslačních modifikací je i u glykosylací potřeba oddělit malé množství modifikovaných proteinů od těch nemodifikovaných, které jsou zastoupeny mnohonásobně více. Kvůli velké variabilitě navázaných glykanů je však analýza glykosylací poněkud složitější.

Glykované proteiny se nejprve našťepí na směs peptidů. Pro obohacení glykosylovaných peptidů je používána řada metod. Jsou používány lektiny imobilizované na filtru, které mají

vysokou afinitu k sacharidovým jednotkám. Glykované peptidy jsou tedy zadrženy, zatímco ostatní promyty [83]. Další možností je použití boronové kyseliny, která s cukry vytváří reverzibilní vazbu. Na rozdíl od laktonů je jí možné použít pro různé glykany.

Pro zjištění místa glykosylace se nejčastěji používá enzymů, které glykan odštěpí a na jeho místě nechají hmotnostní značku pro MS. Kvůli různorodosti glykanů ale neproběhne toto odštěpení u všech glykanů. Následuje tandemová MS [84].

4. Posttranslační modifikace proteinů v expresních systémech

Existuje množství různých expresních systémů pro přípravu rekombinantních proteinů, které se liší ve schopnosti poskytovat určité posttranslační modifikace. Ty mají zásadní vliv na funkce proteinu. Zvláště pak u biofarmak, kde může změna v posttranslační modifikaci hrát významnou roli v účinku a biologickém poločase léčiva, případně i způsobovat nechtěnou imunologickou odpověď [85].

Zatímco savčích expresních systémech mohou všechny posttranslační modifikace probíhat tak, jak bylo popsáno v předchozí kapitole, v ostatních expresních systémech tomu tak není. Velké rozdíly mezi jednotlivými systémy jsou hlavně v glykosylaci a při volbě expresního systému je třeba zvážit jeho možnosti požadavky daného proteinu na posttranslační modifikace. Existuje snaha o zavádění některých savčích PTMs do jednodušších systémů, převážně do *E. coli*. Přesto pro přípravu komplexních lidských proteinů jsou savčí buněčné linie stále nejjistější volbou.

4.1. Posttranslační modifikace v prokaryotických expresních systémech

Prokaryotní systémy, jako *E. coli*, jsou oblíbenou volbou pro přípravu rekombinantních proteinů. Nicméně nejsou vhodné pro přípravu komplexních proteinů, protože nejsou schopny většiny posttranslačních modifikací proteinů.

Přestože Ser/Thr fosforylační mechanismy posttranslační modifikace jsou v prokaryotických buňkách přítomny, nezaručují správnou modifikaci rekombinantních eukaryotických proteinů [86]. Autoři Hee-Sung Park a Michael J. Hohn vyvinuli postup, který umožňuje cílenou kotranslační inkorporaci fosfoserinu (Sep) do proteinu produkovaném v *E. coli*. Využívá Sep-tRNA syntetázy která se vyskytuje v některých archea, a tvoří Sep-tRNA modifikovanou pro rozpoznání UAG kodonu. UAG kodon běžně funguje jako stop kodon a pro zvýšení výtěžku proteinu je potřeba inhibovat release factor 1 [87,88].

Podobná situace je u acetylací, kdy jsou sice v bakteriích přítomny enzymy katalyzující acetylaci, ale ta je mnohem méně častá a její regulace a specifita nejsou dobře známé. Jedna z využívaných možností pro Nt-acetylaci eukaryotických proteinů exprimovaných v bakteriích je koexprese kvasinkových NAT spolu s rekombinantním proteinem. Při koexpresi ale může docházet k tvorbě směsi acetylovaných a neacetylovaných proteinů, protože N-acetylace probíhá hlavně kotranslačně a NAT nemusí být správně složený před expresí

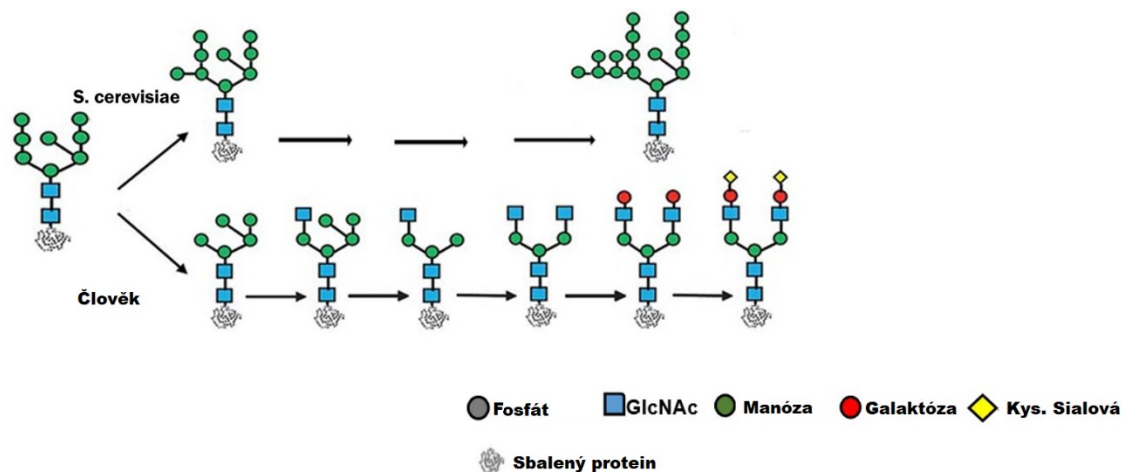
cílového proteinu [89]. Do proteinu lze také inkorporovat acetyllysin podobným způsobem jako fosfoserin [90].

V některých prokaryotických buňkách byly popsány také N- a O-glykosylace. Některé připomínající ty probíhající v eukaryotických buňkách a některé jsou zcela unikátní pro bakterie. Tradiční N-glykosylační reakce byly popsány pouze u několika gram negativních bakterií, modelovým organismem je pro ně *Campylobacter jejuni*. Syntéza oligosacharidu se podobá té v eukaryotických buňkách. Aktivované monosacharidy jsou v cytoplazmě vázány na undecaprenyl fosfát a vzniklý oligosacharid je translokován přes vnitřní membránu do periplazmy. Následně je analogem oligosacharyltransferázy připojen na asparaginový zbytek proteinu. Tuto N-glykosylační dráha se podařilo vložit do *E. coli*. Výťažky glykosylovaného proteinu z takto upravené *e. coli* jsou ale velmi malé [91].

4.2. Posttranslační modifikace v kvasinkových expresních systémech

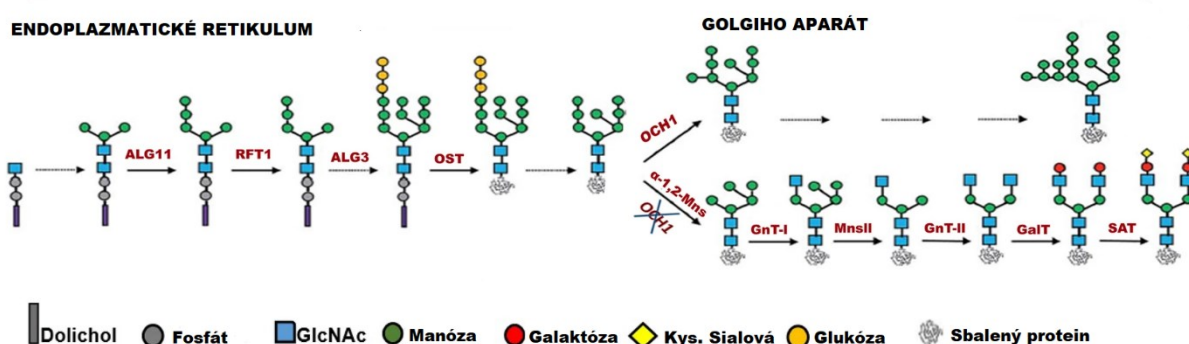
Kvasinky jsou schopné provádět většinu PTMs které se vyskytují v lidských buňkách a poskytovat funkční správně sbalené proteiny. Jsou schopné N- i O-glykosylace, nicméně tady se od savčích buněk liší produkcí vysoce manosylovaných glykanů.

N-glykosylační dráha v ER je shodná se savčími buňkami, a zahrnuje tvorbu Glc3Man9GlcNAc2-PP-dolicholových prekurzorů, ze kterých je pak glykan připojen na protein. U většiny kvasinek následuje i odštěpení tří glukóz a jedné manózy. Výjimku zde tvoří *Kluyveromyces lactis* *Schizosaccharomyces pombe*. Dráhy se liší po transportu glykoproteinu do Golgiho aparátu, kde u kvasinek dochází k navázání manózových zbytků, jak je naznačeno na obrázku 9. Počet navázaných manóz se u jednotlivých druhů liší. Nejvýraznější hyperglykosylaci provádí *S. cerevisiae*, která připojuje 50-150 manózových zbytků. *P. pastoris* a *H. polymorpha* přidávají až 20 a *Y. lipolytica* přidává 8–10 manóz [92,93].



Obrázek 9: Rozdíl mezi glykosylací v lidských buňkách a buňkách *S.cerevisiae*. Převzato a upraveno z [92].

Až 70 % terapeutických proteinů jsou glykoproteiny a správná glykosylace je důležitá pro jejich správné působení. Jednak kvůli dostatečnému poločasů života proteinu a jednak proto, že jiné než lidské typy glykosylací mohou způsobovat imunologickou odpověď organismu. Proto byly vyvinuty kmeny kvasinek produkující glykany podobné lidským [24]. V těchto kmenech je omezena, nebo úplně potlačena hypermannozylace, jak je znázorněno na obrázku. Provádí se to inaktivací nebo delecí genu *OCH1*, který kóduje 1,6-mannosyltransferázu působící v GA. Zároveň je exprimována α -1,2-mannosidáza jejímž působením vznikají $\text{Man}_5\text{GlcNAc}_2$ glykany [94].

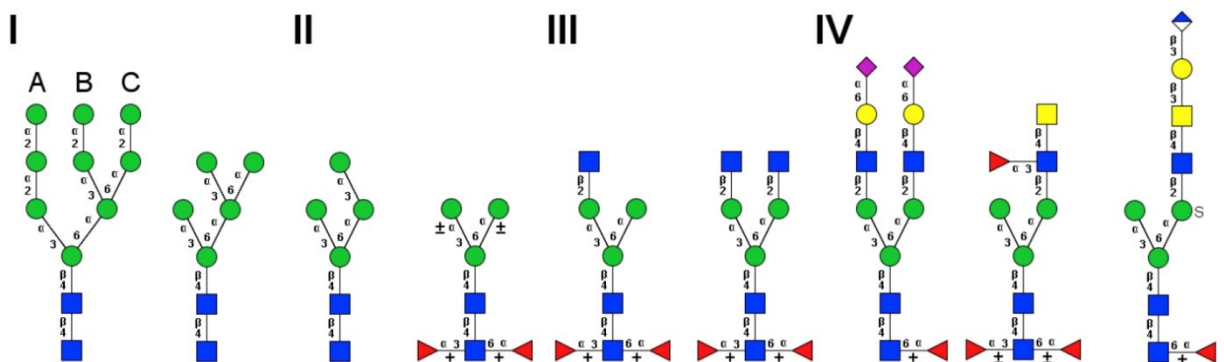


Obrázek 10: Syntéza glykanu lidského typu v kvasinkách, pomocí delecí genu *OCH1*. Převzato a upraveno z [92].

4.3. Posttranslační modifikace v expresních systémech hmyzích buněk

Hmyzí buňky jsou schopny produkovat proteiny obsahující posttranslační modifikace, které se velmi blíží těm přítomným v savčích buňkách. Jsou schopny N- i O-glykosylace, nicméně struktura produkovaných N-glykanů se od savčích liší. Jsou zde produkovány hlavně glykany manózového a paucimanózového typu. Paucimanózové glykany obsahují méně než pět manóz a nemají terminální GlcNAc.

Glykosylační kroky v ER jsou stejné jako u savců, rozdíl nastává po transportu do GA. V GA pak N-glykan může být zkrácen na $\text{Man}_5\text{GlcNAc}_2$ odštěpením manózových zbytků pomocí I α -mannosidázy. Pro velkou část glykoproteinů v této fázi úpravy končí a jsou dále distribuovány po buňce, případně sekretovány. Další úpravy v GA vedou k produkci paucimanózových glykanů, často s navázanými fukózami. $\text{Man}_5\text{GlcNAc}_2$ může být prodloužen o GlcNAc což pak dovoluje manosidáze odstranit z řetězce dvě manózy za vzniku $\text{GlcNAcMan}_3\text{GlcNAc}_2$, což je hlavní intermediát pro paucimanózové, fukosylované i komplexní glykany. Jedna z částí glykosylační dráhy, která odlišuje hmyzí buňky od savčích je β -N-Acetylglucosaminidáza, která odstraňuje terminální GlcNAc. Tím je tvorba glykanů směřovaná k tvorbě paucimanózových glykanů a vzniká menší množství komplexních glykanů [95,96]. Jednotlivé typy N-glykanů produkované v hmyzích buňkách jsou na obrázku 11.



Obrázek 11: Typy N-glykanů produkované v hmyzích buňkách. ■ - GalNAc, ● - Gal, ■ - GlcNAc, ▲ - Fuc, ◆ - GlcA, ● - Man, ◆ - SA. I) Glykany manózového typu s pěti až devíti manózami, II) Glykany paucimanózového typu obsahující méně než pět manóz, bez terminální GlcNAc, některé s navázanou fukózou, III) Běžné hmyzí glykany s terminální GlcNAc, IV) v hmyzích buňkách vzácné komplexní glykany. Převzato z [95].

Několik postupů ukázalo, že je v hmyzích buňkách možné produkovat i proteiny se savčími glykosylacemi. Využívá se koexprese savčích glykosyltransferáz a je tak možné hmyzí buňky

upravit pro produkci komplexních glykanů. Koexprimovány jsou enzymy β -1,2-*N*-acetylglucosaminyltransferáza I, β -1,4-galactosyltransferáza, a sialyltransferáza. Také inhibice β -*N*-Acetylglucosaminidázy se ukázala jako účinná [97,98].

5. Závěr

Rekombinantní proteiny mají nezastupitelnou roli ve výzkumu i v medicíně. Neustále jsou vyvíjeny nové postupy a navrhovány nové vektory pro optimalizaci růstu buněčných kultur a expresi proteinů. Rekombinantní proteiny se používají v diagnostice, jako terapeutika i jako vakcíny. K tomu, aby mohly proteiny správně fungovat je potřeba, aby byly správně sbalené ve své nativní formě. To vyžaduje, aby u nich byly správně provedeny posttranslační modifikace. Každý expresní systém má jiné možnosti provádění těchto posttranslačních modifikací a řadu dalších výhod a nevýhod. Při volbě expresního systému je tedy nutné vzít v úvahu vlastnosti exprimovaného proteinu, z jakého organismu pochází, jaké má posttranslační modifikace a také požadavky na jeho výslednou čistotu.

Prokaryotní expresní systémy jsou oblíbené pro svou snadnou kultivaci vysoké výtěžky proteinu a nízké náklady. Nejsou ale schopny posttranslačních modifikací a protein může být kontaminován endotoxinem. Kvasinkové expresní systémy jako eukaryotické jednobuněčné organismy nabízejí snadnou přípravu expresních vektorů a kultivaci buněk, a zároveň jsou schopny některých PTMs. Hmyzí expresní systémy nabízejí většinu PTMs jsou ale náročnější na kultivaci. Savčí systémy jsou schopné plnohodnotných PTMs a jsou schopné poskytovat správně sbalené proteiny. K jejich nevýhodám patří pomalý růst buněk, náročná kultivace a vysoké náklady.

E. coli je nejpoužívanějším systémem v oblasti průmyslu, i přesto že není schopna provádět většinu PTMs, zejména pro nenáročnost kultivace a vysoké výtěžky. Komplexní lidské proteiny určené k terapeutickému použití jsou nejčastěji produkovány v eukaryotických, nejlépe v savčích expresních systémech, které jsou schopny PTMs provádět. Největší rozdíly v posttranslačních modifikacích jednotlivých systémů jsou u glykosylace, která se výrazně liší i mezi eukaryotickými organismy. Neustále jsou také vyvíjeny nové metody, které umožňují provádět savčí PTMs v ostatních expresních systémech.

Literatura

1. Khan, S. *et al.* Role of Recombinant DNA Technology to Improve Life. *Int J Genomics* **2016**:2405954, (2016).
2. Jackson, D. A., Symons, R. H. & Berg, P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proc Natl Acad Sci U S A* **69**, 2904–2909 (1972).
3. Nora, L. C. *et al.* The art of vector engineering: towards the construction of next-generation genetic tools. *Microb Biotechnol* **12**, 125–147 (2018).
4. Medicine, I. of M. (US) C. on T. I. in, Rosenberg, N., Gelijns, A. C. & Dawkins, H. *Incentives and Focus in University and Industrial Research: The Case of Synthetic Insulin*. 157–186 (National Academies Press (US), 1995)
5. Recombinant protein production in bacterial hosts. *Drug Discovery Today* **19**, 590–601 (2014).
6. Gibson, D. G. Enzymatic Assembly of Overlapping DNA Fragments. *Methods Enzymol* **498**, 349–361 (2011).
7. Celie, P. H., Parret, A. H. & Perrakis, A. Recombinant cloning strategies for protein expression. *Current Opinion in Structural Biology* **38**, 145–154 (2016).
8. Kant, R. *et al.* Synchronized Electromechanical Shock Wave-Induced Bacterial Transformation. *ACS Omega* **4**, 8512–8521 (2019).
9. Liu, J. *et al.* An Improved Method of Preparing High Efficiency Transformation Escherichia coli with Both Plasmids and Larger DNA Fragments. *Indian J Microbiol* **58**, 448–456 (2018).
10. Bacterial expression systems for recombinant protein production: E. coli and beyond. *Biotechnology Advances* **30**, 1102–1107 (2012).

11. Wingfield, P. T. Overview of the Purification of Recombinant Proteins. *Curr Protoc Protein Sci* **80**, 6.1.1-6.1.35 (2015).
12. Kimple, M. E., Brill, A. L. & Pasker, R. L. Overview of Affinity Tags for Protein Purification. *Curr Protoc Protein Sci* **73**, Unit-9.9 (2013).
13. Růčková, E., Müller, P. & Vojtěšek, B. Expresa a purifikace proteinů. *Klinická onkologie* **27**, (2014).
14. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* **98**, 5301–5317 (2014).
15. Jia, B. & Jeon, C. O. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biol* **6**, (2016).
16. Wang, T.-Y. & Guo, X. Expression vector cassette engineering for recombinant therapeutic production in mammalian cell systems. *Appl Microbiol Biotechnol* **104**, 5673–5688 (2020).
17. Khoo, O. & Suntrarachun, S. Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pac J Trop Biomed* **2**, 159–162 (2012).
18. Ferrer-Miralles, N. & Villaverde, A. Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microb Cell Fact* **12**, 113 (2013).
19. Lipinszki, Z. *et al.* Enhancing the Translational Capacity of *E. coli* by Resolving the Codon Bias. *ACS Synth. Biol.* **7**, 2656–2664 (2018).
20. Jeong, H., Kim, H. J. & Lee, S. J. Complete Genome Sequence of *Escherichia coli* Strain BL21. *Genome Announc* **3**, (2015).
21. Yoon, S. H. *et al.* Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. *Genome Biol* **13**, R37 (2012).
22. Burdette, L. A., Leach, S. A., Wong, H. T. & Tullman-Ercek, D. Developing Gram-negative bacteria for the secretion of heterologous proteins. *Microb Cell Fact* **17**, (2018).

23. Song, A. A.-L., In, L. L. A., Lim, S. H. E. & Rahim, R. A. A review on *Lactococcus lactis*: from food to factory. *Microb Cell Fact* **16**, (2017).
24. Çelik, E. & Çalık, P. Production of recombinant proteins by yeast cells. *Biotechnology Advances* **30**, 1108–1118 (2012).
25. Waegeman, H. & Soetaert, W. Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering. *Journal of Industrial Microbiology and Biotechnology* **38**, 1891–1910 (2011).
26. Liu, Z., Tyo, K. E. J., Martínez, J. L., Petranovic, D. & Nielsen, J. Different Expression Systems for Production of Recombinant Proteins in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* **109**, 1259–1268 (2012).
27. Adrio, J.-L. & Demain, A. L. Recombinant organisms for production of industrial products. *Bioeng Bugs* **1**, 116–131 (2010).
28. Barrero, J. J., Casler, J. C., Valero, F., Ferrer, P. & Glick, B. S. An improved secretion signal enhances the secretion of model proteins from *Pichia pastoris*. *Microb Cell Fact* **17**, (2018).
29. Cox, M. M. J. Recombinant protein vaccines produced in insect cells. *Vaccine* **30**, 1759–1766 (2012).
30. Drugmand, J.-C., Schneider, Y.-J. & Agathos, S. N. Insect cells as factories for biomanufacturing. *Biotechnology Advances* **30**, 1140–1157 (2012).
31. van Oers, M. M. Opportunities and challenges for the baculovirus expression system. *Journal of Invertebrate Pathology* **107**, S3–S15 (2011).
32. Yazdani, Y., Azari, S. & Kalhor, H. R. Expression of Functional Recombinant Human Tissue Transglutaminase (TG2) Using the Bac-to-Bac Baculovirus Expression System. *Adv Pharm Bull* **6**, 49–56 (2016).

33. Scholz, J. & Suppmann, S. A new single-step protocol for rapid baculovirus-driven protein production in insect cells. *BMC Biotechnol* **17**, (2017).
34. Insect Cell–Based Protein Expression — Protein Expression Handbook - CZ. at <http://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/protein-expression-handbook/pex-handbook-insect-cell-based-protein-expression.html> online 24.5.2021
35. Contreras-Gómez, A., Sánchez-Mirón, A., García-Camacho, F., Molina-Grima, E. & Chisti, Y. Protein production using the baculovirus-insect cell expression system. *Biotechnology Progress* **30**, 1–18 (2014).
36. Xiao, S., Shiloach, J. & Betenbaugh, M. J. Engineering Cells to Improve Protein Expression. *Curr Opin Struct Biol* **0**, 32–38 (2014).
37. Hacker, D. L. *et al.* Polyethyleneimine-based transient gene expression processes for suspension-adapted HEK-293E and CHO-DG44 cells. *Protein Expression and Purification* **92**, 67–76 (2013).
38. Mayrhofer, P. *et al.* Accurate comparison of antibody expression levels by reproducible transgene targeting in engineered recombination-competent CHO cells. *Appl Microbiol Biotechnol* **98**, 9723–9733 (2014).
39. Chi, X., Zheng, Q., Jiang, R., Chen-Tsai, R. Y. & Kong, L.-J. A system for site-specific integration of transgenes in mammalian cells. *PLoS One* **14**, (2019).
40. Dumont, J., Euwart, D., Mei, B., Estes, S. & Kshirsagar, R. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit Rev Biotechnol* **36**, 1110–1122 (2016).
41. McCue, J. *et al.* Validation of the manufacturing process used to produce long-acting recombinant factor IX Fc fusion protein. *Haemophilia* **20**, e327–e335 (2014).

42. Kunert, R. & Reinhart, D. Advances in recombinant antibody manufacturing. *Appl Microbiol Biotechnol* **100**, 3451–3461 (2016).
43. Ward, O. P. Production of recombinant proteins by filamentous fungi. *Biotechnology Advances* **30**, 1119–1139 (2012).
44. Loh, H.-S., Green, B. J. & Yusibov, V. Using transgenic plants and modified plant viruses for the development of treatments for human diseases. *Curr Opin Virol* **26**, 81–89 (2017).
45. Hunter, P. The prospects for recombinant proteins from transgenic animals. *EMBO Rep* **20**, (2019).
46. Burgenson, D. *et al.* Rapid recombinant protein expression in cell-free extracts from human blood. *Sci Rep* **8**, (2018).
47. Carlson, E. D., Gan, R., Hodgman, C. E. & Jewett, M. C. Cell-Free Protein Synthesis: Applications Come of Age. *Biotechnol Adv* **30**, 1185–1194 (2012).
48. Craveur, P., Narwani, T. J., Rebehmed, J. & de Brevern, A. G. Investigation of the impact of PTMs on the protein backbone conformation. *Amino Acids* **51**, 1065–1079 (2019).
49. Martin, L., Latypova, X. & Terro, F. Post-translational modifications of tau protein: implications for Alzheimer's disease. *Neurochem Int* **58**, 458–471 (2011).
50. Kamath, K. S., Vasavada, M. S. & Srivastava, S. Proteomic databases and tools to decipher post-translational modifications. *Journal of Proteomics* **75**, 127–144 (2011).
51. Adam, K. & Hunter, T. Histidine kinases and the missing phosphoproteome from prokaryotes to eukaryotes. *Lab Invest* **98**, 233–247 (2018).
52. Nishi, H., Hashimoto, K. & Panchenko, A. R. Phosphorylation in protein-protein binding: effect on stability and function. *Structure* **19**, 1807–1815 (2011).
53. Arrington, J. V., Hsu, C.-C., Elder, S. G. & Tao, W. A. Recent Advances in Phosphoproteomics and Application to Neurological Diseases. *Analyst* **142**, 4373–4387 (2017).

54. Huttlin, E. L. *et al.* A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. *Cell* **143**, 1174–1189 (2010).
55. Engholm-Keller, K., Hansen, T. A., Palmisano, G. & Larsen, M. R. Multidimensional Strategy for Sensitive Phosphoproteomics Incorporating Protein Prefractionation Combined with SIMAC, HILIC, and TiO₂ Chromatography Applied to Proximal EGF Signaling. *J. Proteome Res.* **10**, 5383–5397 (2011).
56. Tichy, A. *et al.* Phosphoproteomics: Searching for a needle in a haystack. *Journal of Proteomics* **74**, 2786–2797 (2011).
57. Weinert, B. T. *et al.* Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol Cell* **51**, 265–272 (2013).
58. Drazic, A., Myklebust, L. M., Ree, R. & Arnesen, T. The world of protein acetylation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1864**, 1372–1401 (2016).
59. Kalvik, T. V. & Arnesen, T. Protein N-terminal acetyltransferases in cancer. *Oncogene* **32**, 269–276 (2013).
60. Rope, A. F. *et al.* Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *Am J Hum Genet* **89**, 28–43 (2011).
61. Van Damme, P. *et al.* NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. *PLoS Genet* **7**, e1002169 (2011).
62. Dinh, T. V. *et al.* Molecular identification and functional characterization of the first N α -acetyltransferase in plastids by global acetylome profiling. *PROTEOMICS* **15**, 2426–2435 (2015).
63. Nguyen, K. T., Mun, S.-H., Lee, C.-S. & Hwang, C.-S. Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Exp Mol Med* **50**, 1–8 (2018).

64. Hwang, C.-S., Shemorry, A. & Varshavsky, A. N-terminal acetylation of cellular proteins creates specific degradation signals. *Science* **327**, 973–977 (2010).
65. Shemorry, A., Hwang, C.-S. & Varshavsky, A. Control of Protein Quality and Stoichiometries by N-Terminal Acetylation and the N-End Rule Pathway. *Molecular Cell* **50**, 540–551 (2013).
66. Forte, G. M. A., Pool, M. R. & Stirling, C. J. N-Terminal Acetylation Inhibits Protein Targeting to the Endoplasmic Reticulum. *PLOS Biology* **9**, e1001073 (2011).
67. Scott, D. C., Monda, J. K., Bennett, E. J., Harper, J. W. & Schulman, B. A. N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. *Science* **334**, 674–678 (2011).
68. Holmes, W. M., Mannakee, B. K., Gutenkunst, R. N. & Serio, T. R. Loss of N-terminal Acetylation Suppresses A Prion Phenotype By Modulating Global Protein Folding. *Nat Commun* **5**, 4383 (2014).
69. Aksnes, H., Drazic, A., Marie, M. & Arnesen, T. First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. *Trends in Biochemical Sciences* **41**, 746–760 (2016).
70. Zhao, S. *et al.* Regulation of Cellular Metabolism by Protein Lysine Acetylation. *Science* **327**, 1000–1004 (2010).
71. Wagner, G. R. & Payne, R. M. Widespread and Enzyme-independent N ϵ -Acetylation and N ϵ -Succinylation of Proteins in the Chemical Conditions of the Mitochondrial Matrix \blacklozenge . *J Biol Chem* **288**, 29036–29045 (2013).
72. Hirschey, M. D. *et al.* SIRT3 deficiency and mitochondrial protein hyperacetylation accelerate the development of the metabolic syndrome. *Mol Cell* **44**, 177–190 (2011).
73. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381–395 (2011).

74. Wang, S.-J. *et al.* Acetylation Is Crucial for p53-mediated Ferroptosis and Tumor Suppression. *Cell Rep* **17**, 366–373 (2016).
75. Ree, R., Varland, S. & Arnesen, T. Spotlight on protein N-terminal acetylation. *Exp Mol Med* **50**, 1–13 (2018).
76. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* **13**, 448–462 (2012).
77. Zhang, X. & Wang, Y. Glycosylation quality control by the Golgi structure. *J Mol Biol* **428**, 3183–3193 (2016).
78. Aebi, M. N-linked protein glycosylation in the ER. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1833**, 2430–2437 (2013).
79. Mohanty, S., P Chaudhary, B. & Zoetewey, D. Structural Insight into the Mechanism of N-Linked Glycosylation by Oligosaccharyltransferase. *Biomolecules* **10**, (2020).
80. Breitling, J. & Aebi, M. N-Linked Protein Glycosylation in the Endoplasmic Reticulum. *Cold Spring Harb Perspect Biol* **5**, (2013).
81. Tran, D. T. & Ten Hagen, K. G. Mucin-type O-Glycosylation during Development. *J Biol Chem* **288**, 6921–6929 (2013).
82. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat Rev Nephrol* **15**, 346–366 (2019).
83. Zielinska, D. F., Gnad, F., Wiśniewski, J. R. & Mann, M. Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. *Cell* **141**, 897–907 (2010).
84. Xiao, H., Suttapitugsakul, S., Sun, F. & Wu, R. Mass Spectrometry-Based Chemical and Enzymatic Methods for Global Analysis of Protein Glycosylation. *Acc Chem Res* **51**, 1796–1806 (2018).

85. Post-translational modifications of protein biopharmaceuticals. *Drug Discovery Today* **15**, 773–780 (2010).
86. Dworkin, J. Ser/Thr phosphorylation as a regulatory mechanism in bacteria. *Curr Opin Microbiol* **24**, 47–52 (2015).
87. Park, H.-S. *et al.* Expanding the Genetic Code of Escherichia coli with Phosphoserine. *Science* **333**, 1151–1154 (2011).
88. Heinemann, I. U. *et al.* Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett* **586**, 3716–3722 (2012).
89. Rovere, M., Powers, A. E., Patel, D. S. & Bartels, T. pTSara-NatB, an improved N-terminal acetylation system for recombinant protein expression in E. coli. *PLoS One* **13**, (2018).
90. Venkat, S. *et al.* Genetically Incorporating Two Distinct Post-translational Modifications into One Protein Simultaneously. *ACS Synth Biol* **7**, 689–695 (2018).
91. Strutton, B. *et al.* Engineering Pathways in Central Carbon Metabolism Help to Increase Glycan Production and Improve N-Type Glycosylation of Recombinant Proteins in E. coli. *Bioengineering (Basel)* **6**, (2019).
92. Baghban, R. *et al.* Yeast Expression Systems: Overview and Recent Advances. *Mol Biotechnol* **61**, 365–384 (2019).
93. Vieira Gomes, A. M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F. & Parachin, N. S. Comparison of Yeasts as Hosts for Recombinant Protein Production. *Microorganisms* **6**, (2018).
94. De Pourcq, K. *et al.* Engineering the yeast *Yarrowia lipolytica* for the production of therapeutic proteins homogeneously glycosylated with Man₈GlcNAc₂ and Man₅GlcNAc₂. *Microb Cell Fact* **11**, 53 (2012).

95. Walski, T., De Schutter, K., Van Damme, E. J. M. & Smagghe, G. Diversity and functions of protein glycosylation in insects. *Insect Biochemistry and Molecular Biology* **83**, 21–34 (2017).
96. Miyazaki, T., Kato, T. & Park, E. Y. Heterologous expression, purification and characterization of human β -1,2-N-acetylglucosaminyltransferase II using a silkworm-based *Bombyx mori* nucleopolyhedrovirus bacmid expression system. *Journal of Bioscience and Bioengineering* **126**, 15–22 (2018).
97. Miyazaki, T., Miyashita, R., Mori, S., Kato, T. & Park, E. Y. Expression and characterization of silkworm *Bombyx mori* β -1,2-N-acetylglucosaminyltransferase II, a key enzyme for complex-type N-glycan biosynthesis. *Journal of Bioscience and Bioengineering* **127**, 273–280 (2019).
98. Mabashi-Asazuma, H., Kuo, C.-W., Khoo, K.-H. & Jarvis, D. L. Modifying an Insect Cell N-Glycan Processing Pathway Using CRISPR-Cas Technology. *ACS Chem. Biol.* **10**, 2199–2208 (2015).