

Univerzita Karlova v Praze

Přírodovědecká fakulta



Bakalářská práce

Purifikace rekombinantních proteinů pomocí afinitní chromatografie

Ondřej Zemek

školitel: RNDr. Jaroslav Mácha

Praha 2010

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně, pod vedením RNDr. Jaroslava Máchy a na základě uvedené literatury.

V Praze, srpen 2010

.....

Ondřej Zemek

## **Abstrakt**

Izolace a purifikace rekombinantních proteinů je základním předpokladem pro studium jejich strukturních a funkčních vlastností. Mezi nejdůležitější metody, které jsou k tomuto účelu využívány patří afinitní chromatografie. V této práci jsou shrnuty v literatuře publikované vlastnosti a zkušenosti s využitím nejpoužívanějších afinitních tagů. Zde probrané tagy zahrnují: CBP, MBP, GST, polyhistidinové a polyargininové tagy, FLAG-tag, Strep-tag II a SpA a některé jejich modifikace. U jednotlivých tagů je probrán jejich původ a vlastnosti, vliv na formu a lokalizaci fúzního proteinu, způsob jeho vazby a eluce ze substrátu a možnosti jeho odstranění nebo využití v jiných metodách.

**Klíčová slova:** afinitní chromatografie, rekombinantní protein, purifikace, afinitní tag

## **Abstract**

The isolation and purification of recombinant proteins is essential for further study of their structural and functional properties. The affinity chromatography is usually the method of choice for this task. In this paper the most used affinity tags are reviewed for their properties and experience with their application. The tags covered here include CBP, MBP, GST, polyhistidine and polyarginine tags, FLAG-tag, Strep-tag II and SpA. The origin and properties of the tags, their influence on form and localization of fusion protein as well as binding, elution and removal are discussed.

**Keywords:** affinity chromatography, recombinant protein, purification, affinity tag

## **Poděkování**

Rád bych poděkoval svému školiteli RNDr. Jaroslavu Máchovy za pomoc při zpracování této práce a také členům mé rodiny za všestrannou podporu.

# Obsah

Cíl práce.....	6
Úvod.....	7
FLAG-tag.....	8
MBP.....	9
Polyhistidinový tag.....	11
Strep-tag II.....	13
CBP.....	14
GST.....	15
Polyargininový tag.....	17
Stafylokokový protein A a Z.....	18
Závěr.....	20
Seznam použitých zkratk.....	21
Seznam použité literatury.....	22

## **Cíl práce**

V této práci se pokusím přiblížit současný stav a možnosti využití afinitní chromatografie při separaci a purifikaci rekombinantních proteinů. Protože je tato oblast velmi široká, zaměřím se na nejpoužívanější a běžně dostupné systémy, které mohou být nejnáze využity při praktické aplikaci. U jednotlivých afinitních tagů shrnu jejich vlastnosti a dosavadní zkušenosti s jejich využitím popsané v literatuře. Popíši jejich vliv na množství a strukturu exprimovaných fúzních proteinů, možnosti jejich odstranění nebo naopak využití v dalších metodách. Dále rozeberu ligandy používané k jejich vazbě a podmínky imobilizace a eluce.

## Úvod

Se stále rostoucím počtem známých sekvencí genů i celých genomů a automatizací jejich získávání stoupá také potřeba metod pro rychlou a ekonomickou produkci jimi kódovaných proteinů pro funkční a strukturní analýzu. Pro produkci proteinů pro farmakologické aplikace je pak klíčová jejich vysoká čistota, absence antigenních faktorů a ekonomičnost produkce. Právě pro tyto účely je velmi výhodné využít metody afinitní chromatografie.

O objevení a počáteční vývoj metody afinitní chromatografie se zasloužily Pedro Cuatrecasas a Meir Wilchek, kteří byli za tento počin oceněni roku 1987 Wolfovou cenou za Medicínu.

Oproti jiným chromatografickým metodám používaných při separaci proteinů, jako je SEC, IEX nebo HIC, je interakce proteinů se stacionární fází mnohem specifitější a pevnější. Tím je umožněno důkladné odmytí ostatních proteinů a nečistot a následná desorbce specifickým činidlem nebo výraznou změnou prostředí. Díky tomu je možné jedním krokem připravit i z hrubého lysátu poměrně čistý protein. Nevýhodou však je skutečnost, že pouze velmi málo neupravených proteinů vykazuje tak specifickou vazbu na snadno dostupné substráty. Proto se afinitní chromatografie uplatňuje především při přípravě rekombinantních proteinů, kdy se cílový protein nechá exprimovat spolu s aminokyselinovou sekvencí, která vykazuje specifickou vazbu na příslušnou náplň kolony. Tyto sekvence, označované jako afinitní tagy nebo značky, je po purifikaci možné opět oddělit obvykle uměle zařazeným cílovým místem pro specifickou endopeptidázu nebo u kratších tagů umístěných na N-konci pomocí aminopeptidázy a na C-konci pomocí karboxypeptidázy.

Afinitní tagy je možné rozdělit do dvou základních skupin: několika aminokyselin dlouhé, které se váží na imobilizované proteiny a delší vazující menší neproteinové ligandy. Do první skupiny spadají především systémy protilátka-antigen, jejichž hlavní nevýhodou je vysoká cena a nižší kapacita náplně kolony. Druhé pak představují vyšší metabolickou zátěž pro exprimující buňku a mohou spíše ovlivnit výslednou konformaci proteinu.

Žádný z dosud v literatuře popsáných tagů nelze považovat za univerzální. Naopak, jejich aplikaci je třeba zvolit na základě dosud známých vlastností cílového proteinu, druhu exprimující kultury, skutečnosti zda potřebujeme protein v nativní konformaci nebo postačí denaturovaný, požadavků na jeho čistotu a množství a mnoha dalších faktorech. Také je třeba zvážit zda bude nutné afinitní tag po purifikaci odstranit, protože by překážel při další aplikaci, nebo jej bude naopak možné využít například v analytických metodách jako Western blot, dot blot či ELISA. V literatuře je dostupných několik srovnání některých afinitních systémů (Joakim Nilsson et al. 1997) (Waugh 2005) ale vzhledem k širší aplikaci afinitních tagů a rychlosti vývoje na tomto poli není snadné vybrat pro konkrétní potřebu ten nejvhodnější.

## FLAG-tag

Tento oktapeptid se sekvencí AspTyrLysAspAspAspAspLys byl vyvinut a patentován jako první epitop určený pro široké použití při purifikaci jako součást fúzního proteinu (Thomas P. Hopp et al. 1988). Jeho sekvence byla navržena na základě poznatků o antigenitě proteinů především v závislosti na jejich hydrofobitě (TP Hopp & Woods 1981). První čtyři aminokyseliny tvoří nejpodstatnější část epitopu, protože interakce mezi aromatickými jádry je klíčová ve vazbě protilátka – antigen, je zde tyrosin, jehož antigenita je zvýšena okolními polárními aminokyselinami. Pět koncových aminokyselin je pak minimální rozpoznávaná cílová sekvence pro enterokinázu, používanou k oddělení tagu. Zároveň je i tento úsek velmi hydrofilní a přispívá k antigenitě a lokalizaci tagu na povrchu molekuly proteinu v nativní konformaci. Vzhledem k malé délce sekvence nepředstavuje její produkce významnější metabolickou zátěž a díky hydrofilní povaze zvyšuje rozpustnost fúzního proteinu. Pomocí FLAG tagu je možné exprimovat a separovat i velmi krátké peptidy. U velkých proteinů představuje hlavní omezení špatná rozpustnost.

Díky své jednoduchosti je FLAG-tag využitelný pro expresi v mnoha typech prokaryotických i eukaryotických kultur. Nejčastěji se používá v buňkách *E.coli*, *S. cerevisiae*, *P. pastoris*, *S. frugiperda* infikovaných bakulovirem, CHO buňkách a COS transformovaných opičích buňkách (Einhauer & Jungbauer 2001b).

Vlastní chromatografická separace probíhá obvykle na koloně s jedním ze tří typů imobilizovaných monoklonálních protilátek. Nejstarší protilátka označovaná M1 váže jen sekvenci FLAG-tagu umístěnou na úplném N-konci proteinu. Není tedy použitelná ani pro proteiny s methioninem před tagem, váže se však závisle na vápníku, čehož se s výhodou využívá při eluci cílového proteinu roztokem chelatačního činidla (EDTA) (Einhauer & Jungbauer 2001a). Protilátka M2 váže tag umístěný na obou koncích proteinu a je možné jej použít i na sekvenci s methioninem na začátku, její vazba je však nezávislá na vápníku a eluce je tak možná jen snížením pH nebo vytěsněním proteinu syntetickým FLAG peptidem. Nově byla také navržena eluce roztokem argininu (Futatsumori-Sugai et al. 2009). Poslední běžně používanou protilátkou je M5, která váže jen sekvenci Met-FLAG, zato silněji než M2. Podmínky eluce jsou zde stejné. Protilátky jsou citlivé na denaturační činidla a podmínky a jejich působením ztrácí afinitu. Nevýhodou je vysoká cena náplně kolony a poměrně nízká vazná kapacita (přibližně 0,6 mg/ml pro 30 kDa protein), proto tento systém není příliš vhodný pro produkce vyšších množství proteinu (Lichty et al. 2005). Pro mnoho aplikací ve výzkumu a vývoji není nutné tag oddělovat, ale při jakémkoli farmakologickém použití proteinu je kvůli jeho vysoké antigenitě nutné jej vždy separovat. K oddělení tagu z N-konce slouží enterokináza, jejíž cílová sekvence je integrální součástí tagu. Pro štěpení z C-konce je nutno před tag zařadit sekvenci cílového místa pro specifickou endopeptidázu.

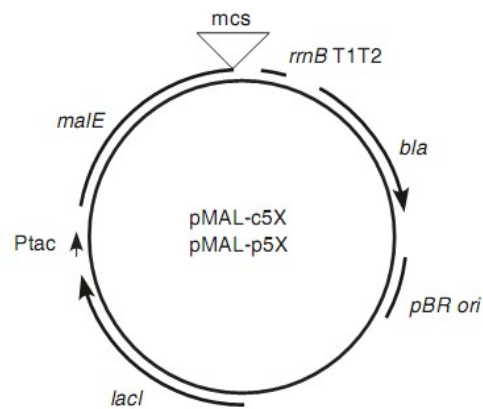


## MBP

Maltose binding protein je přirozeným proteinem *E.coli*, který je vylučován do periplazmatického prostoru, kde se váže na maltózu nebo maltodextriny a následně vstupuje přes cytoplazmatickou membránu zpět do buňky. Bakteriím slouží jednak jako transportní protein, ale také jako regulační faktor chemotaxe buněk ve směru gradientu vázaného oligosacharidu. Přenos signálu se uskutečňuje přes Tar transmembránový protein, který aktivuje motory bičíků. MBP se přirozeně vyskytuje až v padesátinásobném přebytku oproti příslušným receptorům, což jednak zajišťuje rychlou odpověď na změnu koncentrace ligandu, ale protože také napomáhá renaturaci některých proteinů, uvažuje se, že může mít i funkci chaperonu pro periplazmatické proteiny. V buňce vzniká translací genu *malE* nejprve jako 396 aminokyselinový prekurzor, kde prvních 26 aminokyselin slouží jako značka pro transport do periplazmatického prostoru a je během něj odštěpeno. Výsledný protein tedy obsahuje 370 aminokyselin. V nativním stavu vytváří MBP dvě kulovité domény spojené závěsem, na jejichž povrchu je vazné místo pro maltodextrin. Vázat se mohou i jiné oligosacharidy tvořené řetězcem  $\alpha(1-4)$  D glukózových jednotek od maltózy až do počtu osmi jednotek a to jak lineární tak i cyklické. Disociační konstanta se u různých oligosacharidů pohybuje v rozmezí  $K_d = 1,6$  až  $40 \times 10^{-7}$ , kromě toho se MBP dobře váže i na amylózu. Obsazením vazebného místa substrátem dochází ke změně konformace proteinu sevřením úhlu mezi dvěma doménami přibližně o  $35^\circ$ , což zároveň vede k odhalení hydrofobních oblastí proteinu (Millet et al. 2003). Vlastní vazebné místo je od C i N konce přibližně stejně vzdáleno a jejich modifikace navázáním cílového proteinu nemá většinou za následek výraznou změnu vazebných vlastností. MBP se vyznačuje velice dobrou rozpustností ve vodě, což se přenáší i na fúzní proteiny a umožňuje tak jejich produkci ve vysokém množství až kolem 40% celkového obsahu buněčných proteinů (Pryor & Leiting 1997), přičemž se výrazně snižuje jejich tendence k precipitaci (Kapust & Waugh 1999). Další výhodnou vlastností MBP je jeho vysoká termostabilita, která se částečně přenáší i na fúzní proteiny a může tak usnadnit jejich čištění (Huang et al. 2006). Jeho struktura také neobsahuje žádné cysteinové zbytky, takže nehrozí nebezpečí změny struktury fúzního proteinu tvorbou nevhodných disulfidických můstků.

Exprimace fúzních proteinů MBP s cílovým proteinem v sobě spojuje několik výhod a je s úspěchem využívána. Nejsnadnějším způsobem využití je použití předem připraveného vektoru pMAL, který kromě vlastního genu *malE* modifikovaném na vazebném místě pro zvýšení afinity k amylóze obsahuje i silný tac promotor a *lacI<sup>q</sup>* gen kódující Lac represor pro potlačení exprese za nepřítomnosti IPTG. Na gen *malE* navazuje ve směru translace polylinker a stop kodony ve všech třech čtecích rámcích. Součástí polylinkeru je restriční místo pro postpurifikační oddělení MBP z cílového proteinu pomocí proteázy. V současné době jsou k dispozici tři varianty vektoru s různými

štěpicími enzymy: Faktor Xa, enterokináza a Genenáza I. Polylinker je koncipován tak, aby restrikce proběhla hned za MBP a bylo tak možné získat cílový protein zcela bez, nebo jen několika málo aminokyselinami navíc. K dispozici jsou dvě varianty vektoru, které se liší přítomností 26 aminokyselinové signální sekvence před MBP. Její přítomnost zajišťuje, že výsledný fúzní protein je transportován do periplazmatické oblasti, bez ní zůstává v cytoplazmě. Správný výběr mezi těmito variantami je klíčový pro vytvoření proteinu v nativní konformaci.



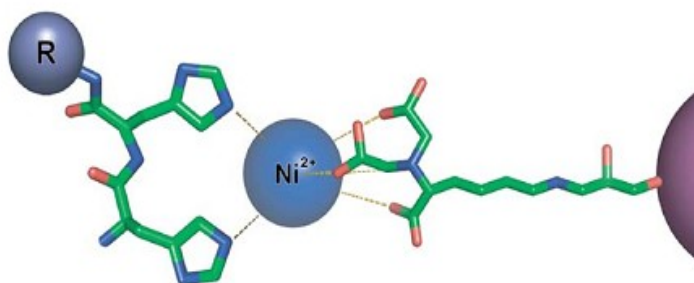
Obr.1 Schéma pMAL plasmidu, (převzato z New England BioLabs 2009)

Varianta pMAL-p, která exportuje protein do periplazmatické oblasti je vhodná především pro proteiny s intramolekulárními disulfidickými můstky a extracelulární části transmembránových proteinů. Množství exprimovaného proteinu se obvykle pohybuje mezi 1-20 % všech buněčných proteinů, což je podstatně méně než v případě, kdy protein zůstává v cytoplazmě. Výhodou však může být, že není potřeba provádět kompletní lýzu buněk, ale stačí chladovým a osmotickým šokem uvolnit pouze proteiny z periplazmatické oblasti. Extracelulární exprese představuje obecně zajímavou možnost při přípravě rekombinantních proteinů, protože umožňuje zbavit se značného množství interferujících proteinů ještě před vlastní purifikací a díky velkému naředění nedochází zpravidla k tvorbě inkluzních tělísek, ne vždy je však predikovatelná (Ni & Chen 2009). Pro některé proteiny, které vyžadují redukční prostředí je výhodnější varianta vektoru, kdy zůstávají v cytoplazmě. Dosahuje se u ní až 20-40 % zastoupení rekombinantního proteinu. Celkově se u tohoto vektoru udává kolem 75 % úspěšné exprese a izolace proteinu. Příčiny které mohou způsobit neúspěch zahrnují štěpení rekombinantního proteinu proteázami *E.coli*, interakci cílového proteinu s vazným místem MBP, kdy nedojde k zachycení proteinu na koloně při purifikaci. Dále může být na vině tvorba nerozpustných agregátů rekombinantního proteinu, jeho kumulace v membránách nebo toxicita pro buňku, například u proteáz nebo velmi hydrofobních sekvencí. Vlastní purifikace afinitní chromatografií probíhá na koloně s amylózovou náplní za podmínek, kdy se MBP vyskytuje v nativní konformaci. Tedy za neutrálního až lehce bazického pH (pH = 6,5 – 8,5) bez přítomnosti denaturačních činidel a detergentů. Pro snížení nespecifické sorbce ostatních proteinů je doporučováno zvýšení iontové síly roztoku, například na 200 mM pomocí NaCl. Po promytí je MBP fúzní protein uvolněn 10mM roztokem maltózy, která se následně odstraní nejlépe pomocí SEC. Oddělení MBP tagu od cílového proteinu v nativní konformaci je u faktoru Xa úspěšné asi v 50% případech, pokud však faktor Xa neštěpí přímo cílový protein často se tag podaří oddělit alespoň po denuraci (Ausubel et al. 2001). Štěpení pomocí faktoru Xa přímo na koloně se nedoporučuje kvůli velmi malé rychlosti a uvolňování proteinu amylázou.

## Polyhistidinový tag

Krátké polyhistidinová sekvence aminokyselin patří k nejčastěji používaným afinitním tagům pro purifikaci rekombinantních proteinů. Princip His-tagu spočívá v schopnosti histidinu vázat se koordinačně s velkou afinitou na ionty některých kovů jako je  $\text{Ni}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ .

Výhodou polyhistidinového tagu je skutečnost, že může dobře fungovat nezávisle na umístění v rámci sekvence exprimovaného proteinu. Může být zařazen na C i N konec, případně i mezi dva funkční bloky pokud je zachováno jeho umístění na povrchu molekuly složeného proteinu. Jeho afinita také není podmíněna nativní konformací výsledného proteinu. Nejčastěji



Obr.2 Vazba proteinu se dvěma terminálními histidiny na imobilizovaný komplex Ni-NTA. Upraveno z (Bolanos-Garcia & Davies 2006)

používané uspořádání His-tagu je šestinásobné opakování (6xHis), které má nejsilnější vazbu na nikl  $K_d = 1,4 \times 10^{-8}$ . Další silně se vázající sekvence jsou opakování s 1 nebo 4 jinými aminokyselinami mezi histidiny (Knecht et al. 2009). Oproti jiným systémům nepředstavuje sám o sobě, vzhledem k tomu, že se jedná o nejkratší běžně používaný afinitní tag, pro exprimující buňku téměř žádnou metabolickou zátěž. Za fyziologických podmínek nenesou žádný náboj a také konformace cílového proteinu jím nebývá výrazně pozměněna. Na druhou stranu sám výrazně nezvyšuje expresi ani nepodporuje rozpustnost fúzního proteinu. To je však možno doplnit kombinací s jiným vhodným tagem (NusA, SET).

Jako náplň kolony pro afinitní separaci bývá nejčastěji použita agaróza s kovalentně vázanou polydentátní skupinou, která váže ion těžkého kovu, nejčastěji niklu. Mezi běžně dostupné patří náplně s nitril-trioctovou kyselinou (Ni-NTA), imindioctovou kyselinou (Ni-IDA) nebo triskarboxymethylethylendiaminem (Ni-TED). Tyto náplně jsou kromě niklu se používá také náplň s ionty kobaltu (komerční produkt Talon, BD Biosciences Clontech), který vykazuje podstatně menší nespecifickou vazbu proteinů, má však také menší vaznou kapacitu.

Jednoduchá afinitní chromatografie pomocí His-tagu na Ni-NTA obvykle neposkytuje tak čistý výsledný protein jako použití jiných tagů. Je to dáno vazbou dalších proteinů v lyzátu, které se přes histidin nebo specifické kov vázající místa zachycují na náplni kolony. Některé nejčastější z *E.coli* byly charakterizovány, například SlyD (Parsy et al. 2007), Fur, Crp, ArgE, GlnS, GlgA, ODO1, ODO2, YadF a bylo nalezeno několik postupů pro jejich eliminaci (Bolanos-Garcia & Davies 2006). Protože se v mnoha případech jedná o stresové proteiny lze jejich množství významně ovlivnit podmínkami kultivace, další z variant je použití speciálních kmenů *E.coli* s nízkou nebo nulovou

expresí konkrétních genů. Obecně použitelnou metodou je pak zvýšení iontové síly pufru a použití promývacího pufru s malým obsahem histidinu, který vytěsňuje slaběji se vážající necílové proteiny. V poslední době se také používá His-tag v kombinaci dvou afinitních tagů (TAP), což umožňuje velmi dokonalé čištění dvěma následnými afinitními chromatografiemi (Pryor & Leiting 1997). Případně je možné afinitní chromatografii doplnit jinou separační metodou jako SEC (Parsy et al. 2007).

Eluce cílového proteinu z kolony je možná několika způsoby: vytěsnění proteinu roztokem imidazolu o koncentraci kolem 100-250 mM, snížením pH pod 5 čímž se histidin naprotonuje a přerušuje vazbu na kov nebo chelataci kovu pomocí EDTA, která má v komplexu s niklem mnohem nižší disociační konstantu než His-tag. Při použití imidazolu, což je nejčastěji doporučovaná metoda, je možné náplň zregenerovat promytím roztokem NaOH a využít ji přibližně k pětkrát. Pokud je k eluci použit roztok EDTA nebo jiného chelatacího činidla je nutné před dalším použitím kolony regenerovat obsah iontů kovu v matici.

Možnosti odstranění His-tagu jsou závislé na jeho umístění v proteinu. V případě označení C-konce se využívá specifických endopeptidáz, jako faktor Xa nebo enterokináza, s předem vloženou cílovou sekvencí (Abdullah & Chase 2005). Vzhledem k tomu, že většina endopeptidáz štěpí na cílové sekvenci blíže k N-konci, zůstává v tomto případě na cílovém proteinu několik aminokyselin navíc. U His-tagu připojeném na N-konec se využívá obvykle štěpení exopeptidázou s vhodně umístěnou stop aminokyselinou. Za tímto účelem byl zkonstruován systém TAGZyme (Pedersen et al. 1999), který odštěpuje pomocí 6xHis-tagu značené DAPázy (dipeptidyl-aminopeptidázy I) dvojice aminokyselin z N-konce tagu až k první stop aminokyselině (Arg, Lys, Gln jako první z dvojice nebo Pro na libovolné pozici) v cílovém proteinu. Pokud na posledním místě v sekvenci cílového proteinu není žádná ze stop aminokyselin je možné mezi něj a His-tag vložit Gln a štěpení DAPázou provádět za přítomnosti nadbytku 6xHis-tagem značené Qcyklázy (glutamin-cyklotransferázy). Tím se štěpení zastaví na Gln, který je převeden na pyroglutamát. Poté jsou oba enzymy spolu se zbytky odštěpeného His-tagu odstraněny zachycením na koloně s náplní Ni-NTA a pyroglutamát odštěpen 6xHis-tagem značenou pGAPázou (pyroglutamylaminopeptidázou) (Abdullah & Chase 2005). Po dalším promytí kolonou s náplní Ni-NTA je získán cílový protein bez připojených aminokyselin na N-konci. Pro množství aplikací však není nutné His-tag vzhledem k jeho krátkosti odstraňovat. Celé fúzní proteiny se dají například využít k přípravě polyklonálních protilátek (Fei Yan et al. 2007)

## Strep-tag II

Strep-tag II je uměle zkonstruovaný oktapeptid se sekvencí Asn-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys se schopností vázat se s vysokou afinitou na streptavidin. Při konstrukci systému vyšli autoři z myšlenky najít peptid, který by se vázal na streptavidin obdobně specificky a silně jako v mnoha aplikacích využívaný biotin a byl přitom dlouhý jen několik aminokyselin. Expresí náhodných peptidů v *E.coli* a jejich selekcí na schopnost vázat se na streptavidin byla získána sekvence Ala-Trp-Arg-His-Pro-Gln-Phe-Gly-Gly, označovaná původně jako Strep-Tag(Skerra & Schmidt 1999), která však pro silnou ( $K_d=37 \mu\text{M}$ )(Schmidt et al. 1996) vazbu vyžadovala umístění na úplném C-konci. To bylo způsobeno klíčovou úlohou karboxylové skupiny terminálního glycinu. Jeho náhradou kyselinou glutamovou byla získána sekvence jejíž afinita není na umístění v rámci primární struktury proteinu závislá. Tato nová sekvence, označovaná pro odlišení také jako Strep-Tag II, však vykazovala poněkud nižší afinitu k streptavidinu ( $K_d=72 \mu\text{M}$ )(Schmidt et al. 1996), proto byl mutací jeho čtyř aminokyselin a následnou selekcí získán pozměněný analog streptavidinu používaný pod názvem StrepTactin. Ten se vyznačuje téměř o dva řády vyšší afinitou k Strep-tagu s  $K_D$  přibližně  $1 \mu\text{M}$ . Strep-tag představuje díky krátké sekvenci jen minimální metabolickou zátěž, je chemicky značně inertní a zpravidla nemění výslednou strukturu proteinu. Je možné jej použít i k purifikaci proteinů s více podjednotkami, přičemž Strep-Tag může být přítomen jen na jedné z nich. Další výhodou je široké spektrum organismů, které jej mohou exprimovat. Na druhou stranu však nijak nezvyšuje rozpustnost rekombinantního proteinu(Lichty et al. 2005).

Vlastní afinitní chromatografie probíhá obvykle na agaróze s navázaným Strep-Tactinem, kam se nalije buněčný lysát, dojde ke specifické vazbě Strep-tagu a ostatní proteiny jsou odmyty promývacím pufrem. Následně je rekombinantní protein kompetitivně eluován roztokem desthiobiotinu. Regenerace náplně kolony se provádí pomocí roztoku kyseliny

2-(4'-hydroxy phenylazo)benzoové (HABA), která nejprve vytěsňuje desthiobiotin z vazného místa pro biotin a následně ji lze snadno vymýt. Protože se jedná o azobarvivo, které vazbou na Strep-Tactin přechází ze žluté na červenou, je tak i indikován stav kolony. Kolona může být regenerována přibližně 3 až 5-krát, je však nutné se vyvarovat přítomnosti biotinu v lysátu. Toho se dá dosáhnout přidávkou malého množství avidinu, který zachytí biotin, Strep-Tag však téměř neváže. Vazná kapacita Strep-Tactinu je 50-100 nmol/ml nebo asi 3mg 30 kDa proteinu na ml. Podmínky celé purifikace jsou velmi mírné a výhodou je zpravidla velmi vysoká čistota získaného proteinu, zajištěná velmi nízkou nespecifickou vazbou ostatních proteinů na náplň kolony. Cena StrepTactinu je pak podstatně vyšší než u náplně většiny nejpoužívanějších systémů afinitní chromatografie jako His-tag, MBP či GST. Rekombinantních proteinů se Strep-tagem se také široce využívá v různých analytických metodách jako Western blot, dot blot a ELISA.

## CBP

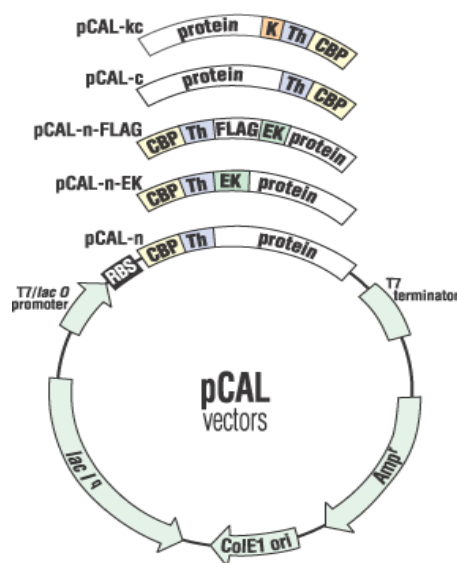
CBP (calmodulin binding peptide) je 26 aminokyselinová C-koncová část kinázy lehkého řetězce myosinu (MYLK), která za přítomnosti iontů vápníku váže velmi pevně kalmodulin ( $K_d = 10^{-9}$ ). Přirozeně je MYLK lokalizována ve svalových buňkách, kde po influxu vápníku a jeho tvorbě komplexu s kalmodulinem dojde k vazbě na C-konec MYLK a tím k její aktivaci. Aktivovaná MYLK potom fosforyluje lehký řetězec myosinu čímž zprostředkuje jeho vazbu na aktin a tím svalovou kontrakci.

Na tomto principu je založený komerčně dostupný systém Affinity<sup>TM</sup> firmy Stratagene, který má několik variant pCAL(Zheng et al. 1997) vektoru pro jednoduchou purifikaci. Tři připravené pro připojení CBP tagu na C- nebo N-konci a jedna pro C-koncové umístění kombinace CBP se sekvencí kemptidu, substrátu pro značení radioaktivním izotopem <sup>32</sup>P pomocí PKA. Pro TAP je dostupná také verze kombinovaná s FLAG-tagem. Ve všech případech je možné tag oddělit thrombinem, pro něž je vložena rozpoznávací sekvence.

Vlastní afinitní chromatografie probíhá na náplni s imobilizovaným kalmodulinem kam se nalije roztok buněčného lysátu s přidavkem iontů vápníku o koncentraci 0,2 – 2 mM za neutrálního pH. Dojde k vazbě rekombinantního proteinu za CBP tag a ostatní proteiny a

nečistoty mohou být odmyty promývacím pufrům s malým přidavkem vápenatých iontů. Eluce proteinu probíhá opět za neutrálního pH roztokem EGTA, která váže velmi pevně vápenaté ionty a tak ruší vazbu kalmodulinu na CBP(Stofko-Hahn et al. 1992).

CBP představuje díky poměrně krátké délce (26 AA) a nízké molekulární hmotnosti (4 kDa) sekvence pro exprimující buňku pouze malou metabolickou zátěž a nemusí být také pro některé aplikace rekombinantního proteinu odstraňován. Podmínky eluce jsou velmi mírné a díky vysoké specifitě vazby je cílový protein obvykle získán v dobré čistotě. Jistou nevýhodou představuje vyšší cena náplně kolony a poměrně nízká kapacita (asi 2 mg proteinu na ml).



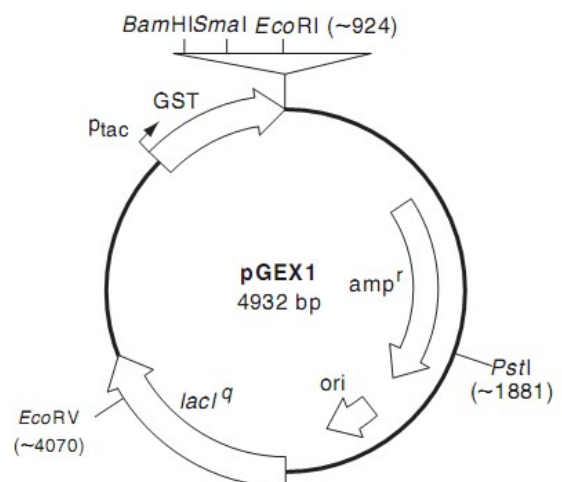
Obr.3 Schéma variant vektoru pCAL (převzato z Biocompare 2010)

## GST

Glutathion S-transferázy (EC 2.5.1.18) jsou skupinou enzymů, schopné vázat redukovanou formu glutathionu ( $\gamma$ -glutamyl-cysteinyl-glycin) a přenášet jej na substrát. Díky široké škále substrátů mají významnou detoxifikační úlohu a podílí se na likvidaci jak endogenních tak xenobiotických látek. GST byla objevena poprvé v roce 1961 v cytosolickém extraktu z jater, díky její schopnosti modifikovat 1,2-dichloro-4-nitrobenzen. Od té doby bylo objeveno nejméně pět tříd tohoto proteinu v cytoplasmě u obratlovců, označovaných jako alfa, pi, mu, theta a kappa. Další, označená sigma, se pak vyskytuje v cytosolu bezobratlých a mimo to existují i mikrosomální varianty. Cytosolické GST tvoří v nativním stavu homo- nebo heterodimery, v případě heterodimerů jsou však podjednotky vždy ze stejné třídy. Velikost podjednotek se pohybuje mezi 23 – 27 kDa (Armstrong 1997). V rámci prokaryot se GST nachází jen u některých skupin, především u sinic (*Cyanobacteria*), proteobakterií a některých fototrofních a některých gram-pozitivních bakterií (Vuilleumier 1997). U eukaryot je hlavní úloha GST v druhé fázi biotransformace xenobiotik, kdy navázáním glutathionu pře sulfidickou vazbu cysteinu výrazně zvýší rozpustnost, zpravidla lipofilního, substrátu a umožní tak jeho eliminaci z těla (Leaver & George 1998). GST nejprve váže glutathion a následně zprostředkuje jeho reakci s jedním ze tří typů substrátů: na aromatických kruzích dochází k nukleofilní substituci, u konjugovaných dvojných vazeb k adici a u oxiranových kruhů k jejich otvírání se současnou adicí glutathionu. Navázání glutathionu na GST se uskutečňuje prostřednictvím vodíkových vazeb a přestože se mezi jednotlivými třídami poněkud liší nalezneme u všech velmi konzervovanou strukturu složenou ze dvou  $\beta$ -barelů a jednoho  $\alpha$ -sheetu, která je zodpovědná za vazbu  $\gamma$ -glutamylu (Armstrong 1997). Vazba glutathionu je za nedenedaturujících podmínek a neutrálním pH velmi silná s disociační konstantou  $K_d = 2$  až  $6 \times 10^{-7}$ .

### Využití GST v afinitní chromatografii

Pro účely afinitní chromatografie se využívá gen kódující GST z krevničky *Schistosoma japonicum*. Ten má délku 220 aminokyselin a molekulární hmotnost 26 kDa a patří do mu třídy cytosolických GST. Nejčastěji se pro expresi fúzních proteinů s GST v *E.coli* využívá plasmid pGEX, který obsahuje silný *tac* promotor v kombinaci s *lac* represorem. Tato kombinace umožňuje vypěstovat dostatečně koncentrovanou kulturu bakterií a až následně zahájit intenzivní expresi přidáním IPTG, který inhibuje *lac*



repressor. Dále se na plasmidu vyskytuje selekční marker a polylinker s restričními místy v každém čtecím rámci a sekvenci pro oddělení GST z cílového proteinu po dokončení purifikace specifickou proteázou. U Plasmidu pGEX se používá buďto thrombinu nebo Factor Xa. (Smith & Johnson 1988) Úroveň exprese u tohoto systému dosahuje obvykle 1-3 mg, za příznivých podmínek až k 10 mg na litr kultury. V některých případech však může být exprese podstatně nižší, na příklad jen 50µg, pokud je protein špatně rozpustný. To je obvykle problémem u hydrofobních nebo větších proteinů (od 50 kDa). Shlukování takových proteinů do inkluzních tělísek je možné snížit kultivací za nižší teploty (20-30°C), zvýšením provzdušnění kultury nebo zkrácením doby indukce. Fúzní protein se shromažďuje v cytosolu a pro jeho extrakci je potřeba buňky lysovat pomocí detergentu. V případě nerozpustných proteinů lze s výhodou využít anionického detergentu sarkosylu spolu s běžně používaným Tritonem<sup>TM</sup> X-100, což v mnoha případech zvyšuje výtěžky purifikovaného proteinu, aniž by docházelo k jeho denaturaci nebo interferenci s vazebnými vlastnostmi GST (Saluta & Bell 1998). Vlastní afinitní chromatografie probíhá na agarózovém gelu s glutathionem navázaným přes linker za atom síry. Tak je umožněna silná vazba GST aktivního centra přes vodíkové můstky za nedenedurujících podmínek a po vymytí ostatních látek následné uvolnění cílového proteinu jedním z následujících způsobů. Nejčastěji používané je vytěsnění roztokem glutathionu, které je k proteinu velmi šetrné. Další možností je odštěpení cílového proteinu příslušnou proteázou přímo na koloně, kde však rychlost štěpení značně závisí na dostupnosti cílové sekvence a poslední variantou je výrazné snížení pH (např. 0,1 M glycin-HCl pufr na pH 2,5), které přerušuje vodíkové můstky mezi glutathionem a GST. Oddělení GST-tagu od cílového proteinu se provádí zpravidla endopeptidázou, jejíž cílová sekvence je kódována ve vektoru (thrombin nebo Factor Xa). V některých případech je však možné použít rekombinantní protein i s GST-tagem. Například při přípravě polyklonálních protilátek proti cílovému proteinu, kdy se k imunizaci používá celý fúzní protein (Zhang et al. 2008) (J. Yang et al. 2007). Protilátky proti GST části nevznikají příliš často a pokud jsou na závadu v další aplikaci je možné je odstranit na koloně s imobilizovaným GST.

Cena náplně kolony je příznivá a rovněž její kapacita je poměrně vysoká, kolem 10 mg/ml. Nevýhodou GST systému je, že tvoří homodimery a není tak vhodný pro získávání oligomerních proteinů, v jeho molekule se také vyskytují dva cysteinové zbytky, které jsou vystaveny okolnímu prostředí a mohou být snadno oxidovány, což může vést k agregaci. Další nevýhodou představuje značná velikost tagu, který jednak s vyšší pravděpodobností ovlivní funkci cílového proteinu a samozřejmě je pro exprimující buňky další metabolickou zátěží.



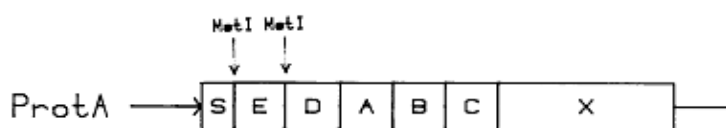
## **Polyargininový tag**

C-koncová sekvence několika argininů byla poprvé využita k purifikaci proteinů Sassenfeldem a Brewerem již v roce 1984 (Brewer & Sassenfeld 1985). Nejčastěji se využívá řetězce pěti či šesti argininů, jejichž guanidinové postranní řetězce jsou kladně nabitý od kyselého až do silně bazického prostředí ( $pK_a = 12,48$ ). Díky tomu je možné tag velmi pevně vázat na sloupce s kationtově výměnnou náplní jako je například SP-Sephadex. Silně nabitý Arg-tag umožňuje kromě toho i reverzibilní vazbu na celou řadu povrchů, jako je slída (Nock et al. 1997), sklo či různé plasty (Chico et al. 2003). Toho se využívá například k přípravě vzorků pro analýzu interakcí proteinů s ligandy nebo studium jejich struktury pomocí mikroskopie skenující sondou (SPM) nebo elektronové mikroskopie (Nock et al. 1997). Polyargininový tag o délce alespoň šest aminokyselin může také sloužit k usnadnění internalizace rekombinantního proteinu do buněk (Mitchell et al. 2000) (Vivès et al. 1997). Děje se tak patrně vazbou kladně nabitých argininových zbytků na záporně nabitý fosfolipidový povrch cílové buňky (Jonathan B. Rothbard et al. 2000). Aby se odstranily ostatní proteiny, které se mohou na kolonu vázat díky vyššímu zastoupení bazických aminokyselin v jejich řetězci, provádí se eluce lineárním koncentračním gradientem roztoku NaCl, obvykle od 0 do 400 mM při alkalickém pH. Tím dojde k postupnému uvolnění vázaných proteinů v závislosti na jejich bazicitě. Delší než šesti argininové tagy pak vyžadují ještě vyšší koncentrace pufru a jejich vazbu na některé substráty je možné považovat pro mnoho aplikací za ireverzibilní (Fuchs & Raines 2005).

Odstranění Arg-tagu se provádí pomocí karboxypeptidázy B, která postupně hydrolyzuje aminokyseliny z C-konce proteinu, přičemž preferuje ty bazické. Výhodou Arg-tagu je jeho krátká délka, nízká cena náplně chromatografické kolony a možnost vazby na některé substráty pro další využití rekombinantního proteinu. Nevýhodou pak je, že silný náboj tagu může ovlivňovat strukturu a případnou enzymatickou aktivitu výsledného proteinu (Fuchs & Raines 2005).

## Stafylokokový protein A a Z

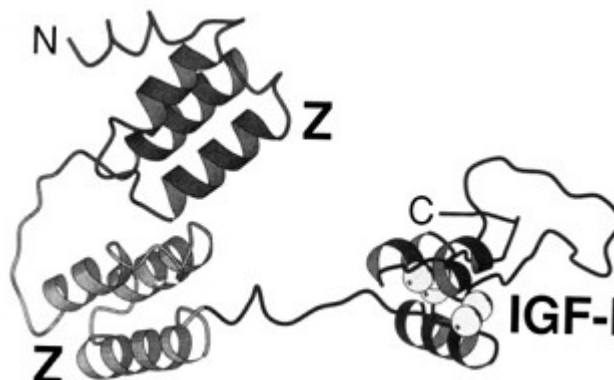
Stafylokokový protein A (SpA) je přirozeným produktem většiny kmenů *Staphylococcus aureus*. Zpravidla je vázán v buněčné stěně na peptidoglykan, nebo u některých mutantů sekretován do média (M Uhlén et al. 1984). SpA vykazuje silnou vazbu na imunoglobuliny G (IgG) mnoha savčích druhů, včetně lidského a myšího. Mimo to je také schopen vázat některé IgM a IgA. K navázání dochází v konstantní části těžkého řetězce (Fc) imunoglobulinu, tedy obráceným koncem než při protilátkové reakci. Tím si stafylokok zajišťuje ochranu proti fagocytóze a opsonizaci. Kromě toho váže i některé Ig se specifickou variabilní oblastí přes Fab, děje se tak ale přes jiné části proteinu než u Fc. Přirozeně produkovaný protein (viz obr.) má na počátku signální sekvenci (S), která u *S. aureus* zajišťuje transport z cytoplasmy do buněčné stěny, kde se váže na peptidoglykan. Při expresi v *E. coli* je pak takový protein



lokalizován především v periplasmě, odkud je možné jej uvolnit (L Abrahmsén et al. 1985))

osmotickým šokem. Každá z dalších pěti sekvencí (E,D,A,B,C) má délku kolem 58 aminokyselin, sestává se ze tří antiparalerních  $\alpha$ -helixů s hydrofobním jádrem a je nezávisle na ostatních schopna vazby na IgG. Poslední část (X) zajišťuje navázání proteinu na peptidoglykan v buněčné stěně (L Abrahmsén et al. 1985). SpA neobsahuje žádné cysteinové funkční skupiny, které by narušovaly tvorbu disulfidických můstků v cílovém proteinu. Při exprimaci v bakteriálních buňkách je možné jej využít jak k intracelulární produkci tak i k extracelulárnímu vylučování rekombinantního proteinu, pokud je zachována signální sekvence (Ni & Chen 2009).

Cílenou mutací B domény SpA byl získán Protein Z, který má také 58 aminokyselin a liší se jen náhradou Ala1 a Gly29 za Val a Ala. Touto změnou je zajištěna stabilita proti chemickému sekvenčně specifickému štěpení hydroxylaminem, který má u přirozené B domény cílové místo Asn-Gly(B. Nilsson et al. 1987). Schopnost vázat Fc oblast u IgG je srovnatelná s nemutovaným proteinem, poněkud nižší je vazba Fab. Protein Z má velikost 7 kDa a je často využíván na místo SpA (velikost 31 kDa) ve formě monomeru nebo lineárního dimeru ZZ, který má asi desetinásobnou afinitu k IgG. Další



Obr.6 Fúzní protein ZZ s IGF-I, 3 antiparalerní  $\alpha$ -helixy proteinu Z (převzato z (Samuelsson et al. 1994))

multiplikace nepřináší další výhody. Díky kompaktní struktuře se protein Z i jeho dimer ZZ dobře renaturuje a tato vlastnost se může přenést i na fúzní protein, zároveň se takto může výrazně zvýšit jeho rozpustnost (Samuelsson et al. 1994). Exprese je možná nejen v bakteriích, ale také v kvasinkách, CHO nebo hmyzích buňkách (Nygren et al. 1994).

Afinity mezi SpA a IgG je možné využít jednak k purifikaci proteinů s SpA nebo Z proteinovým tagem na koloně s IgG, ale také obráceně k separaci IgG na kolonách s náplní SpA.

Náplň kolony pro separaci fúzních proteinů s SpA, Z a ZZ proteiny tvoří imobilizované lidské polyklonální IgG. U aplikací, kde je nežádoucí možná přítomnost stopových množství sérových proteinů se používá rekombinantních Fc fragmentů.

Eluce se obvykle provádí výrazným snížením pH (kolem pH=3, kyselina octová nebo citronová) nebo u pH sensitivejších fúzních proteinů kompetitivním vytěsněním pomocí rekombinantního proteinu ZZ nebo proteinem BiO-Ox (Strambio-de-Castilla et al. 2005). K odštěpení afinitního tagu se používá některá z běžných endopeptidáz a cílová sekvence vložená před cílový protein.

## **Závěr**

Se stále rostoucími možnostmi genomiky a zvyšujícím se množstvím osekvenovaných genomů z různých taxonů stoupá potřeba systémů umožňujících snadnou a spolehlivou expresi a izolaci rekombinantních proteinů. Fúzní proteiny s vhodným afinitním tagem umožňují zpravidla v jednom kroku afinitní chromatografie dosáhnout dobré čistoty při výtěžcích a škálovatelnosti neporovnatelně lepších než jiné metody. Navíc vhodná volba afinitního tagu může žádoucím směrem modifikovat vlastnosti fúzního proteinu. Především může zvyšovat rozpustnost, usnadňovat refolding a umožňovat sekreci do periplasmy či média. Mimo to je možné afinitní tag využít i při některých dalších aplikacích rekombinantního proteinu. V této práci jsem zpracoval informace o nejpoužívanějších, zpravidla komerčně dostupných afinitních systémech, které by měly čtenáře seznámit s problematikou a usnadnit výběr vhodného tagu pro konkrétní aplikaci.

## Seznam použitých zkratek

CHO	chinese hamster ovary
DAPáza	dipeptidyl-aminopeptidáza I
EGTA	ethylenglykol tetraoctová kyselina
ELISA	enzyme-linked immunosorbent assay
GST	glutathion S-transferáza
HABA	kyselina 2-(4'-hydroxy phenylazo)benzoová
HIC	hydrophobic interaction chromatography
IDA	imindioctová kyselina
IEX	ion-exchange chromatography
IGF I	insulin-like growth factor I
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl- $\alpha$ -D-1-thiogalacto-pyranosid
MBP	maltose binding protein
MCS	multiple restriction/cloning site
NTA	nitril trioctová kyselina
pGAPáza	pyroglutamylaminopeptidáza
Qcykláza	glutamin-cyklotransferáza
SEC	size exclusion chromatography
SPM	scanning probe microscopy
TAP	tandem affinity purification
TED	triskarboxymethylethylendiamin

## Seznam použité literatury

- Abdullah, N. & Chase, H.A., 2005. Removal of poly-histidine fusion tags from recombinant proteins purified by expanded bed adsorption. *Biotechnology and Bioengineering*, 92(4), 501-513.
- Abrahmsén, L. et al., 1985. Analysis of signals for secretion in the staphylococcal protein A gene. *The EMBO Journal*, 4(13B), 3901-3906.
- Armstrong, R.N., 1997. Structure, Catalytic Mechanism, and Evolution of the Glutathione Transferases. *Chemical Research in Toxicology*, 10(1), 2-18.
- Ausubel, F.M. et al. eds., 2001. *Current Protocols in Molecular Biology*, Hoboken, NJ, USA: John Wiley & Sons, Inc. Available at: <http://www.currentprotocols.com/protocol/mb1606> [Accessed June 3, 2010].
- Bolanos-Garcia, V.M. & Davies, O.R., 2006. Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *BBA-General Subjects*, 1760(9), 1304–1313.
- Brewer, S.J. & Sassenfeld, H.M., 1985. The purification of recombinant proteins using C-terminal polyarginine fusions. *Trends in Biotechnology*, 3(5), 119-122.
- Einhauer, A. & Jungbauer, A., 2001a. Affinity of the monoclonal antibody M1 directed against the FLAG peptide. *Journal of Chromatography A*, 921(1), 25-30.
- Einhauer, A. & Jungbauer, A., 2001b. The FLAG™ peptide, a versatile fusion tag for the purification of recombinant proteins. *Journal of Biochemical and Biophysical Methods*, 49(1-3), 455-465.
- Fuchs, S.M. & Raines, R.T., 2005. Polyarginine as a multifunctional fusion tag. *Protein Science: A Publication of the Protein Society*, 14(6), 1538-1544.
- Futatsumori-Sugai, M. et al., 2009. Utilization of Arg-elution method for FLAG-tag based chromatography. *Protein Expression and Purification*, 67(2), 148-155.
- Hopp, T.P. et al., Synthesis of protein with an identification peptide.
- Hopp, T.P. et al., 1988. A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. *Nat Biotech*, 6(10), 1204-1210.
- Hopp, T. & Woods, K., 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 78(6), 3828, 3824.
- Huang, H., Liu, J. & Marco, A., 2006. Induced fit of passenger proteins fused to Archaea maltose binding proteins. *Biochemical and Biophysical Research Communications*, 344(1), 25–29.
- Chico, D.E., Given, R.L. & Miller, B.T., 2003. Binding of cationic cell-permeable peptides to plastic and glass. *Peptides*, 24(1), 3-9.
- Kapust, R.B. & Waugh, D.S., 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science*, 8(8), 1668-1674.
- Knecht, S. et al., 2009. Oligohis-tags: mechanisms of binding to Ni<sup>2+</sup>-NTA surfaces. *Journal of Molecular Recognition*, 22(4), 270-279.
- Leaver, M.J. & George, S.G., 1998. A piscine glutathione S-transferase which efficiently conjugates the end-products of lipid peroxidation. *Marine Environmental Research*, 46(1-5), 71-74.

- Lichty, J.J. et al., 2005. Comparison of affinity tags for protein purification. *Protein Expression and Purification*, 41(1), 98-105.
- Millet, O., Hudson, R.P. & Kay, L.E., 2003. The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy. *Proceedings of the National Academy of Sciences*, 100(22), 12700–12705.
- Mitchell, D. et al., 2000. Polyarginine enters cells more efficiently than other polycationic homopolymers. *The Journal of Peptide Research*, 56(5), 318-325.
- New England Biolab 1991. pMAL™ Protein Fusion and Purification System.
- Ni, Y. & Chen, R., 2009. Extracellular recombinant protein production from *Escherichia coli*. *Biotechnology Letters*, 31(11), 1661-1670.
- Nilsson, B. et al., 1987. A synthetic IgG-binding domain based on staphylococcal protein A. *Protein Engineering*, 1(2), 107-113.
- Nilsson, J. et al., 1997. Affinity Fusion Strategies for Detection, Purification, and Immobilization of Recombinant Proteins\*1. *Protein Expression and Purification*, 11(1), 1-16.
- Nock, S., Spudich, J.A. & Wagner, P., 1997. Reversible, site-specific immobilization of polyarginine-tagged fusion proteins on mica surfaces. *FEBS Letters*, 414(2), 233-238.
- Nygren, P., Stefan, S. & Uhlén, M., 1994. Engineering proteins to facilitate bioprocessing. *Trends in Biotechnology*, 12(5), 184-188.
- Parsy, C.B. et al., 2007. Two-step method to isolate target recombinant protein from co-purified bacterial contaminant SlyD after immobilised metal affinity chromatography. *Journal of Chromatography B*, 853(1-2), 314–319.
- Pedersen, J. et al., 1999. Removal of N-Terminal Polyhistidine Tags from Recombinant Proteins Using Engineered Aminopeptidases\*1. *Protein Expression and Purification*, 15(3), 389-400.
- Pryor, K.D. & Leiting, B., 1997. High-Level Expression of Soluble Protein in *Escherichia coli* Using a His6-Tag and Maltose-Binding-Protein Double-Affinity Fusion System. *Protein Expression and Purification*, 10(3), 309-319.
- Rothbard, J.B. et al., 2000. Conjugation of arginine oligomers to cyclosporin A facilitates topical delivery and inhibition of inflammation. *Nat Med*, 6(11), 1253-1257.
- Saluta, M. & Bell, P.A., 1998. Troubleshooting GST fusion protein expression in *E. coli*. *Life Science News*, 1, 1–3.
- Samuelsson, E. et al., 1994. Enhanced in vitro Refolding of Insulin-like Growth Factor I Using a Solubilizing Fusion Partner. *Biochemistry*, 33(14), 4207-4211.
- Schmidt, T.G.M. et al., 1996. Molecular Interaction Between the Strep-tag Affinity Peptide and its Cognate Target, Streptavidin. *Journal of Molecular Biology*, 255(5), 753-766.
- Skerra, A. & Schmidt, T.G.M., 1999. Applications of a peptide ligand for streptavidin: the Strep-tag. *Biomolecular Engineering*, 16(1-4), 79-86.
- Smith, D.B. & Johnson, K.S., 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*, 67(1), 31-40.

- Stofko-Hahn, R.E., Carr, D.W. & Scott, J.D., 1992. A single step purification for recombinant proteins Characterization of a microtubule associated protein (MAP 2) fragment which associates with the type II cAMP-dependent protein kinase. *FEBS Letters*, 302(3), 274-278.
- Strambio-de-Castillia, C. et al., 2005. A Method for the Rapid and Efficient Elution of Native Affinity-Purified Protein A Tagged Complexes. *J. Proteome Res*, 4(6), 2250–2256.
- Uhlén, M. et al., 1984. Expression of the gene encoding protein A in *Staphylococcus aureus* and coagulase-negative staphylococci. *Journal of Bacteriology*, 159(2), 713-719.
- Vivès, E., Brodin, P. & Lebleu, B., 1997. A Truncated HIV-1 Tat Protein Basic Domain Rapidly Translocates through the Plasma Membrane and Accumulates in the Cell Nucleus. *Journal of Biological Chemistry*, 272(25), 16010 -16017.
- Vuilleumier, S., 1997. *Bacterial glutathione S-transferases: what are they good for?*, Am Soc Microbiol.
- Waugh, D.S., 2005. Making the most of affinity tags. *Trends in Biotechnology*, 23(6), 316-320.
- Yan, F. et al., 2007. A novel pro-apoptosis protein PNAS-4 from *Xenopus laevis*: cloning, expression, purification, and polyclonal antibody production. *Biochemistry. Biokhimič*72(6), 664-671.
- Yang, J. et al., 2007. Prokaryotic expression and polyclonal antibody preparation of a novel Rab-like protein mRabL5. *Protein Expression and Purification*, 53(1), 1-8.
- Zhang, P. et al., 2008. Prokaryotic expression of a novel mouse pro-apoptosis protein PNAS-4 and application of its polyclonal antibodies. *Brazilian Journal of Medical and Biological Research = Revista Brasileira De Pesquisas Médicas E Biológicas / Sociedade Brasileira De Biofísica ... [et Al*, 41(6), 504-511.
- Zheng, C. et al., 1997. A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins. *Gene*, 186(1), 55-60.
- Innovative Vector System For High-Level Protein Expression In E. Coli - Biocompare Buyer's Guide For Life Scientists. Available at: <http://www.biocompare.com/Articles/ApplicationNote/195/Innovative-Vector-System-For-High-Level-Protein-Expression-In-E-Coli.html> [Accessed August 1, 2010].