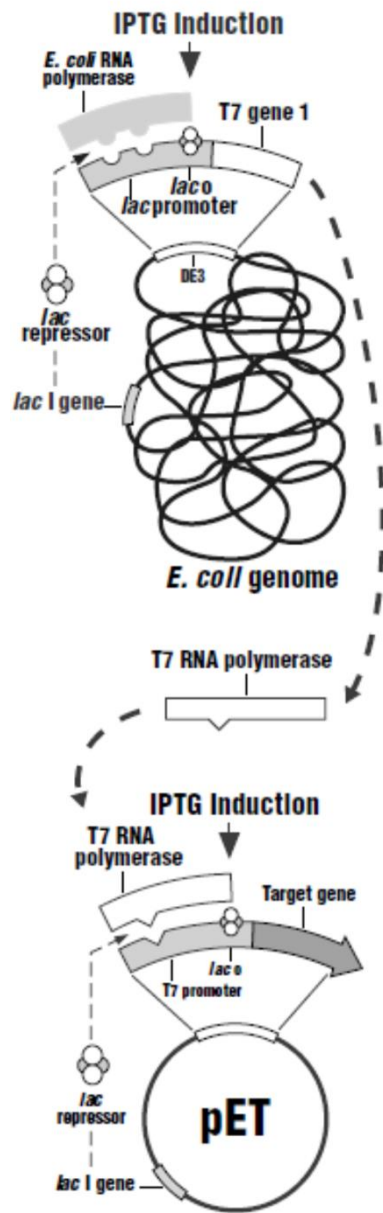


# Expresse rekombinantních protein



# Rekombinantní proteiny

**Rekombinantní DNA** je uměle vytvořená DNA sekvence, která vznikla novou kombinací dvou nebo více různých DNA sekvencí.

**Rekombinantní proteiny** jsou proteiny získané vnesením rekombinantní DNA do heterologního hostitele (např. mikroorganismus, kvasinky), ve kterém dojde k expresi genu.

# Využití rekombinantních proteinů

## 1. V základním výzkumu:

É**Biochemická funkce a charakteristika proteinu** (určení přesných kinetických parametrů  $K_m$ ,  $k_{cat}$  pro enzymy se substrátem,  $K_i$  pro enzymy s inhibítorem,  $K_d$  pro protein - proteinové interakce a ligand - proteinové interakce)

É**Strukturní analýza** (NMR, krystalografie, kryo-EM)

É**Proteinové inženýrství** (zlepšení vlastností proteinů o aktivita, stabilita)

## 2. V praxi:

É**např. léky, vakcíny, proteiny pro diagnostiku, potravinové doplňky**

**Cíl:** Vysoký výtěžek homogenního proteinu (mg o kg proteinu)

Zachování biologické aktivity

# Pro výrobu rekombinantní proteiny?

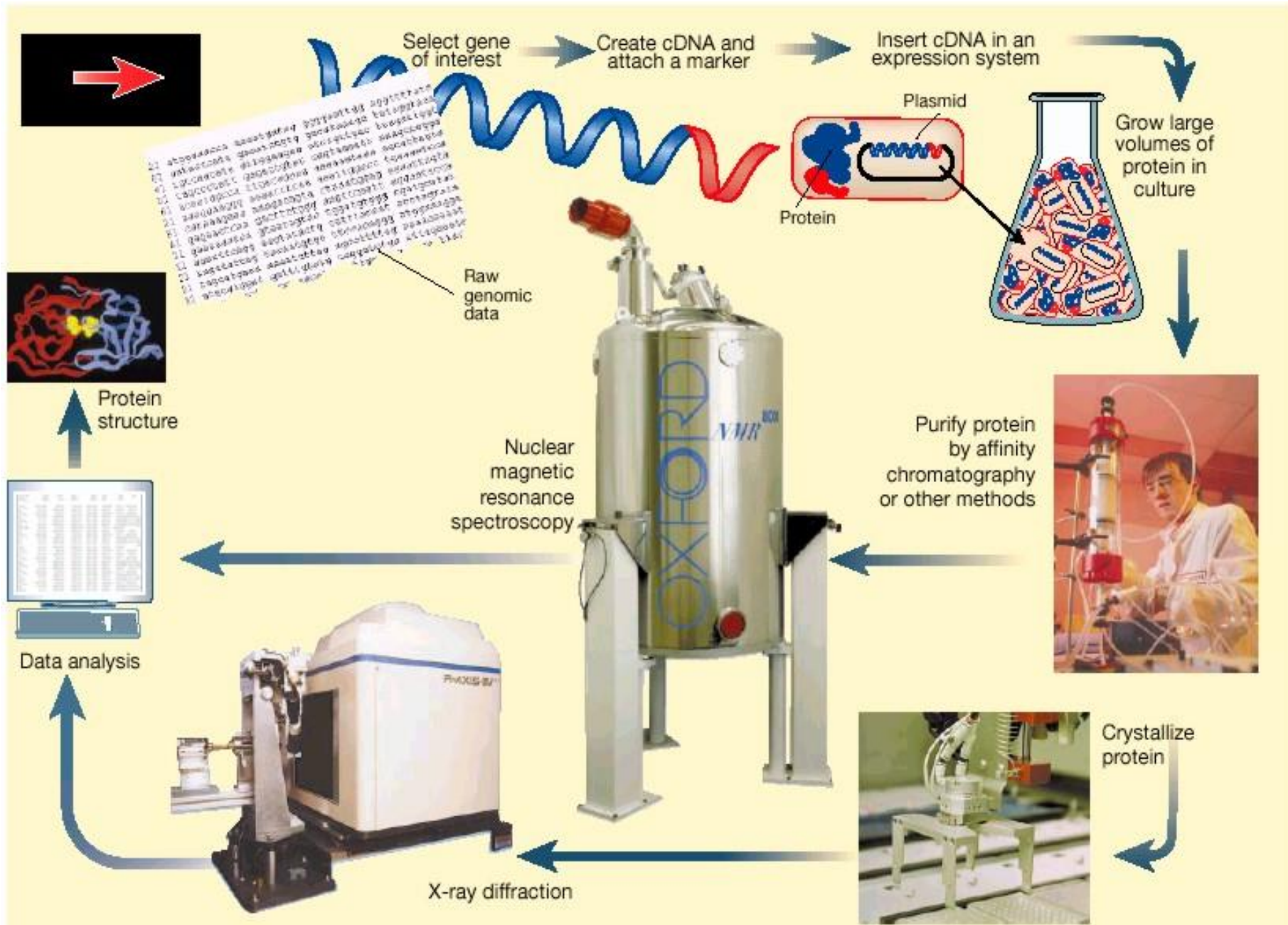
- Pro širozený zdroj:
  - É Obtífn se získává (tkán , orgány).
  - É Obtífn se kultivuje (bakterie, viry, tká ové kultury).
  - É Limitovaná exprese
  - Éasto obtífná purifikace proteinu (mnohakroková)
  - É Mofná infek ní rizika v pr b hu izolace nebo ve výsledném produktu

TABLE 1.2. Examples of low-abundance proteins and peptides isolated from natural biological sources

Protein	Source	Yield ( $\mu\text{g}$ )	Reference
Multipotential colony-stimulating factor	pokeweed mitogen-stimulated mouse spleen-cell-conditioned medium (10 liters)	1	Cutler et al. (1985)
Human A33 antigen	human colon cancer cell lines ( $10^{10}$ cells)	2.5	Catimel et al. (1996)
Platelet-derived growth factor (PDGF)	human serum (200 liters)	180	Heldin et al. (1981)
Granulocyte-colony-stimulating growth factor (G-CSF)	mouse lung-conditioned medium (3 liters)	40	Nicola et al. (1983)
Granulocyte-macrophage colony-stimulating growth factor (GM-CSF)	mouse lung-conditioned medium (3 liters)	12	Burgess et al. (1986)
Coelenterate morphogen	sea anemone (200 kg)	20	Schaller and Bodenmuller (1981)
Peptide YY (PYY)	porcine intestine (4000 kg)	600	Tatemoto (1982)
Tumor necrosis factor (TNF)	HL60 tissue culture medium (18 liters)	20	Wang and Creasy (1985)
Murine transferrin receptor	NS-1 myeloma cells ( $10^{10}$ cells)	20	van Driel et al. (1984)
Fibroblast growth factor (FGF)	bovine brain (4 kg)	33	Gospodarowicz et al. (1984)
Transforming growth factor- $\beta$ (TGF- $\beta$ )	human placenta (8.8 kg)	47	Frolik et al. (1983)
Human interferon	human leukocyte-conditioned medium (10 liters)	21	Rubinstein et al. (1979)
Muscarinic acetylcholine receptor	porcine cerebrum (600 g)	6	Haga and Haga (1985)
$\beta_2$ -adrenergic receptor	rat liver (400 g)	2	Graziano et al. (1985)

Adapted with permission from Simpson and Mize (1989)

# Technologie rekombinantních protein



# Hostitelský organismus pro expresi rekombinantních protein

ÉProkaryotní expresní systémy (*Escherichia coli*, *Bacillus subtilis*, í )

” Kvasinky (*Sacharomyces cerevisiae*, *Pichia pastoris*, í ..)

ÉSav í bu ky (linie ovariálních bun k k e ka ínského- CHO bu ky, linie lidských embryonálních ledvinových bun k-HEK, í )

ÉHmyzí bu ky s bakuloviry

ÉExpresse proteinu *in vitro* (lyzáty z králi ích retikulocyt , p-ení ných klí k , *E.coli*)



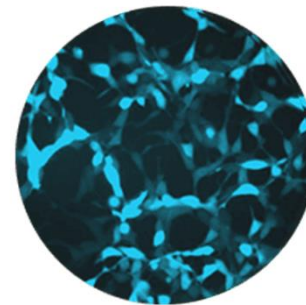
Bacteria  
Expression  
System



Yeast  
Expression  
System



Insect  
Expression  
System



Mammalian  
Expression  
System



## Kritéria pro výběr hostitelského organismu pro expresi rekombinantních proteinů

- Lokalizace (intracelulární, sekretovaný, membránový protein)
- Velikost/hmotnost a zda je protein jedno nebo více doménový
- Přítomnost disulfidických můstků
- Post-translační modifikace
- Možnosti laboratoře (náklady na vybavení)
- Rychlost
- Požadované množství proteinu



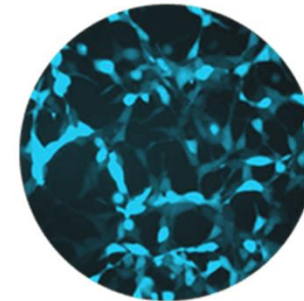
**Bacteria  
Expression  
System**



**Yeast  
Expression  
System**



**Insect  
Expression  
System**



**Mammalian  
Expression  
System**

# Kritéria pro výběr hostitelského organismu pro expresi rekombinantních proteinů

Expresní systém	Posttranslační modifikace					
	N-glykosylace	O-glykosylace	Fosforylace	Acetylace	Acylace	$\gamma$ -karboxylace
<i>E. coli</i>	chybí	-	-	-	-	-
<b>Kvasinky</b>	vysoce manosilované glykany	+	+	+	+	-
<b>Hmyzí buňky</b>	jednoduchá, bez sializace	+	+	+	+	-
<b>Savčí buňky</b>	komplexní	+	+	+	+	+

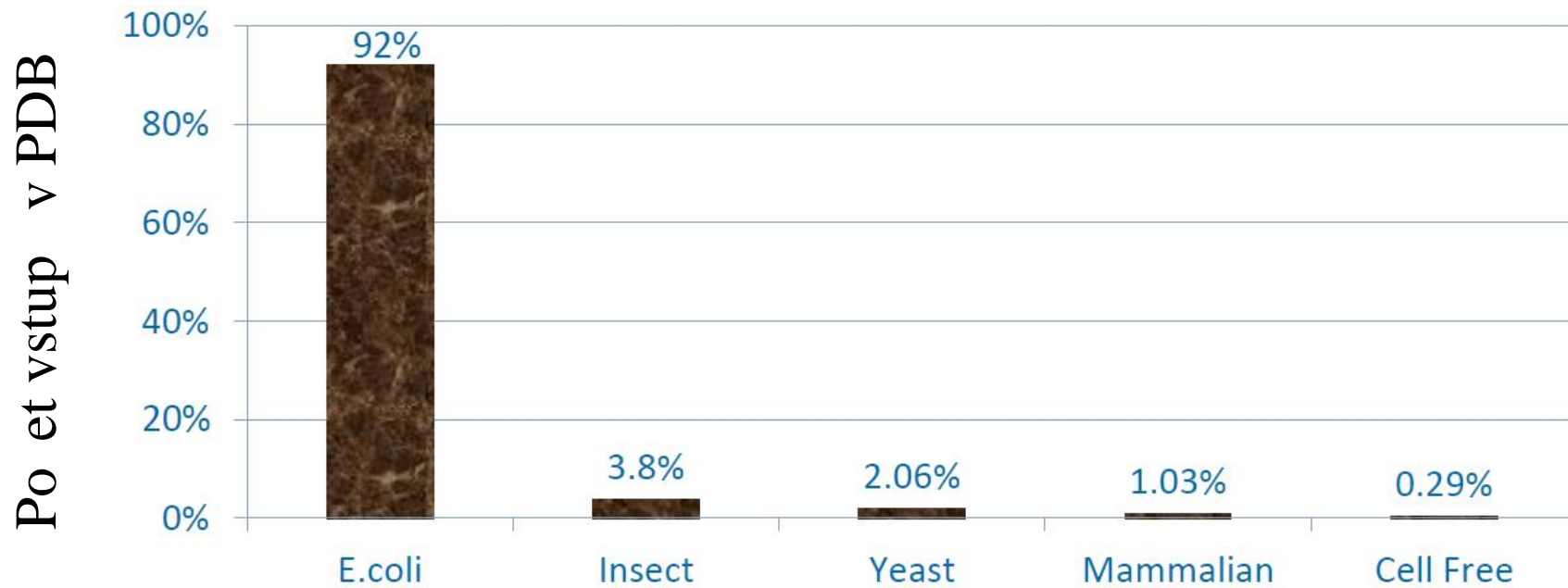
- “ Post-translační modifikace mají vliv na folding proteinu, rozpustnost, stabilitu, transport, imunogenicitu a funkční aktivitu proteinu.
- “ Jeden z hlavních rozdílů mezi eukaryotickými expresními hostiteli je v typu glykosylace (N- a O-glykosylace), kterou mohou poskytnout.



# Kritéria pro výběr hostitelského organismu pro expresi rekombinantních protein

Expresní systém	Náklady na kultivaci	Rychlost růstu	Úroveň exprese	Konformace proteinu
<i>E. coli</i>	nízké	vysoká	vysoká	často nutný refolding
<b>Kvasinky</b>	nízké	vysoká	nízká a vysoká	ne vždy nutný refolding
<b>Hmyzí buňky</b>	vysoké	nízká	nízká a vysoká	většinou správná
<b>Savčí buňky</b>	vysoké	nízká	většinou nízká	většinou správná

## Statistika v PDB



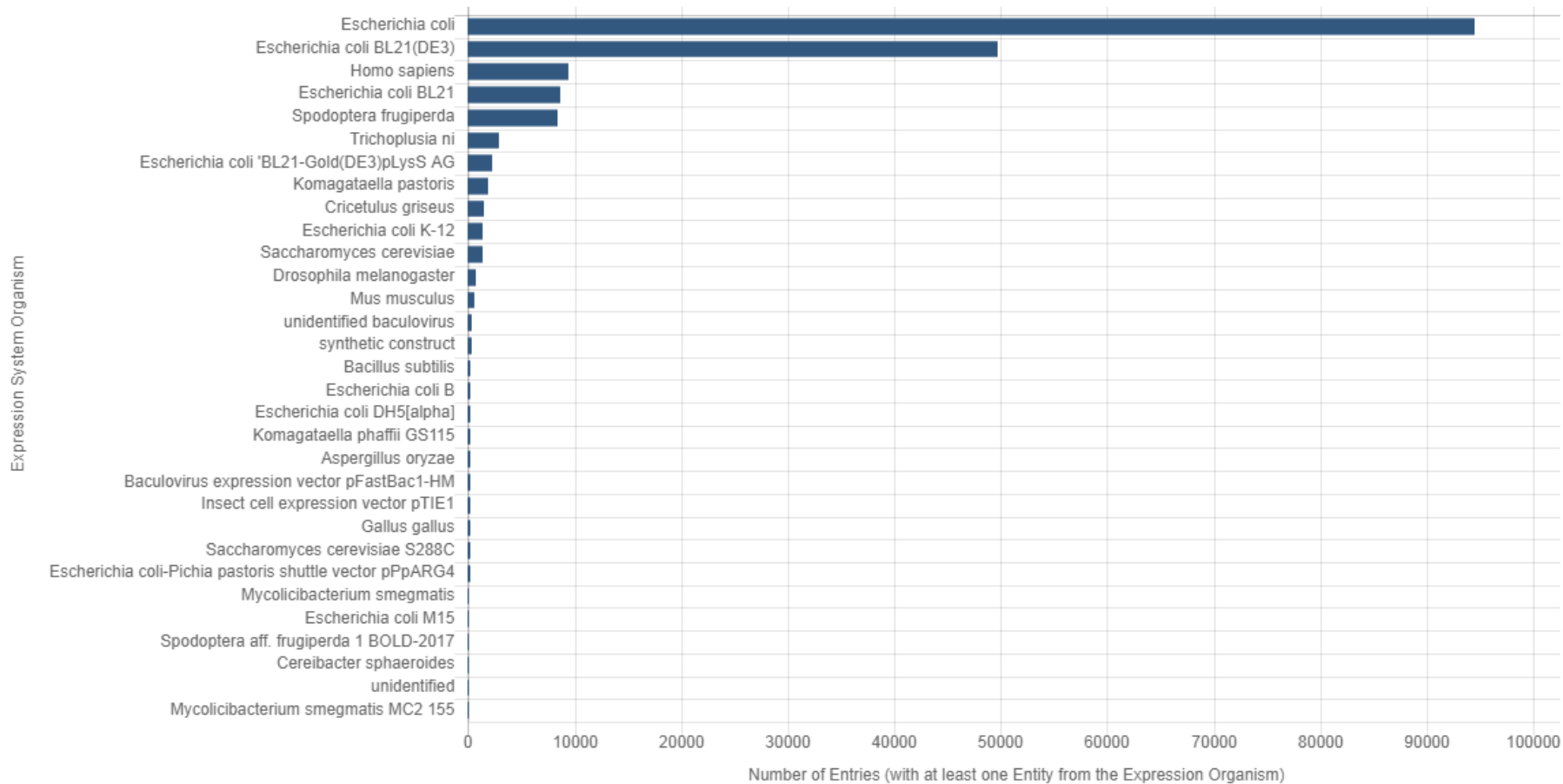
Organismus použitý pro expresi

Jedná se o průměrné hodnoty z roku 2014

<https://www.rcsb.org/stats/distribution-expression-organism-gene>

# PDB Statistics: PDB Data Distribution by Expression System Organism

Expression organism for these structures is from a genetically manipulated source.



<https://www.rcsb.org/stats/distribution-expression-organism-gene>

# Produkce heterologních proteinů v *E. Coli*

## VÝHODY :

ÉJednoduchost

ÉVysoká produkce rekombinantních proteinů

Dobře prostudovaný genom a proteom-usnadnění  
genových manipulací

ÉDesign vektorů usnadňuje klonování a expresi cizích genů

ÉRychlý růst v poměrně levném médiu

ÉMinimální požadavky na vybavení

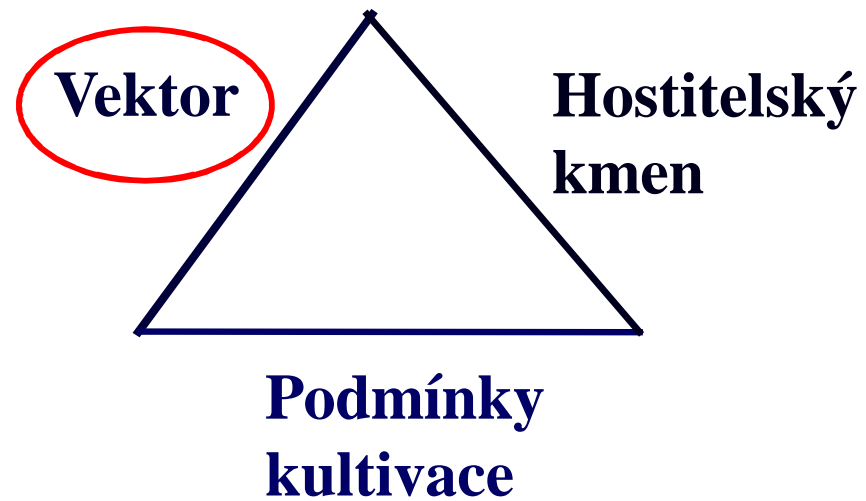
ÉPřizpůsobivost systému

# Produkce heterologních proteinů v *E. Coli*

## NEVÝHODY:

- É Potřeba cDNA zkoumaného proteinu
  - É Absence eukaryotických posttranslačních systémů (posttranslačních modifikací)
  - É Tvorba inkluzních tělísek (nerozpustná forma proteinu)
  - É Omezená schopnost tvorby disulfidických vazeb
  - É Preferují jiný genetický kód než eukaryota
  - É Kontaminace proteinů lipopolysacharidem (endotoxin)
  - É Chybí sekreční mechanismus pro účinné uvolnění proteinu do kultivačního média
- “ často selhává při produkci proteinových komplexů, které obsahují disulfidické vazby, při produkci eukaryotických membránových proteinů, více-doménových proteinů a proteinů složených z více podjednotek

# Expresní systém *E.coli*



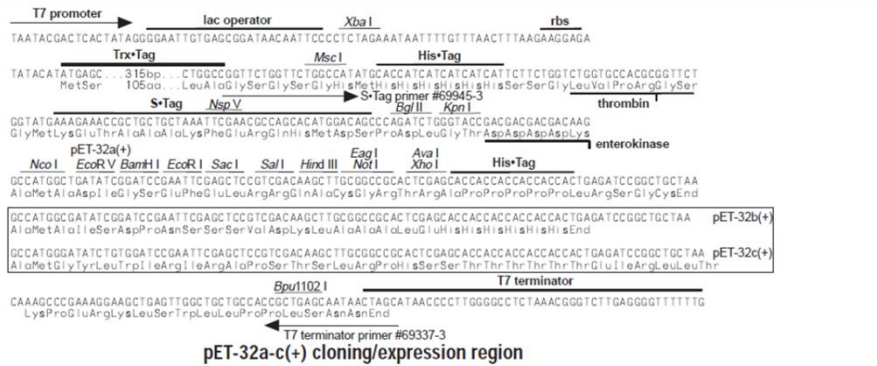
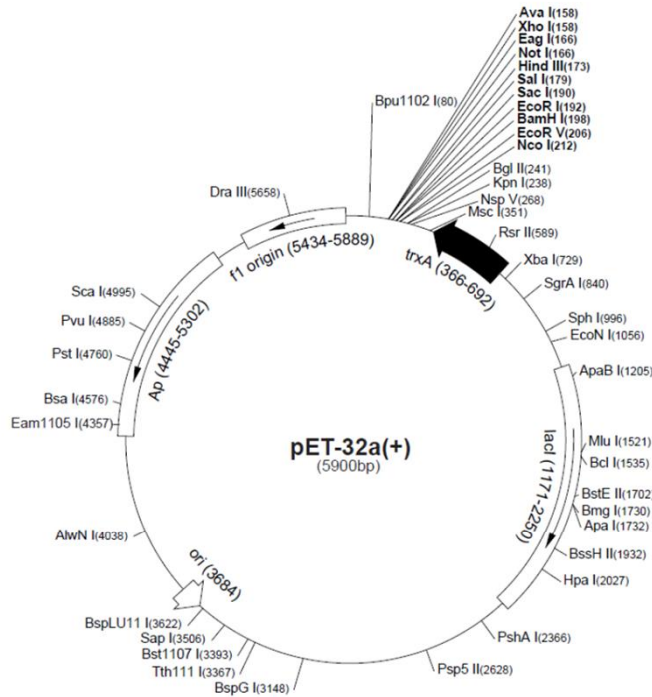


# Expresní vektor

= klonovací vektor, který obsahuje nezbytné regulační sekvence k tomu, aby podporoval expresi inzertního genu.

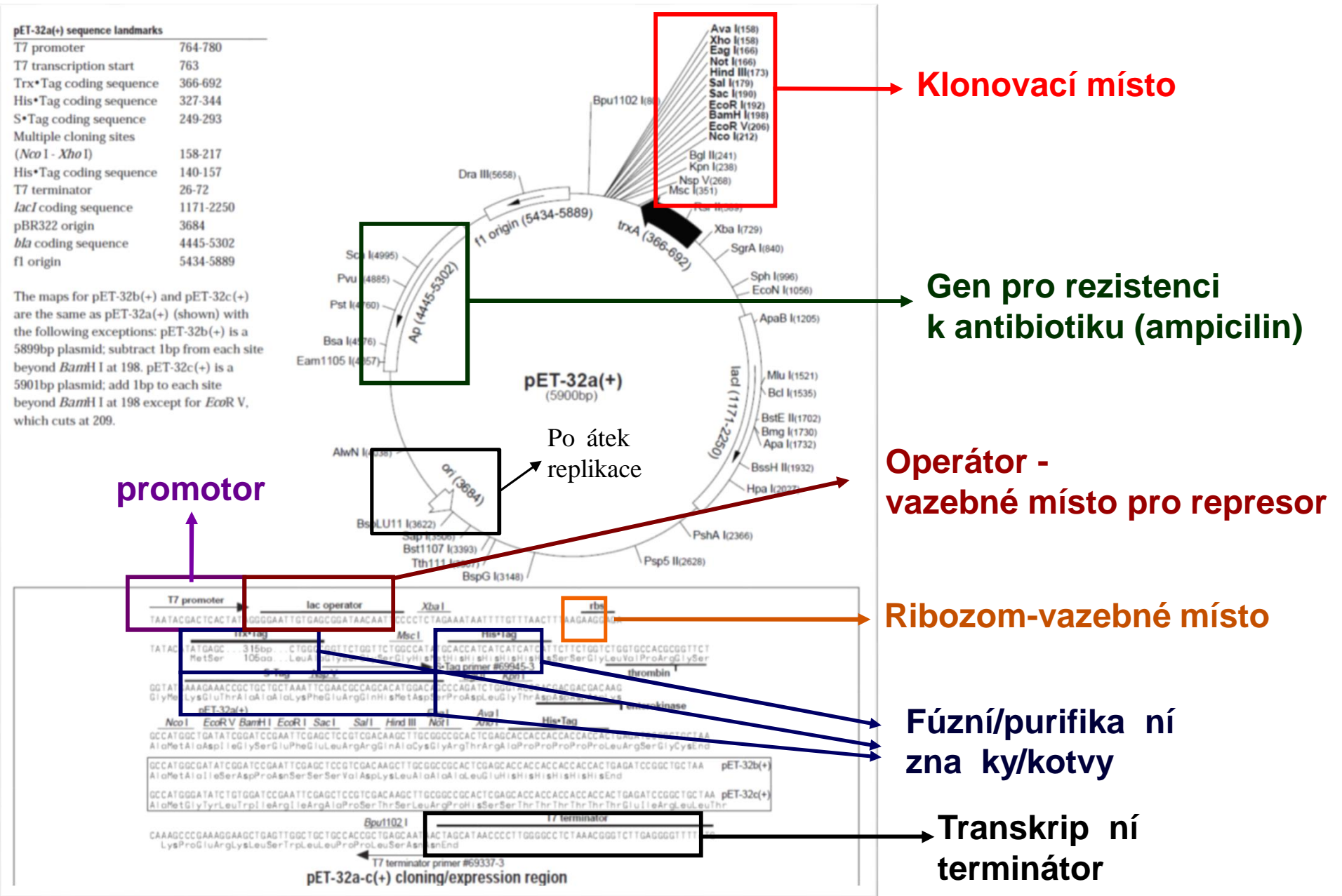
pET-32a(+) sequence landmarks	
T7 promoter	764-780
T7 transcription start	763
Trx*Tag coding sequence	366-692
His*Tag coding sequence	327-344
S*Tag coding sequence	249-293
Multiple cloning sites ( <i>Nco</i> I - <i>Xho</i> I)	158-217
His*Tag coding sequence	140-157
T7 terminator	26-72
<i>lac</i> I coding sequence	1171-2250
pBR322 origin	3684
<i>bla</i> coding sequence	4445-5302
f1 origin	5434-5889

The maps for pET-32b(+) and pET-32c(+) are the same as pET-32a(+) (shown) with the following exceptions: pET-32b(+) is a 5899bp plasmid; subtract 1bp from each site beyond *Bam*HI at 198. pET-32c(+) is a 5901bp plasmid; add 1bp to each site beyond *Bam*HI at 198 except for *Eco*R V, which cuts at 209.



pET-32a-c(+) cloning/expression region

# Struktura vektoru pro úinnou expresi rekombinantních proteinů v *E.coli*

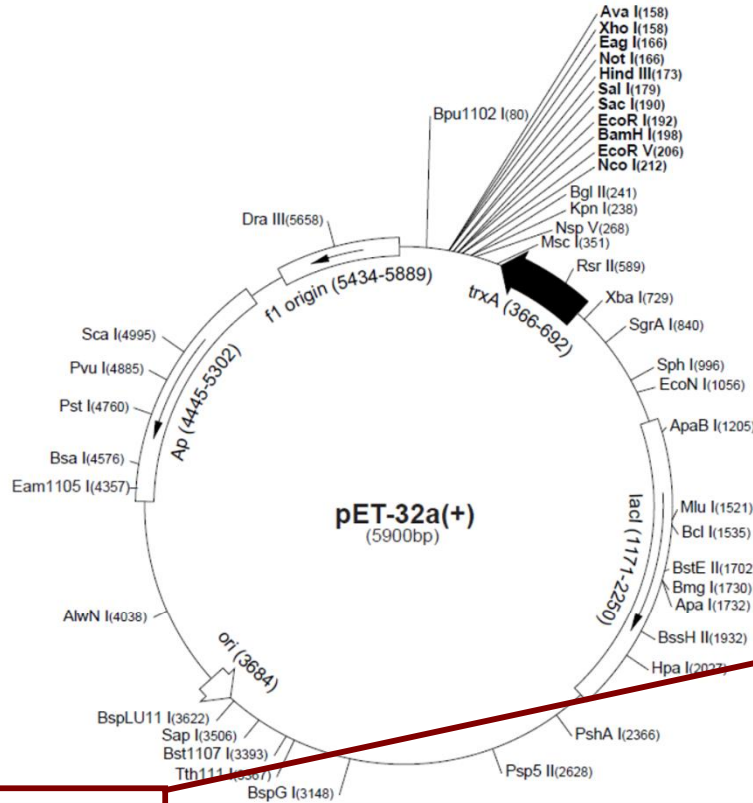


# Struktúra vektoru pro ú innou expresi rekombinantních proteinů v *E.coli*

## pET-32a(+) sequence landmarks

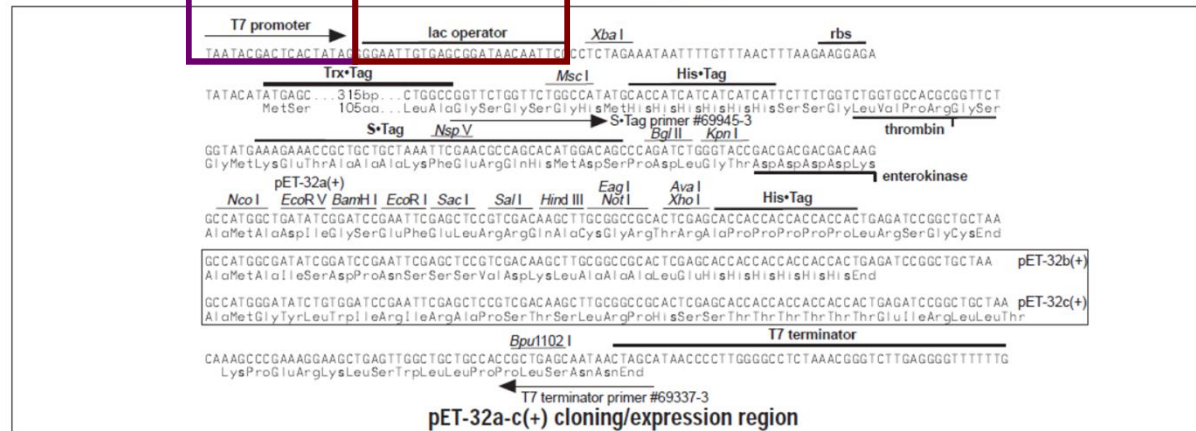
T7 promoter	764-780
T7 transcription start	763
Trx*Tag coding sequence	366-692
His*Tag coding sequence	327-344
S*Tag coding sequence	249-293
Multiple cloning sites ( <i>Nco</i> I - <i>Xho</i> I)	158-217
His*Tag coding sequence	140-157
T7 terminator	26-72
<i>lacI</i> coding sequence	1171-2250
pBR322 origin	3684
<i>bla</i> coding sequence	4445-5302
f1 origin	5434-5889

The maps for pET-32b(+) and pET-32c(+) are the same as pET-32a(+) (shown) with the following exceptions: pET-32b(+) is a 5899bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-32c(+) is a 5901bp plasmid; add 1bp to each site beyond *Bam*H I at 198 except for *Eco*R V, which cuts at 209.



promotor

Operátor -  
vazebné místo pro represor



## **Vlastnosti promotoru:**

**ÉSilný promotor** (ptac, ptrp,  $\lambda$ pL, pT7)

- Protein zájmu by m l tvo it 10-30% a více z celkového bakteriálního proteinu.

**ÉP enositelný do r zných kmen *E.coli***

**ÉJednodu-e a levn inducibilní**

- Teplotní indukce ( $\lambda$ pL)

- Chemická indukce (ptac, ptrp, pT7): IPTG (isopropyl- $\beta$ -D-thiogalaktopyranosid)

**ÉVykazuje minimální hladinu bazální exprese**

- Pokud jsou proteiny netoxické dosahuje se vysokých výt flk protein r stem bun k do vysokých hustot s následnou indukcí aktivity promotoru.

- U toxických protein je nutná minimalizace bazální transkripce p ed p ídavkem induk ního inidla pomocí vhodného represoru.

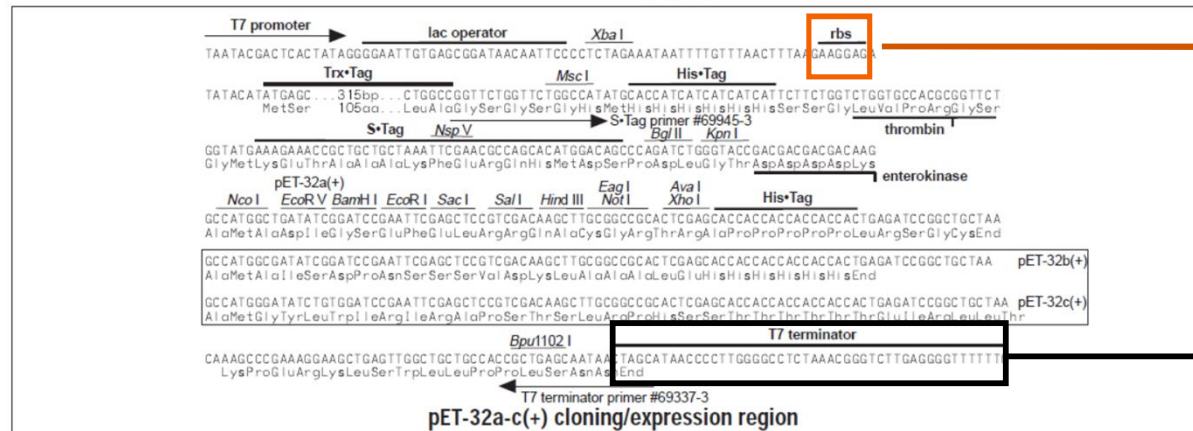
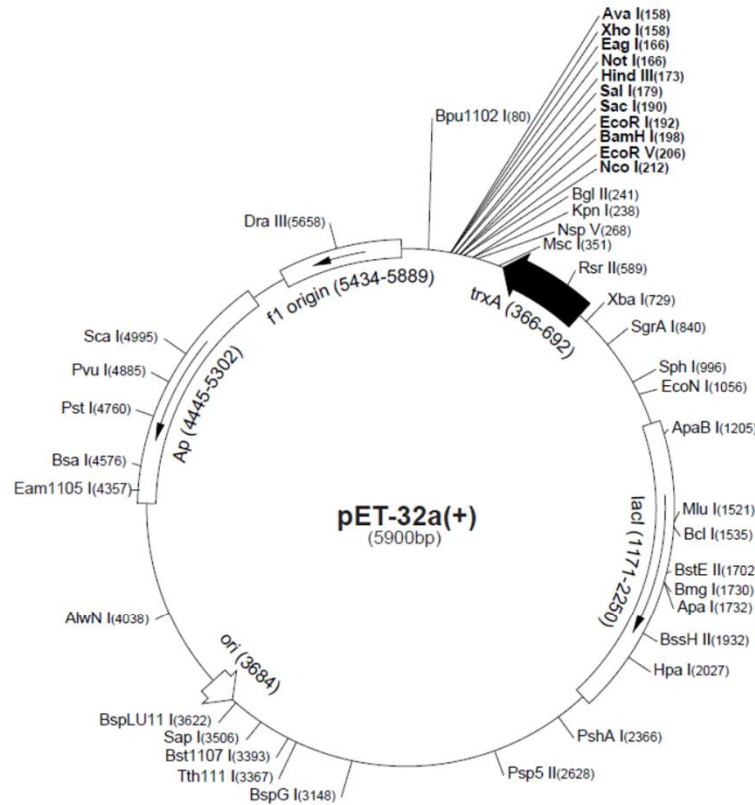


# Struktura vektoru pro ú innou expresi rekombinantních proteinů v *E.coli*

## pET-32a(+) sequence landmarks

T7 promoter	764-780
T7 transcription start	763
Trx*Tag coding sequence	366-692
His*Tag coding sequence	327-344
S*Tag coding sequence	249-293
Multiple cloning sites ( <i>Nco</i> I - <i>Xho</i> I)	158-217
His*Tag coding sequence	140-157
T7 terminator	26-72
<i>lacI</i> coding sequence	1171-2250
pBR322 origin	3684
<i>bla</i> coding sequence	4445-5302
f1 origin	5434-5889

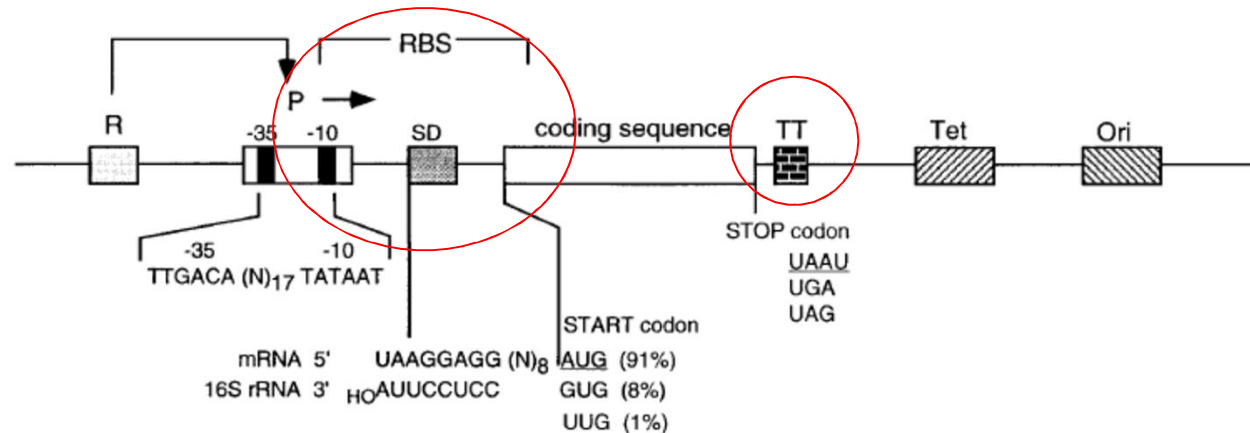
The maps for pET-32b(+) and pET-32c(+) are the same as pET-32a(+) (shown) with the following exceptions: pET-32b(+) is a 5899bp plasmid; subtract 1bp from each site beyond *Bam*HI at 198. pET-32c(+) is a 5901bp plasmid; add 1bp to each site beyond *Bam*HI at 198 except for *Eco*R V, which cuts at 209.



Ribozom-vazebné miesto

Transkrip ní terminátor

# Struktura vektoru pro účinnou expresi rekombinantních proteinů v *E.coli*



**Ribozom-vazebné místo:** zahrnuje Shine-Dalgarnova (SD) sekvenci a translační iniciační kodon

Vzdálenost mezi SD sekvencí a iniciačním kodonem AUG je 4-13 nukleotidů, tato vzdálenost ovlivňuje účinnost iniciace translace (**optimální vzdálenost je 4-8 nukleotidů**), oblast bohatá na AT páry.

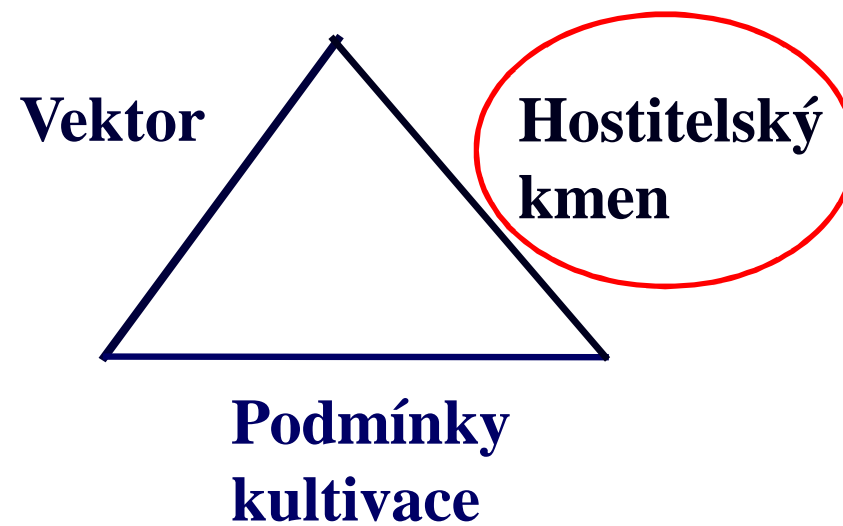
**Transkripční terminátor**

T<sub>7</sub> term, rrnT1, T2

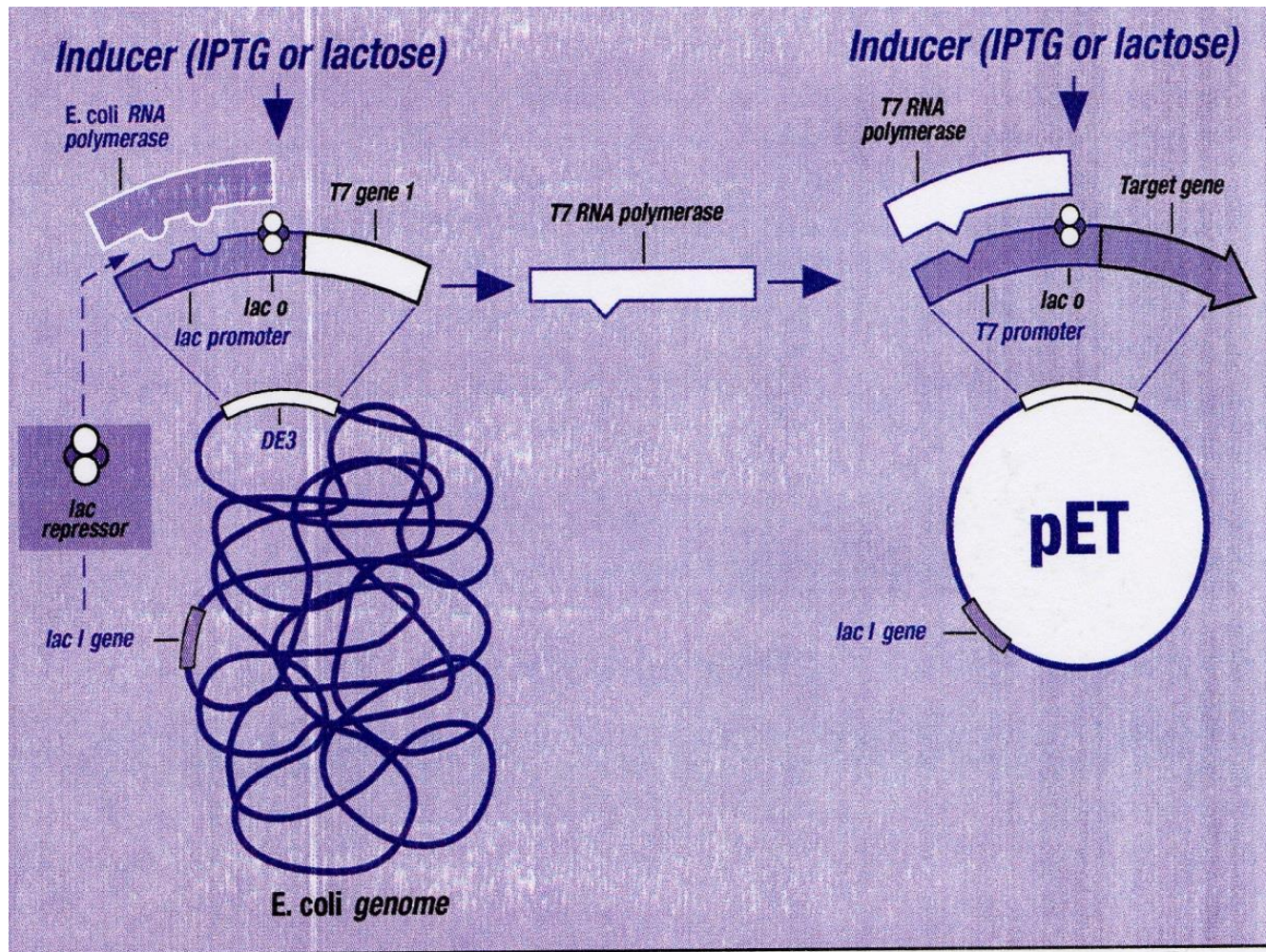
(zabraňuje okluzi promotoru, zvyšuje stabilitu mRNA)



# Výběr hostitelského kmene *E.coli*



# Indukce exprese rekombinantního proteinu v hostitelském kmeni *E. coli* (BL21(DE3))



## Výběr hostitelského kmene *E.coli* s ohledem na toxicitu proteinu pro hostitele



É Toxicita není omezena na pouhý fakt, že je protein je pro buňku cizí, ale může být způsobena i tím, že je nadprodukován určitý nativní gen.

É Nutná přesná regulace expresního systému

Komerčně dostupné bakteriální kmeny s různými úrovněmi regulace exprese, zajišťující snížení bazální exprese.

BL21(DE3)

BL21(DE3)pLysS/E

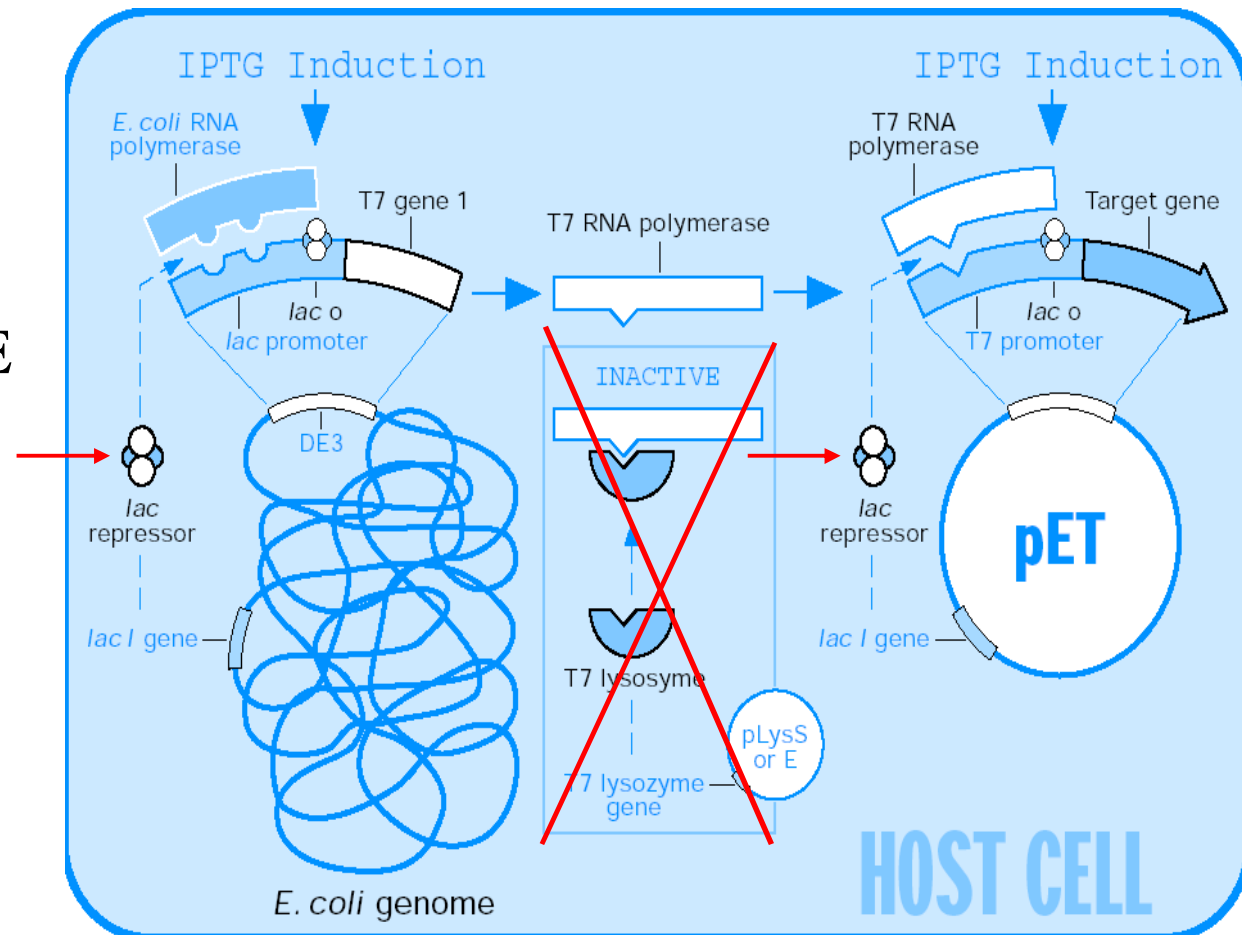
BL21

# R zné úrovn minimalizace bazální exprese

✓ **BL21(DE3)**

BL21(DE3)pLysS/E

BL21



Cca 10 % hladina bazální exprese (p ed indukcí exprese) klonovaného genu.

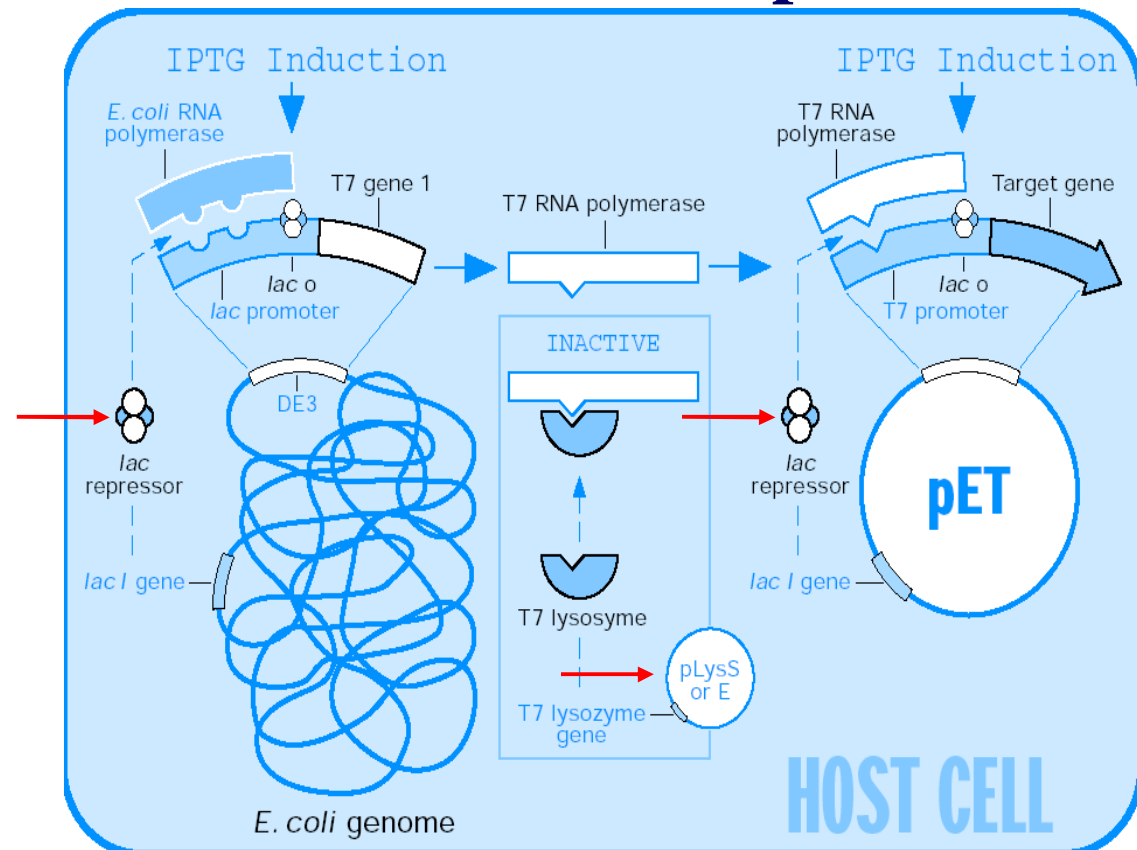


## Různé úrovně minimalizace bazální exprese

BL21(DE3)

✓ **BL21(DE3)pLysS/E**

BL21



Expresní kmen obsahující plazmidy pLysS nebo pLysE umožní účinnou regulaci expresního systému využívající T7 promotor. Tyto plazmidy kódují lysozym, který se váže na T7 RNA polymerasu a inaktivuje ji a snižuje tak bazální expresi.

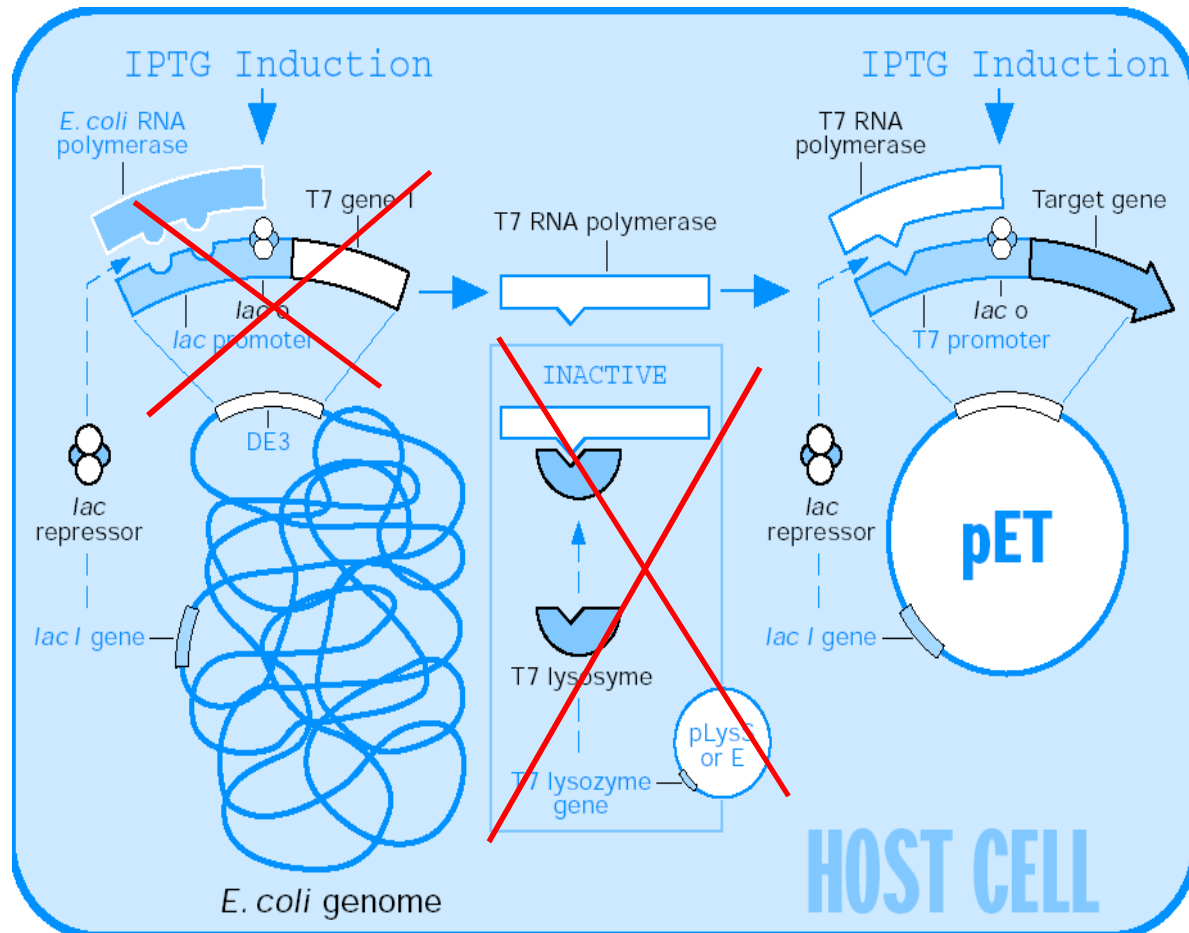
Cca 1-3% hladina bazální (před indukci exprese) exprese klonovaného genu.

# R zné úrovn minimalizace bazální exprese

BL21(DE3)

BL21(DE3)pLysS/E

✓ **BL21**



ÉIndukce exprese infekcí bakteriofágem CEG (gen pro T7 RNA polymerasu)

Nejvyší úroveň exprese!!

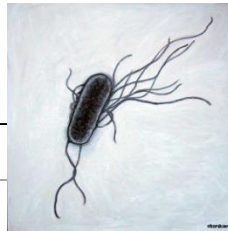


# Využívání kodon *E.coli* (codon usage)

ÉGeny u prokaryot a eukaryot se vyzna ují nenáhodným využíváním synonymních kodon .

ÉKodony málo využívané u *E. Coli* se mohou hojn vyskytovat u heterologních gen pocházejících z eukaryot, archebakterií aj.

ÉFrekvence využití synonymních kodon obvykle odráží zastoupení jejich tRNA v cytoplazm .



<i>Escherichia coli</i> K12 [gbbet]: 14 CDS's (5122 codons)								<i>Arabidopsis thaliana</i> [gbpln]: 80395 CDS's (31098475 codons)							
fields: [triplet] [frequency: per thousand] ([number])								fields: [triplet] [frequency: per thousand] ([number])							
UUU 19.7 ( 101)	UCU 5.7 ( 29)	UAU 16.8 ( 86)	UGU 5.9 ( 30)	UUC 15.0 ( 77)	UCC 5.5 ( 28)	UAC 14.6 ( 75)	UGC 8.0 ( 41)	UUU 21.8 (678320)	UCU 25.2 (782818)	UAU 14.6 (455089)	UGU 10.5 (327640)	UUC 20.7 (642407)	UCC 11.2 (348173)	UAC 13.7 (427132)	UGC 7.2 (222769)
UUA 15.2 ( 78)	UCA 7.8 ( 40)	UAA 1.8 ( 9)	UGA 1.0 ( 5)	UUA 12.7 (394867)	UCA 18.3 (568570)	UAA 0.9 ( 29405)	UGA 1.2 ( 36260)	UUA 12.7 (394867)	UCA 18.3 (568570)	UAA 0.9 ( 29405)	UGA 1.2 ( 36260)	UUA 12.7 (394867)	UCA 18.3 (568570)	UAA 0.9 ( 29405)	UGA 1.2 ( 36260)
UUG 11.9 ( 61)	UCG 8.0 ( 41)	UAG 0.0 ( 0)	UGG 10.7 ( 55)	UUG 20.9 (649150)	UCG 9.3 (290158)	UAG 0.5 ( 16417)	UGG 12.5 (388049)	UUG 20.9 (649150)	UCG 9.3 (290158)	UAG 0.5 ( 16417)	UGG 12.5 (388049)	UUG 20.9 (649150)	UCG 9.3 (290158)	UAG 0.5 ( 16417)	UGG 12.5 (388049)
CUU 11.9 ( 61)	CCU 8.4 ( 43)	CAU 15.8 ( 81)	CGU 21.1 ( 108)	CUU 10.5 ( 54)	CCC 6.4 ( 33)	CAC 13.1 ( 67)	CGC 26.0 ( 133)	CUU 24.1 (750114)	CCU 18.7 (580962)	CAU 13.8 (428694)	CGU 9.0 (280392)	CUU 16.1 (500524)	CCC 5.3 (165252)	CAC 8.7 (271155)	CGC 3.8 (117543)
CUA 5.3 ( 27)	CCA 6.6 ( 34)	CAA 12.1 ( 62)	CGA 4.3 ( 22)	CUA 9.9 (307000)	CCA 16.1 (502101)	CAA 19.4 (604800)	CGA 6.3 (195736)	CUA 9.9 (307000)	CCA 16.1 (502101)	CAA 19.4 (604800)	CGA 6.3 (195736)	CUA 9.9 (307000)	CCA 16.1 (502101)	CAA 19.4 (604800)	CGA 6.3 (195736)
CUG 46.9 ( 240)	CCG 26.7 ( 137)	CAG 27.7 ( 142)	CGG 4.1 ( 21)	CUG 9.8 (305822)	CCG 8.6 (268115)	CAG 15.2 (473809)	CGG 4.9 (151572)	CUG 9.8 (305822)	CCG 8.6 (268115)	CAG 15.2 (473809)	CGG 4.9 (151572)	CUG 9.8 (305822)	CCG 8.6 (268115)	CAG 15.2 (473809)	CGG 4.9 (151572)
AUU 30.5 ( 156)	ACU 8.0 ( 41)	AAU 21.9 ( 112)	AGU 7.2 ( 37)	AUU 21.5 (668227)	ACU 17.5 (544807)	AAU 22.3 (693344)	AGU 14.0 (435738)	AUU 21.5 (668227)	ACU 17.5 (544807)	AAU 22.3 (693344)	AGU 14.0 (435738)	AUU 21.5 (668227)	ACU 17.5 (544807)	AAU 22.3 (693344)	AGU 14.0 (435738)
AUC 18.2 ( 93)	ACC 22.8 ( 117)	AAC 24.4 ( 125)	AGC 16.6 ( 85)	AUC 18.5 (576287)	ACC 10.3 (321640)	AAC 20.9 (650826)	AGC 11.3 (352568)	AUC 18.5 (576287)	ACC 10.3 (321640)	AAC 20.9 (650826)	AGC 11.3 (352568)	AUC 18.5 (576287)	ACC 10.3 (321640)	AAC 20.9 (650826)	AGC 11.3 (352568)
AUA 3.7 ( 19)	ACA 6.4 ( 33)	AAA 33.2 ( 170)	AGA 1.4 ( 7)	AUA 12.6 (391867)	ACA 15.7 (487161)	AAA 30.8 (957374)	AGA 19.0 (589788)	AUA 12.6 (391867)	ACA 15.7 (487161)	AAA 30.8 (957374)	AGA 19.0 (589788)	AUA 12.6 (391867)	ACA 15.7 (487161)	AAA 30.8 (957374)	AGA 19.0 (589788)
AUG 24.8 ( 127)	ACG 11.5 ( 59)	AAG 12.1 ( 62)	AGG 1.6 ( 8)	AUG 24.5 (762852)	ACG 7.7 (240652)	AAG 32.7 (1016176)	AGG 11.0 (340922)	AUG 24.5 (762852)	ACG 7.7 (240652)	AAG 32.7 (1016176)	AGG 11.0 (340922)	AUG 24.5 (762852)	ACG 7.7 (240652)	AAG 32.7 (1016176)	AGG 11.0 (340922)
GUU 16.8 ( 86)	GCU 10.7 ( 55)	GAU 37.9 ( 194)	GGU 21.3 ( 109)	GUU 27.2 (847061)	GCU 28.3 (880808)	GAU 36.6 (1139637)	GGU 22.2 (689891)	GUU 27.2 (847061)	GCU 28.3 (880808)	GAU 36.6 (1139637)	GGU 22.2 (689891)	GUU 27.2 (847061)	GCU 28.3 (880808)	GAU 36.6 (1139637)	GGU 22.2 (689891)
GUC 11.7 ( 60)	GCC 31.6 ( 162)	GAC 20.5 ( 105)	GGC 33.4 ( 171)	GUC 12.8 (397008)	GCC 10.3 (321500)	GAC 17.2 (535668)	GGC 9.2 (284681)	GUC 12.8 (397008)	GCC 10.3 (321500)	GAC 17.2 (535668)	GGC 9.2 (284681)	GUC 12.8 (397008)	GCC 10.3 (321500)	GAC 17.2 (535668)	GGC 9.2 (284681)
GUA 11.5 ( 59)	GCA 21.1 ( 108)	GAA 43.7 ( 224)	GGA 9.2 ( 47)	GUA 9.9 (308605)	GCA 17.5 (543180)	GAA 34.3 (1068012)	GGA 24.2 (751489)	GUA 9.9 (308605)	GCA 17.5 (543180)	GAA 34.3 (1068012)	GGA 24.2 (751489)	GUA 9.9 (308605)	GCA 17.5 (543180)	GAA 34.3 (1068012)	GGA 24.2 (751489)
GUG 26.4 ( 135)	GCG 38.5 ( 197)	GAG 18.4 ( 94)	GGG 8.6 ( 44)	GUG 17.4 (539873)	GCG 9.0 (280804)	GAG 32.2 (1002594)	GGG 10.2 (316620)	GUG 17.4 (539873)	GCG 9.0 (280804)	GAG 32.2 (1002594)	GGG 10.2 (316620)	GUG 17.4 (539873)	GCG 9.0 (280804)	GAG 32.2 (1002594)	GGG 10.2 (316620)
Coding GC 52.35% 1st letter GC 60.82% 2nd letter GC 40.61% 3rd letter GC 55.62%								Coding GC 44.59% 1st letter GC 50.84% 2nd letter GC 40.54% 3rd letter GC 42.38%							

<http://www.kazusa.or.jp/codon/>

## Málo preferované kodony u *E.coli*

Codon(s)	Amino acid
AGA, AGG, CGA, CGG.....	Arg
UGU, UGC.....	Cys
GGA, GGG.....	Gly
AUA.....	Ile
CUA, CUC.....	Leu
CCC, CCU, CCA.....	Pro
UCA, AGU, UCG, UCC .....	Ser
ACA .....	Thr

Makrides, 1996

Expres heterologní gen obsahující málo preferované kodony vede k translačním chybám !

ÉP ed asné ukonění translace (zkrácený produkt)

ÉPosunutí tečího rámce (posun aflu 2 AK v místě kodonu AGA)

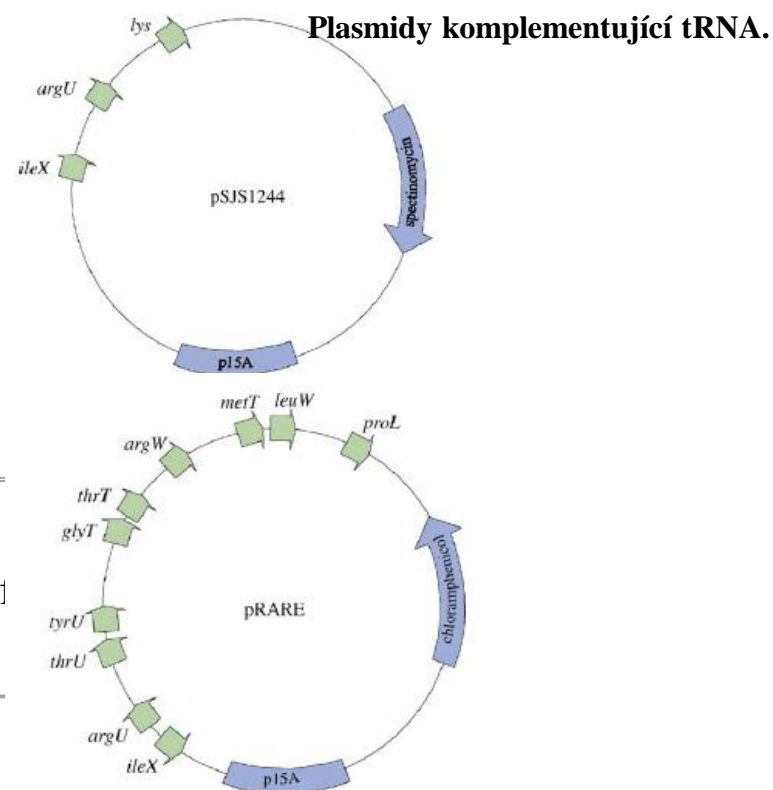
ÉZámna aminokyseliny - místo arginin (kodon AGA) za lyzin

## Výběr hostitelského kmene *E.coli*



Pro produkci genů obsahující velké množství kodonů, které jsou málo využívané *E.coli*, je možné použít komerčně dodávané kmeny, které produkují tRNA málo využívaných kodonů.

<b>BL21 (DE3) CodonPlus-RIL</b>	<b>AGG/AGA</b> (arginine, <b>R</b> ), <b>AUA</b> (isoleucine, <b>I</b> ) and <b>CUA</b> (leucine, <b>L</b> )
<b>BL21 (DE3) CodonPlus-RP</b>	<b>AGG/AGA</b> (arginine, <b>R</b> ) and <b>CCC</b> (proline, <b>P</b> )
<b>Rosetta or Rosetta (DE3)</b>	<b>AGG/AGA</b> (arginine, <b>R</b> ), <b>CGG</b> (arginine, <b>R</b> ), <b>AUA</b> (isoleucine, <b>I</b> ) <b>CUA</b> (leucine, <b>L</b> ) <b>CCC</b> (proline), and <b>GGA</b> (glycine, <b>G</b> )



NEBO: Místní řízená mutageneze - zaměřena na málo využívaného kodonu nebo syntéza optimalizovaného genu (optimalizace pro produkci v jednom nebo více organismech s ohledem na stabilitu RNA a vznik sekundárních struktur)

# Degradace proteinu

## Bakteriální proteolytický systém:

- *E. coli* obsahuje velké množství proteas, nejvíce v cytoplazm

- Selektivn odstra uje šabnormálníõ proteiny:

ÉNekompletní polypeptidy

ÉProteiny se zam n nými AK

ÉNadm rn syntetizované podjednotky multimerních protein

ÉProteiny po-kozené oxidací nebo volnými radikály

ÉCizí, rekombinantní proteiny (problémem jsou proteiny < 8 kDa)

# Aminokyseliny redukující stabilitu heterologního proteinu

## 1. N-koncové pravidlo

É Stabilita proteinu je ovlivněna aminokyselinou, která následuje první aminokyselinou polypeptidového řetězce (methionin).

É Aminokyseliny na této pozici redukují stabilitu proteinu:

**Arg, Lys, Leu, Phe, Tyr a Trp**

É Poločas rozpadu proteinu pouze 2 min

## 2. Lysin ve vnitřní sekvenci poblíž N-konce proteinu

É Degradace ubiquitin-dependentními proteasami

## 3. PEST hypotéza

É Oblasti bohaté na Pro (P), Glu (E), Ser (S), Thr (T)

É Po fosforylaci PEST degradace proteinu Ca-dependentními proteasami

## Výběr hostitelského kmene *E.coli*



### Kmeny deficientní na proteasy

ÉMutace, eliminující produkci proteas a tím proteolytickou degradaci rekombinantních proteinů .

Expresní kmeny **BL21** jsou deficientní na:

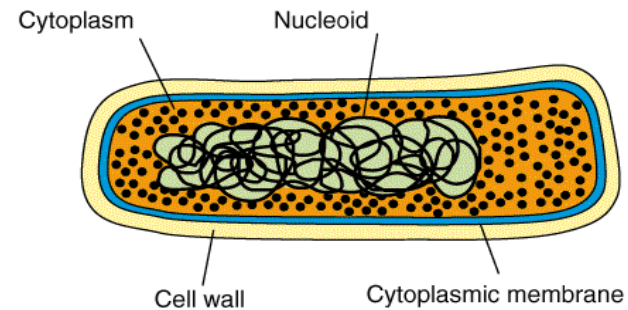
cytoplazmatickou proteasu *lon*  
periplazmatickou proteasu *ompT*

# Cílená exprese proteinu

**Cytoplazmatická exprese**

**Periplazmatická exprese**

**Extracelulární exprese  
(do kultiva ního média)**





# Cytoplazmatická exprese

ÉPreferovaný zp sob

## Výhody

ÉVysoký výt flek proteinu

ÉJednodu—í plazmidové konstrukty

ÉInkluzní t líska

## Nevýhody

ÉInkluzní t líska

ÉReduk ní prost edí

ÉProteolýza

ÉVíce komplexní purifikace

# Inkluzní tělíska

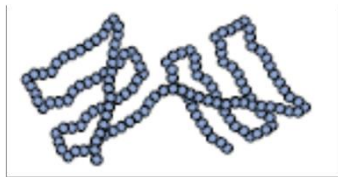
Nerozpustné shluky (cca  $2\mu\text{m}^3$ ) složené z nativních proteinů s nízkou rozpustností, z nesložených nebo z částečně poskládaných proteinů

## Co způsobuje jejich tvorbu?

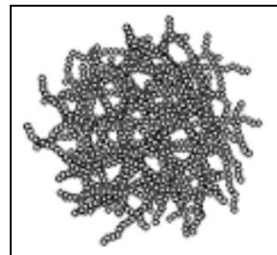
1. Prostředí v *E.coli* se liší od pirozeného prostředí ve smyslu redoxního potenciálu (redukující prostředí v cytoplazmě *E.coli*), pH, osmolarity, absence chaperonů, kofaktorů, absence posttranslačních modifikací

2. Vysoká hladina exprimovaných proteinů

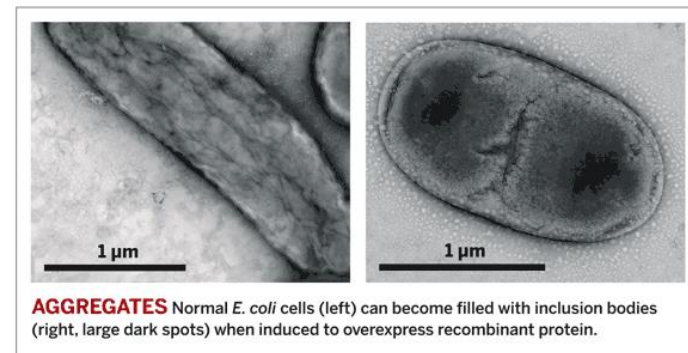
- exponované hydrofobní domény nesbalených polypeptidových řetězců navzájem (intramolekulárně) asociují



Rozpustný protein



Inkluzní tělíska z nerozpustného proteinu



# Inkluzní t líska

## Výhody

ÉSnadná izolace ve vysoké istot a koncentraci

ÉOchrana p ed proteasami

ÉPro produkci protein , jejichfl aktivita je pro bu ku letální

## Nevýhody

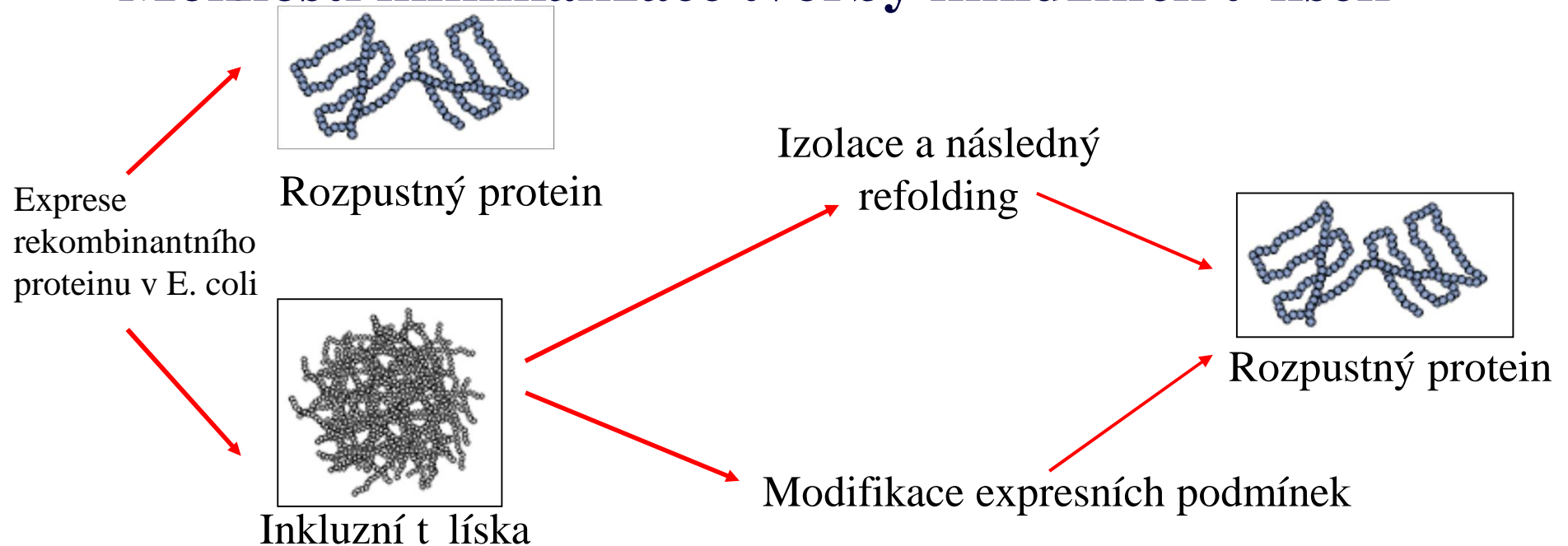
ÉProteinová nerozpustnost

ÉRefolding pro op tné získání biologické aktivity

“Refolding nemusí vést k zaktivování proteinu

“ Redukce výt flku proteinu

# Možnosti minimalizace tvorby inkluzních tělíček



Možnosti modifikace expresních podmínek:

• Snížení teploty kultivace bakteriální kultury

• Koprodukce chaperon

• Použití fúzního partnera zlepšujícího rozpustnost (thioredoxin, GST, MBP, NusA, ...)

• Výběr bakteriálního hostitele, např. kmene deficientního na thioredoxin reduktasu

• a další .

# Možnosti minimalizace tvorby inkluzních tělíček

## Výběr hostitelského kmene *E.coli*



ÉPokud protein obsahuje jeden nebo více disulfidických vazeb, správné poskládání proteinu je stimulováno v hostiteli s více oxidujícím prostředím cytoplazmy.

ÉAD494	ÉMutace v genu pro thioredoxinreduktasu ( <b>trxB</b> )
ÉOrigami	ÉMutace v genech pro thioredoxinreduktasu ( <b>trxB</b> ) and glutathionreduktasu ( <b>gor</b> )

# Periplazmatická exprese

É Periplazma obsahuje jen 4% všech buněčných proteinů (cca 100 proteinů)

É Transmembránový transport je zprostředkován signálním peptidem na N-konci proteinu

É Prokaryotické signální peptidy úspěšně použité v *E.coli* (ompA, ompT z *E.coli*, protein A z *S. Aureus*, endoglukanasa z *B.subtilis*)

“ Signální peptid je po translokaci odštěpen

## Výhody

É Jednodušší purifikace

É Není zde tak rozsáhlá proteolýza

É Oxidace probíhá v periplazmu umožňuje tvorbu disulfidických můstků /foldingu

## Nevýhody

É Signální peptid nezajistí vždy transport do periplazmy (níže výtěk)

É Mohou se také tvořit inkluzní tělíska



# Extracelulární exprese

É Sekrece proteinů do kultivačního média

É Chybí účinná cesta pro transport skrz vnější membránu (*E. coli* sekretuje velmi málo proteinů).

É Některé proteiny sekretované do periplazmy pasivně difundují do média.

É Zatím spíše neúspěšná manipulace s různými transportními cestami, které by usnadnily sekreci cizího proteinu.

## Výhody

É Minimální kontaminace ostatními proteiny (jednodušší purifikace)

É Nejmenší hladina proteolýzy

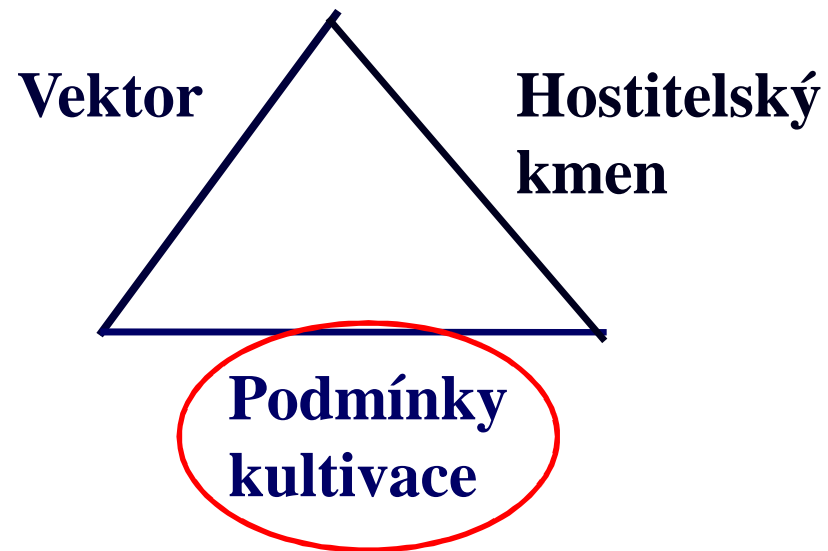
É Zlepšení foldingu

## Nevýhody

É Často velmi nízká sekrece

É Hodně ztracený protein

# Modifikace r stových podmínek



# Podmínky kultivace

Možnosti modifikace podmínek kultivace bakteriální kultury *E.coli* za účelem zvýšení produkce rozpustného proteinu:

## **Experimentálně se testuje:**

• Hodnota hustoty (OD) buněk kultury

• Složení kultivačního média (pH, přísady specifických substrátů, kofaktorů, složení živin-bohaté i minimální média)

• Teplota růstu bakterií, zejména teplota při indukci exprese

• Koncentrace indukčního induktoru

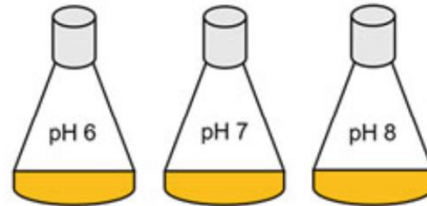
• Délka indukce exprese

# Experimentální postup pro testování exprese

Overnight culture in LB medium, 3 mL



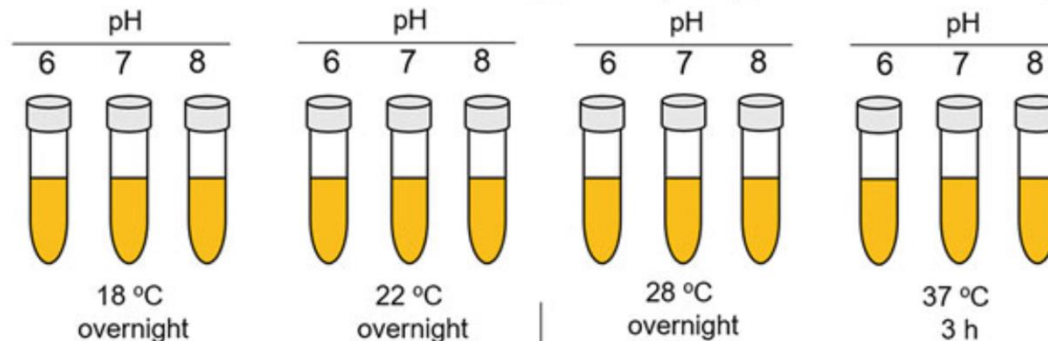
Inoculate 800  $\mu$ L into each flask containing 20 mL of TB medium.



Agitate at 37  $^{\circ}$ C until  $OD_{600} = 0.6-0.8$ .

Sample 1 mL of non-induced culture from each flask,  
spin down, remove the supernatant and store the pellet at -20  $^{\circ}$ C.  
Induce protein expression with IPTG (20  $\mu$ L of 1 M IPTG).

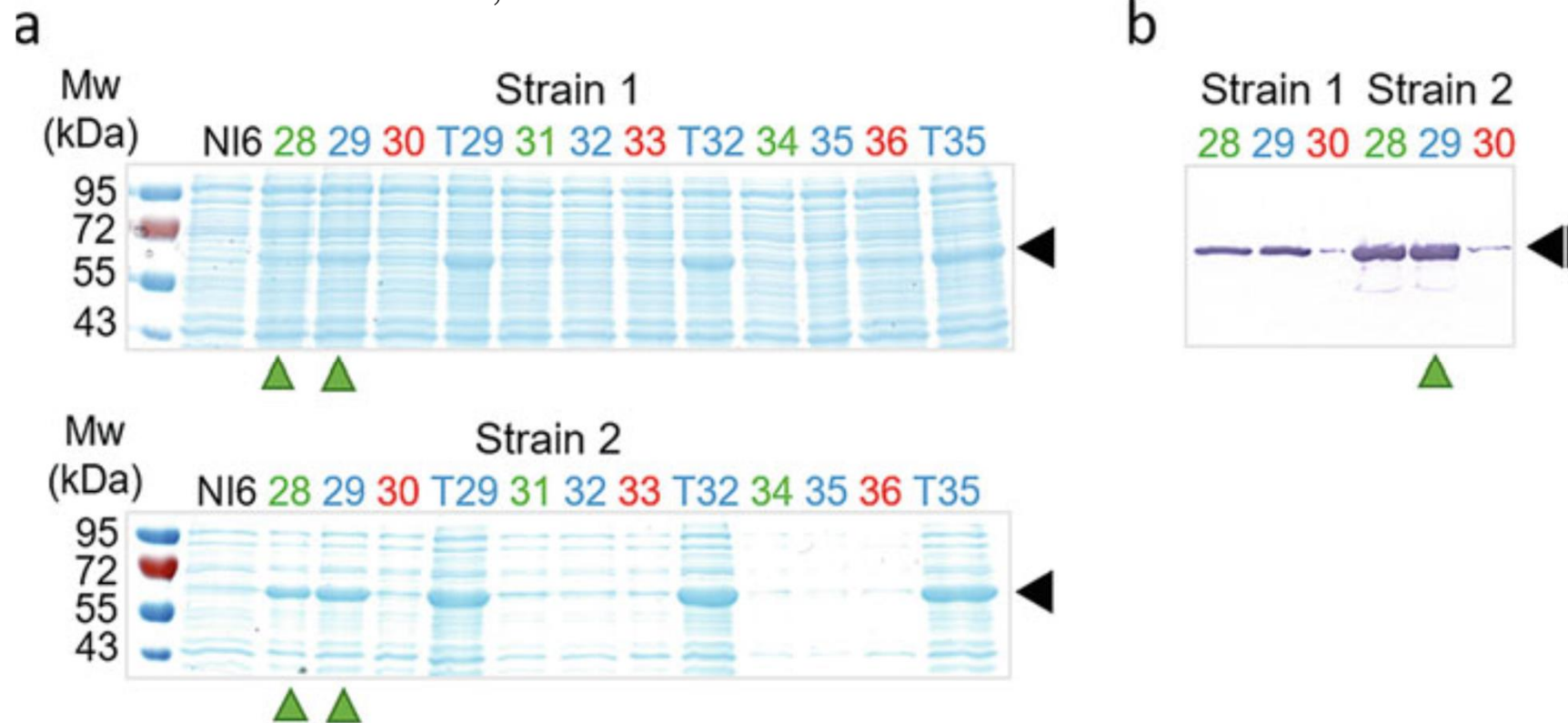
Split the content of each flask into 4 bacterial tubes (4 mL per tube) and grow them at different temperatures/times.



Sample 1 mL from each tube and place separately in TissueLyser adapter according to the scheme in Table 2.  
Centrifuge at 3220  $\times g$  at 4  $^{\circ}$ C for 2 minutes, remove supernatant with aspirator.

## Příklad výsledku testování exprese

Strain 1: Rosetta 2; Strain 2: ER2566



*E. coli* ER2566, TB medium pH 6.0, 3 h induction at 37°C, lysis buffer: MES pH 6.0 (**28**) or Tris-HCl pH 7.5 (**29**)

# Testování exprese AHP protein v *E. coli* -optimalizace teploty kultivace

Médium: LB médium, bakteriální kmen BL21(DE3)pLysS  
*R st* ( $OD_{600} \sim 0.5-0.6$ ) Indukce 0,4 mM IPTG/3hodiny



Procenta produkce AHP protein v rozpustné form						
<i>t</i> (°C) <i>R st</i> /indukce	AHP1	AHP2	AHP3	AHP4	AHP5	AHP6
37°C/28°C	8 %	85 %	100%	0	76 %	0
37°C/22°C	82 %	73 %	100%	0	81 %	51 %
22°C/22°C	71 %	78 %	100%	30 %	81 %	73 %



# Produkce heterologních proteinů v kvasinkách

## VÝHODY :

• Snadné genové manipulace

• Rychlý růst do vysokých hustot (fermentor), nízká cena

• Schopnost posttranslačně modifikovat produkovaný protein (glykosylace, tvorba disulfidických můstků)

• Schopnost tvořit správnou konformaci proteinu

• Schopnost extracelulárně sekretovat produkovaný protein

• Konečný produkt bez kontaminace endotoxiny (lipopolysacharidy)

## NEVÝHODY:

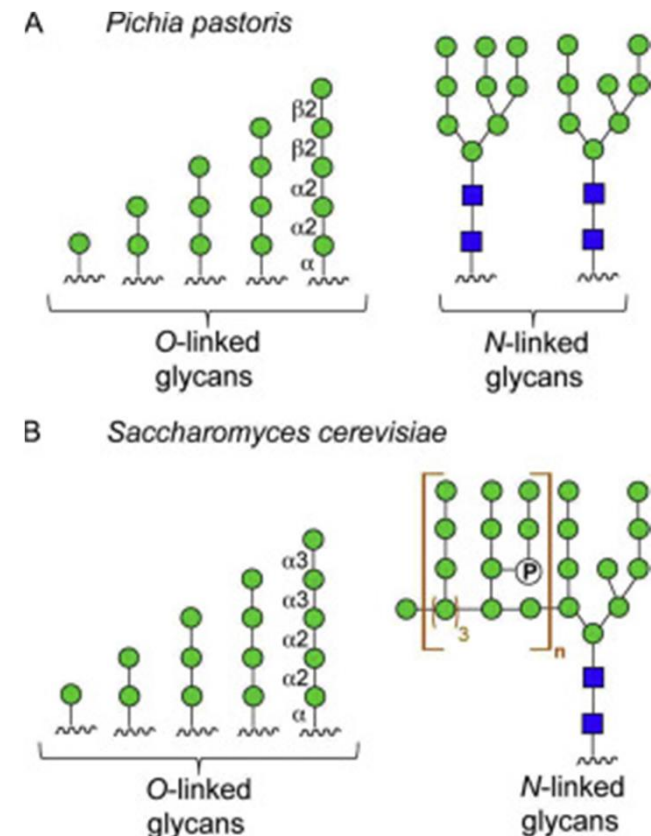
• Používají strukturálně odlišný typ N-oligosacharidů než savci

• hyperglykosylace

# Kvasinkový expresní systém

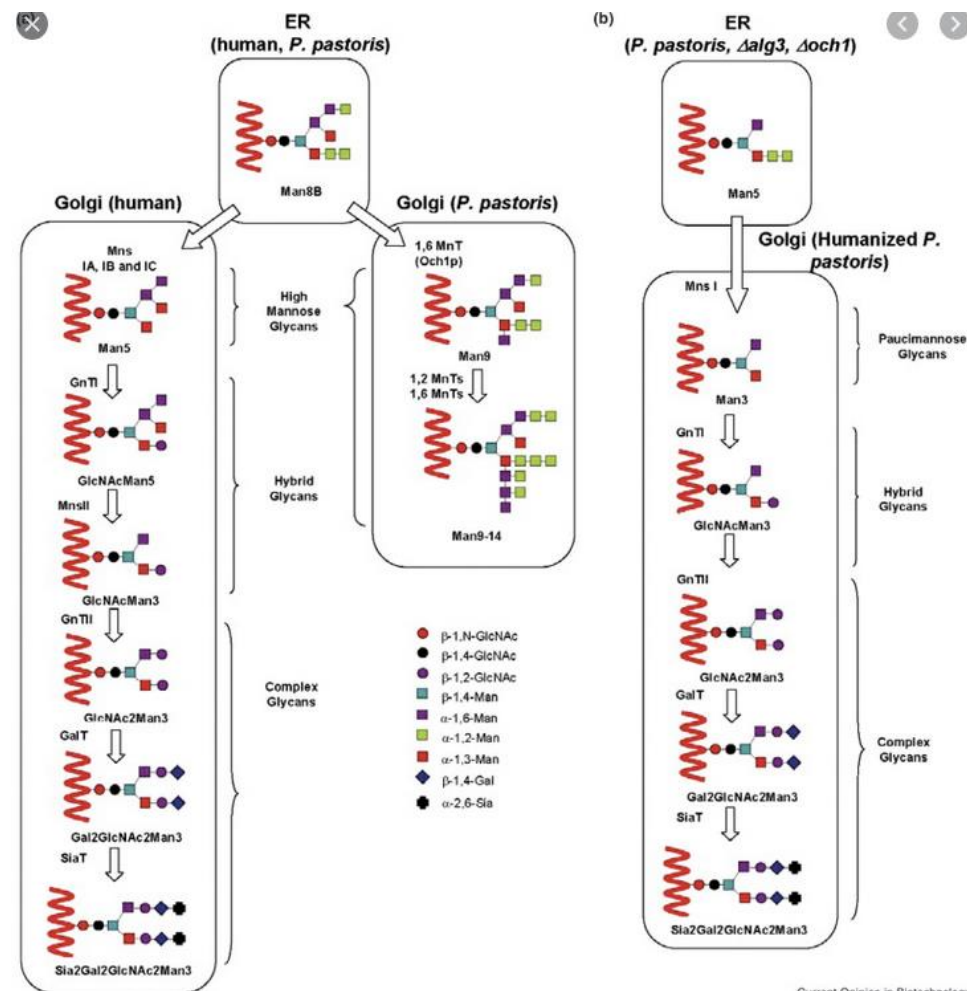
- *Sacharomyces cerevisiae* (první kvasinkový systém použitý pro produkci rekombinantních proteinů), *Pichia pastoris* (nejvyužívanější), *Yarrowia lipolytica*, *Kluyveromyces lactis*
- *P. pastoris* - schopnost kultivace do vysokých hustot (*S. cerevisiae* produkuje velké množství sekundárních metabolitů, které limitují dosažení vysoké buněčné hustoty)

- Od *S. cerevisiae* se liší typem N-glykosylace: glykany u *P. pastoris* obsahují 8-17 molekul manózy (spojené vazbou  $\alpha$  1,2), zatímco glykany *S. cerevisiae* obsahují 50-150 molekul manózy (terminální manózy spojené vazbou  $\alpha$  1,3) - tzv. hyperglykozylace (hypermanozylace)

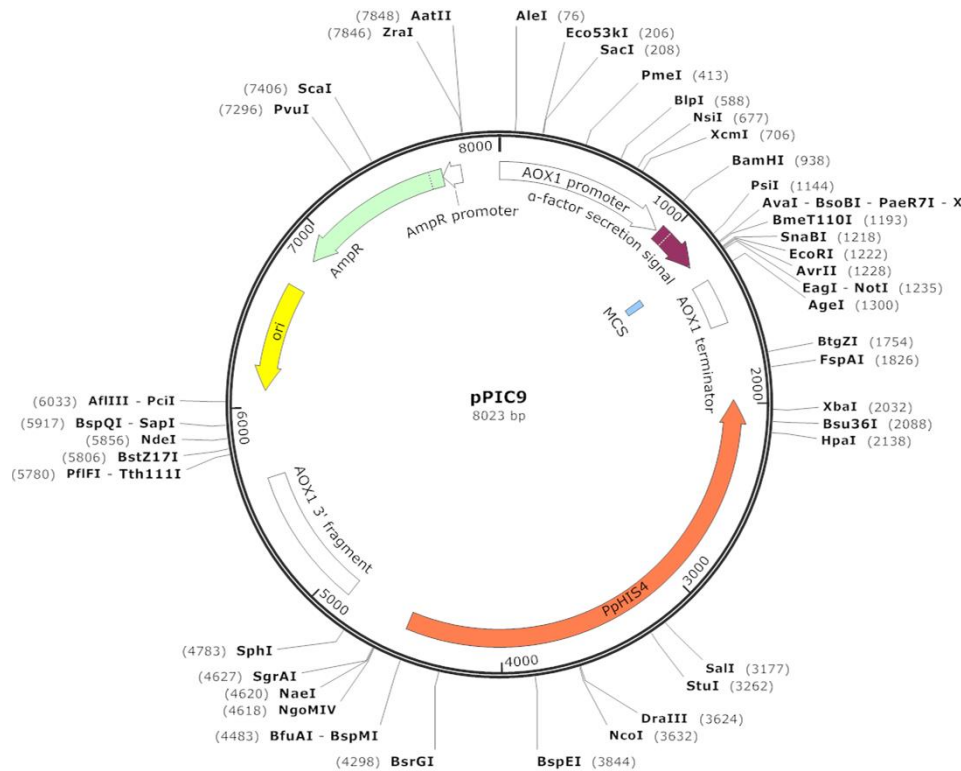


# Odlišnost glykosylace v *Pichia pastoris*

- “ vysoký obsah manózy je imunogenní pro lov k
- “ Savci vyvíjejí komplexnější glykanové struktury k modifikaci (N-acetylglukosamin, galaktóza, fukóza, kyselina sialová)
- “ pro produkce terapeutických proteinů byly vytvořeny *P. pastoris* (např. kmen SuperMan5 a další) s uniformním lidským způsobem N-glykozylace
- “ dodatečné enzymy: manosidáza I a II, transferáza galaktózy, N-acetylglukosamin a kyseliny sialové
- “ rozšíření použití *P. pastoris* například pro expresi protilátek



# Struktura vektoru pro ú innou expresi v *P. pastoris*



- transformace se provádí pomocí integrujících se vektor (vektory replikující se po integraci do chromozomu)

## Promotory:

### “ Inducibilní

**Promotor genu pro alkohol oxidázu 1 (AOX1) ó**  
silný a striktn regulovaný

-je pln potla en p i r stu na glukóze i glycerolu  
a siln indukován p i r stu na zdroji uhlíku- metanolu

### “ Konstitutivní

**Promotor genu pro glyceraldehyd-3 ó fosfát  
dehydrogenasu (GAP)**

- konstitutivn aktivní, aktivita p i r stu na glukose

- 20-30 % celkového intracelulárního protein  
(u obou typ; promotor )

## Selek ní markery:

### Auxotrofní *HIS4*

Komplementace auxotrofní mutace His

- selekce v kvasinkách

### Rezistence v i antibiotik m

Ampicillin ó pro selekci v bakteriích

## Sekre ní signály

PHO1 (acid phosphatase) z *P. pastoris*

$\alpha$ -mating faktor ( $\alpha$ -MF) z *S. cerevisiae*

SUC2 (invertase) z *S. cerevisiae*

# Expese rekombinantních protein v *Pichia pastoris*

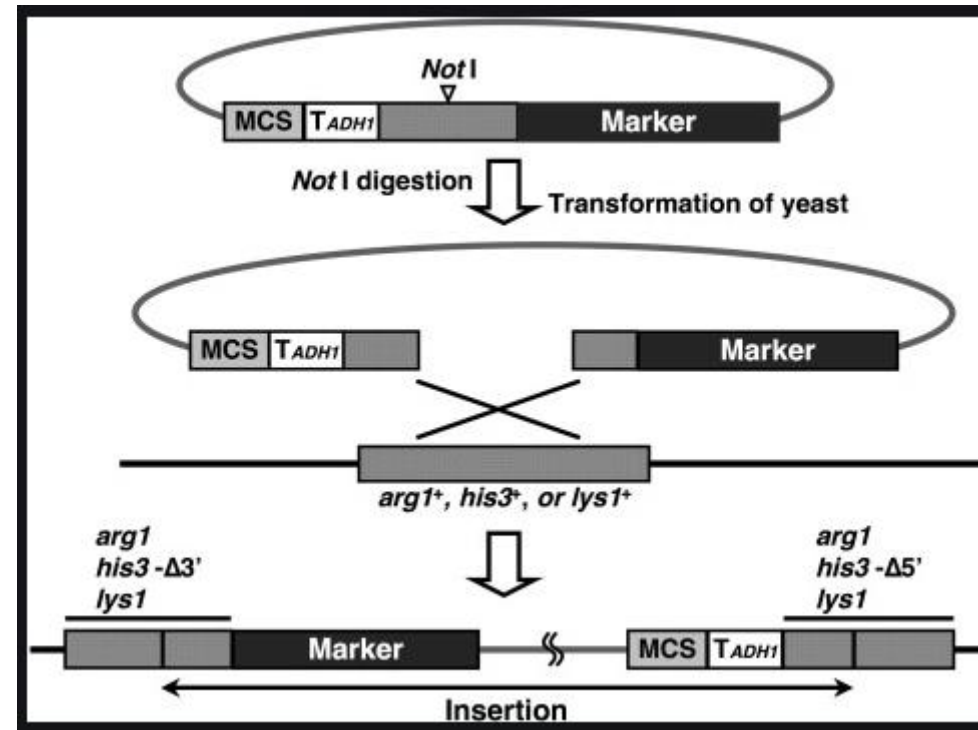
Klonování genu do vektoru, selekce konstruktů v *E.coli*, izolace DNA konstruktů

Linearizace konstruktů

Transformace kompetentních buněk *P. pastoris*

Selekce transformantů

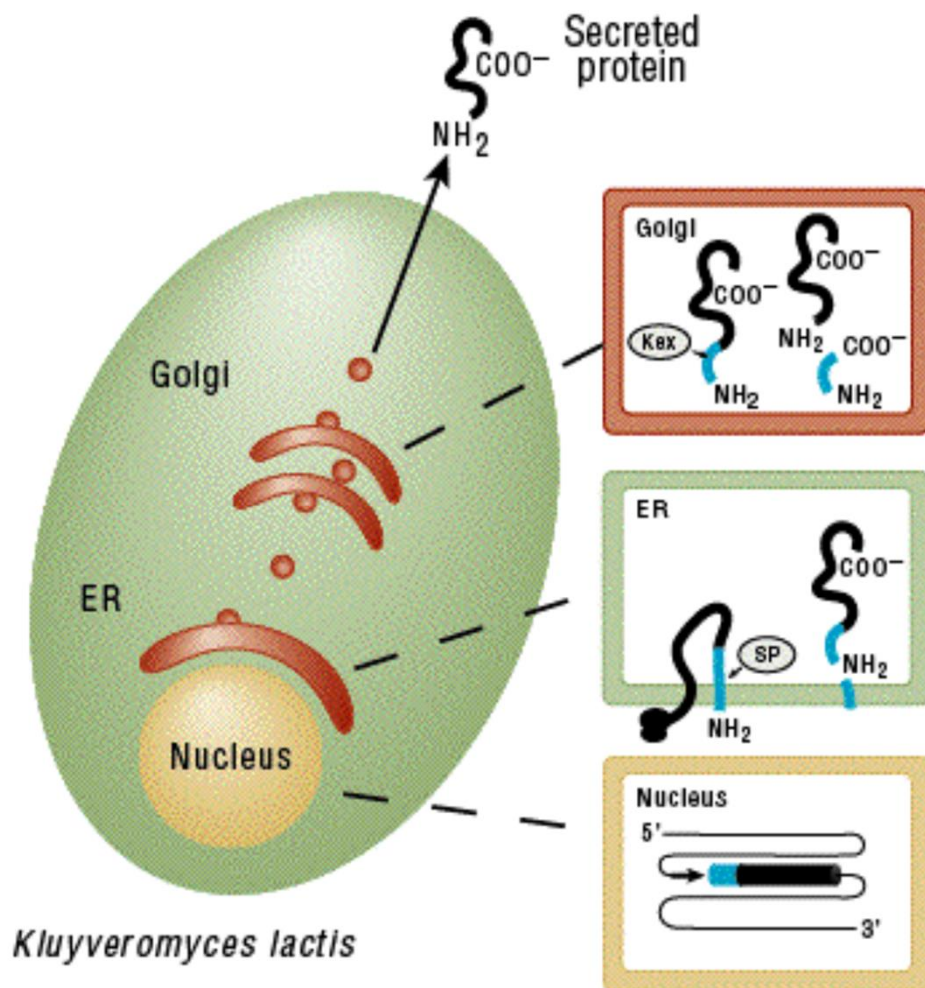
Expese rekombinantního proteinu





## Sekrece proteinů u kvasinek

” Bu intracelulární exprese nebo sekrece do média



Sekreční dráha je velmi podobná dráze vyších eukaryot.

N-terminální peptid pro kotranslační translokaci sekretovaného proteinu do ER je odštěpen signální peptidázou.

Příklady signálních sekvencí: Pho5, Suc2 a  $\alpha$ -mating factor

### Sekrece pomocí $\alpha$ -mating faktoru

Protein je společně se signálním peptidem  $\alpha$ -MF ( $\alpha$ -mating factor) doménou exprimován v jádře.

1. Signální peptid v  $\alpha$ -MF doméně navede protein do ER, kde je odštěpen signální peptidázou.
2. Fúzní protein je transportován do Golgiho aparátu, kde Kex proteáza štěpí zbytek  $\alpha$ -MF domény a protein je sekretován do média.

# **Produkce heterologních proteinů v hmyzích buňkách (Hmyzí buňky s bakulovirovými vektory)**

## **VÝHODY:**

ÉZajiť ní nativní konformace proteinu

ÉPosttranslační modifikace

ÉNení kontaminace endotoxiny (lipopolysacharidy)

ÉProtein může být produkován intracelulárně nebo sekretován do média

## **NEVÝHODY:**

“ Negativní efekt infekce bakulovirem na životnost buněk

“ Heterologní geny nejsou produkovány kontinuálně (každá exprese vyžaduje novou infekci buněk bakulovirem)

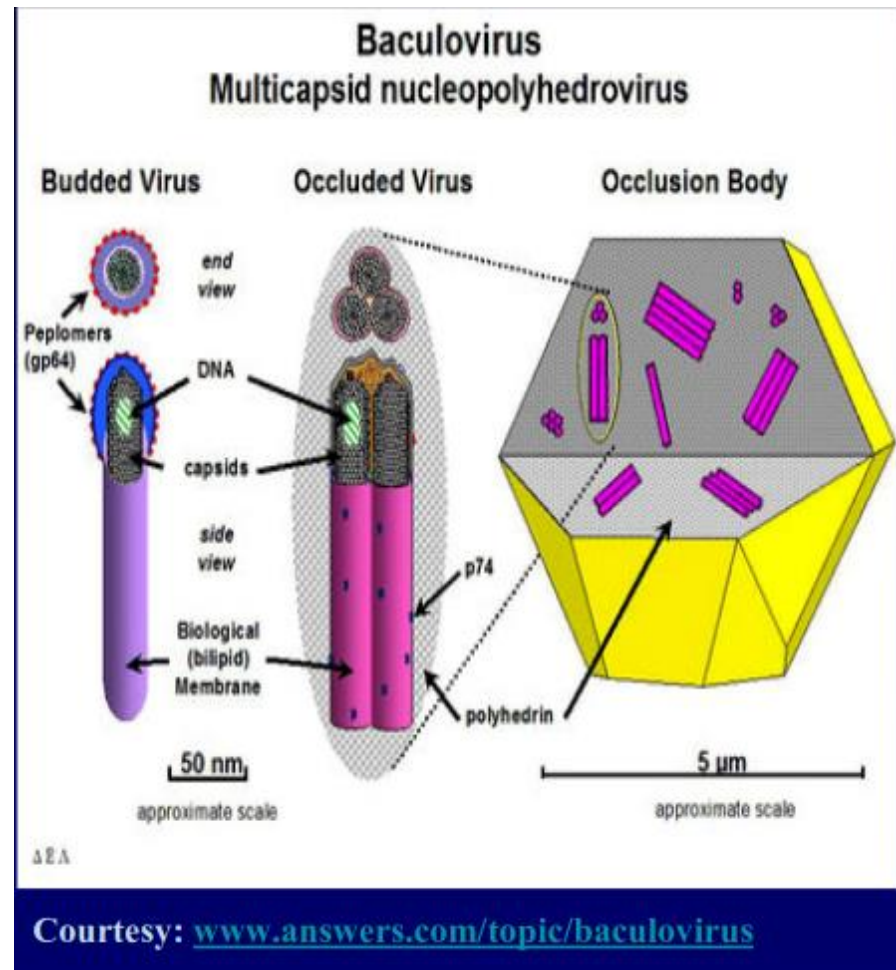
“ Způsob glykozylace odlišný od savčích buněk (není tak komplexní)

“ Nízká výťažnost, časově náročné, drahé média, potřeba na manipulaci (možnost kontaminace)



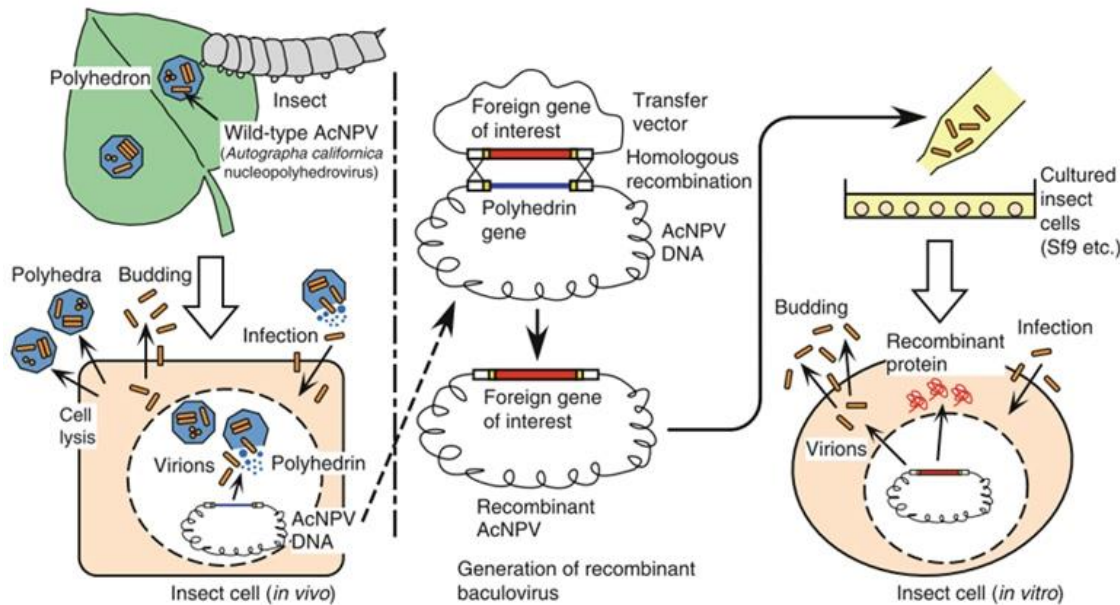
# Co je bakulovirus?

- “ obalený, ds DNA virus s kapsidou ty inkovitého tvaru
- “ b hem fivotního cyklu existuje ve dvou rozdílných formách ōpu ící virusō a opouzdený virus
- “ vysoce druhov specifické - infikují jen bezobratlý hmyz
- “ mezi nej ast j í hostitele pat í nedosp lé larvální formy hmyzu
- “ k nejvíce prostudovaným a vyuffivaným bakulovir m pat í virus AcMNPV (*Autographa californica multiple nuclear polyhedrosis virus*)



# Bakulovirový expresní systém

- založený na infekci kultivovaných hmyzích bun k rekombinantním virem AcMNPV (Autographa californica multiple nuclear polyhedrosis virus) nesoucím gen pro tvorbu cílového proteinu.

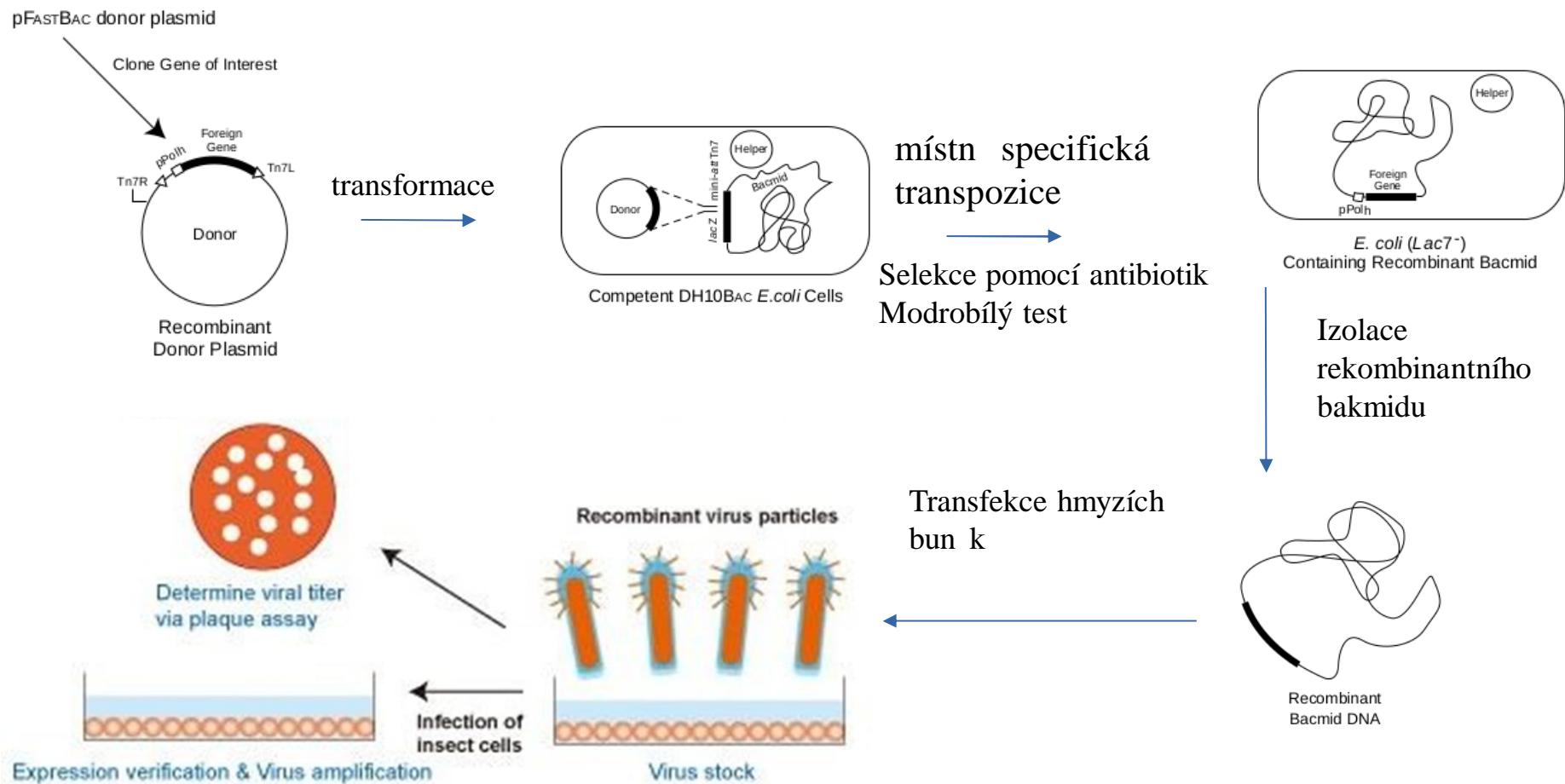


- “ hostitelské hmyzí bu ky
  - ovariální bu ky motýl druhu *Spodoptera frugiperda* (Sf9, Sf-21)
- “ bakulovirový vektor (tzv. bakmid)
  - velký, kyvadlový vektor (AcMNPV)
  - obsahuje všechny geny nezbytné k produkci virových částic
- “ velmi silný promotor polyhedrinového genu (polyh)

## Produkce heterologních proteinů v hmyzích buňkách (Hmyzí buňky s bakulovirovými vektory)

- “ Hlavní technologie pro přípravu membránových proteinů, proteinových komplexů, složených z více podjednotek.
- “ Bakulovirus je neprokatelný, pokud jde o velikost šnázkladu DNA, který musí nést a přenášet do hmyzích buněk (afl 17 podjednotek multiproteinového komplexu bylo úspěšně koexprimováno v hmyzích buňkách).
- “ Nové technologie molekulárního klonování jako Zlatá Gate (GoldenBac), Gibsonovo sestavení PCR fragmentů (biGBac), Cre-lox rekombinace (MultiBac) umožňují efektivní multigenové sestavení v genomu bakuloviru.

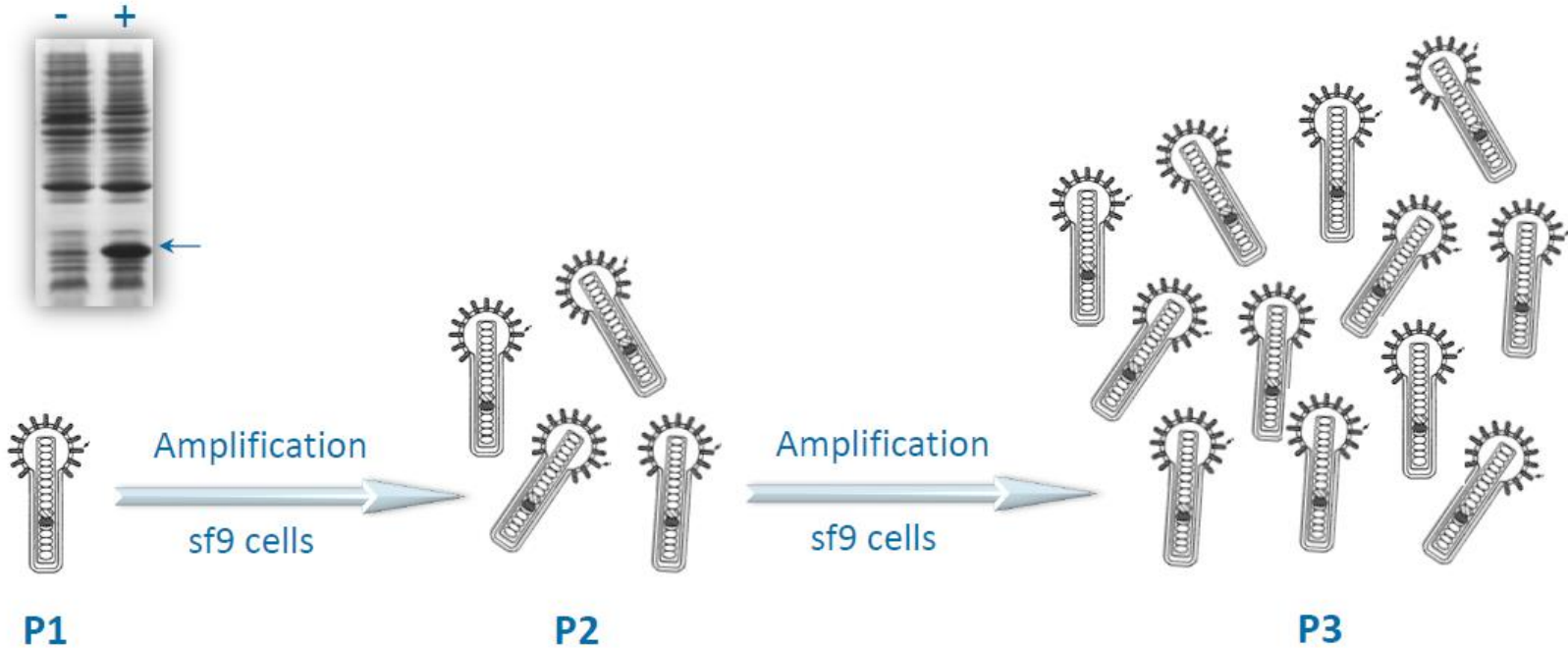
# Bac-to-Bac expresní systém : šfrom Bacterium to Baculovirusů



“ Po úspěšné transfekci, rekombinantní virové částice jsou získány z média (virový zásobní stock) a mohou být amplifikovány do vyšších množství, aby mohly být použity pro infekci velkého množství bun k.

# Bac-to-Bac expresní systém: šfrom Bacterium to Baculovirusõ

Expression evaluation



Po dvou generacích je dosaženo titru  $10^6 - 10^8$  pfu/ml





# WORST

# BEST

SPEED



Transgenics



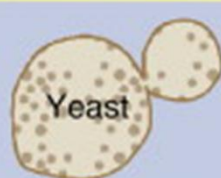
Plants



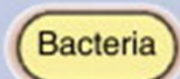
Mammalian



Insect



Yeast



Bacteria

COST



Transgenics



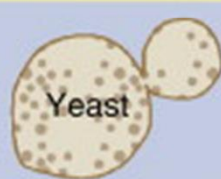
Mammalian



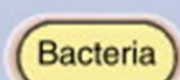
Plants



Insect

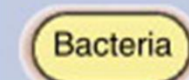


Yeast

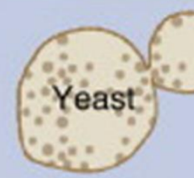


Bacteria

GLYCO-SYLATION



Bacteria



Yeast



Plants



Insect

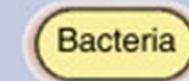


Transgenics

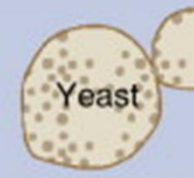


Mammalian

FOLDING



Bacteria



Yeast



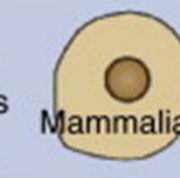
Plants



Insect



Transgenics



Mammalian

GOVERNMENT REGULATION



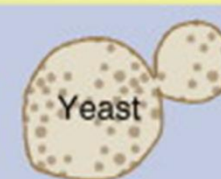
Transgenics



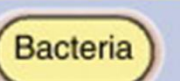
Plants



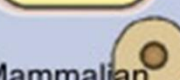
Insect



Yeast



Bacteria



Mammalian