

Zajištění exprese klonovaných genů a její optimalizace

Faktory ovlivňující expresi klonovaných genů

A. Regulační sekvence pro genovou expresi

1. *Transkripce*

- Síla promotoru
- Terminátor transkripce
- Stabilita mRNA

2. *Translace*

- struktura vazebného místa pro ribozom
- využívání kodonů
- stabilita proteinu

3. *Transport proteinu*

- charakter signální sekvence

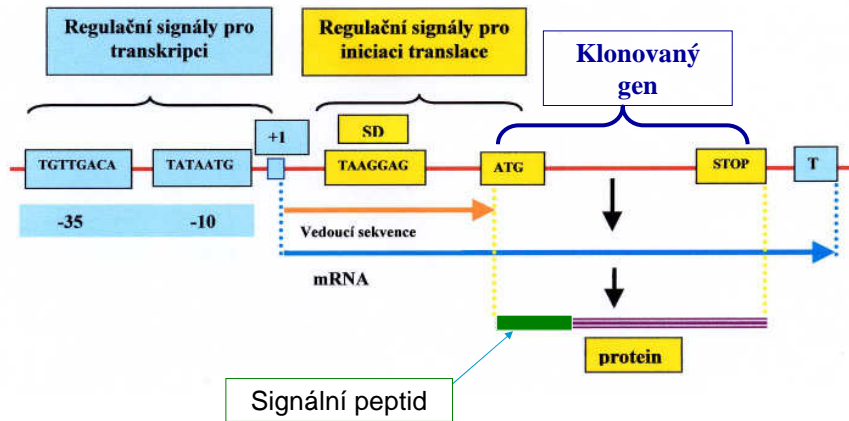
B. Vlastnosti vektorů

- Počet kopií vektoru
- Stabilita vektoru

C. Fyziologie hostitelské buňky

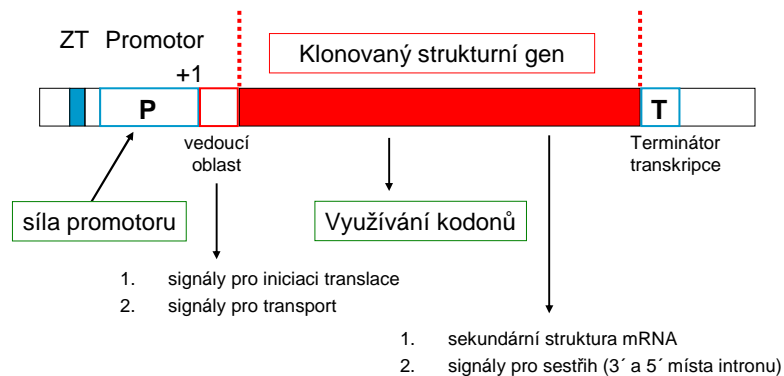
- růstové podmínky
- enzymový aparát hostitelské buňky

Signály ovlivňující transkripci a translaci strukturního genu (bakterie *E. coli*)

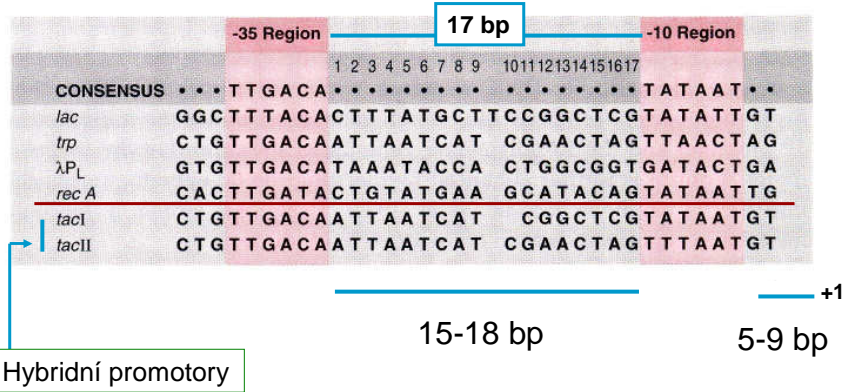


Faktory ovlivňující expresi klonované DNA

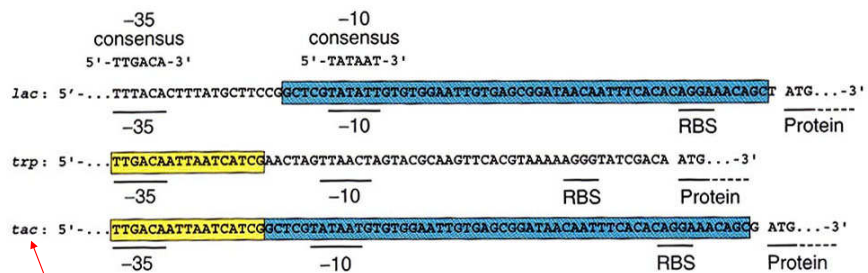
1. Regulační sekvence vektoru, vlastnosti inzertu



Sekvence přirozených a hybridních promotorů



Sekvence lac, trp a tac promotoru



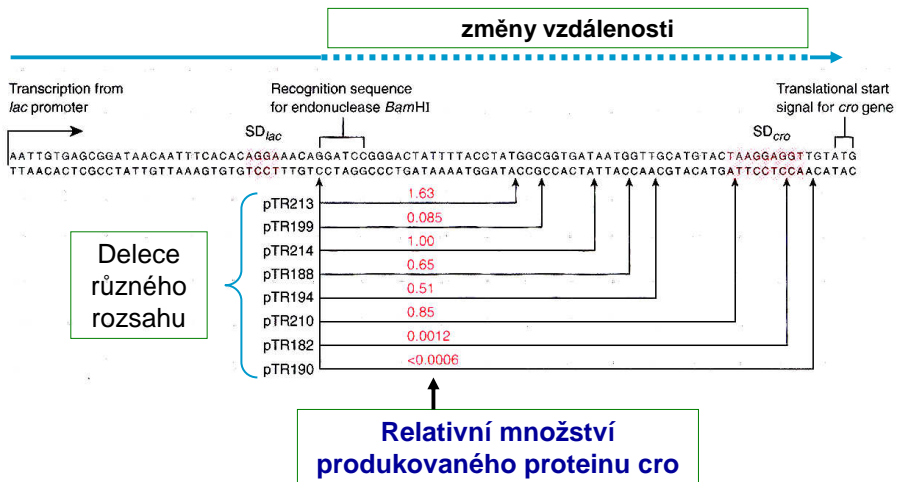
Regulovatelné promotory používané v expresních vektorech

	Promotor	Zdroj	Způsob regulace	
<i>E. coli</i>	λ pL, λ pR	Leftward and rightward early promoters of λ	Off 30°C	On >37°C (in <i>c₁₈₅₇</i> host)
lac-UV5 nezávislost na katabolické represi	<i>lac</i>	<i>E. coli lac</i> operon	—	IPTG in medium
	<i>trp</i>	<i>E. coli trp</i> operon	Tryptophan in medium	Indoleacetic acid in medium
	<i>tac</i>	<i>trp</i> -35 region <i>lac</i> -10 region hybrid	—	IPTG in medium
	<i>phoA</i>	<i>E. coli</i> alkaline phosphatase operon	Excess phosphate in medium	Phosphate-limited medium
	<i>recA</i>	<i>E. coli recA</i> gene	—	Mitomycin C in medium
	<i>tet</i>	Tn10 tetracycline-resistance gene	—	Tetracyclines in medium

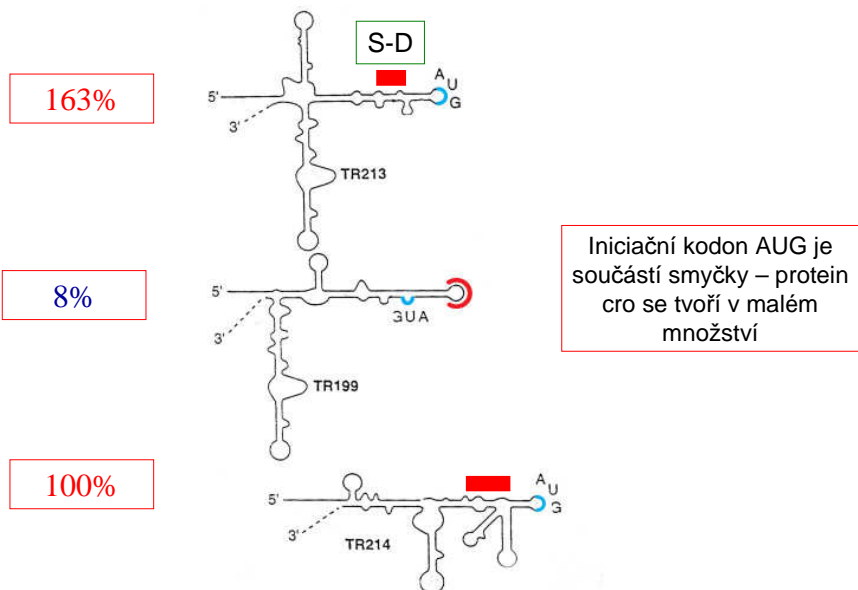
Promotory fága lambda vykazují velmi striktní kontrolu transkripce s krátkým lagem mezi neindukovaným a indukovaným stavem. V neindukovaném stavu nedochází vůbec k transkripci, na rozdíl od promotorů „metabolických“ operonů, kdy částečná transkripce probíhá pořád.

	Promotor	Zdroj	Způsob regulace	
<i>B. subtilis</i>	<i>bla</i>	<i>Bacillus licheniformis</i> β -lactamase gene	Off —	On β -lactams in medium
	<i>cat</i>	<i>Bacillus pumilus</i> chloramphenicol acetyl transferase	—	Chloramphenicol in medium
<i>Streptomyces</i>	<i>gyl</i>	<i>Streptomyces coelicolor</i> glycerol operon	Glucose in medium	Glycerol in medium
<i>S. cerevisiae</i>	<i>ADH</i>	Yeast repressible alcohol dehydrogenase (<i>ADR</i>) gene	High glucose in medium	Low glucose in medium
	<i>GAL1</i>	Yeast galactose utilisation operon	Glucose in medium	Galactose in medium
	<i>GPD-PH05</i>	Hybrid between yeast glyceraldehyde 3-phosphate dehydrogenase and alkaline phosphatase gene promoters	Excess phosphate in medium	Phosphate-limited medium

Vliv vzdálenosti mezi promotorem a startem translace na množství vytvářeného proteinu cro



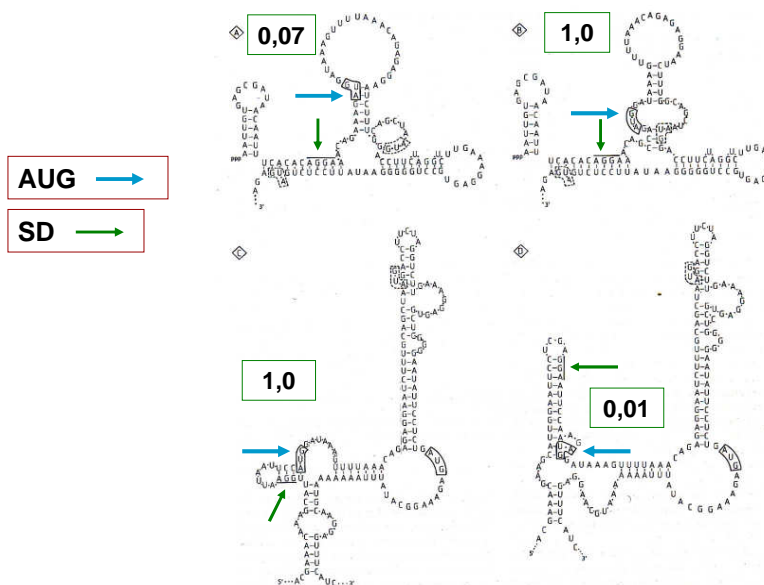
Sekundární struktura cro-mRNA



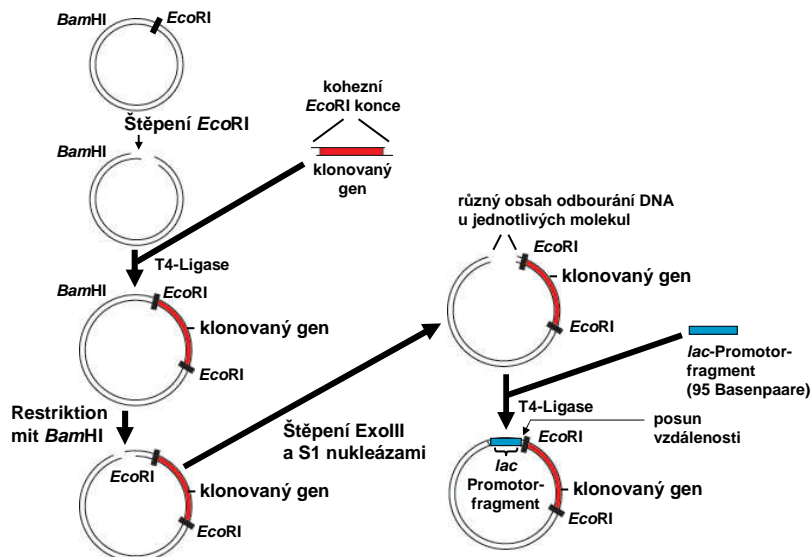
Vliv regulačních sekvencí na výtěžek t-antigenu SV40 z různých plazmidových konstruktů

% celkového proteinu	Vzdálenost SD-ATG	Sekvence oblasti pro iniciaci translace	vektor
0.068	9	AGGAACAGAAAGATGGAT	pTR436
1.0	9	AGGAACAGCCAGATGGAT	HP1
0.01	5	GTCGAGGAATTCCATGGAT	pPLcSVt5-372
0.1	5	ATTGGAATTCCATGGAT	pPLcSVt5-37
2.5	8	TTGGAATTATTCCATGGAT	pPLcSVt5-379
1.0	9	TTGGAATTAATCCATGGAT	pPLcSVt5-374
0.01	9	AGGAATTCCAAGATGGAT	pPLcSVt5-72

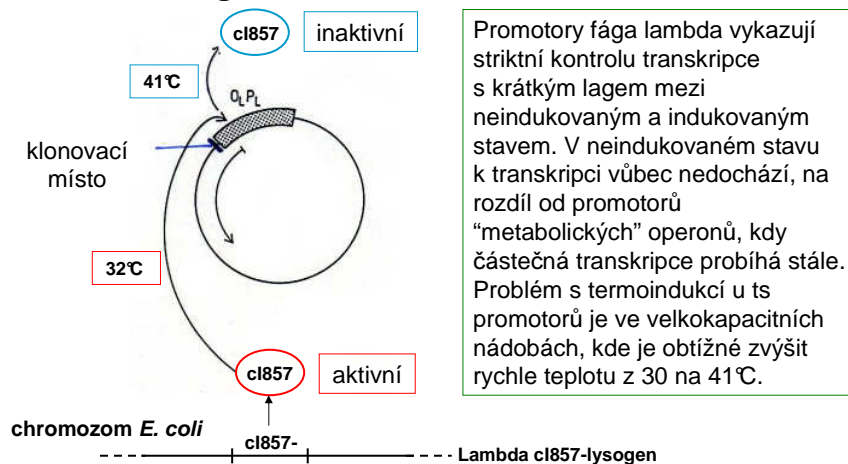
Sekundární struktura mRNA (pro t-antigen) v oblasti iniciačního kodonu v různých plazmidových konstruktech



Posun vzdálenosti mezi promotorem a klonovaným genem



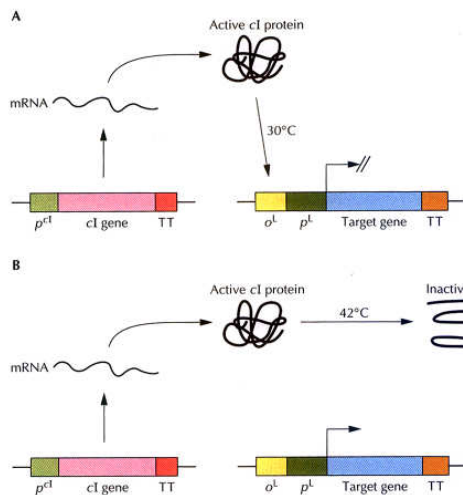
Použití teplotně senzitivního represoru pro kontrolu promotoru PL fága λ



Promotory fága lambda vykazují striktní kontrolu transkripce s krátkým lagem mezi neindukovaným a indukovaným stavem. V neindukovaném stavu k transkripci vůbec nedochází, na rozdíl od promotorů "metabolických" operonů, kdy částečná transkripce probíhá stále. Problém s termoindukcí u ts promotorů je ve velkokapacitních nádobách, kde je obtížné zvýšit rychle teplotu z 30 na 41°C.

Při 32°C se termolabilní represor *ci857* kódovaný genem *ci* na chromozomu váže na operátor O_L na plazmidu a zabraňuje transkripci z promotoru P_L . Zvýšením teploty na 41°C je represor inaktivován, uvolní se z operátoru a transkripce probíhá.

Regulace genové exprese promotorem pL fága λ

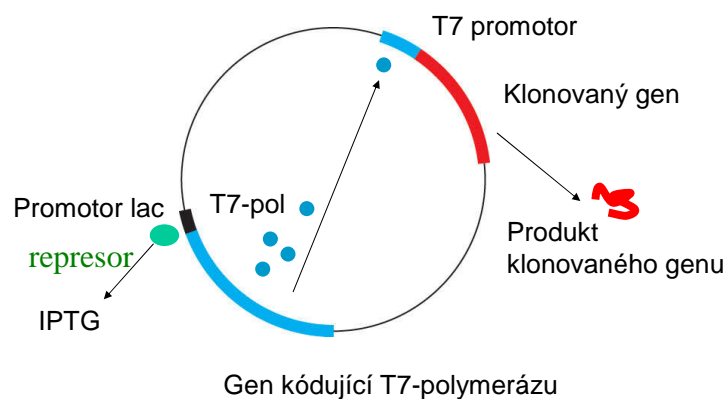


Při 30°C se represor cI, který je konstitutivně syntetizován pod kontrolou vlastního promotoru p_{cI} , váže na operátor o_L promotoru p_L a tím zabraňuje expresi cílového genu

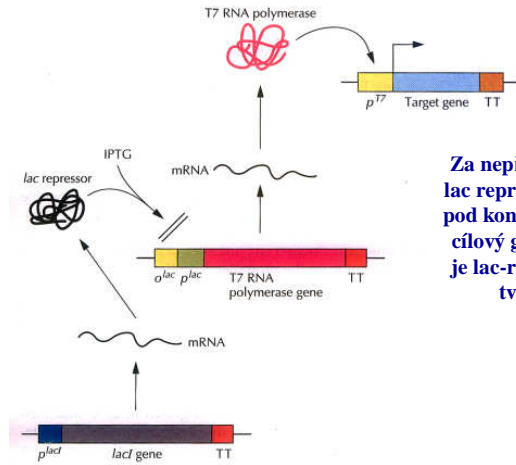
Při 42°C je ts represor cI inaktivován a dochází k transkripci cílového genu

Expresní vektory obsahující T7-promotor

RNA-polymeráza T7 rozpoznává pouze promotory genů fága T7

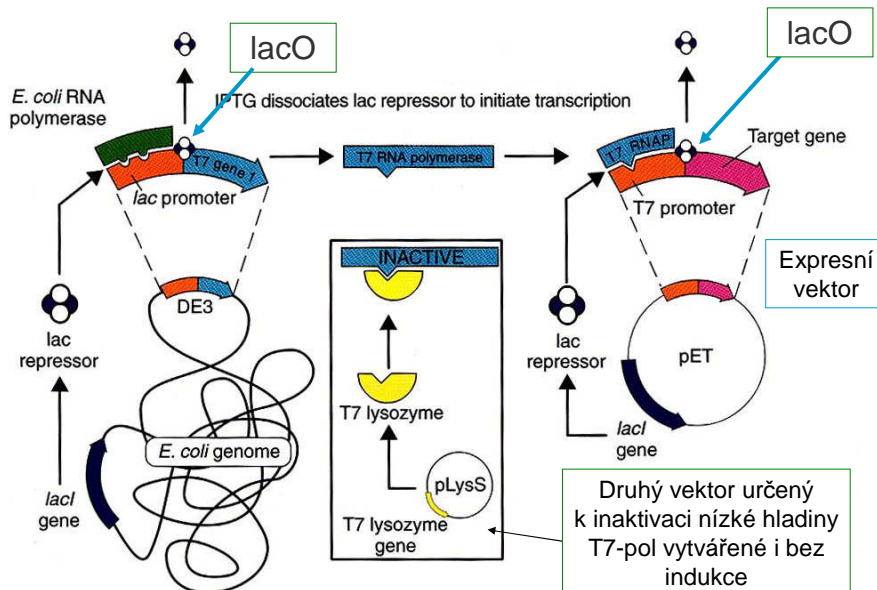


Regulace genové exprese cílového genu kontrolovanané promotorem pT7 fága T7



Za nepřítomnosti induktoru (IPTG) represor lac reprimuje syntézu T7-polymerázy, která je pod kontrolou lac promotoru a lac operátoru a cílový gen se netranskribuje. Po přidání IPTG je lac-represor inaktivován, T7-polymeáza se tvoří a je transkribován cílový gen

Systém T7 pro expresi proteinů v E. coli



Terminátory transkripce používané v expresních vektorech u *E. coli*

- a) terminátory T1 a T2 bakteriofága lambda
- b) terminátory T1 a T2 z operonu *rrnB* rRNA *E. coli*
používají se v tandemu

**Účinná terminace transkripce je esenciální pro
dosažení vysoké hladiny exprese:**

- zvyšuje stabilitu mRNA,
- zvyšuje hladinu akumulovaných proteinů.

Stabilita mRNA

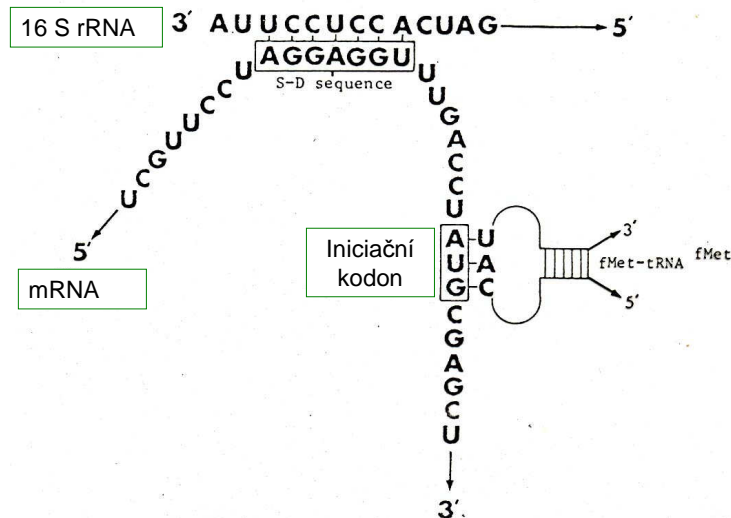
- Rychlost syntézy proteinu závisí na množství mRNA v buňce
- Existuje rovnováha mezi syntézou a rozkladem daného druhu
Mrna /turnover/

**Snížení rozkladu mRNA (kombinované působení endonukleázy
a 3' exonukleázy)**

- u *E. coli* činí poločas rozpadu molekul mRNA 1-2 minut
- poločas rozpadu mRNA genu 32 bakteriofága T4 je 20 minut a
více – za zvýšenou stabilitu jsou odpovědné specifické
sekvence, které se nacházejí před iniciačním kodonem genu
32 – díky této 5' nepřekládané sekvenci mohou být také
stabilizovány jiné mRNA molekuly.
- Konstrukce plazmidu s expresní kazetou genu 32, pomocí níž
je možné v buňkách *E. coli* syntetizovat velká množství cizích
proteinů. Vzniklé hybridní transkripty mají dlouhou životnost.
Poločas rozpadu se podle klonované sekvence pohybuje od 4
do 10 minut.

Zajištění účinné translace




Interakce mRNA s 16S rRNA při iniciaci translace

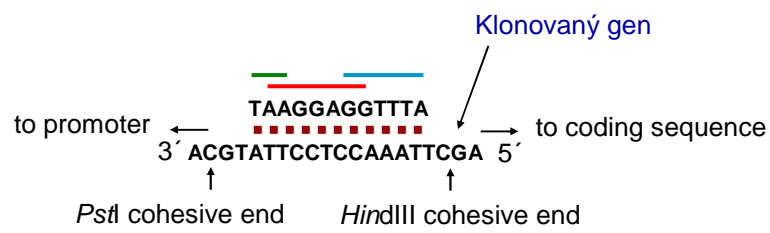


Kromě iniciačního kodonu AUG existují v RBS ještě tři další oblasti, jejichž sekvence jsou více či méně konzervovány:

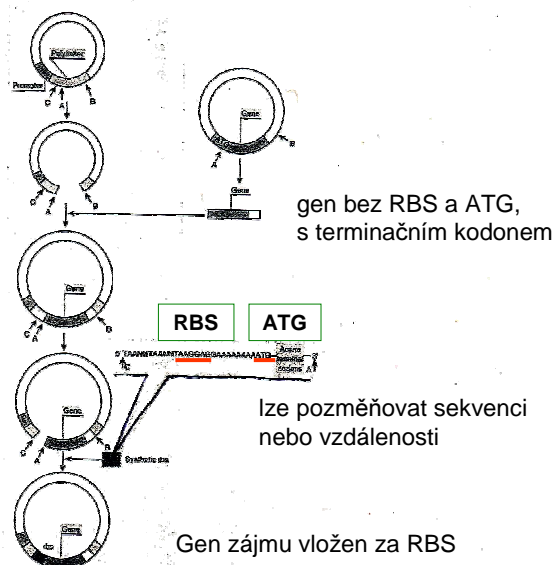
1. Shine-Dalgarnova sekvence, níž se obvykle vyskytuje sekvence 5'UAAGGAGGU 3'.
2. U mRNA (alespoň polycistronické) jsou součástí RBS jeden nebo více terminačních kodonů.
3. U genů, které jsou silně exprimovány (např. geny pro fágové kapsidy nebo ribozomální proteiny), se v RBS nachází sekvence PuPuUUUPuPu (nebo sekvence jí podobná). Bývá označována též jako RRUUURR sekvence. Může se vyskytovat vedle SD-sekvence nebo místo ní. Ukazuje se, že přítomnost této sekvence je nezbytná pro translaci eukaryotických genů v *E. coli*, jak bylo zjištěno při expresi malého T antigenu SV40.

Synteticky připravené ribozomové vazebné místo

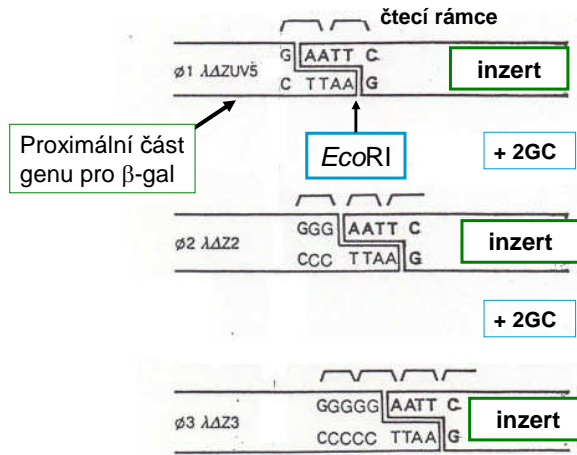
1. S-D sekvence 
2. RRUUURR (GGTTTAA) 
3. Terminační kodon TAA 



Zajištění účinné translace použitím optimalizované RBS

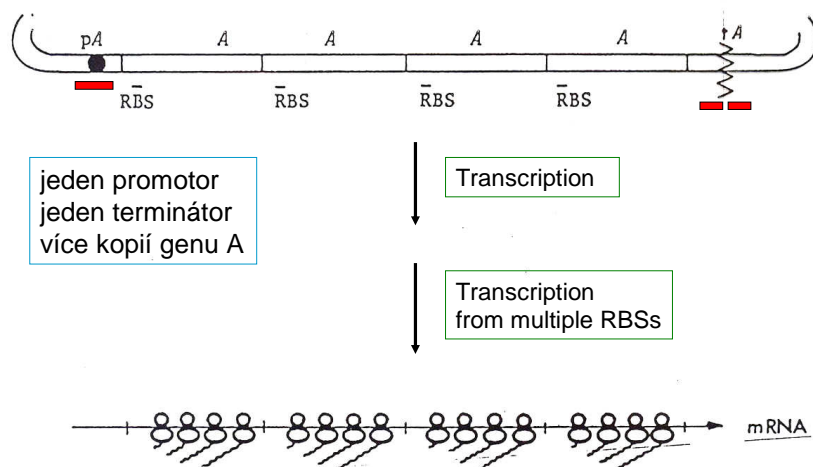


Různé čtecí rámce vzhledem k iniciaci translace genů lacZ u tří různých vektorů



V devátém kodonu genu pro β-galaktozidázu je jedinečné místo pro *EcoRI*. Čtecí rámec, který tímto *EcoRI* místem začíná, byl označen jako Φ 1. Byly připraveny λΔZ vektory se zabudovaným fragmentem v čtecích rámcích Φ 2 a Φ 3 připojením 2GC.

Zvýšení genové exprese konstrukcí homopolycistronicke sekvence



Srovnání využívání kodonů silně a slabě exprimovaných genů u *E. coli*

	U		C		A		G		
	silně/slabě		silně/slabě		silně/slabě		silně/slabě		
U	Phe	39 151	Ser	93 36	Tyr	34 96	Cys	13 34	U
		113 102		87 49		98 48		23 39	
	Leu	12 71	5 37	ochr. ochr.	opel	25 98	A		
		16 54	12 94	ochr. ochr.				G	
C	Leu	26 73	Pro	21 29	His	19 95	Arg	223 99	U
		33 68		2 48		75 59		101 133	
	→ 3 22	35 45	38 30	→ 3 27	A				
		345 294	100 101	Gln	169 169	→ 1 42	G		
A	Ile	87 158	Thr	103 48	Asn	13 101	Ser	10 58	U
		262 118		137 119		153 68		49 61	
	→ 2 27	15 32	259 133	→ 3 20	A				
		146 136	26 76	Lys	106 44	→ 1 17	G		
G	Val	192 108	Ala	173 87	Asp	116 133	Gly	228 124	U
		41 99		48 178		204 108		174 140	
	119 49	119 107	333 210	→ 4 42	A				
		83 123	129 149	Glu	108 88	→ 15 65	G		

Silně exprimované geny představuje 24 druhů mRNA s celkovým počtem 5253 kodonů. Mezi tyto geny patří gen pro RNA-polymerázu, geny pro dvanáct ribozomových proteinů, několik proteinů vnější membrány a geny pro elongační translační faktory.

Slabě exprimované geny představuje 18 druhů mRNA s 5231 kodony. Patří sem několik represorových genů, gen pro transponázu a β -laktamázu.

Kodony, které jsou čteny jen jedinou tRNA a jejichž výběr je závislý na povaze a síle interakcí mezi kodonem a antikodonem, jsou v rámečku. Šipkami jsou označeny kodony, které jsou používány jen zřídka a mohou se podílet na regulaci genové exprese.

Řešení problému rozdílného využívání kodonů

1. Koexprese genů pro tRNA pro alternativní kodony

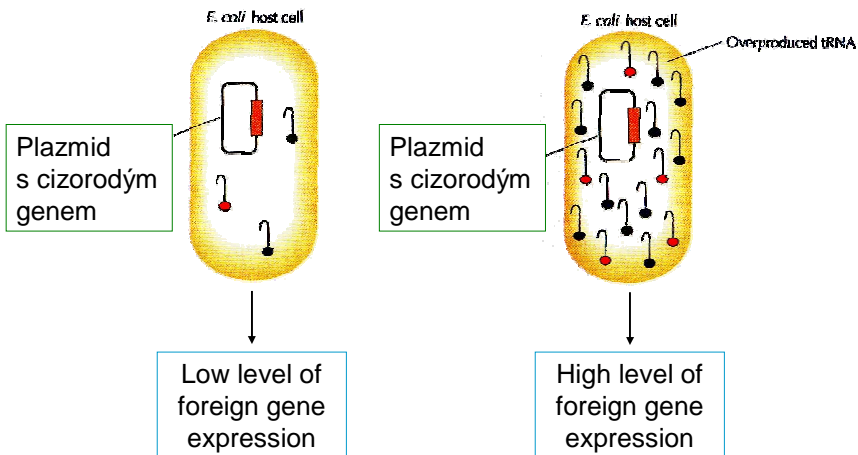
- příprava kmenů s klonovanými geny pro tRNA na samostatných vektorech
- kmen *E. coli* Rosetta má geny pro tyto tRNA na plazmidu, který je kompatibilní s expresním vektorem.

2. Změna méně často používaných kodonů mutagenézí *in vitro* za kodony používané častěji (pracnější postup)

Dosažení vysoké exprese cizorodého genu v kmenech *E. coli* obsahujících geny pro vzácné tRNA

Standardní kmen

Upravený kmen



Zvýšení stability cizích proteinů v *E. coli*

- **Změna lokalizace** (poločas krysího proinzulinu v *E. coli* je v cytoplasmě 2 min, v periplasmě 10 x vyšší)
- **Tvorba fúzních proteinů** (bakteriální + eukaryotická část = beta-galaktosidáza + somatostatin, pak štěpení fúzního proteinu)
- **Expresa v mutantách *E. coli* s nižší aktivitou intracelulárních proteáz** (lon-proteáza – zabraňuje akumulaci denaturovaných nebo jinak pozměněných polypeptidů).
- **Snížení degradace proteinů produktem genu *pin* fága T4 (protease inhibition)** – stabilizace eukaryotických proteinů (interferon)

Zvýšení stability proteinů změnou sekvence jeho aminokyselin

Stabilita β -galaktosidázy po přidání aminokyselin k jejímu N-konci

Amino acid added	Half-life
Met, Ser, Ala	>20 h
Thr, Val, Gly	>20 h
Ile, Glu	>30 min
Tyr, Gln	~10 min
Pro	~7 min
Phe, Leu, Asp, Lys	~3 min
Arg	~2 min

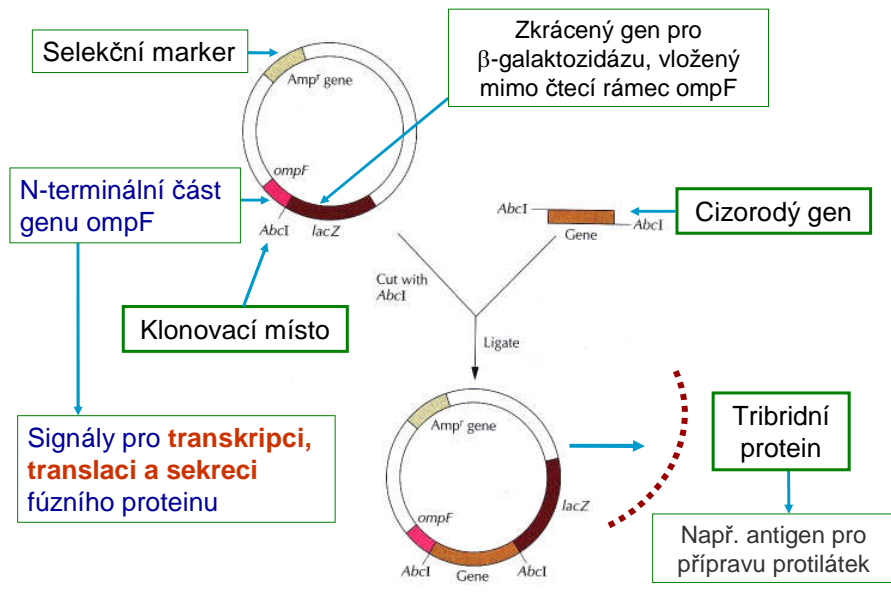
PEST = aminokyseliny (prolin P, glutamová kys. E, serin S, treonin T), jejichž přítomnost v určitých vnitřních oblastech proteinu zvyšuje jeho citlivost k proteolytické degradaci

Vytváření fúzních proteinů

Fúzní protein: produkt vytvořený po spojení dvou nebo více genů/sekvencí:

1. Původní gen hostitelského organismu = stabilizující partner (cizí proteiny jsou v heterologních systémech často nestabilní)
2. Cizorodý gen (gen zájmu)
3. +/- spojující sekvence (oligonukleotidový linker), kódující krátké úseky aminokyselin rozpoznávané **nebakteriálními** proteázami
 - umožňují dodatečné odštěpení cílového produktu z fúzního proteinu
 - používají se k purifikaci rekombinantních proteinů

Klonovací vektor pro přípravu fúzních proteinů



Některé fúzní systémy používané k purifikaci cizorodých proteinů vytvářených v *E. coli*

Fusion partner	Size	Ligand	Elution condition
ZZ	14 kDa	IgG	Low pH
His tail	6-10 aa	Ni ²⁺	Imidazole
Strep-tag	10 aa	Streptavidin	Iminobiotin
PinPoint	13 kDa	Streptavidin	Biotin
MBP	40 kDa	Amylose	Maltose
β-Lactamase	27 kDa	Phenyl-boronate	Borate
GST	25 kDa	Glutathione	Reducing agent
Flag	8 aa	Specific MAb	Low calcium

ZZ = fragment proteinu A (*S. aureus*); His = histidin; Strep-tag = peptid s afinitou ke streptavidinu; PinPoint = fragment proteinu biotinylovaný *in vivo* v *E. coli*; MBP = protein vázající maltózu; GST = glutation S-transferáza; Flag = peptid rozpoznávaný enterokinázou; Mab = monoklonální protilátka.

Proteolytické štěpení fúzního proteinu krevním koagulačním faktorem Xa

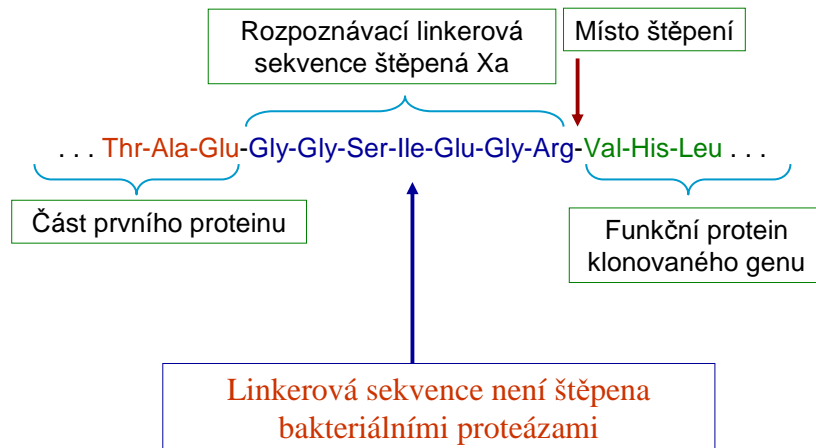
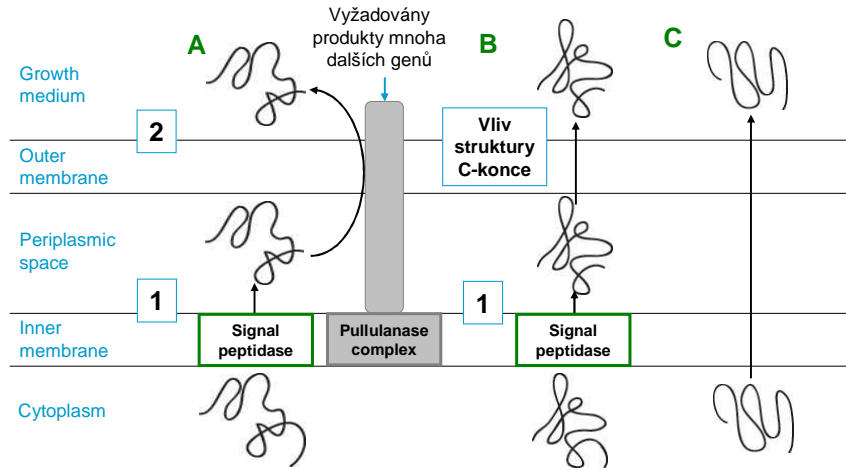


Fig. 7-79. Thin sections of *E. coli* bacteria showing granules of β -galactosidase/proinsulin fusion peptides (fixed in formaldehyde/glutaraldehyde according to Karnofsky; embedded in Epon; stained with uranylacetate/lead hydroxide; 38 000:1; Courtesy of Dr. W. Wetekam, Hoechst AG).

Tři možné způsoby transportu sekretovaných proteinů



- A. obecná exportní dráha (general export pathway, GEP) – SP + Sec proteiny (chaperony)
- B. dráha IgA-proteázy (SP + C-konec proteinu)
- C. dráha nezávislá na SP – vyžaduje ABC-transportery (ATP-dependenční transportní proteiny)

Příklady signálních sekvencí různých proteinů

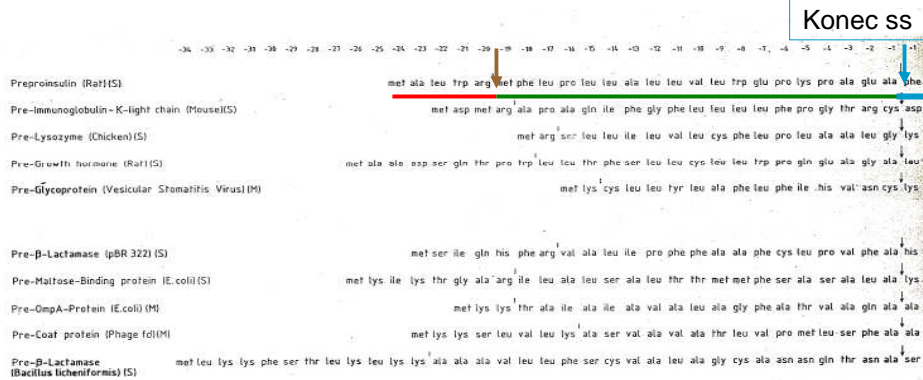
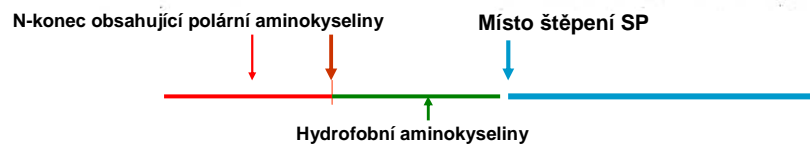


Fig. 7-78. Amino acid sequences of N-terminal signal sequences of various precursors for membrane proteins (M), and a variety of secretory proteins (S). The vertical line indicates the transition from hydrophilic to hydrophobic regions within the signal sequences, the arrow the start of the mature proteins.

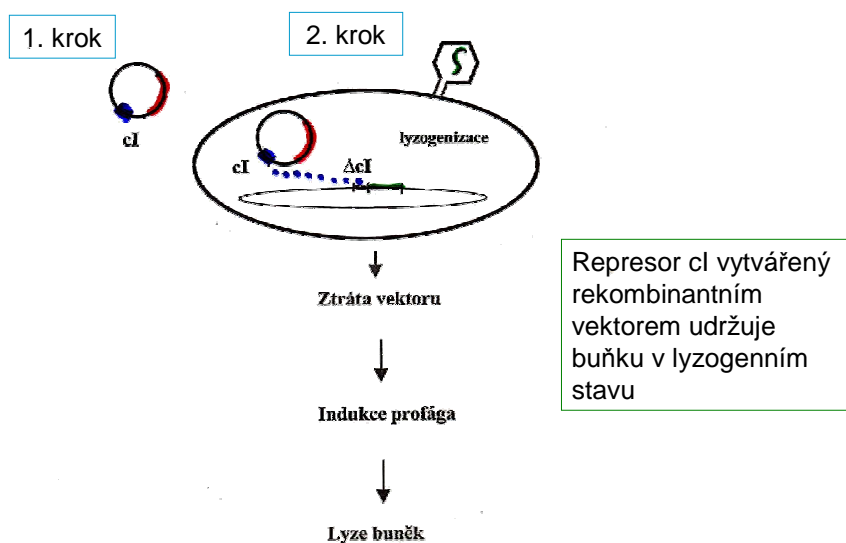


Instabilita vektorů

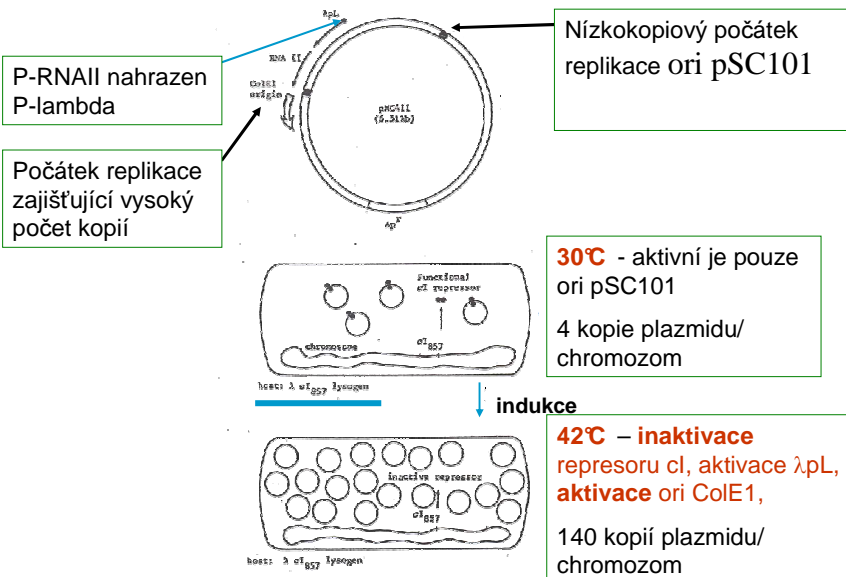
1. **Segregační instabilita:** Ztráta plazmidu, ke které může dojít při dělení buněk v důsledku
 - **chybění funkce par** (partitioning) u některých vektorů (pBR322), která zaručuje rovnoměrné dělení kopií do dceřinných buněk. Problém lze řešit selekcí antibiotiky, nebo lze oblast **par** klonovat, např. z pSC101 do pBR322 a tím plazmidy stabilizovat.
 - **vytváření multimerních forem plazmidu** a následné nerovnoměrné dělení kopií do dceřinných buněk. Multimerní plazmidy nevznikají u ColE1, který využívá rekombinační systém **cer xer**, který rozkládá multimery. Toto místo lze klonovat do plazmidů typu pBR322 a eliminovat problémy s multimerizací.

*Možnou strategií pro překonání segregační instability je **eliminace buněk**, které plazmid ztratily (např. použití genu *ci* fága v klonovacím vektoru a využití lyzogenních hostitelů)*
2. **Strukturní instabilita:** Důsledek delecí, inzercí nebo translokací v chromozomech, plazmidech nebo virech (homologní rekombinace a transpozice).

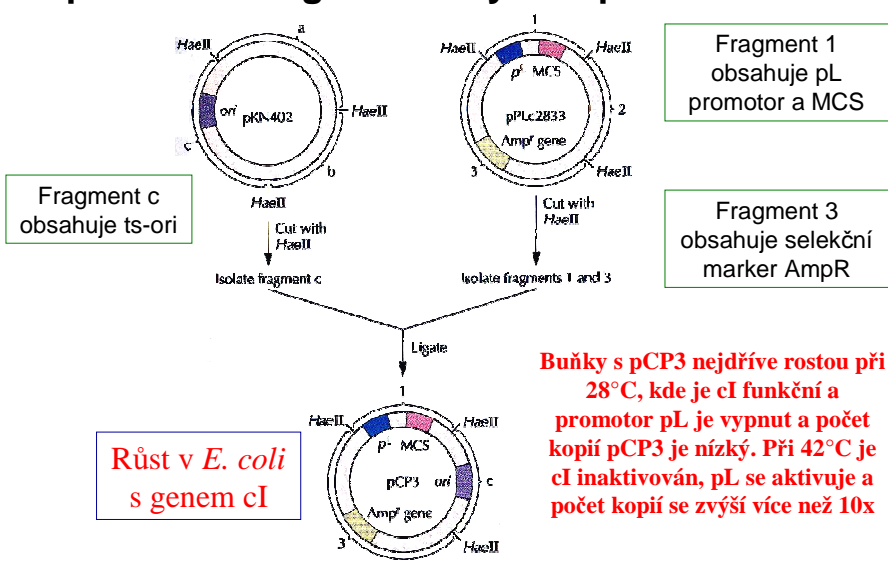
Překonání segregační instability vektoru eliminací buněk, které vektor ztratily



Vektory s dvojím počátkem replikace pro regulaci počtu kopií



Příprava vektoru s regulovatelným ts-počátkem replikace a s regulovatelným ts-promotorem



Mol Biotechn 127

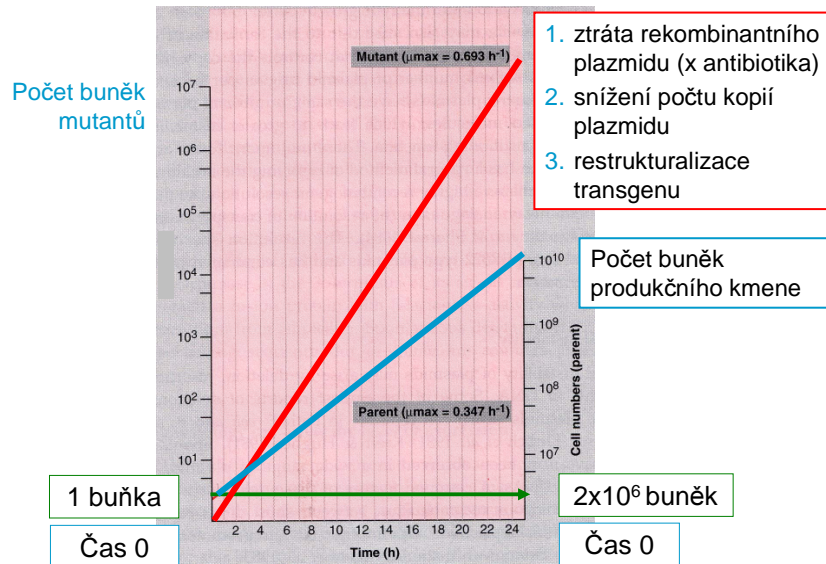
Vliv teploty na počet kopií plazmidů u tří expresních vektorů

Plasmid	Plasmids/cell at:		p ^L present
	28°C	→ 42°C	
pKN402	82	512	No
pPLc2833	38	42	Yes
pCP3	60	713	Yes

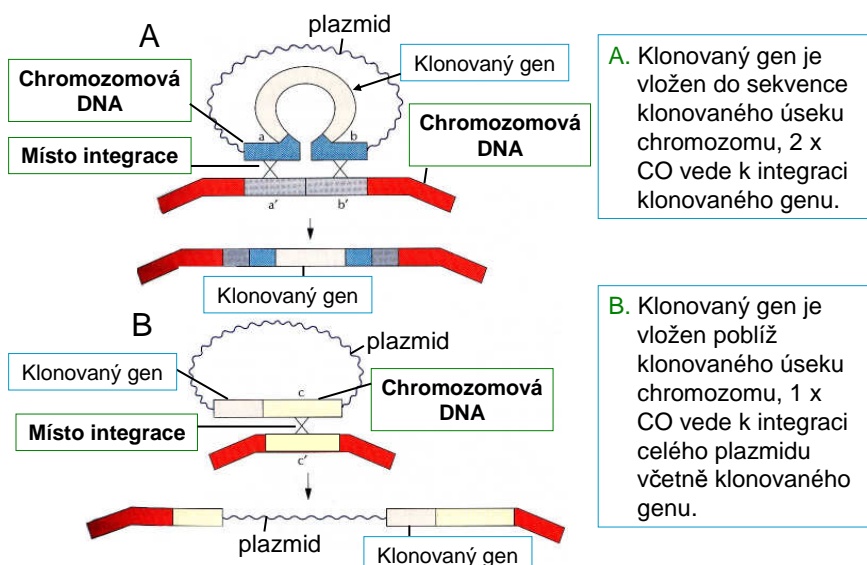
Důsledky vyplývající z přítomnosti rekombinantního plazmidu v buňkách

1. Snížení růstové rychlosti buněk
2. Změny morfologie buněk, nebo zvýšená fragilita buněk
3. Restrukturalizace klonovaného genu (rekombinace)
 - výběr vhodného plazmidu RC x theta

Kompetice mezi pomalu rostoucími buňkami produkčního kmene (obsahuje vektor) a rychle rostoucími mutantami (které ztratily vektor)



Integrace klonovaného genu do bakteriálního chromozomu



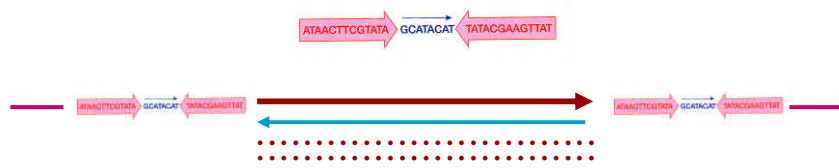
Počet kopií klonovaného genu pro α -amylázu v *B. subtilis* a její aktivita v buňkách

Copies/genome	Activity (U/mL of mid-log cells)
2	500
5	2,300
7	3,100
8	3,400
9	4,400
Multicopy plasmid (20-40 kopií)	700

geny na chromozomu

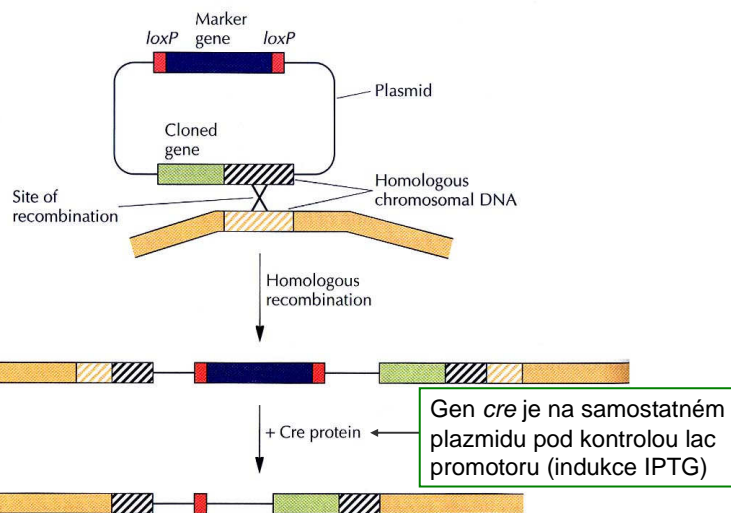
Gen pro α -amylázu z *B. amyloliquefaciens* byl vložen do oblasti pocházející z chromozomu *B. subtilis* naklonované v plasmidu *E. coli*, konstrukt byl vnesen do *B. subtilis*. Plasmid nesl rovněž geny pro rezistenci k chloramfenikolu a ampicilinu. Buňky *B. subtilis* s plasmidem byly pěstovány v prostředí s chloramfenikolem, což vedlo k selekci buněk se **zvýšeným počtem kopií integrovaného plasmidu**, tím zvýšení počtu kopií genu pro amylázu a ke zvýšení jejího množství v buňkách.

Struktura rozpoznávací sekvence lox P fága P1



Podle orientace loxP sekvencí je úsek DNA mezi nimi místně specifickou rekombinací prostřednictvím Cre-rekombinázy buď deletován nebo invertován

Odstranění selekčního markeru po vnesení plazmidu do bakteriálního chromozomu



Vlivy prostředí ovlivňující expresi genů a tvorbu produktů

1. Složení kultivačního media, zdroje živin
2. Teplota
3. pH
4. Koncentrace kyslíku

Důkaz tvorby cizorodého produktu

- Imunologické testy
- Enzymové testy
- SDS-PAGE
- Minibuňky
- Maxibuňky
- Transkripce a translace in vitro