

Zajištění exprese klonovaných genů a její optimalizace

Faktory ovlivňující expresi klonovaných genů

A. Regulační sekvence pro genovou expresi

1. *Transkripční úroveň*

- Síla promotoru a jeho charakter
- Terminátor transkripce
- Stabilita mRNA, její posttranskripční úpravy

2. *Translační úroveň*

- Účinnost vazby mRNA na ribozom
- využívání kodonů
- stabilita proteinu

3. *Transport proteinu*

- charakter signální sekvence

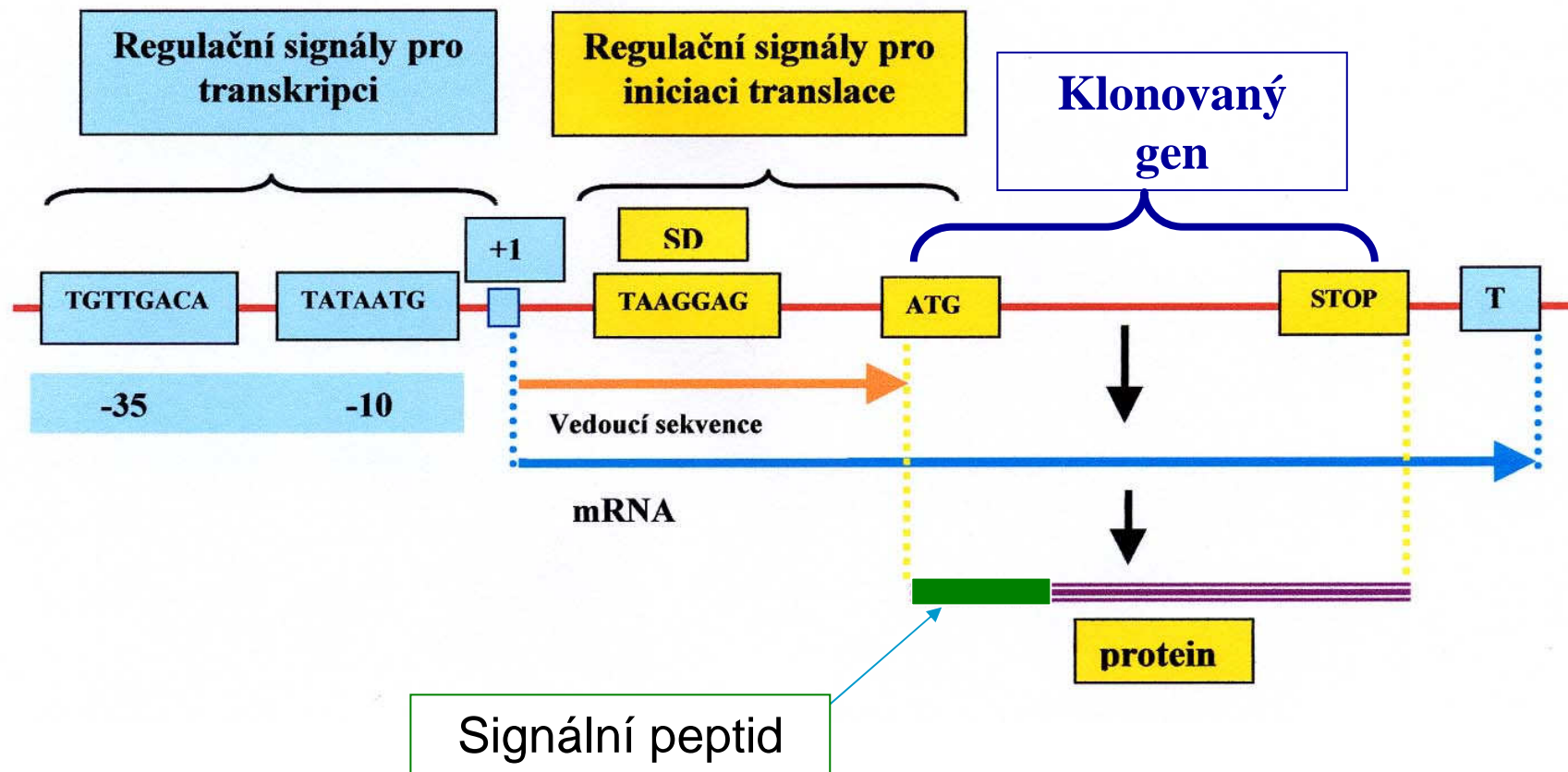
B. Vlastnosti vektorů

- Počet kopií vektoru
- Stabilita vektoru

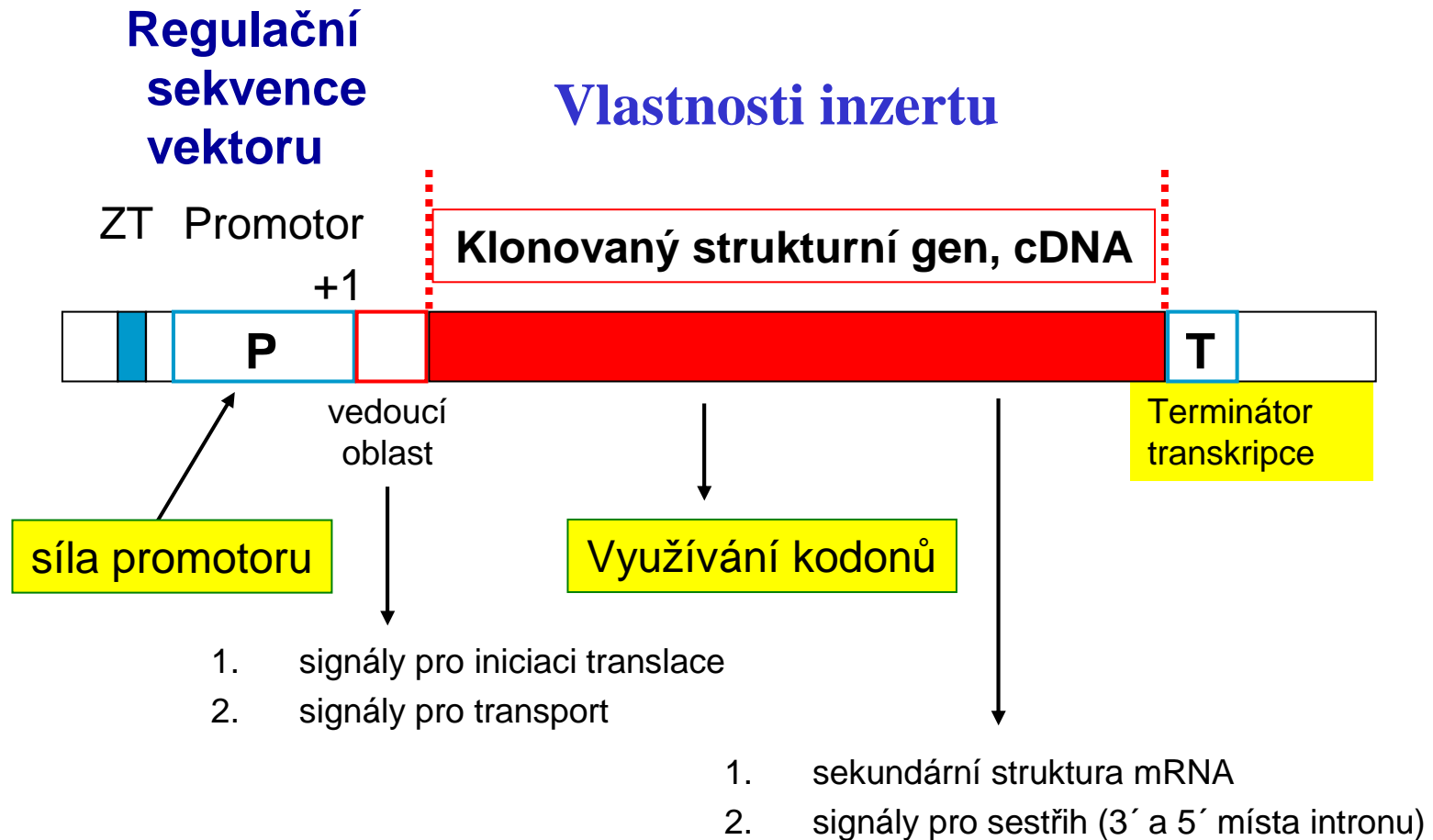
C. Fyziologie hostitelské buňky

- růstové podmínky
- enzymový aparát hostitelské buňky

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



Faktory ovlivňující expresi klonované DNA



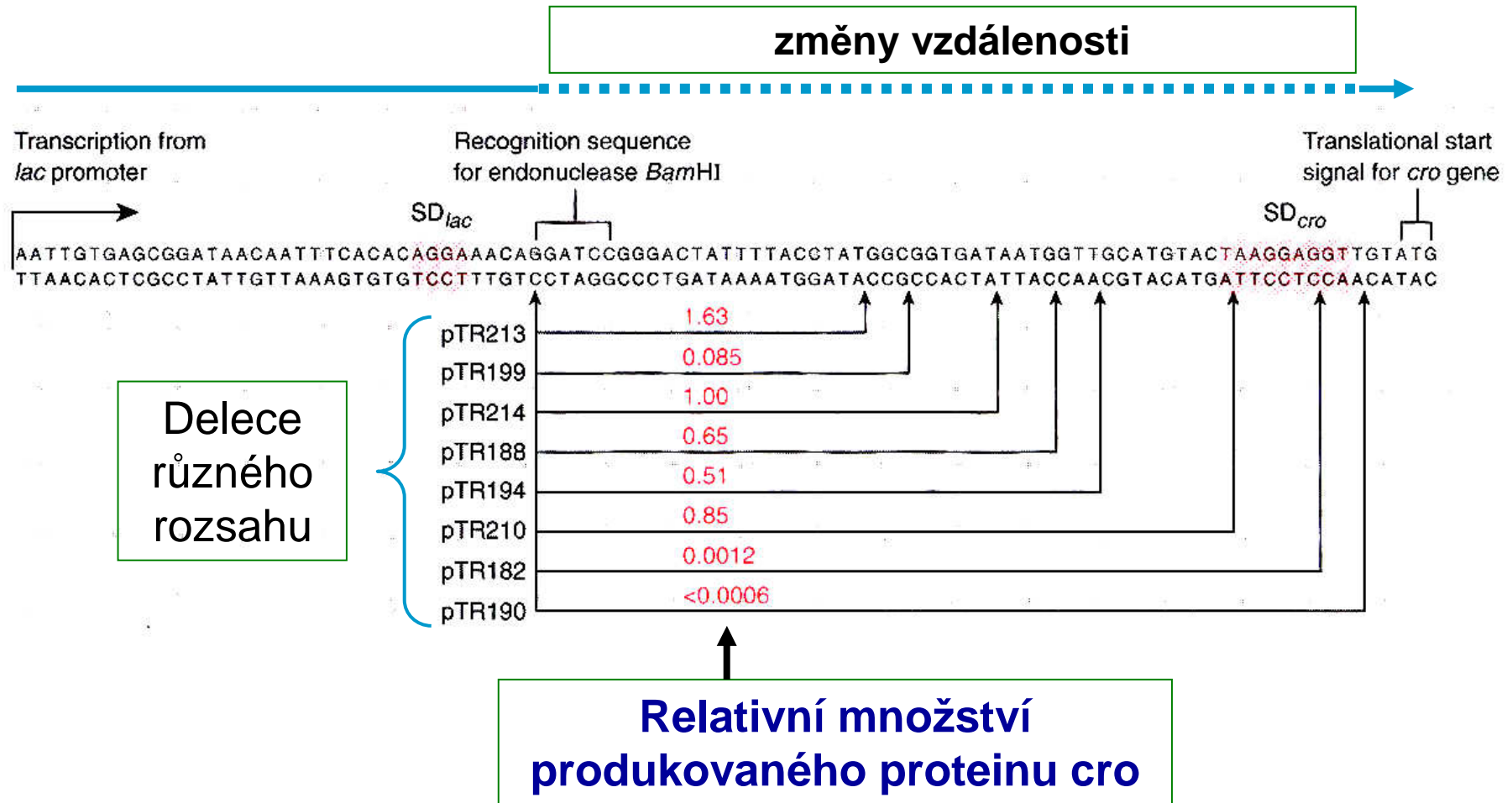
Regulovatelné promotory používané v expresních vektorech

		Promotor	Zdroj	Způsob regulace	
				Off	On
<i>E. coli</i>		λ pL, λ pR	Leftward and rightward early promoters of λ	30°C	>37°C (in <i>cl₈₅₇</i> host)
<div style="border: 1px solid green; padding: 5px; width: fit-content;"> <p>lac-UV5 nezávislost na katabolické represi</p> </div>	←	<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	
			Off	On
<i>B. subtilis</i>	<i>bla</i>	<i>Bacillus licheniformis</i> β-lactamase gene	—	β-lactams in medium
	<i>cat</i>	<i>Bacillus pumilis</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

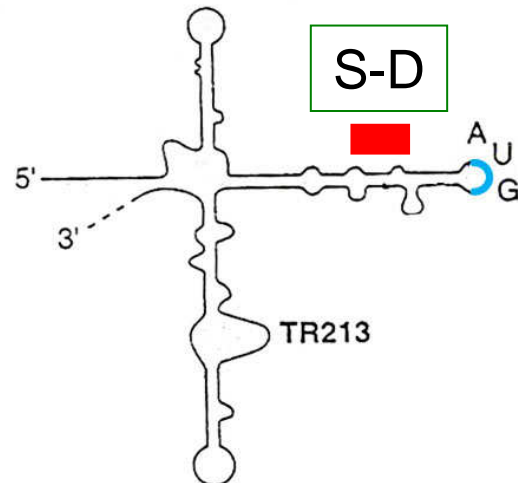
Vliv vzdálenosti mezi promotorem a startem translace na množství vytvářeného proteinu cro



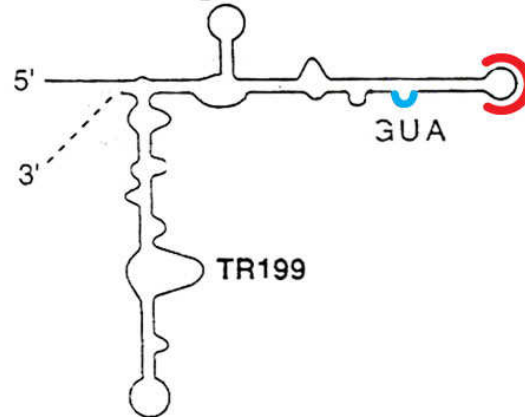
Sekundární struktura cro-mRNA

Množství proteinu

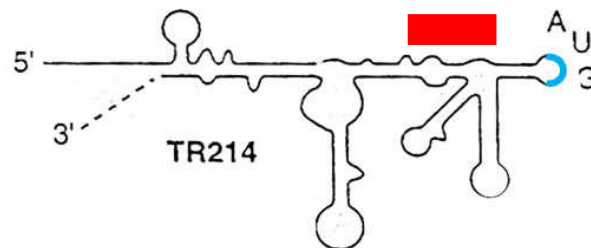
163%



8%



100%



Iniciační kodon AUG je součástí smyčky – protein cro se tvoří v malém množství

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	AGGAAACAGAAAGATGGAT	pTR436
1.0	9	AGGAAACAGCCAGATGGAT	HP1
0.01	5	GTCGAGGAATTCCATGGAT	pPLcSVt5-372
0.1	5	ATTGGAATTCCATGGAT	pPLcSVt5-37
2.5	8	TTGGAATTATTCCATGGAT	pPLcSVt5-379
1.0	9	TTGGAATTAATTCCATGGAT	pPLcSVt5-374
0.01	9	AGGAATTCCAAAGATGGAT	pPLcSVt5-72

SD

Inic.kodon

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

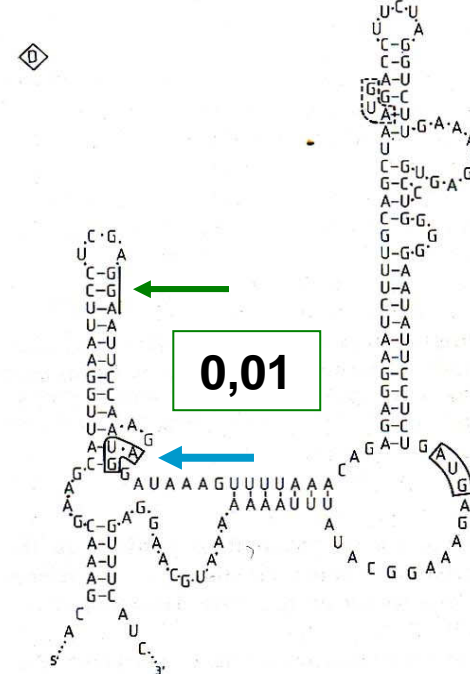
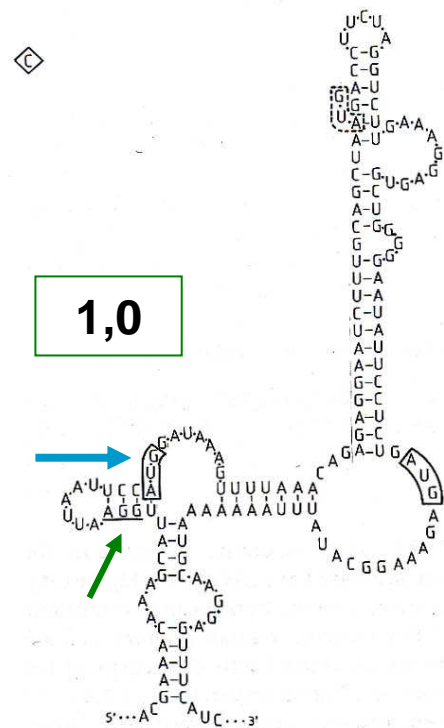
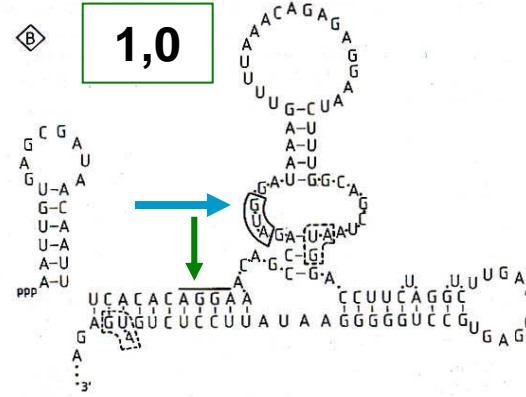
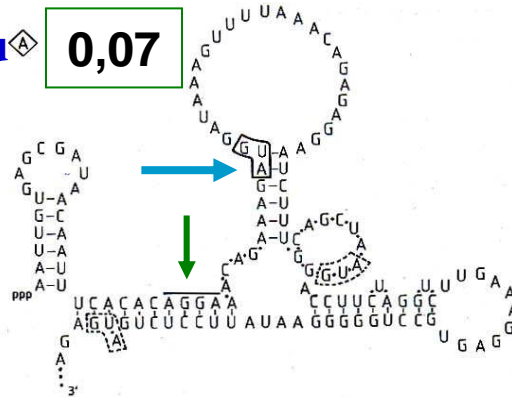
Množství proteinu

0,07

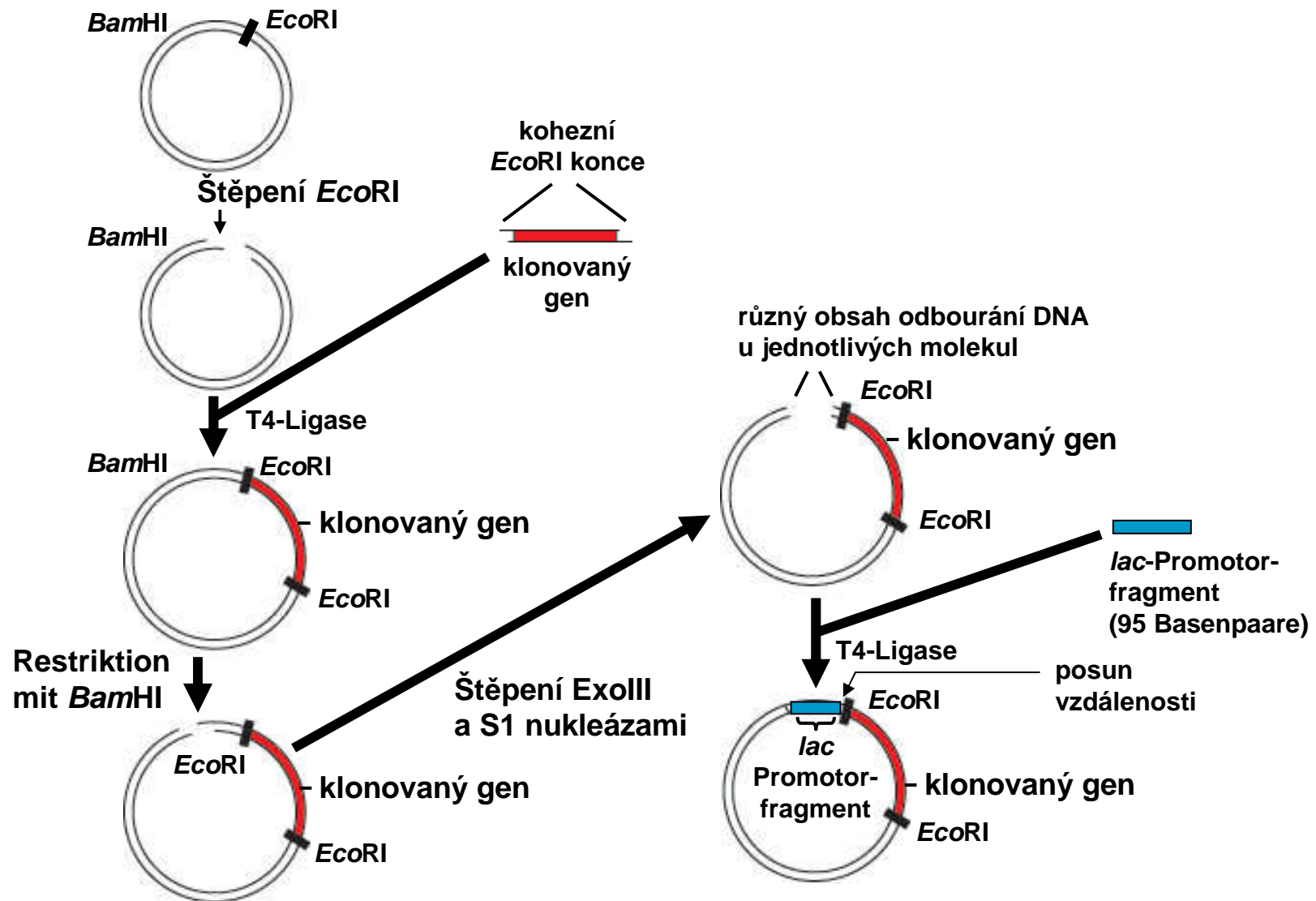
1,0

AUG →

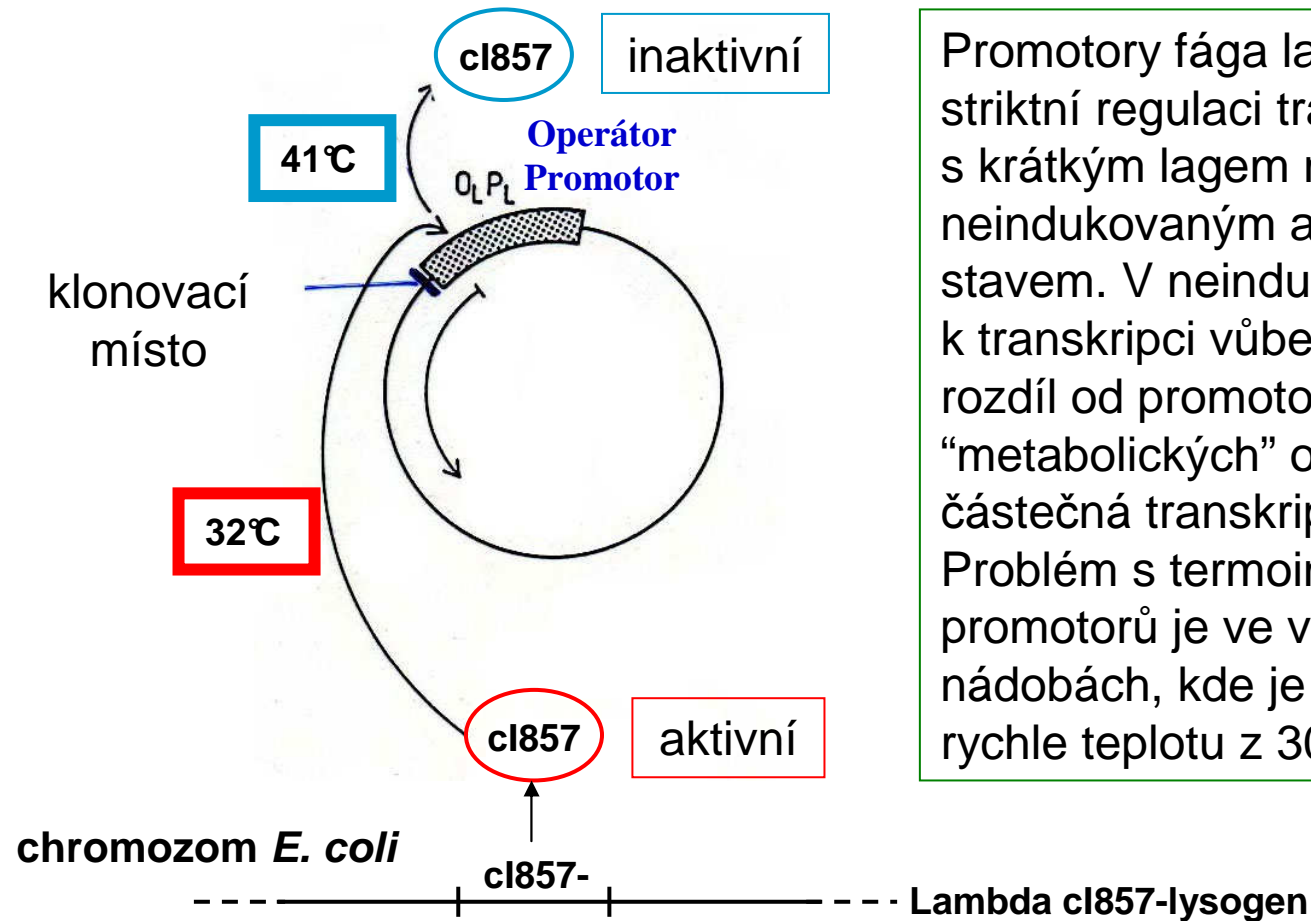
SD →



Posun vzdálenosti mezi promotorem a klonovaným genem



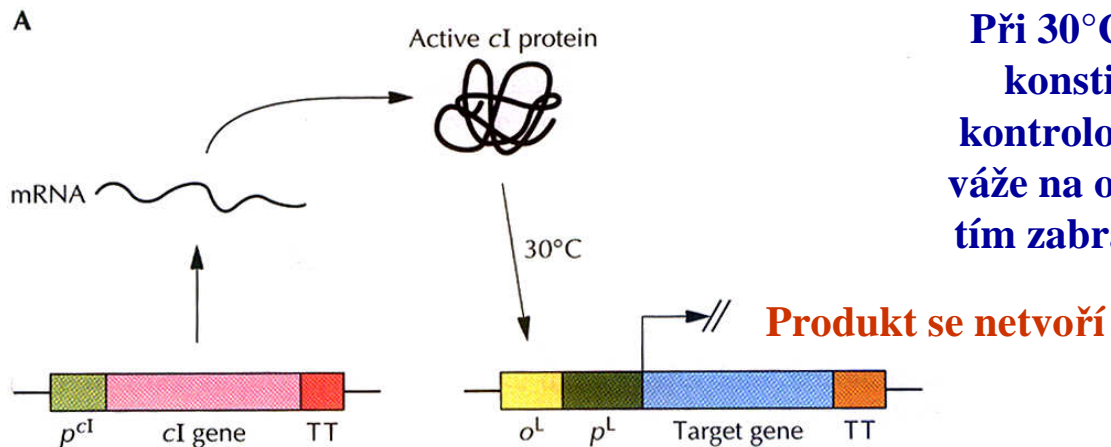
Použití teplotně senzitivního represoru pro regulaci promotoru PL fága λ



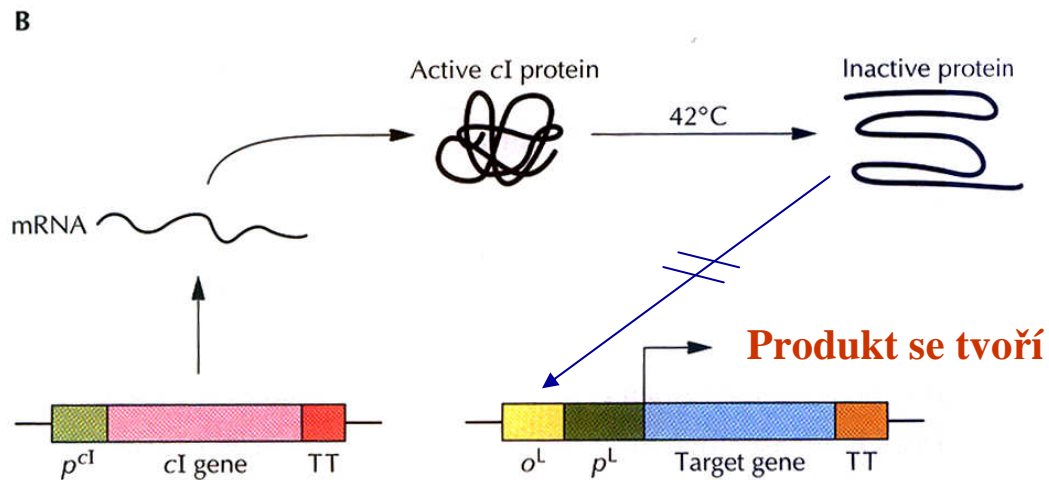
Promotory fága lambda vykazují striktní regulaci transkripce s krátkým lagem mezi neindukovaným a indukovaným stavem. V neindukovaném stavu k transkripci vůbec nedochází, na rozdíl od promotorů “metabolických” operonů, kdy částečná transkripce probíhá stále. Problém s termoindukcí u ts promotorů je ve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

Při 32°C se termolabilní represor cl857 kódovaný genem *cl* 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 klonovaného genu probíhá.

Regulace genové exprese promotorem pL fága λ



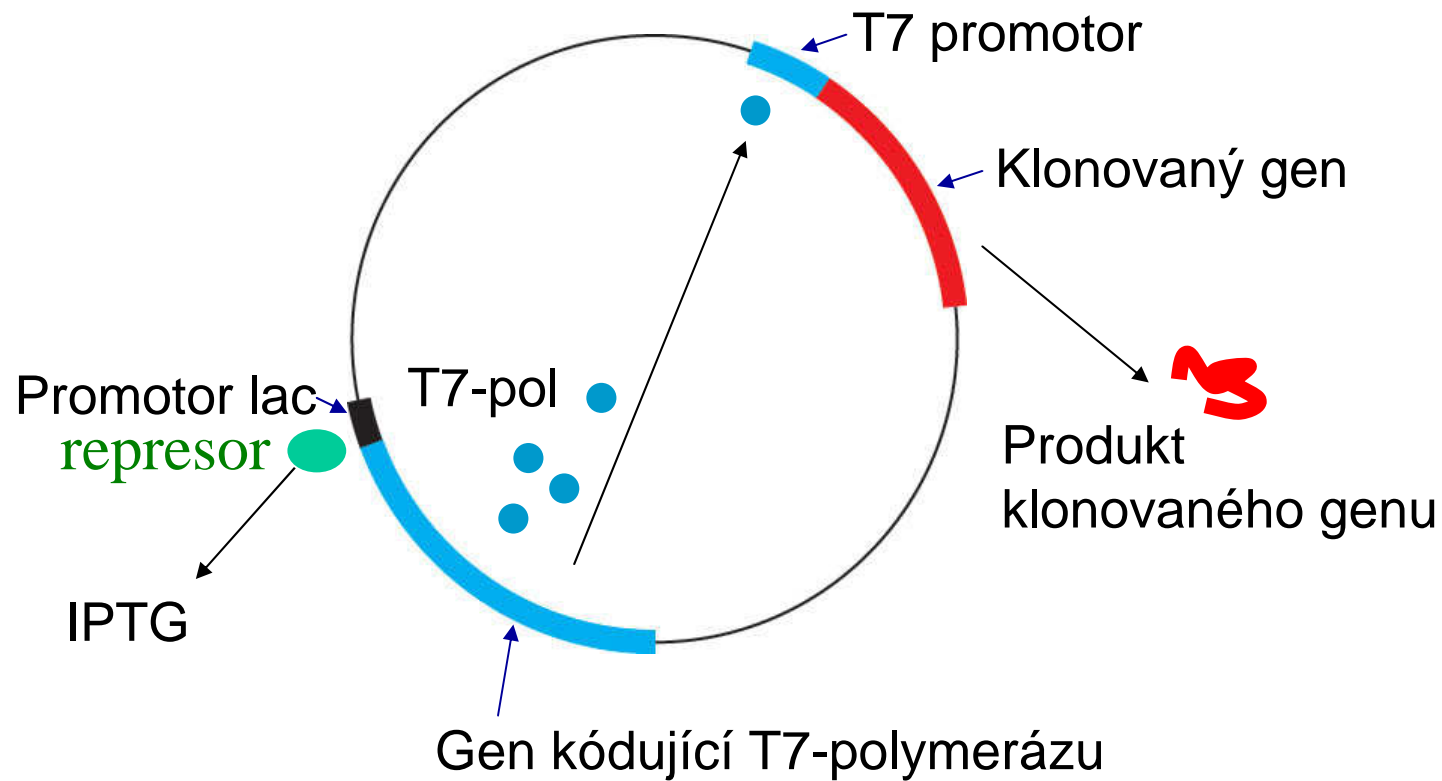
Při 30°C se ts-represor cI, který je konstitutivně syntetizován pod kontrolou vlastního promotoru p^{cI} , váže na operátor o^L promotoru p^L a tím zabraňuje expresi cílového genu



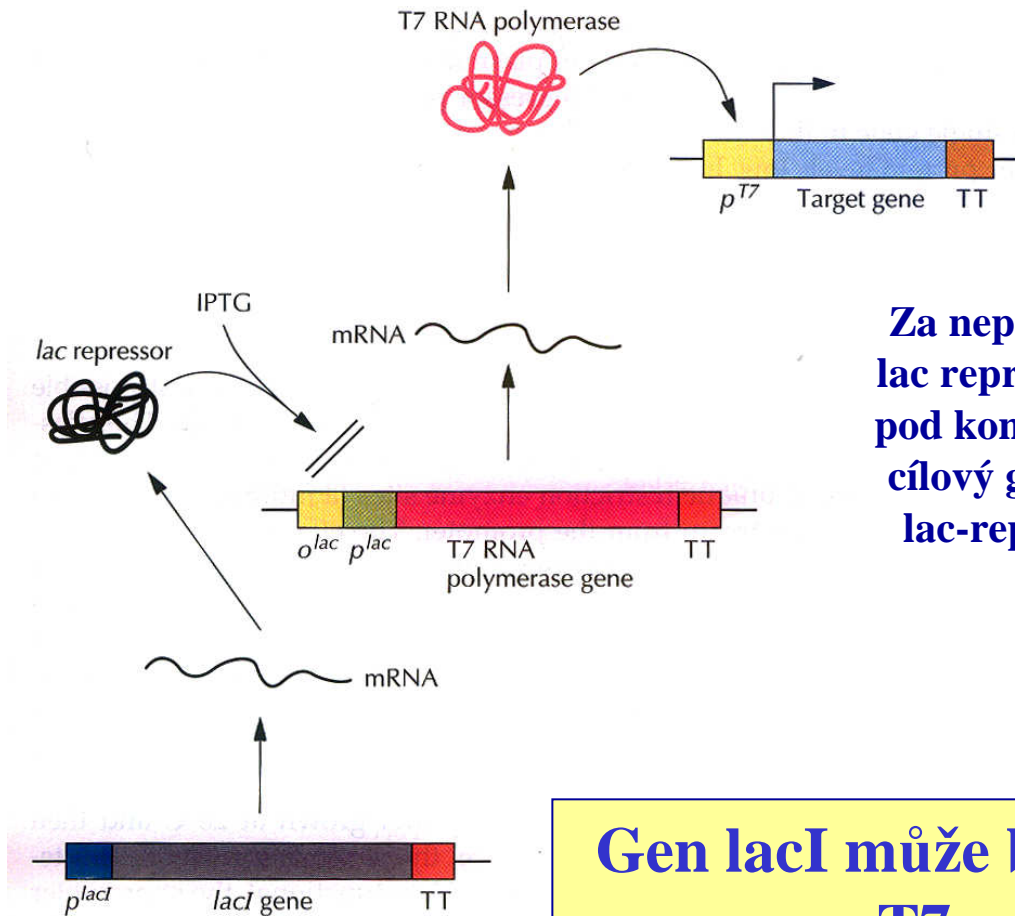
Při 42°C je ts represor cI inaktivován a dochází k transkripci cílového genu

Expresní vektory obsahující T7-promotor

RNA-polymeráza T7 rozpoznává pouze promotory genů fága T7



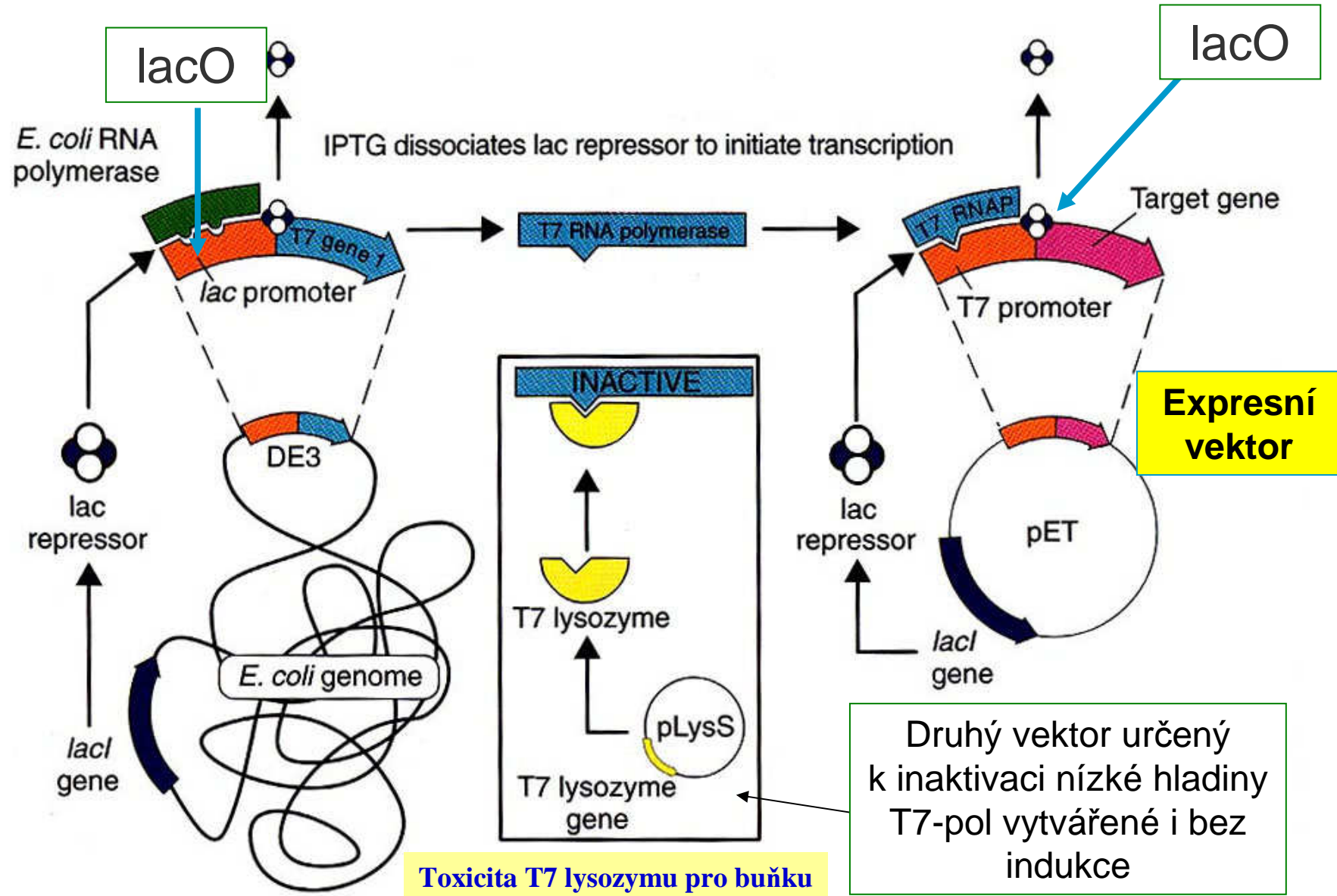
Regulace genové exprese cílového genu kontrolovaná promotorem pT7 fága T7



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

Gen *lacI* může být na jiném vektoru než gen pro T7-polymerázu a cílový gen

System 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.

**Silné terminátory se zařazují rovněž před inducibilní
promotory, aby zabránily transkripci z promotorů
lokalizovaných před klonovaným genem („read-through“)**

Stabilita mRNA

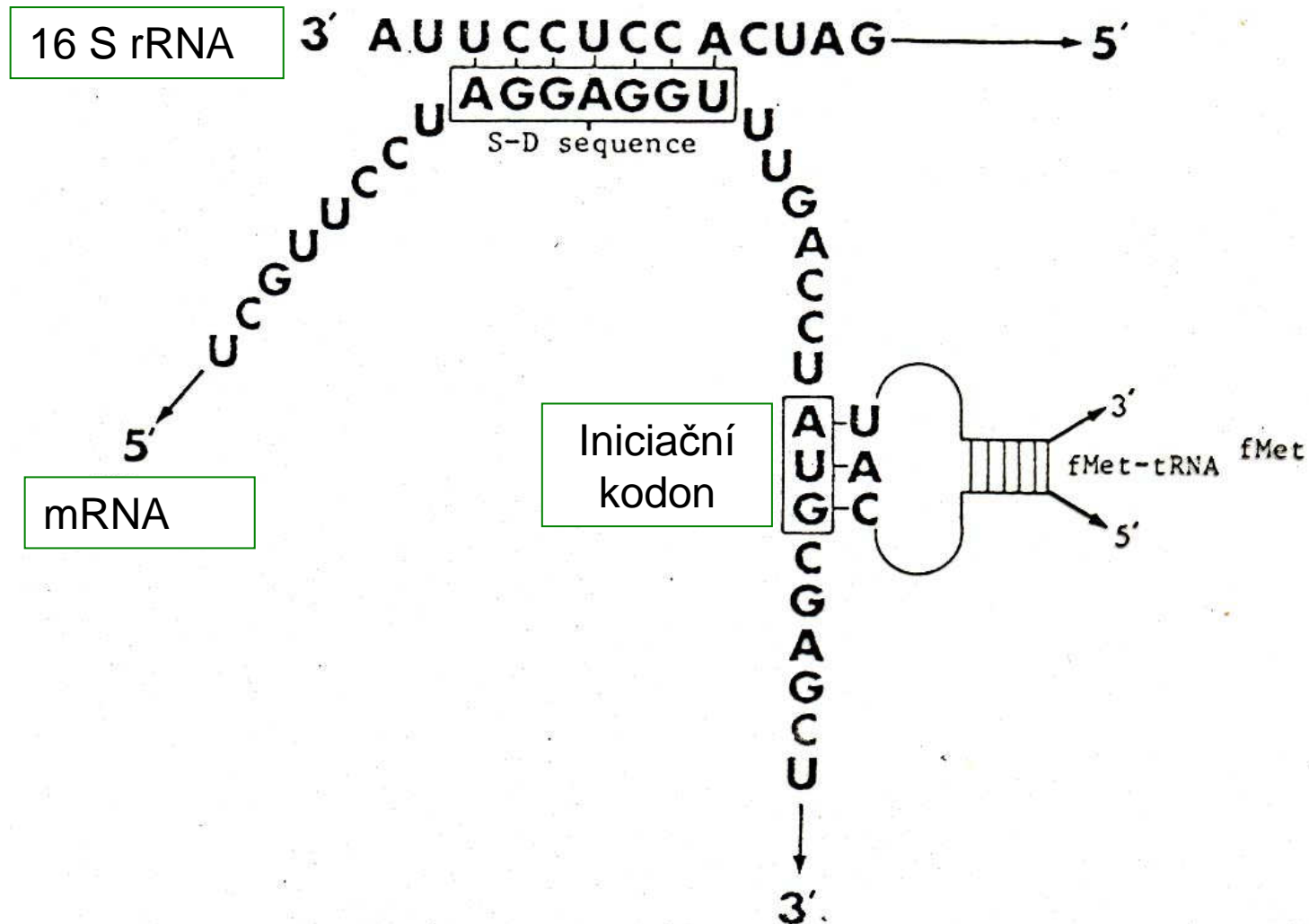
- Rychlost syntézy proteinu závisí na množství mRNA v buňce
- Existuje rovnováha mezi syntézou a rozkladem daného druhu mRNA /turnover/

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

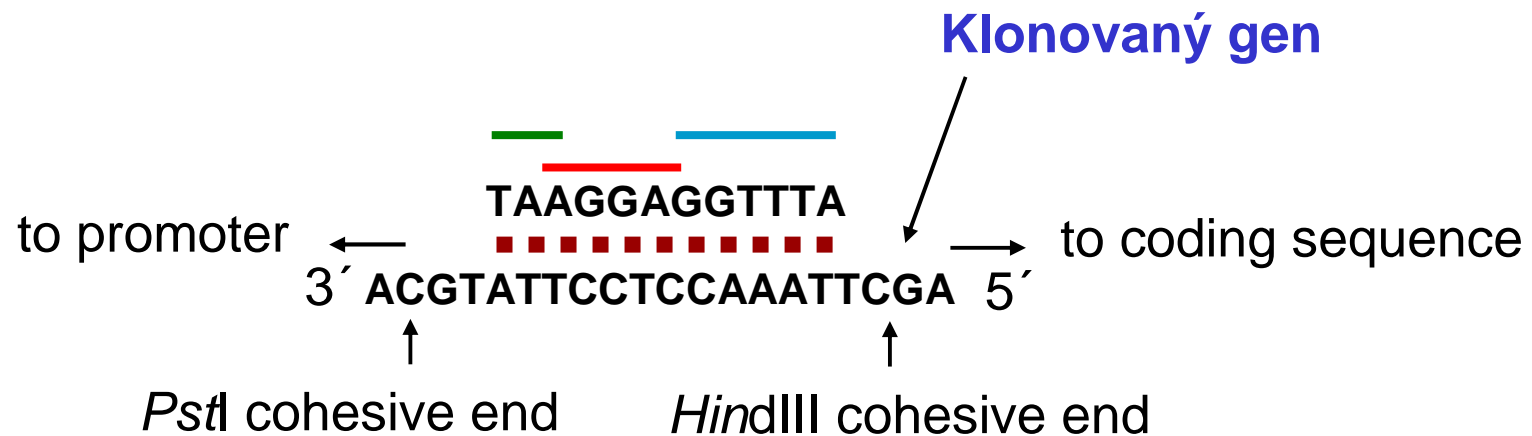


Kromě iniciačního kodonu AUG existují v RBS ještě tři další oblasti, jejichž sekvence jsou více či méně konzervovány:

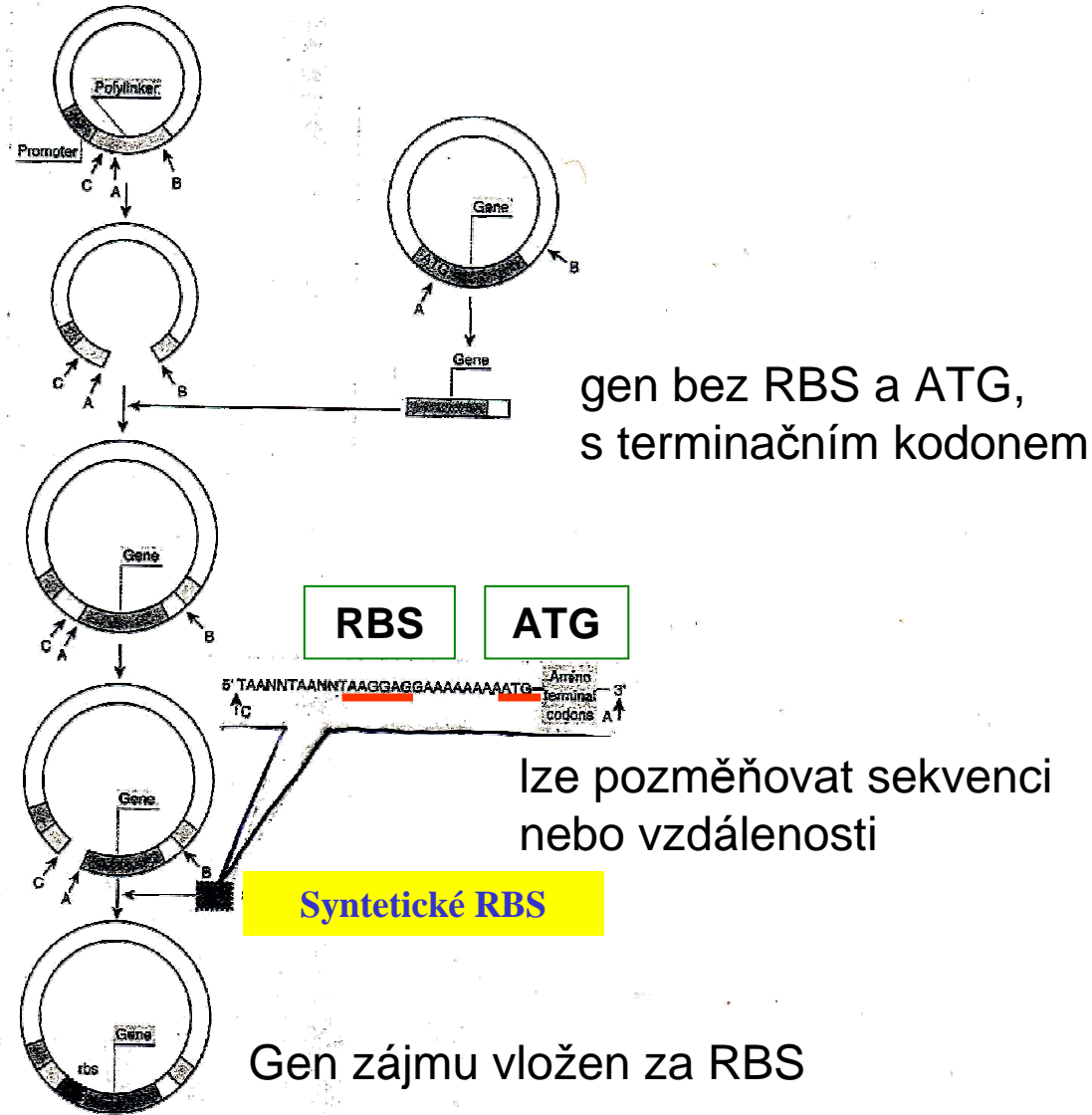
1. Shine-Dalgarnova sekvence, níž se obvykle vyskytuje sekvence 5'UAAGGAGGU 3'.
2. U mRNA (alespoň polycistronické) jsou součástí RBS jeden nebo více terminačních kodonů.
3. U genů, které jsou silně exprimovány (např. geny pro fágové kapsidy nebo ribozomální proteiny), se v RBS nachází sekvence **PuPuUUUPuPu** (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

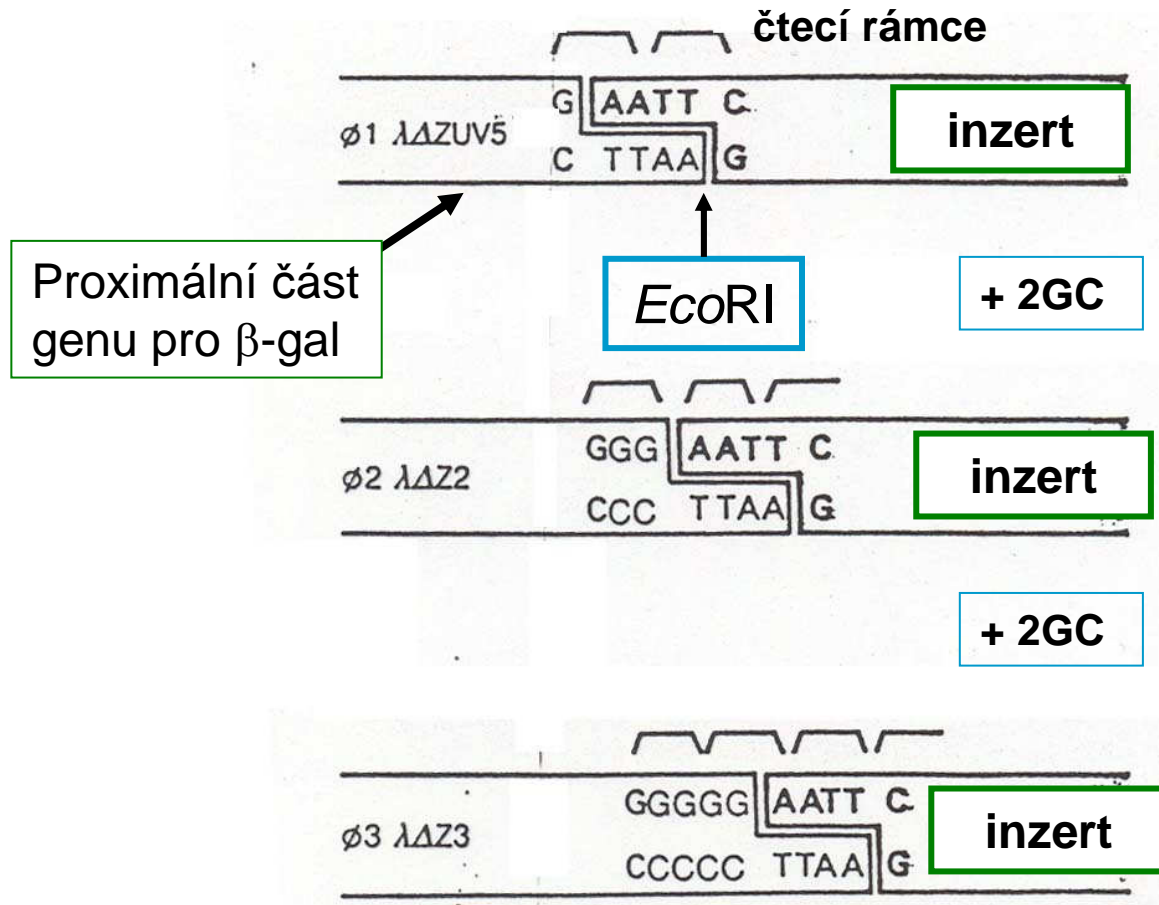
1. S-D sekvence 
2. RRUUURR (GGTTTAA) 
3. Terminační kodon TAA 



Zajištění účinné translace použitím optimalizované RBS

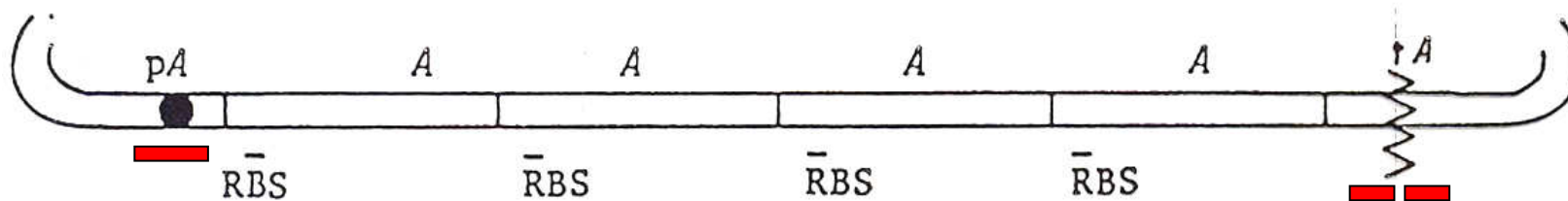


Různé čtecí rámce vzhledem k iniciaci translace genů lacZ u tří různých vektorů



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

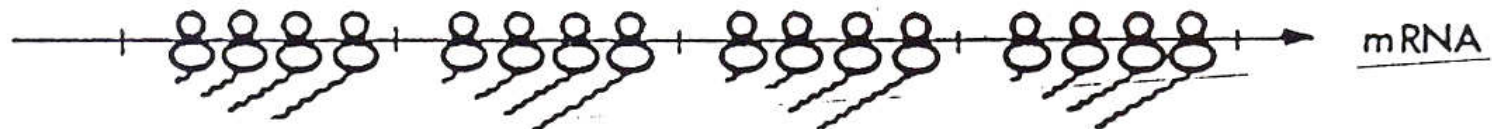
Zvýšení genové exprese konstrukcí homopolycistronické sekvence



jeden promotor
jeden terminátor
více kopií genu A

Transcription

Transcription
from multiple RBSs



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 65		23 39	
C	Leu	12 71	Pro	21 29	His	19 95	Arg	223 99	U
		16 64		2 46		75 59		101 133	
A	Ile	→ 3 22	Thr	26 45	Gln	38 90	Ser	→ 3 27	A
		345 294		162 101		169 166		→ 1 42	
G	Met	67 156	Ala	103 46	Asn	13 101	Arg	10 56	U
		262 118		137 119		159 98		49 61	
G	Val	→ 2 27	Glu	15 32	Lys	259 163	Gly	→ 3 28	A
		140 130		28 76		106 44		→ 1 17	
G	Val	192 108	Glu	173 87	Asp	116 183	Gly	226 124	U
		41 66		48 178		204 106		174 140	
G	Val	119 48	Glu	119 107	Glu	333 210	Gly	→ 4 42	A
		83 123		129 149		106 98		→ 14 66	

Silně exprimované geny představuje 24 druhů mRNA s celkovým počtem 5253 kodonů. Mezi tyto geny patří gen pro RNA-polymerázu, geny pro dvanáct ribozomických 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

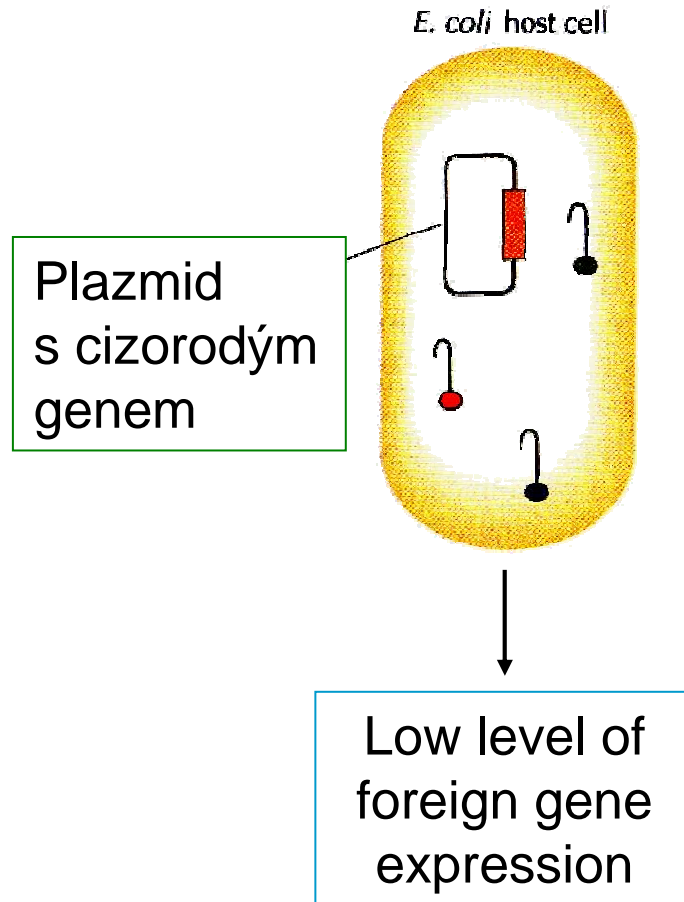
1. Koexprese genů pro tRNA pro alternativní kodony

- 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.

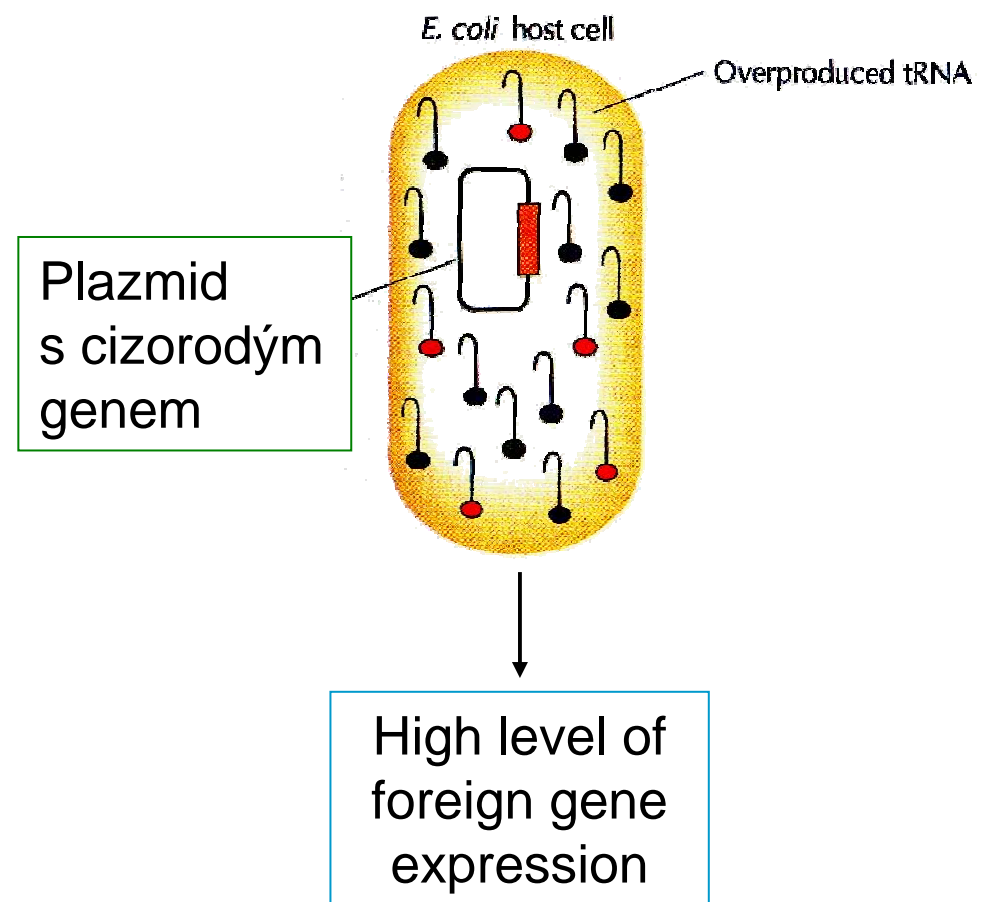
2. Změna méně často používaných kodonů mutagenezí *in vitro* za kodony používané častěji (pracnější postup)

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 cytoplazmě 2 min, v periplazmě 10 x vyšší)
- Tvorba fúzních proteinů (bakteriální + eukaryotická část = beta-galaktozidáza + somatostatin, pak štěpení fúzního proteinu)
- Exprese 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 -galaktozidázy po přidání aminokyselin k jejímu N-konci

Přidané aminokyseliny	Poločas
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

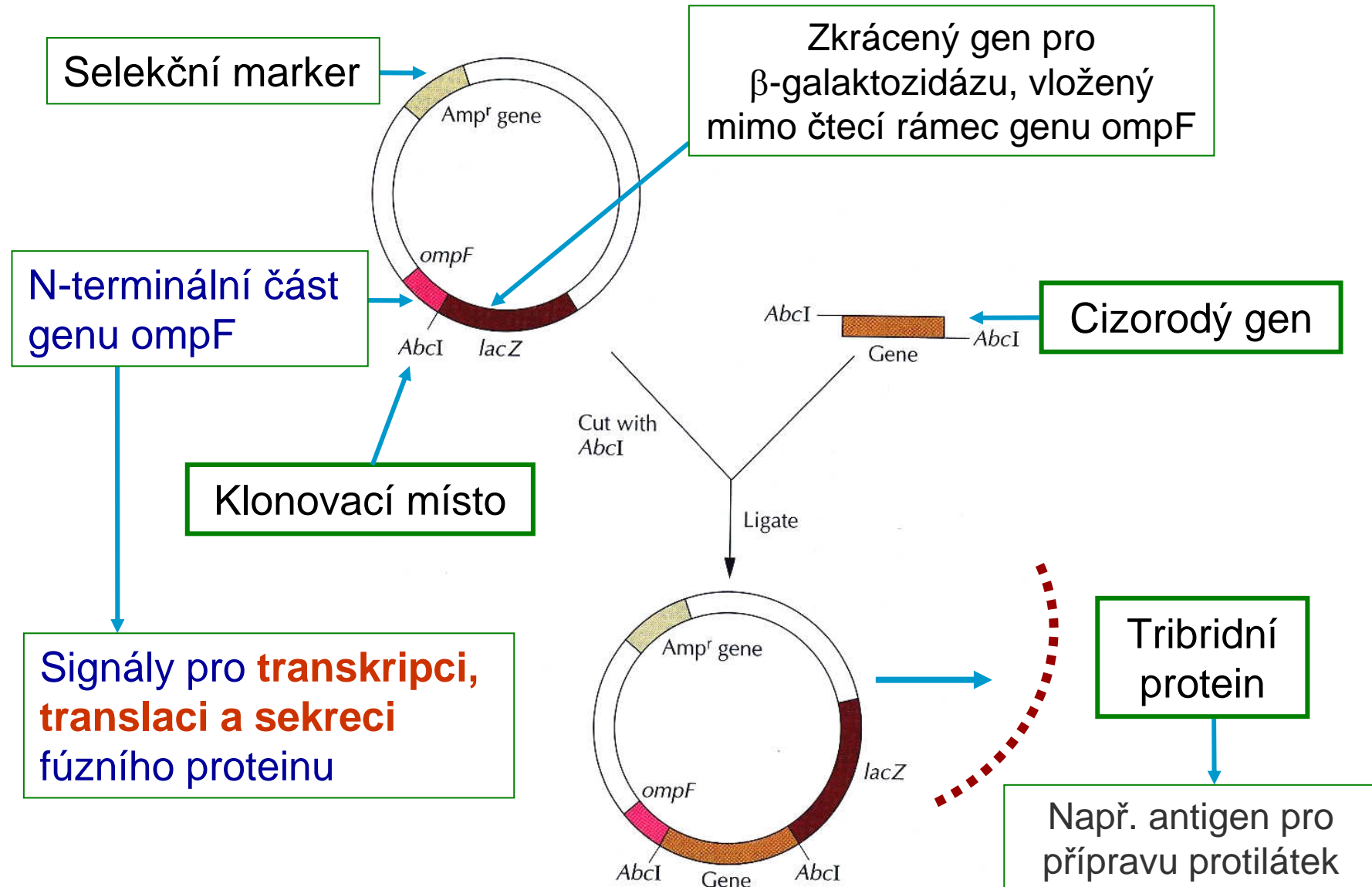
PEST = aminokyseliny (prolin P, glutamová kys. E, serin S, treonin T), jejichž přítomnost v určitých vnitřních oblastech proteinu zvyšuje jeho citlivost k proteolytické degradaci

Vytváření fúzních proteinů

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

- 1. Přirozený gen hostitelského organismu = stabilizující partner (cizí proteiny jsou v heterologních systémech často nestabilní)**
- 2. Cizorodý gen (gen zájmu)**
- 3. +/- spojovací 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

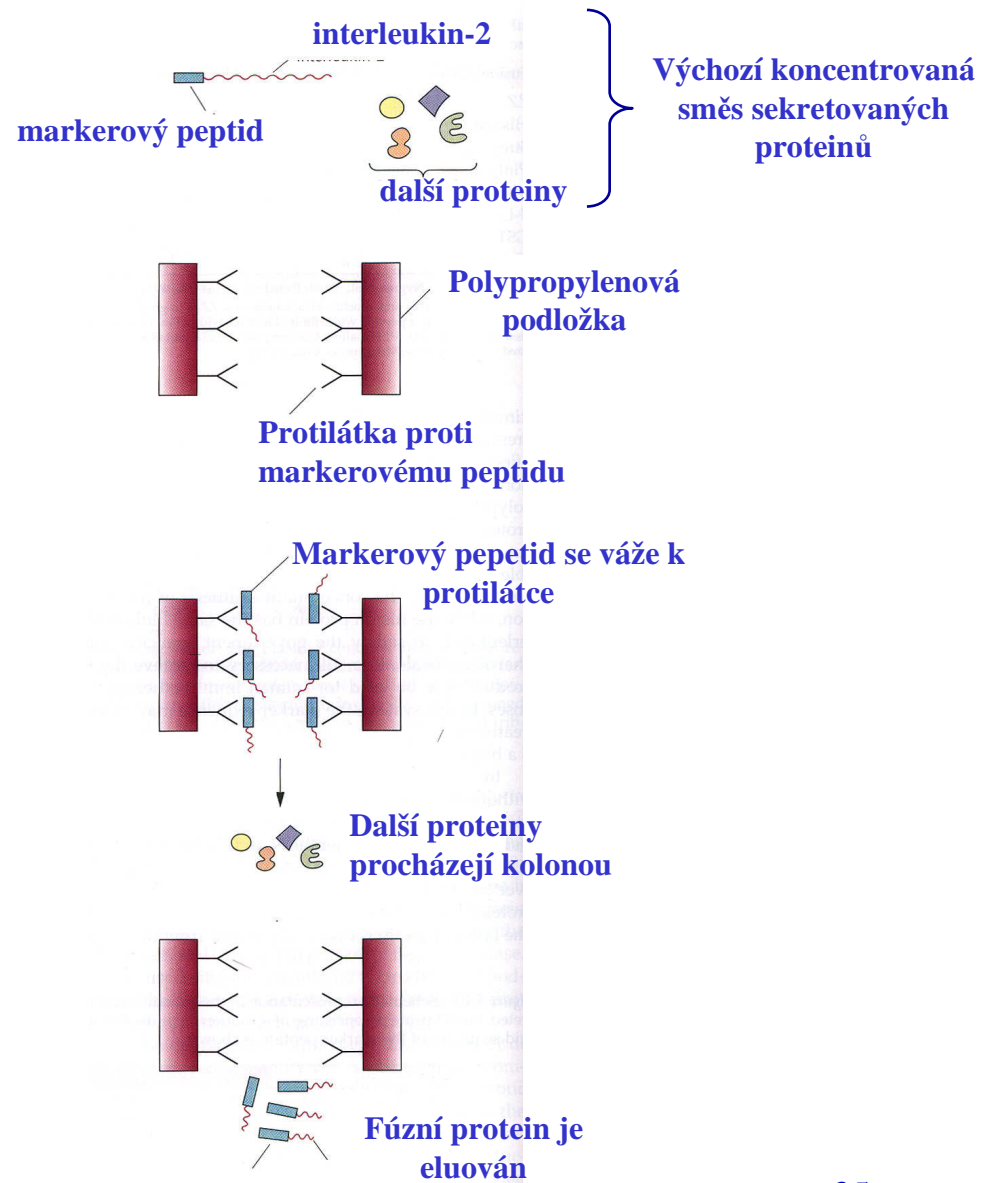


Některé fúzní systémy používané k purifikaci cizorodých proteinů vytvářených v *E. coli*

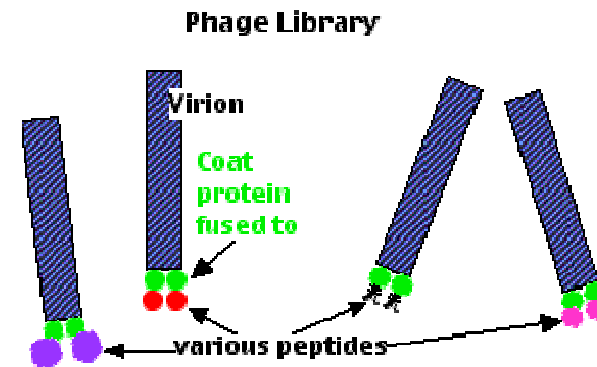
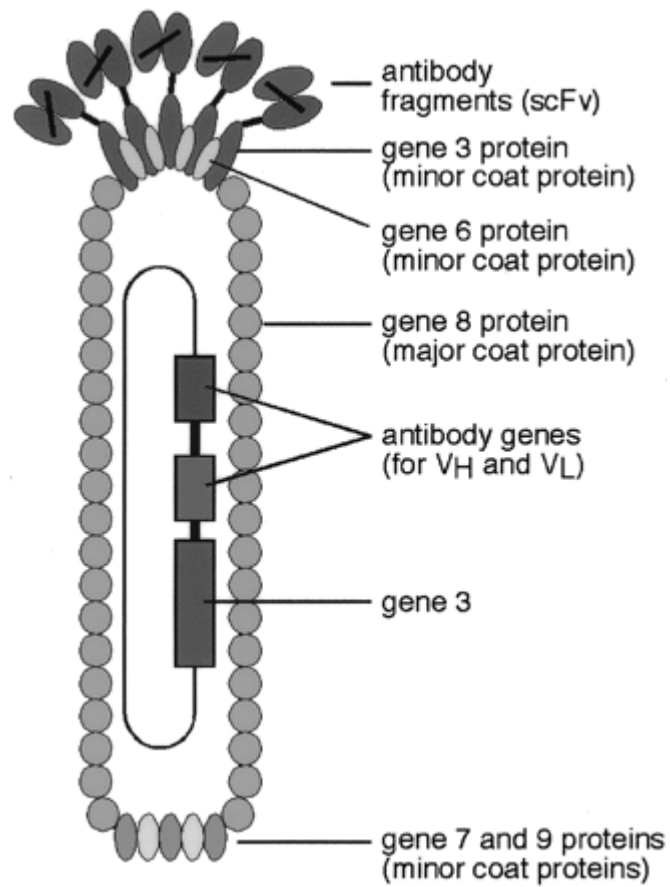
Fúzní partner	Velikost	Ligand	Podmínky eluce
ZZ	14 kDa	IgG	Low pH
His tail (tag)	6-10 aa	Ni ²⁺	Imidazole
Strep-tag	10 aa	Streptavidin	Iminobiotin
PinPoint	13 kDa	Streptavidin	Biotin
MBP	40 kDa	Amylose	Maltose
β-Lactamase	27 kDa	Phenyl-boronate	Borate
GST	25 kDa	Glutathione	Reducing agent
Flag	8 aa	Specific 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.

Purifikace fúzních proteinů imunoafinitní chromatografií



Fágový displej



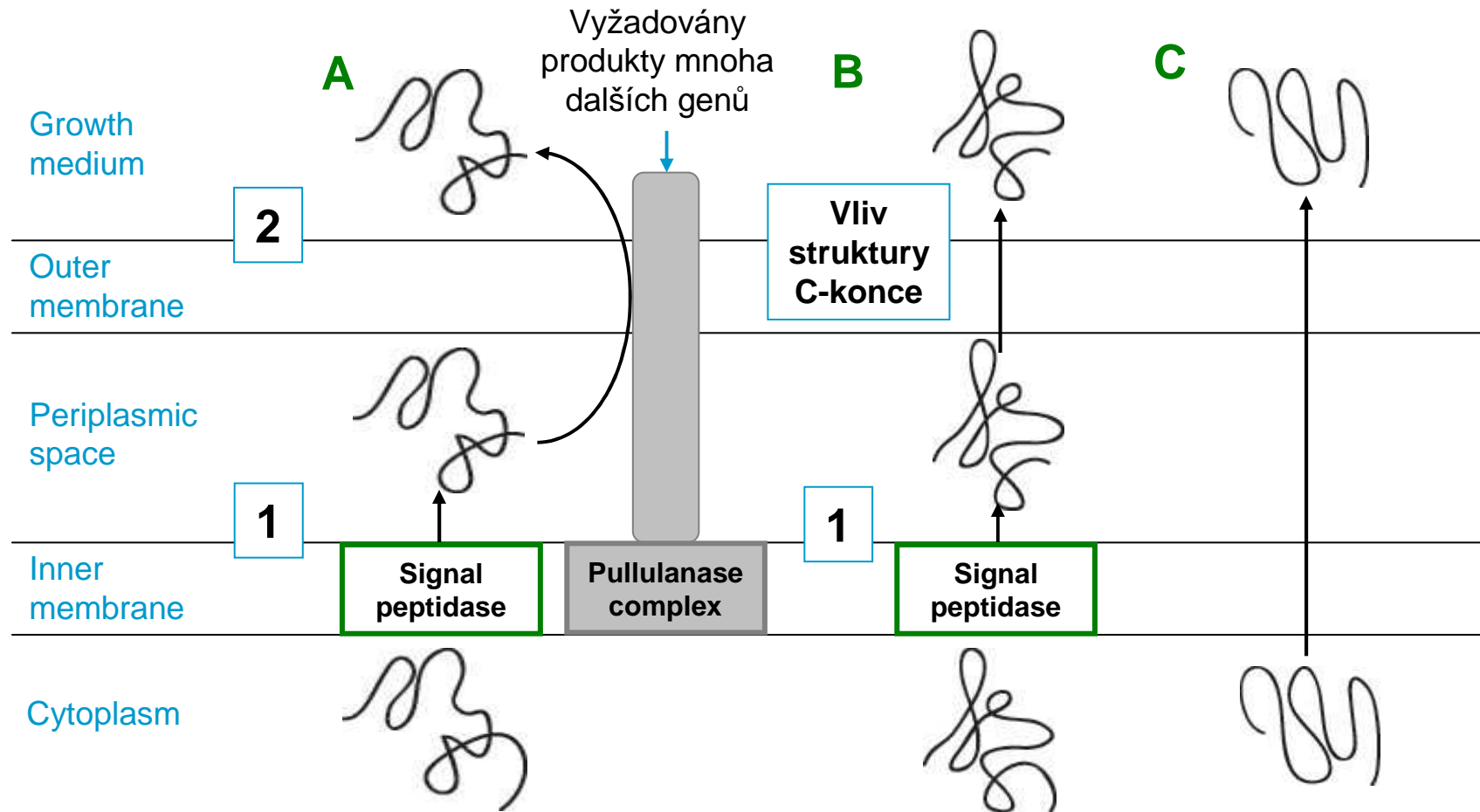
Proteolytické štěpení fúzního proteinu krevním koagulačním faktorem Xa





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-dependentní transportní proteiny)

Příklady signálních sekvencí různých proteinů

Konec ss

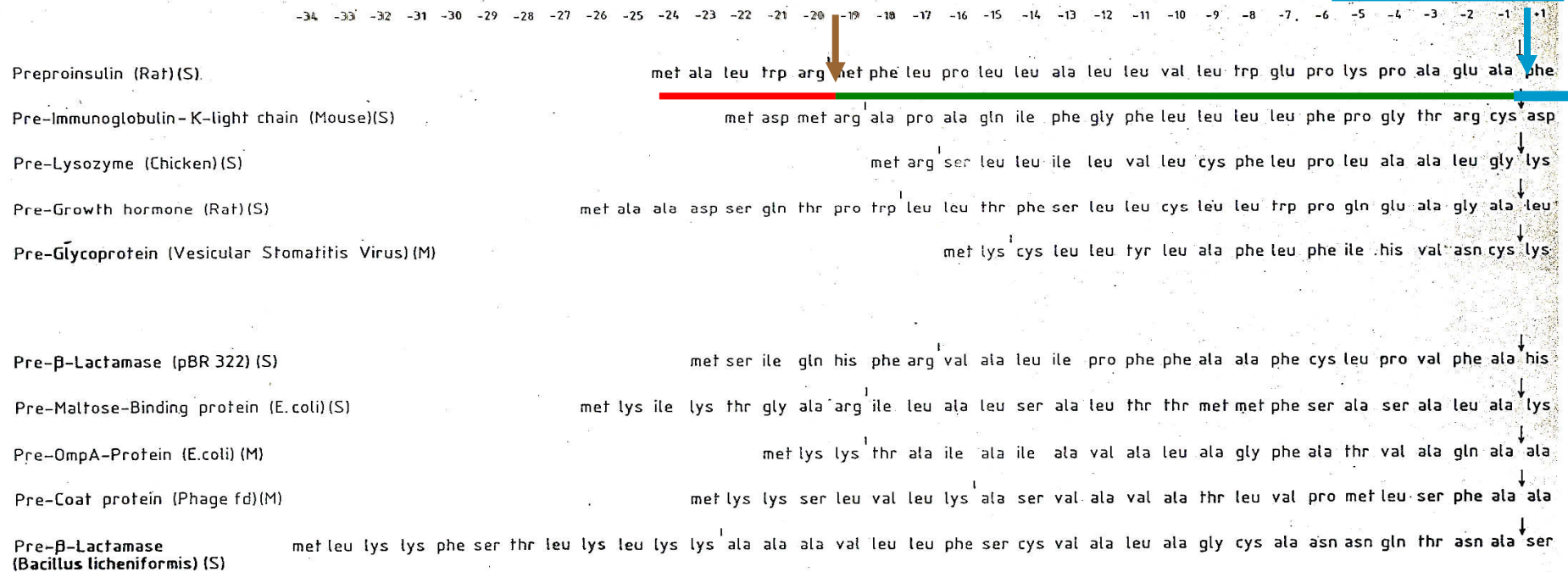
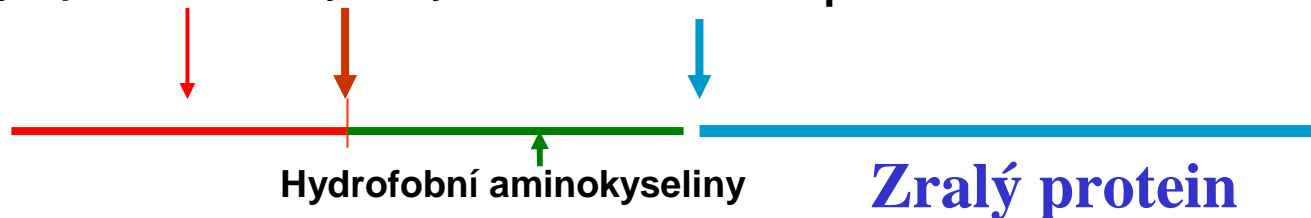


Fig. 7-78. Amino acid sequences of N-terminal signal sequences of various precursors for membrane proteins (M), and a variety of secretory proteins (S).

The vertical line indicates the transition from hydrophilic to hydrophobic regions within the signal sequences, the arrow the start of the mature proteins.

N-konec obsahující polární aminokyseliny

Místo štěpení SP



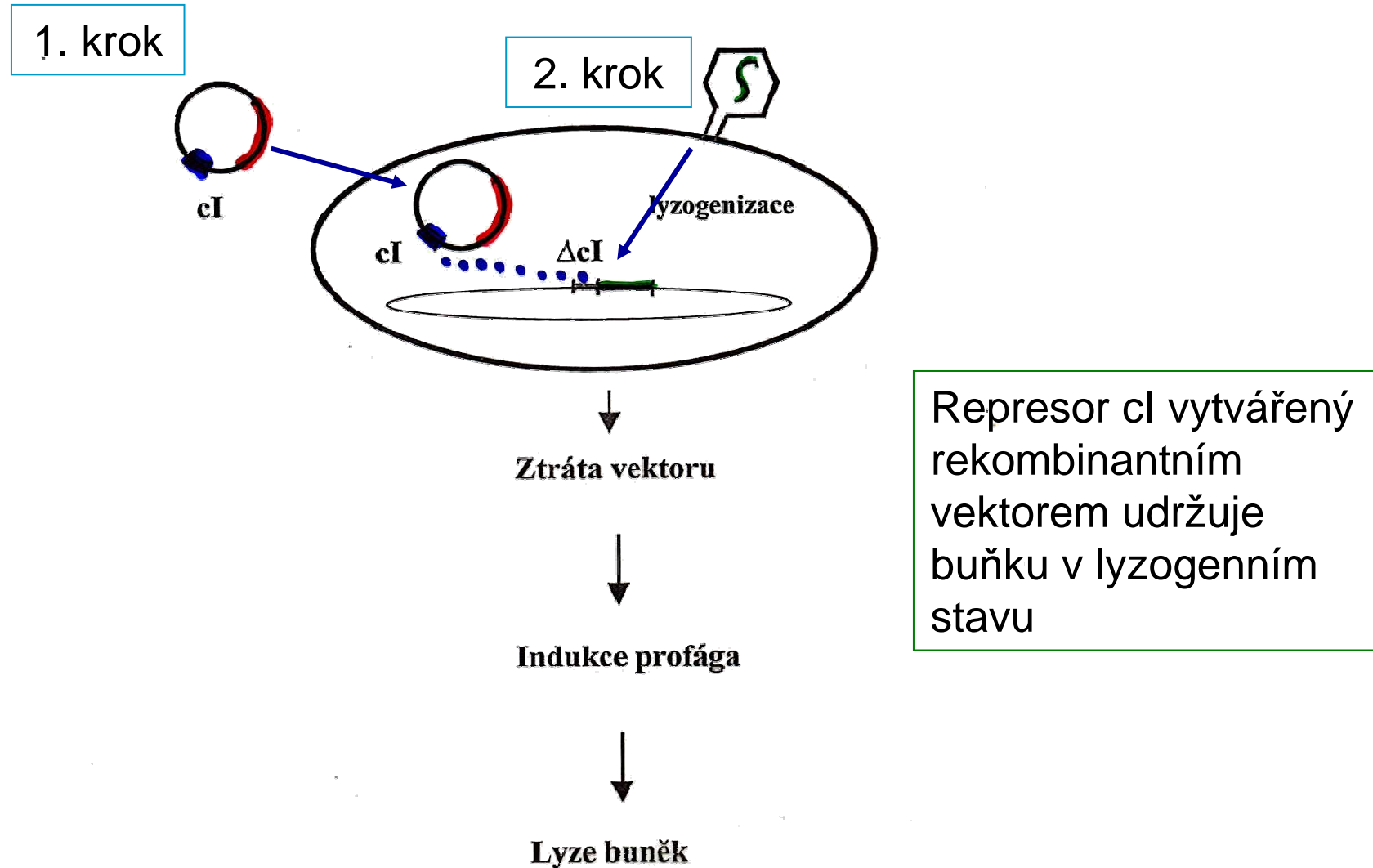
Instabilita vektorů

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.

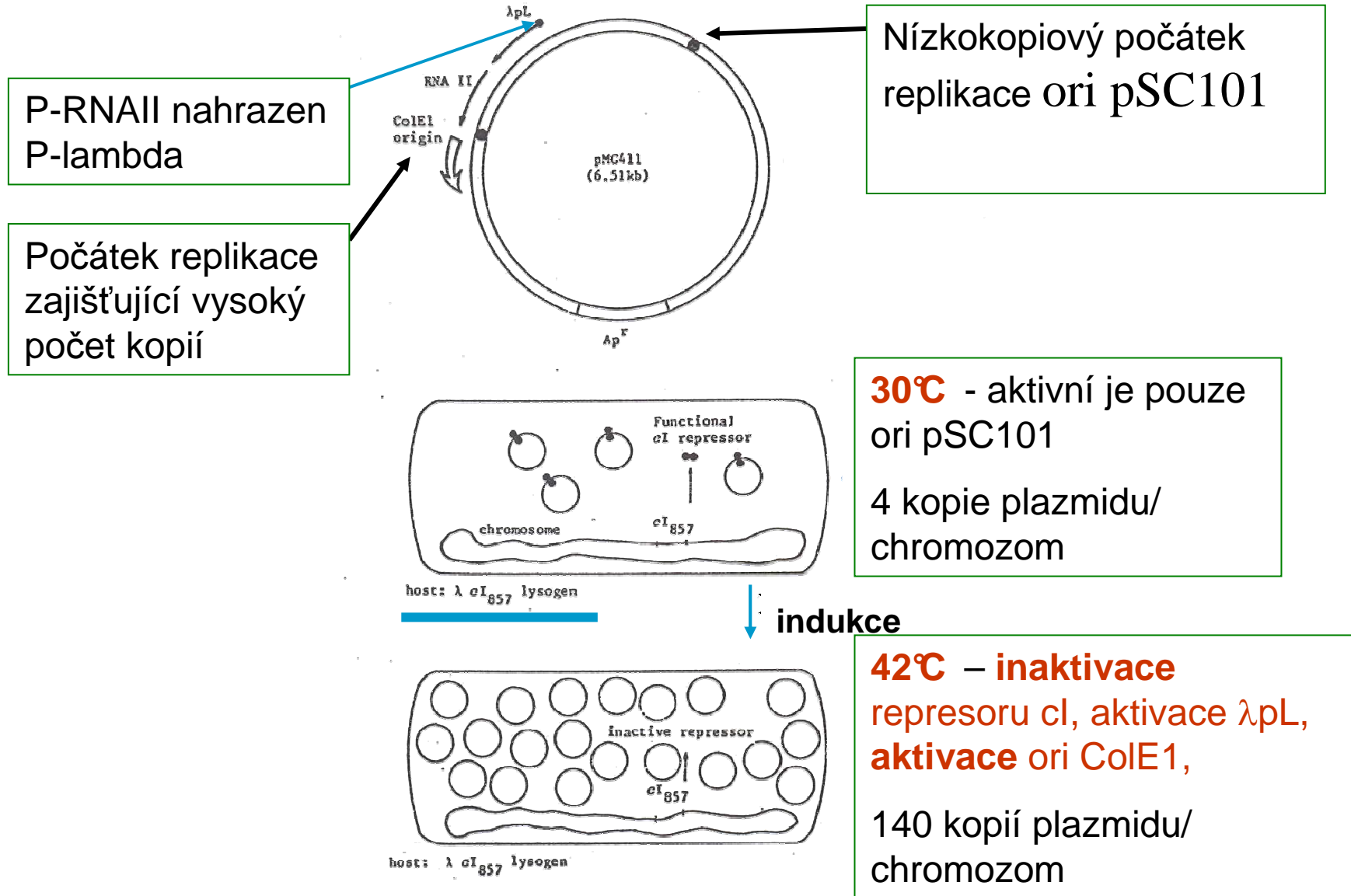
*Možnou strategií pro překonání segregační instability je **eliminace buněk**, které plazmid ztratily (např. použití genu *cl* fága v klonovacím vektoru a využití lyzogenních hostitelů)*

2. **Strukturní instabilita:** Důsledek delecí, inzercí nebo translokací v chromozomech, plazmidech nebo virech (homologní rekombinace a transpozice).

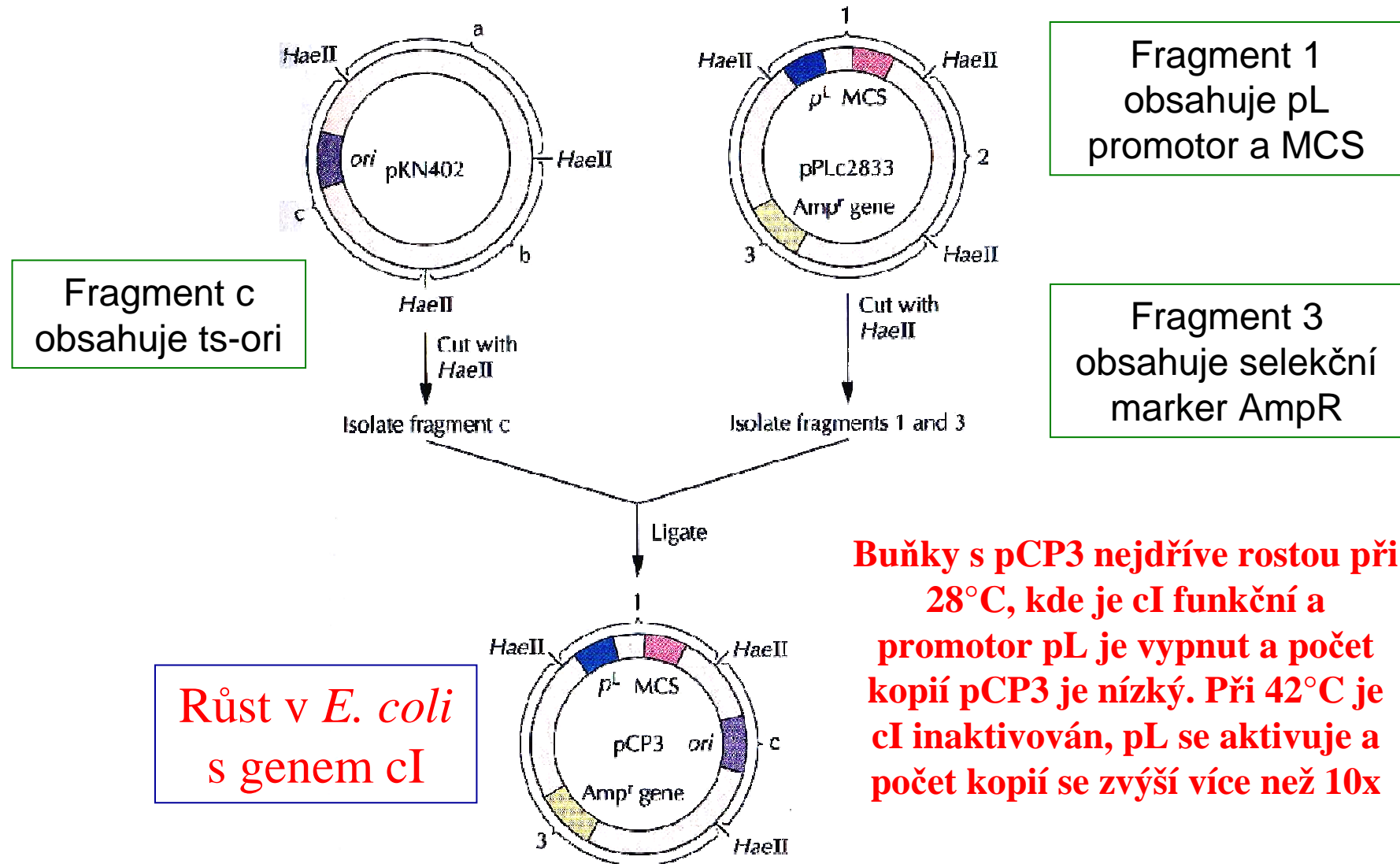
Překonání segregační instability vektoru eliminací buněk, které vektor ztratily



Vektory s dvojm počátkem replikace pro regulaci počtu kopií vektoru



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



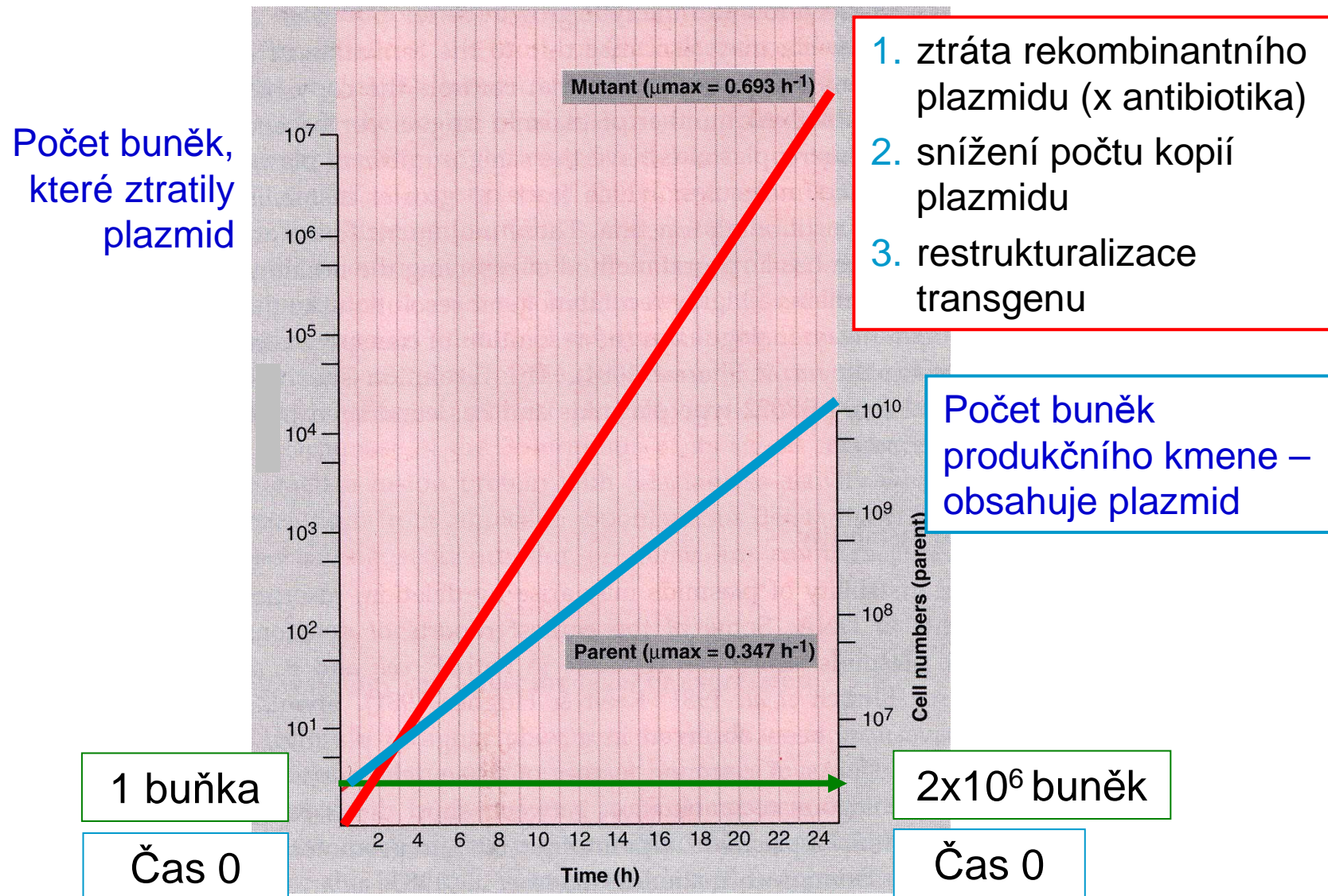
Vliv teploty na počet kopií plazmidů u tří expresních vektorů

Plazmid	Počet kopií plazmidů v buňce		Promotor p ^L
	28°C	42°C	
pKN402	82	512	No
pPLc2833	38	42	Yes
pCP3	60	713	Yes

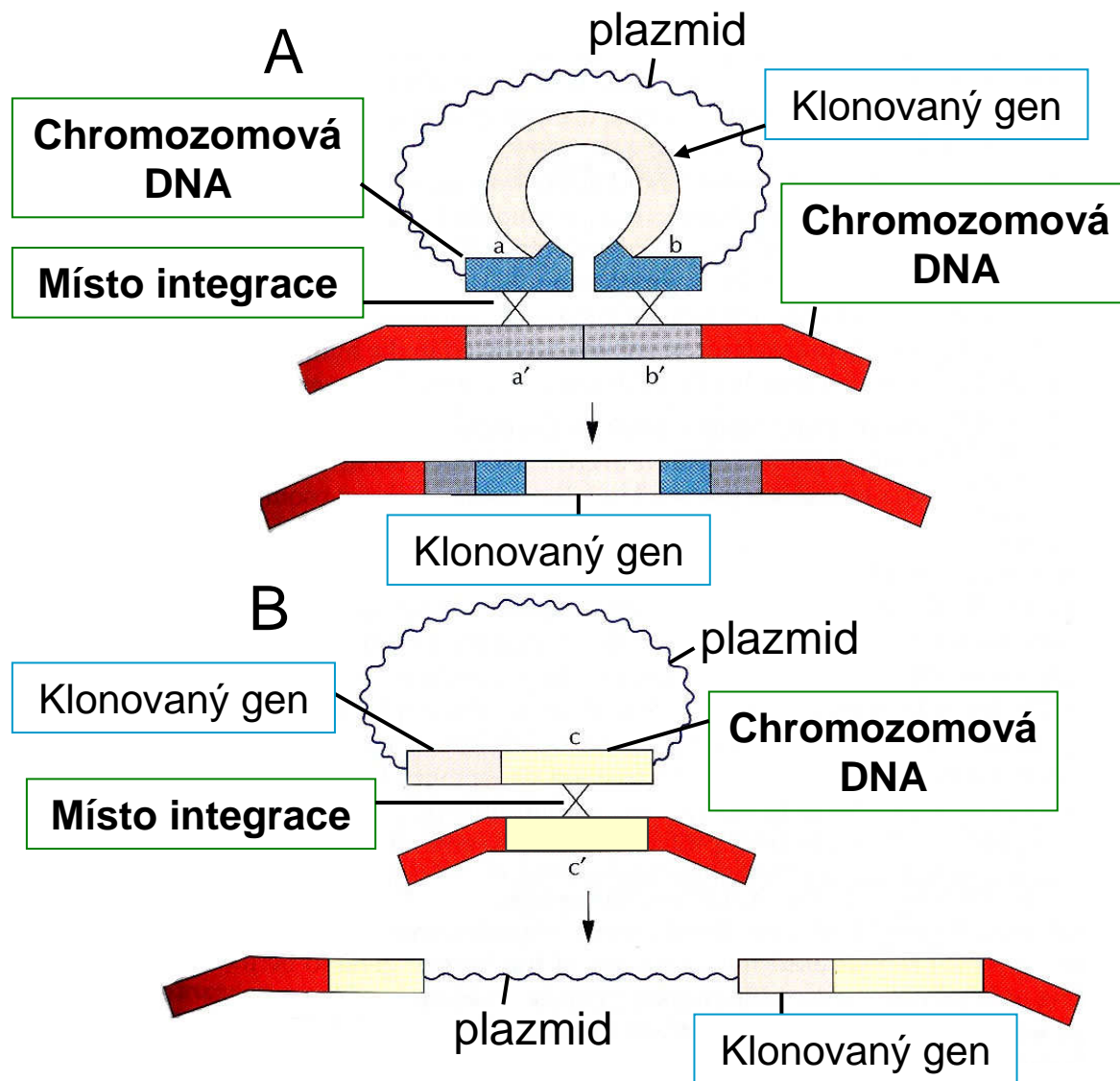
Důsledky vyplývající z přítomnosti rekombinantního plazmidu v buňkách

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



A. Klonovaný gen je vložen do sekvence klonovaného úseku chromozomu, 2 x CO vede k integraci klonovaného genu.

B. Klonovaný gen je vložen poblíž klonovaného úseku chromozomu, 1 x CO vede k integraci celého plazmidu včetně klonovaného genu.

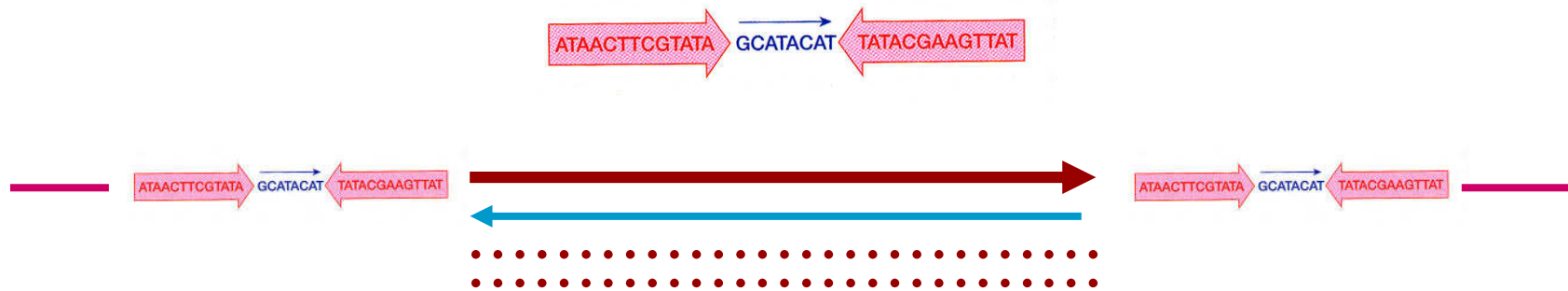
Počet kopií klonovaného genu pro α -amylázu v *B. subtilis* a její aktivita v buňkách

Počet kopií/genom	Activita (U/mL rostoucích buněk)
2	500
5	2,300
7	3,100
8	3,400
9	4,400
Multicopy plasmid (20-40 kopií)	700

geny na chromozomu

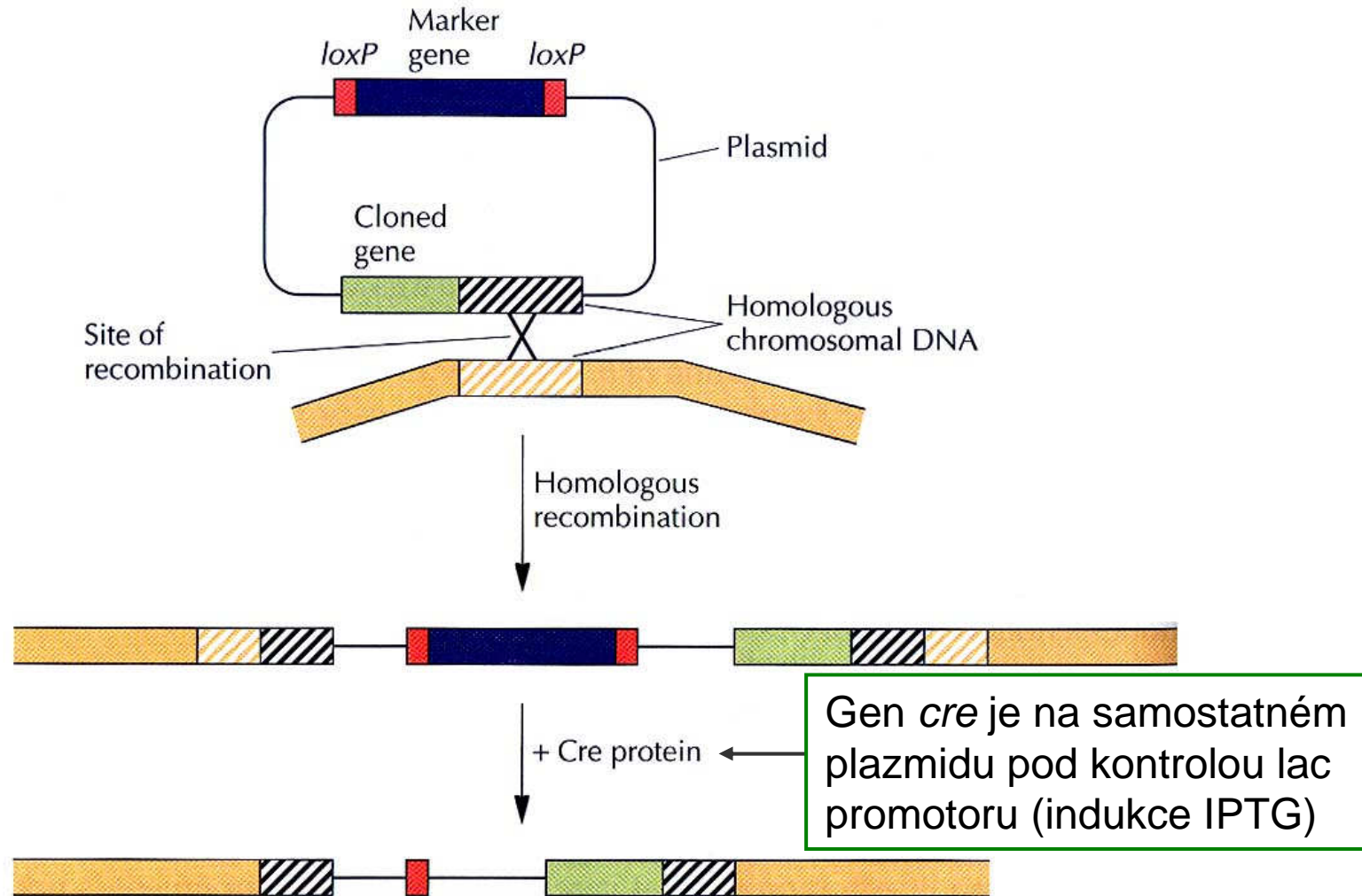
Gen pro α -amylázu z *B. amyloliquefaciens* byl vložen do oblasti pocházející z chromozomu *B. subtilis* naklonované v plazmidu *E. coli*, konstrukt byl vnesen do *B. subtilis*. Plazmid nesl rovněž geny pro rezistenci k chloramfenikolu a ampicilinu. Buňky *B. subtilis* s plazmidem byly pěstovány v prostředí 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 (laktóza x glukóza – ovlivňují inducibilní systémy - lac)**
- 2. Teplota (při nižší teplotě bývají cizí proteiny méně toxické)**
- 3. pH**
- 4. Koncentrace kyslíku**

Důkaz tvorby cizorodého produktu

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

Až 80% problému při klonování cizorodých genů spočívá v toxicitě proteinů, zbývajících 20% je způsobeno jinými faktory (např. využívání kodonů)

- Toxicita jednotlivých typů rekombinantních molekul pro hostitelské buňky:**
 - Rekombinantní DNA není obvykle toxická, pokud neobsahuje repetitivní sekvence (méně než 1% případů)**
 - Funkční rekombinantní RNA může být toxická (asi 15% případů)**
 - Toxické proteiny (více než 80% případů)**
- normálně je protein vytvářen v určitém kompartmentu buňky, během definované časové periody a v určitém množství (prostorová, časová a kvantitativní exprese)**
- rekombinantní proteiny jsou vytvářeny do nefyziologických koncentrací**
- funkce rekombinantních proteinů může být pro buňky škodlivá až toxická – pomalejší růst buněk, nižší hustota buněk**

