

PŘÍRODOVĚDECKÁ FAKULTA UNIVERSITY KARLOVY V PRAZE  
KATEDRA GENETIKY A MIKROBIOLOGIE

BAKALÁŘSKÁ PRÁCE

**Regulace exprese S10 operonu *Escherichia coli***

Zbyněk Černý

Školitel: RNDr. Ján Krahulec, Ph.D.

**2009/2010**

## Obsah

1. Úvod.....	1
2. Struktura a regulace exprese operonu S10 u G- bakterie <i>Escherichia coli</i> .....	2
2.1. Operon S10: strukturní a regulační prvky.....	2
2.2. Regulace translace ribozomálním proteinem L4.....	4
2.3. Vazebné místo proteinu L4 na mRNA.....	6
2.4. Určení vazebného místa L4.....	7
2.5. Zastavení RNA polymerázy.....	9
2.6. Terminátory ATT a ATT' a jejich funkce v závislosti na koncentraci UTP.....	10
2.7. Význam transkripčního faktoru NusA pro regulaci transkripce S10 operonu.....	11
2.8. Vazebné místo transkripčního faktoru NusA.....	13
2.9. Charakteristiky mRNA rozpoznávané proteiny L4, NusA a RNA polymerázou při předčasné terminaci transkripce.....	14
2.9.1. Deteminanty rozhodující pro vazbu proteinu L4.....	15
2.9.2. Oblast komplementárního párování bází a smyčky vlásenky HE v nepřítomnosti NusA.....	16
2.9.3. Komplementace úseků vedoucí sekvence syntetickými oligonukleotidy.....	17
3. Role proteinu L4 z <i>E. coli</i> a jiných bakteriálních organismů v regulaci operonu S10.....	19
3.1. Srovnání proteinů L4 z <i>E. coli</i> a dalších mikroorganismů.....	22
3.2. Regulace S10 operonu <i>E. coli</i> cizími proteiny L4.....	22
4. Genové klastry operonu S10 z <i>E. coli</i> u vybraných druhů mikroorganismů.....	23
4.1. Operon S10 u G+ bakterie <i>Bacillus subtilis</i> .....	23
4.2. Operon S10 u G- bakterie <i>Thermus thermophilus</i> .....	25
4.3. Operon S10 u G- <i>Vibrio cholerae</i> .....	26
4.4. Operon S10 u G- spirochety <i>Leptospira interrogans</i> .....	27
4.5. Operon S10 u G- <i>Sinorhizobium meliloti</i> .....	27
5. Závěr.....	28
6. Seznam použité literatury.....	29

### Seznam použitých zkratk:

<b>DNA</b>	deoxyribonucleic acid	deoxyribonukleová kyselina
<b>kDa</b>	kiloDalton	kiloDalton
<b>mRNA</b>	messenger ribonucleic acid	mediátorová ribonukleová kyselina
<b>RNAP</b>	RNA polymerase	RNA polymeráza
<b>r-protein</b>	ribosomal protein	ribozomální protein
<b>rRNA</b>	ribosomal ribonucleic acid	ribozomální ribonukleová kyselina
<b>UTP</b>	uridintriphosphate	uridintrifosfát

**Abstrakt:** Ribozóm Gram-negativní bakterie *Escherichia coli* se skládá ze tří různých molekul rRNA a z 56 molekul bílkovin. Jejich syntéza v exponenciálně rostoucích buňkách je koordinovaně řízena. Operon S10 kóduje 11 ribozomálních proteinů. Regulace exprese proteinů S10 operonu probíhá nejenom na úrovni translace mechanismem autogenní kontroly běžným u operonů ribozomálních proteinů v *E. coli*, při kterém jeden z proteinů kódovaných operonem má jak strukturní tak i regulační funkci, ale také na úrovni transkripce mechanismem předčasného ukončení transkripce – atenuace.

**Abstract:** A Gram-negative *Escherichia coli* ribosome particle is comprised of 3 different rRNA molecules and 56 protein molecules. Their synthesis in exponentially growing cultures is coordinated. S10 operon codes for 11 ribosomal proteins. The expression of ribosomal proteins clustered in S10 operon is regulated autogenously, not only at the level of translation, as is common with other *E. coli* ribosomal protein operons where one of the proteins in the operon has both structural and regulatory functions, but also at the level of transcription by way of premature transcription termination – attenuation.

**Klíčová slova:** antiterminace transkripce, atenuace, autogenní zpětná vazba, determinanta, genový klastr, L4, NusA, operátor, operon, pauza/pozastavení RNA polymerázy, promotor,  $\rho$  faktor, regulace genové exprese, regulátor, ribozomální proteiny, ribozóm, S10, terminace transkripce, terminátor, transkripční faktor, translační spřažení, vedoucí sekvence mRNA

**Keywords:** autogenous feedback, attenuation, determinant, gene cluster, gene expression regulation, L4, mRNA leader, NusA, operator, operon, promoter, regulator, rho factor, ribosomal proteins, ribosome, RNA polymerase pause, S10, terminator, transcription antitermination, transcription factor, transcription termination, translational coupling

## 1. Úvod

*Escherichia coli*, původním názvem *Bacterium coli*, byla identifikována v roce 1885 německým pediatrem a bakteriologem Theodorem Escherichem. Je to fakultativně anaerobní gram-negativní tyčinkovitá bakterie žijící v tlustém střevě teplokrevných živočichů. Je jedním z nejdůležitějších zástupců střevní mikroflóry a její přítomnost je nezbytná pro správný průběh trávicích procesů ve střevě (Feng, Weagant & Grant, 2002).

*E. coli* je pro svou relativní jednoduchost a nenáročnost významným nástrojem v biotechnologiích a slouží jako modelový organizmus v genetice a mikrobiologii (Wikipedie, 2010). Tato bakterie patří mezi intenzivně studované organizmy mimo jiné i v oblasti regulace exprese ribozomálních proteinů. Tato práce si klade za cíl vytvořit přehled poznatků týkající se regulace exprese operonu S10 kódujícího geny pro ribozomální proteiny (r-proteiny), se zaměřením především na bakterii *E. coli*.

Ribozóm *E. coli* se skládá ze tří různých molekul rRNA (5S, 16S a 23S rRNA) a z 56 molekul ribozomálních proteinů (Lindahl & Zengel, 1986; Arnold & Reilly, 1999; Wikipedie, 2010). Jednotlivé ribozomální proteiny jsou unikátní až na tři výjimky: jeden protein se nachází v obou podjednotkách ribozómu (S20 and L26), L7 a L12 jsou acetylované a metylované formy téhož proteinu, a L8 je komplex L7/L12 a L10. Protein L31 navíc existuje ve dvou formách, v plné velikosti 7.9 kDa a jako fragment o velikosti 7.0 kDa. Syntéza ribozomálních proteinů v exponenciálně rostoucích buňkách je koordinovaně řízena. Geny pro všechny tři typy rRNA jsou součástí operonu *rrn*, který se v genomu *E. coli* vyskytuje v sedmi kopiích (Kaczanowska & Rydén-Aulin, 2007). Geny pro ribozomální proteiny jsou naproti tomu organizovány v několika různých operonech přítomných na chromozómu v jedné kopii. Je tedy nutné, aby syntéza těchto proteinů byla v buňce koordinována s množstvím dostupných rRNA.

Hlavním mechanismem pro řízení syntézy ribozomálních proteinů a jejich koordinaci s dalšími r-proteiny a dostupnými rRNA v *E. coli* je autogenní zpětná vazba (Zengel & Lindahl, 1994; Kaczanowska & Rydén-Aulin, 2007). Tento mechanismus se uplatňuje při represí translace mRNA. Jeden z r-proteinů daného operonu se neváže pouze na cílovou rRNA v průběhu sestavování ribozómu, ale také na svou vlastní mRNA v místě nazvaném operátor. Operátor může být umístěn jak směrem upstream od prvního genu, tak mezi dvěma geny operonu. Při nadbytku represorového r-proteinu v buňce se jeho volné molekuly vážou na operátor a inhibují translaci proteinů kódovaných mRNA operonu. Represor může kompetovat s ribozómy o mRNA, nebo svou vazbou na mRNA vyvolá vznik struktur, které zabrání přístupu ribozomů k místu počátku translace.

Výsledkem této regulace je vyvážená exprese r-proteinových operonů navzájem i vzhledem k syntéze rRNA.

Při mechanismu autogenní zpětné vazby se uplatňuje strukturní homologie mezi vazebným místem represorových r-proteinů na rRNA a jejich vazebným místem na mRNA. Molekuly regulačních proteinů, které buňka vyprodukuje v nadbytku, a které nenaleznou volné vazebné místo na rRNA, se váží na mRNA a inhibují její translaci; v případě operonu S10 je kontrola translace navíc podpořena i kontrolou na úrovni transkripce (Lindahl, Archer & Zengel, 1983; Lindahl & Zengel, 1986; Freedman, Zengel, Archer & Lindahl, 1987; Lindahl, Sor, Archer, Nomura & Zengel, 1990; Lindahl & Zengel, 1990). Zatím bylo identifikováno deset represorových r-proteinů: S1, S4, S7, S8, S15, S20, L1, L4, L10, a L20, regulujících následující operony (v uvedeném pořadí): S1,  $\alpha$ , *str*, *spc*, S15, S20, L11, S10, L10, a L35 (Kaczanowska & Rydén-Aulin, 2007). Výše zmíněná strukturní homologie byla například prokázána u operonů regulovaných proteiny S4 a S7 (Nomura, Yates, Dean & Post, 1980) S8 (Olins & Nomura, 1981) a L1 (Branlant, Krol, Machatt & Ebel, 1981; Gourse, Thurlow, Gerbi & Zimmermann, 1981).

Koregulace všech genů v rámci jednoho operonu je dosaženo mechanismem translačního spřažení genů r-proteinů společně s dalšími mechanismy, jako je například degradace mRNA (Nomura, Gourse & Baughman, 1984; Zengel & Lindahl, 1994). Translačním spřažením může buňka také zvýšit rychlost translace mRNA, když je koncentrace r-proteinů v buňce nízká – geny umístěné proximálně na mRNA způsobí změnu sekundární struktury mRNA a usnadní tak translaci genů umístěných distálně. Alternativně, ribozóm účastnící se translace mRNA může být přenesen přímo z proximálního genu na iniciační kodón distálního genu (Asato, 2005). Koregulace nezávislých operonů r-proteinů v genomu *E. coli* probíhá spřažením translace všech mRNA nesoucích geny pro r-proteiny s jednou velkou reakcí, sestavováním ribozómu. Přestože konkrétní mechanismy represe specifických operonů mají své odlišnosti, obecný princip regulace je jednoduchý a elegantní (Nomura, 1999).

## **2. Struktura a regulace exprese operonu S10 u G- bakterie *Escherichia coli***

### **2.1. Operon S10: strukturní a regulační prvky**

Operon S10 u *E. coli* je situován mezi operony *str* a *spc*, a zahrnuje 11 genů: S10, L3, L4, L23, L2, L22, S19, S3, L16, L29 a S17. Jako regulační protein exprese S10 operonu byl identifikován L4, jak v *in vitro* experimentech, tak *in vivo* (Yates & Nomura, 1980; Lindahl *et al.*, 1983).

Expresce r-proteinů v *E. coli* je autogenně regulována především na translační úrovni (Nomura *et al.*, 1980; Zengel, Archer, Freedman & Lindahl, 1984). Ovšem v případě operonu S10 regulační

protein L4, který je kódován v pořadí třetím genem operonu a skládá se z 201 aminokyselin, inhibuje jak translaci, tak transkripci (Lindahl *et al.*, 1983; Freedman *et al.*, 1987; Zengel & Lindahl, 1990a; Zengel & Lindahl, 1994).

Kromě své regulační funkce při řízení transkripce a translace S10 operonu vykazuje protein L4 také na ribozómech nezávislou funkci regulace degradace RNA v *E. coli*. Bylo prokázáno, že L4 interaguje s místem mimo katalytickou doménu RNázy E, čímž inhibuje endoribonukleolytickou aktivitu RNázy E *in vitro* a stabilizuje mRNA, která je cílovým substrátem RNázy E *in vivo*. Analýzy mRNA stability ukázaly, že průměrný poločas existence transkriptů účastnících se stresové odpovědi buňky *E. coli* se zvýšil při zvýšené expresi L4 (Singh, Chang, Lin, Averina, Kaberdin & Lin-Chao, 2009).

Kontrola transkripce operonu S10 *E. coli* probíhá mechanismem atenuace. Protein L4 v kombinaci s transkripčním faktorem NusA stimuluje předčasnou terminaci transkripce na nepřekládané vedoucí sekvenci operonové mRNA asi 30 bází proximálně od iniciačního kodónu prvního genu (Lindahl *et al.*, 1983; Freedman *et al.*, 1987; Zengel & Lindahl, 1990b). Mediátorová RNA v oblasti vedoucí sekvence se skládá do sekundární struktury (viz. Obr. 2) připomínající terminátor nezávislý na  $\rho$ -faktoru (Stelzl *et al.*, 2003). Tato vlásenková struktura obsahuje devítibázovou sekvenci, která se vyskytuje rovněž ve vazebném místě pro L4 na 23S rRNA. Kromě toho se ve směru 3' od této sekvence nachází další sekvence homologická s 23S rRNA (Zengel & Lindahl, 1992).

Při experimentálním studiu účinku proteinu L4 na terminaci transkripce operonu S10 v *in vitro* transkripčním systému bylo zjištěno, že samotný protein L4 není pro terminaci transkripce dostačující, a že je nutný ještě jeden faktor – protein NusA (Zengel & Lindahl, 1990a; Zengel & Lindahl, 1992; Stelzl *et al.*, 2003). RNA polymeráza rozeznává signál pro zastavení transkripce obsažený ve vedoucí sekvenci mRNA, ale bez dodatečné stimulace pouze krátce zastaví. Protein NusA je faktor, který rozeznává zastavenou RNA polymerázu a stabilizuje ji. Pozastavený transkripční komplex doplněný o NusA protein pak umožňuje vazbu L4 proteinu, který dále zvyšuje jeho stabilitu (Zengel & Lindahl, 1992) a vede k vytvoření „pre-terminačního komplexu“ a nakonec k terminaci transkripce. Ztráta schopnosti regulovat transkripci prostřednictvím L4 může tedy být výsledkem defektu buď na úrovni zastavení RNA polymerázy závislého na přítomnosti NusA nebo na úrovni stabilizace zastaveného komplexu proteinem L4.

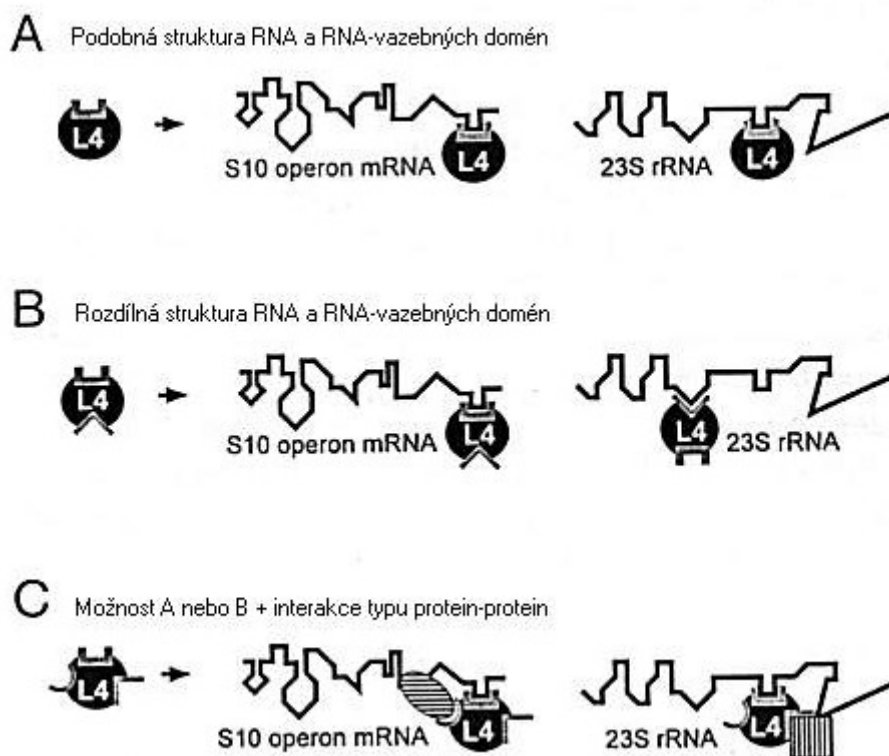
Při regulaci translace S10 operonu dochází k inhibici translace prvního genu operonu a k rozšíření této inhibice na distální geny operonu. mechanismus translačního spřažení (coupling) nedovolí v takovém případě translaci genu, pokud předchozí gen není také translatován. Například v intercistronové oblasti mezi geny pro r-proteiny S10 a L3 je obrácená repetice schopná tvořit

sekundární strukturu. Tato struktura může inhibovat translaci L3; ke translaci následujícího genu může dojít pouze pokud ribozóm překládající distální část genu pro S10 znemožní tvorbu této sekundární struktury (Oppenheim & Yanofsky, 1980; Yates & Nomura, 1980; Lindahl, Archer, McCormick, Freedman & Zengel, 1989).

## **2.2. Regulace translace ribozomálním proteinem L4**

Kanonický model autogenní regulace syntézy ribozomálních proteinů v *E. coli* je založen na předpokladu, že regulační protein rozeznává strukturně podobné cíle na rRNA a na mRNA, která tento protein kóduje (Nomura *et al.*, 1980). Tento model vychází z hypotézy, že regulátor má vazebnou doménu, která se účastní inkorporace proteinu do ribozómu a zároveň je schopná vázat se na vedoucí sekvenci svého vlastního operonu.

Bylo však prokázáno, že tyto dvě vlastnosti lze geneticky od sebe oddělit. Podařilo se izolovat mutantní formy proteinu L4, které se inkorporovaly do ribozómu, ale byly defektní ve své regulační funkci, a naopak. Nepovedlo se ovšem oddělit od sebe oba typy regulace, tedy regulaci transkripce a translace, a zatím neexistují údaje, které by naznačovaly zda je něco takového možné.



**Obr. 1: Modely rozpoznání 23S rRNA a S10 mRNA proteinem L4 (Li, Lindahl & Zengel, 1996).**

- A: Jednoduchý kanonický model – stejné vazebné místo pro mRNA i rRNA.
- B: L4 má dvě na sobě nezávislá RNA vazebná místa pro mRNA a rRNA.
- C: Funkce L4 vyžaduje kromě RNA-L4 interakcí také interakce mezi L4 a dalšími faktory, například ribozomálními proteiny nebo proteiny transkripčního aparátu.

Na Obr. 1 jsou příklady modelů, které by mohly vysvětlit způsob chování L4. Obrázek A představuje klasický jednoduchý model autogenní regulace proteinem L4. Podle obrázku B by L4 mohl mít dvě nezávislé domény, jednu pro interakci s rRNA a druhou pro mRNA. Tato vazebná místa mohou být zcela oddělena, ale také se mohou částečně překrývat. Další možnost spočívá v tom, že funkce proteinu L4 vyžaduje pro své splnění nejenom interakce typu L4-RNA, ale také přítomnost dalších faktorů, například ribozomálních proteinů nebo proteinů transkripčního aparátu.

Operon S10 je dosud jediným operonem pro ribozomální proteiny u kterého bylo možno oddělit schopnost regulace jeho regulačního proteinu od funkce stavební jednotky ribozómu. Genetická analýza r-proteinu S8, který reguluje *spc* operon, ukázala, že mutantní formy proteinu izolované pro jejich sníženou schopnost regulace mají také nízkou afinitu pro 16S rRNA. Protein S8 tedy zřejmě obsahuje doménu společnou pro regulační i stavební funkci a jeho chování lze charakterizovat klasickým kanonickým modelem pro autogenní kontrolu.

Byly identifikovány aminokyselinové motivy společné pro mnoho RNA vazebných proteinů (Burd & Dreyfuss, 1994; Koonin, Bork & Sander, 1994), ale žádný z nich nebyl nalezen v sekvenci

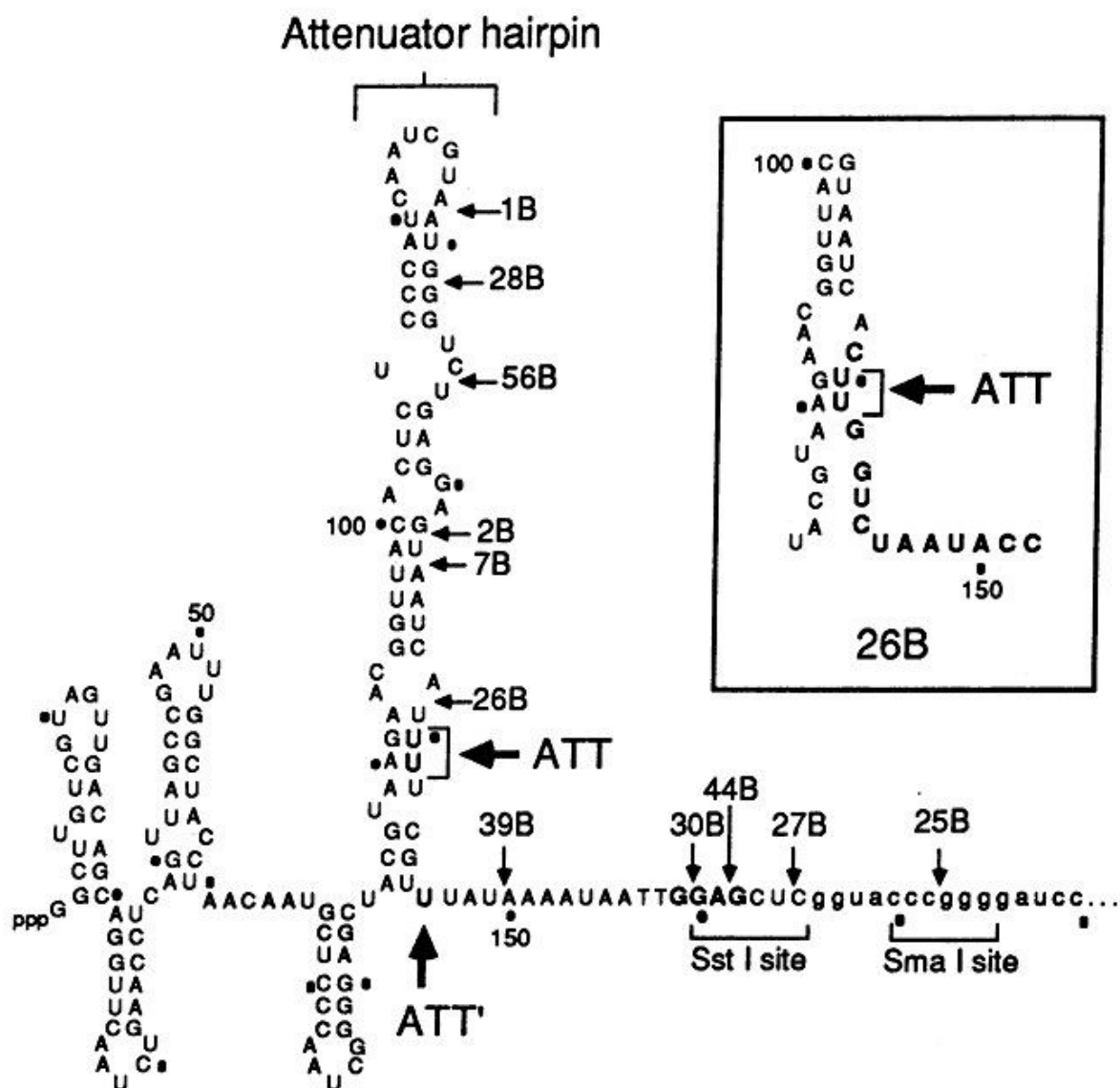


L4 *E. coli* (Zengel, Vorozheikina, Li & Lindahl, 1995). Nejmenší delece, která znemožňuje inkorporaci L4 do ribozómu, obsahuje oblast se třemi lyziny dobře zachovanými v eubakteriálních druzích. Přestože je lákavé považovat tyto lyzinové zbytky za esenciální determinanty pro vestavení L4 do ribozómu prostřednictvím vazby na rRNA, nelze vyloučit jejich účast v kritických interakcích mezi proteiny při stavbě ribozómu.

### **2.3. Vazebné místo proteinu L4 na mRNA**

Regulace transkripce S10 operonu se děje prostřednictvím atenuace, při které dochází ke stimulaci předčasné terminace transkripce v nepřekládané vedoucí sekvenci asi 140 nukleotidů od začátku transkripce. Tento proces začne pozastavením RNA polymerázy v určité části atenuátoru (viz Obr. 2, Obr. 5 a Obr. 6); konkrétně se jedná o sekvence čtyř a tří uridinových zbytků situovaných v sestupné části atenuační vlásenky (viz Obr. 2). Následuje stabilizace pozastaveného transkripčního komplexu obecným transkripčním faktorem NusA a nato proteinem L4. Nakonec dojde k uvolnění atenuovaného transkriptu a RNA polymerázy z templátové DNA. Celkem je pro terminaci nutných pouze prvních 149 bází z celé 172 nukleotidové vedoucí sekvence.

Sekundární struktura vedoucí sekvence operonu S10 sestává z šesti vlásenek označených HA až HG spojených krátkými jednořetězcovými sekvencemi. Největší vlásenka pojmenovaná HE (atenuační vlásenka) je esenciální pro kontrolu transkripce i translace. Jako další vlásenka důležitá pro transkripční kontrolu byla identifikována čtvrtá vlásenka v pořadí – HD. Pokusy s operonem S10, ve kterém byla tato vlásenka odstraněna, ukázaly, že pro regulační funkci L4 je HD naprosto nezbytná. Z dalších experimentů vyplynulo, že významná je především sekundární struktura vlásenky, ne tolik už její nukleotidová sekvence. Naopak proximální tři vlásenky, HA, HB a HC vykazují velmi malý nebo vůbec žádný vliv na regulaci transkripce pomocí L4 (viz Obr. 2).



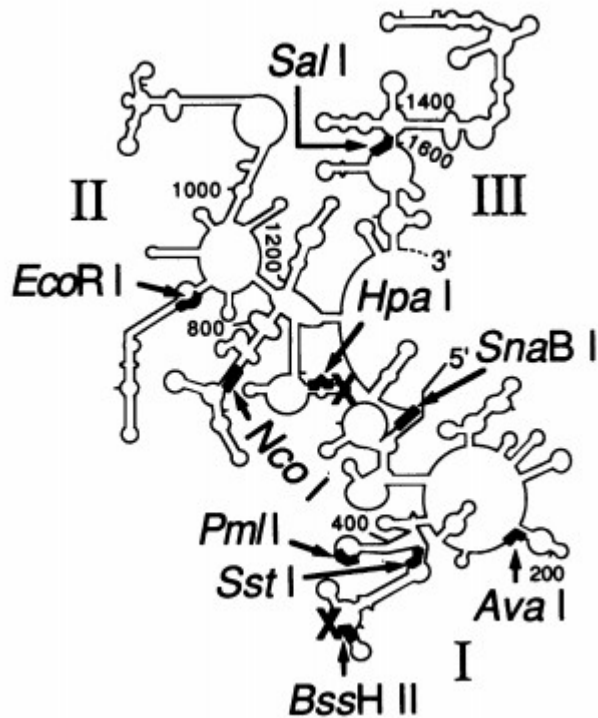
Obr. 2: Vedoucí sekvence S10 operonu *E. coli* (Zengel & Lindahl, 1990).

Sekundární struktura nepřekládané vedoucí sekvence S10 operonu. Vlásoková struktura účastníci se regulace transkripce je označena. Terminační místa ATT a ATT' jsou označena většími šipkami. Shine-Dalgarnova sekvence pro gen S10 (GGAG) je vyznačena tučně. Sekvence vyskytující se rovněž ve 23S rRNA v navrhovaném místě vazby L4 zahrnují nukleotidy 117-125, 146-150, 152-155, a 159-163. Na výřezu vpravo je načrtnuta spodní část atenuační vlásenky delečně mutované vedoucí sekvence S10 označené 26B. Sekvence downstream od delece je vyznačena tučně.

Pro regulaci translace jsou postradatelné všechny čtyři proximální vlásenky, což znamená, že L4 buď rozeznává společnou sekvenci distálně od HD (tedy vzestupnou část vlásenky HE), ale vyžaduje HD pro kontrolu transkripce anebo vedoucí sekvence obsahuje dvě vazebná místa pro L4.

#### 2.4. Určení vazebného místa L4

Jedna z klíčových charakteristik autogenní kontroly S10 operonu proteinem L4 je přesná lokalizace regulačního místa. Vezmeme-li v úvahu primární funkci L4 jako proteinu tvořícího



**Obr. 3: Sekundární struktura tří proximálních domén 23S rRNA (Zengel & Lindahl, 1993).**

Místa vazby L4 na 23S rRNA v intaktní 50S podjednotce byla určena cross-linking experimenty, a jsou vyznačena 'X' (poblíž BssH II restrikčního místa v doméně I a Hpa I místa v doméně II).

součástí větší ribozomální podjednotky a jeho specifitu vzhledem k vlastnímu operonu, lze se domnívat, že jeho vazebné místo alespoň zčásti leží ve vedoucí sekvenci operonu. I současný obecný model autogenní kontroly zprostředkované r-proteiny předpokládá, že regulační protein rozeznává cíl na vlastní mRNA, který svou strukturou napodobuje vazebné místo na rRNA.

V případě L4 bylo zjištěno, že se protein váže ke 110-ti nukleotidové oblasti domény I 23S rRNA (viz Obr. 3). Delece dvou bází v této oblasti 23S rRNA poblíž místa, na které se v intaktních ribozómech váže L4, už silně snižuje její afinitu k L4 (Zengel & Lindahl, 1992).

Zvláštní v tomto případě je, že sekundární struktura S10 mRNA (viz Obr. 2) nevykazuje žádnou nápadnou podobnost se sekundární strukturou odpovídající oblasti na 23S rRNA (viz Obr. 3). Je možné, že sekundární struktura nascentní 23S rRNA, ke které se L4 váže v rané fázi procesu sestavování ribozómu, je odlišná od struktury zralé 23S rRNA a podobá se více S10 mRNA. Druhou možností je, že domény L4 interagující s S10 mRNA nemusí přesně souhlasit s vazebnými doménami pro 23S rRNA. Protein by tak mohl rozpoznat struktury RNA, které mají podobnou primární sekvenci, ale liší se v prostorové konfiguraci.

Byla analyzována také kinetika kompetice o protein L4 mezi 23S rRNA a ternárním transkripčním komplexem. Z výsledků vyplývá, že vazba L4 se zastaveným transkripčním

komplexe není tak stabilní jako jeho vazba na 23S rRNA. Přidáním dalších molekul 23S rRNA do transkripčního systému *in vitro* dojde k vyvázání L4 a k ukončení stabilizace transkripčního komplexu tímto proteinem zprostředkované, třebaže účinek je méně výrazný než když jsou molekuly kompetitoru přítomny už před elongační fází transkripce. Tyto pokusy také ozřejmily, že efekt proteinu L4 na transkripční komplex spočívá v jeho stabilizaci a nikoliv například v přechodné modifikaci některé z komponent komplexu (např. RNA polymerázy) a následném uvolnění z komplexu (Zengel & Lindahl, 1993).

## 2.5. Zastavení RNA polymerázy

Zastavení RNA polymerázy je důležitým krokem v regulaci transkripce. I když je z velké části transkripce regulována při iniciaci, elongace je u prokaryot i eukaryot často narušována sekvenčně specifickými pauzami. Pohyb RNA polymerázy po matrici DNA je tedy diskontinuální, což umožňuje přímo ovlivnit expresi jakéhokoliv genu v genomu organismu (Herbert *et al.*, 2006). Během těchto zastavení dojde ke zpomalení syntézy nascentního mRNA řetězce 100-1000x (Yager & von Hippel, 1987). Buňka používá rozdílných mechanismů zastavení v závislosti na životních podmínkách a dostupnosti živin v buňce. Tento regulační krok je zdokumentován například u *E. coli* při spřažení transkripce s translací prostřednictvím synchronizace syntézy mRNA a pohybem ribozomu v kódující oblasti mRNA (Wong, 2004), při terminaci transkripce nezávislé na  $\rho$ -faktoru (Yager & von Hippel, 1987; Platt, 1994), antiterminaci závislé na faktoru Q u fága  $\lambda$  (Grayhack, Yang, Lau & Roberts, 1985) nebo při atenuaci vázané na translaci (Landick & Yanofsky, 1987; Yager & von Hippel, 1987). Specifický signál pro zastavení není ve všech případech zcela jasný, třebaže často před místem zastavení dojde k vytvoření stabilní vlásenkové struktury (Levin & Chamberlin, 1987; Yang & Gardner, 1991). Mezi další faktory ovlivňující zastavení RNA polymerázy patří sekvence DNA ve směru 3' od zastavení, dostupnost volných nukleotidů a přítomnost transkripčních faktorů jako je například protein NusA (Levin & Chamberlin, 1987; Yager & von Hippel, 1987).

Předpokládá se existence dvou odlišných typů sekvenčně specifického pozastavení RNA polymerázy: pauzy třídy I, jako v případě vedoucí sekvence *his* operonu, které závisí na tvorbě vlásenky s pěti páry bází tvořícími stonek vlásenky a osminukleotidovou smyčkou, a pauzy třídy II jako například *ops* (operon polarity suppressor), při kterých se RNA polymeráza vrátí zpátky po DNA a zablokuje aktivní místo pomocí nascentní mRNA (Artsimovitch & Landick, 2000).

V případě terminace závislé na  $\rho$ -faktoru zastavení transkripčního komplexu umožní  $\rho$ -faktoru dosáhnout ternární komplex a uvolnit mRNA z matrice (Jin, Burgess, Richardson & Gross, 1992).

U terminace nezávislé na  $\rho$  může při zastavení polymerázy dojít ke konformační změně, která připraví komplex pro následnou terminaci (Yang & Roberts, 1989). Korelace mezi zastavením RNA polymerázy a terminací transkripce byla pozorována i v jiných systémech. Pozastavení transkripčního komplexu hraje důležitou roli rovněž při jiných dějích než je terminace transkripce, například antiterminace u  $\lambda$  fága závisí na schopnosti Q proteinu zabránit pozastavení RNA polymerázy (Yang & Roberts, 1989).

## 2.6. Terminátory ATT a ATT' a jejich funkce v závislosti na koncentraci UTP

Pauza RNA polymerázy ve vedoucí sekvenci operonu S10 v *E. coli* je ve standardních podmínkách *in vitro* transkripčního systému nestabilní a přechodná. Pouze při velmi nízkých koncentracích UTP (5  $\mu$ M) je tato pauza trvalejšího charakteru (Sha, Lindahl & Zengel, 1995a). Ze skutečnosti, že koncentrace UTP je kritickým parametrem ve stabilitě zastaveného ternárního komplexu v nepřítomnosti NusA, lze usuzovat na pravděpodobného kandidáta odpovědného za tuto interakci – sekvenci čtyř zbytků U v terminačním místě ATT, kde dochází k ukončení transkripce *in vivo*, a tří v místě ATT' účinkujícího rovněž v *in vitro* systémech (viz Obr. 2). Hlavní důvod pro existenci těchto uridinových zbytků nejspíš spočívá v jejich významu pro terminaci transkripce; rU-dA interakce jsou relativně nestabilní a tím usnadňují uvolnění nascentního řetězce mRNA z matrice (Farnham & Platt, 1980).

Bylo experimentálně prokázáno, že delece dosahující na svém 5' konci téměř k terminačnímu místu ATT' nemá žádný výrazný vliv na kontrolu transkripce zprostředkovanou proteinem L4, i když způsobí malý pokles v množství *in vitro* transkriptů ukončených na ATT'. Delece zasahující do oblasti terminačních míst ATT a ATT' způsobí zeslabení, ale nikoliv inaktivaci atenuátoru, částečně díky skutečnosti, že sekvence oligo U terminátorové oblasti ATT odstraněná touto delecí byla funkčně nahrazena dvěma uridinovými zbytky distálně od 3' konce delece (Zengel & Lindahl, 1992). V tomto případě RNA polymeráza rozpozná místo pro zastavení, ale stabilita komplexu je velmi snížena. Rozpoznávací místo pro L4 je touto delecí nedotčeno.

Uridinové zbytky v oblasti ATT/ATT' hrají svou roli při zastavení RNA polymerázy v nepřítomnosti NusA, ale pro rozpoznání místa pro zastavení je třeba také spolupráce dalších signálů. Tento závěr je doložen skutečností, že polymeráza stabilně nezastavuje na třech uridinových zbytcích nacházejících se v koncové smyčce vlásenky HC ve směru 5' od atenuační vlásenky. Dalším signálem by mohla být struktura smyčky v atenuační vlásence HE. Bylo prokázáno, že podobné struktury bohaté na G+C páry v jiných transkripčních jednotkách instruují RNA polymerázu, aby zastavila nebo i ukončila transkripci (Yager & von Hippel, 1987; Lodish,

Berk, Zipursky, Baltimore & Darnell, 1999). Podobná kumulace signálů vedoucích k zastavení RNA polymerázy zahrnujících jak RNA vlásenku, tak sekvence sousedící s místem zastavení polymerázy, byla rovněž pozorována u vedoucích sekvencí *trp* a *his* operonů (Chan & Landick, 1989; Lee, Phung, Stewart & Landick, 1990; Chan & Landick, 1993).

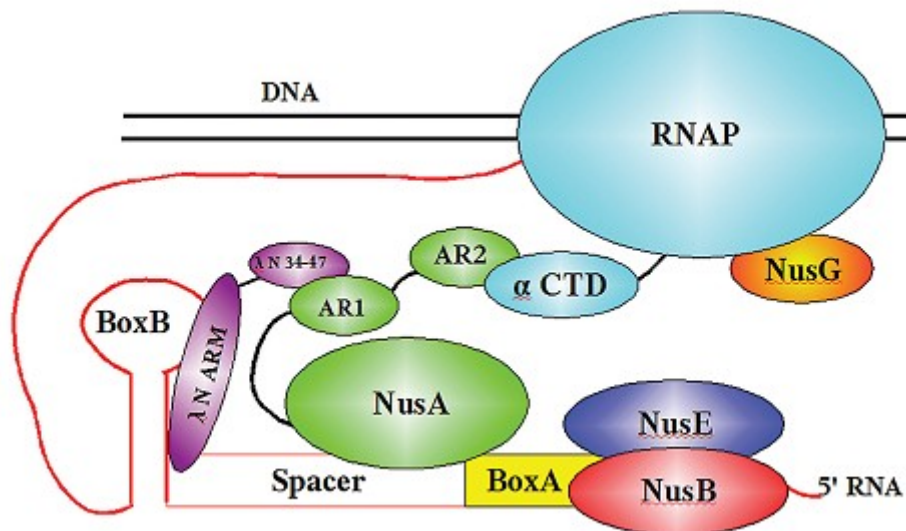
Příspěvek vlásenkové struktury ke stabilitě pauzy RNA polymerázy tedy závisí na koncentraci UTP. Pro zastavení polymerázy při koncentraci UTP 5 $\mu$ M není dokonalá struktura vlásenky třeba, jak ukazují pokusy s mutantními vedoucími sekvencemi.

## 2.7. Význam transkripčního faktoru NusA pro regulaci transkripce S10 operonu

Studie genové exprese bakteriofága  $\lambda$  okamžitě po infekci bakteriální buňky ukázala, že transkript produkovaný z promotoru P<sub>L</sub> kóduje virový protein nazvaný N, který brání terminaci transkripce na terminátorech t<sub>L</sub> a t<sub>R</sub>. Jakmile koncentrace proteinu N dostatečně stoupne, syntéza dlouhých RNA transkriptů zbývajících částí genomu fága  $\lambda$  může pokračovat. Protein N vyžaduje přítomnost specifické sekvence *nut* (N-utilization) mezi promotorem a terminátorem, aby mohl bránit ukončení transkripce (Lodish *et al.*, 1999).

Izolace mutantních kmenů *E. coli*, které nejsou schopné antiterminace prostřednictvím proteinu N, a následně jsou resistantní vůči infekci bakteriofágem  $\lambda$ , vedla k identifikaci několika proteinů nazvaných N-utilization substances (Nus), které N protein vyžaduje ke své funkci (viz Obr. 4). Tyto proteiny – označené NusA, NusB, NusE a NusG – mají další funkce v neinfikovaných buňkách *E. coli*. Například NusA je elongační faktor, a NusE (jiný název pro r-protein S10) je jeden z proteinů tvořících malou podjednotku ribozomu *E. coli*. Tyto proteiny spolupracují na zabránění terminace transkripce genů pro ribozomální RNA (Lodish *et al.*, 1999).

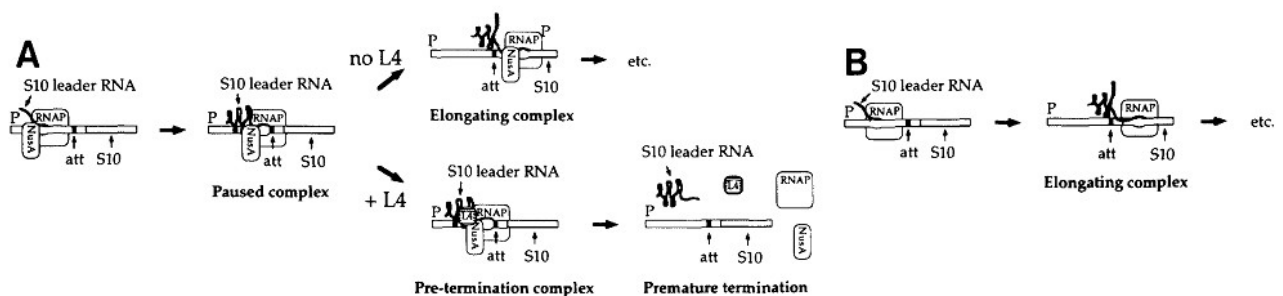
Jiný příklad regulace genové exprese *E. coli* založený na mechanismu terminace transkripce jsou ribozomální operony *rrn*, jejichž transkripce je částečně regulována potlačením terminace (Prasch, Jurk, Washburn, Gottesman, Wöhrl & Rösch, 2009). Antiterminace u *rrn* operonů je zprostředkována rozpoznávanou RNA sekvencí (AT) distálně od tandemu promotorů, blízko 5' konce pre-rRNA transkriptu, a účastní se jí řada faktorů, včetně NusA, NusB, NusE a NusG, které modifikují RNA polymerázu pozastavenou v AT místě. Takto modifikovaná RNA polymeráza nereaguje při transkripci na terminátory závislé na  $\rho$ -faktoru. U obou antiterminačních systémů – *rrn* operony a bakteriofág  $\lambda$  – je modifikovaná RNA polymeráza schopna přepisovat přes více terminátorů, ale *rrn* antiterminace je efektivní pouze v případě terminátorů závislých na  $\rho$ -faktoru zatímco  $\lambda$  antiterminační komplexy jsou vysoce rezistentní vůči  $\rho$ -závislým i  $\rho$ -nezávislým terminátorům (Condon, Squires & Squires, 1995).



Obr. 4: Antiterminace pomocí N proteinu fága  $\lambda$  a proteinů *E. coli* (Prasch *et al.*, 2009).

Znázornění interakce RNAP s různými faktory důležitými pro antiterminaci. Jakmile RNAP nasyntetizuje  $\lambda nut$  RNA, NusE a NusB se vážou na *boxA* a NusA na *spacer*. Vazba NusA je stabilizována interakcí domény AR2 s C-terminální doménou  $\alpha$  podjednotky RNAP.  $\lambda N$  protein se váže na AR1 doménu NusA a vytváří tak na svém N-konci helixovou strukturu, která rozeznává *boxB*.

Transkripční faktor NusA hraje rovněž důležitou roli při kontrole transkripce S10 operonu u *E. coli*. Předpokládá se, že NusA buď indukuje konformační změnu v zastaveném komplexu, která umožní RNA polymeráze interagovat s L4, nebo umožní funkci proteinu L4 přímou cestou interakcí typu protein-protein mezi NusA a L4.



Obr. 5: Atenuace S10 operonu zprostředkovaná proteinem L4 (Zengel & Lindahl, 1992).

- A: Elongační komplex v přítomnosti NusA – RNAP pozastaví na ATT terminačním místě.
- B: Elongační komplex v nepřítomnosti NusA – signál pro pauzu RNAP je ignorován.

NusA se silně váže na RNA polymerázu způsobem, který vylučuje současnou vazbu  $\sigma$ -faktoru (Greenblatt & Li, 1981) a zároveň mění konformaci polymerázy v místě vzájemným 3' konec vznikající mRNA (Zhang & Hanna, 1994). Tato konformační změna částečně vysvětluje některé účinky NusA na transkripci, jako například stabilizaci zastavení polymerázy na určitém typu vlásenek nebo zesílení terminace u  $\lambda t_{r2}$ , *rrnB* T1 *E. coli* (Greenblatt, 1981; Schmidt &

Chamberlin, 1987) a na *trp* terminátoru *E. coli* (Farnham, Greenblatt & Platt, 1982). Efekt NusA na chování RNA polymerázy vůči potenciálním terminačním strukturám v nascentní mRNA je tedy schopen vysvětlit stabilizaci transkripčního komplexu na atenuátoru vedoucí sekvence S10 operonu (Zengel & Lindahl, 1992). V porovnání s jinými *in vitro* transkripčními systémy, ve kterých stimulace zastavení polymerázy nebo terminace transkripce pomocí faktoru NusA není esenciálním krokem, v systému S10 operonu je přítomnost NusA nutná.

V případech, kdy NusA modifikuje chování RNA polymerázy v přítomnosti dalších proteinů, se nabízejí dvě možnosti jeho účinku. NusA může hrát roli bezprostředního adaptoru vázajícího další faktory nebo vnášet konformační změny do RNA polymerázy a tak usnadňovat její interakci s ostatními faktory. Například při antiterminaci zprostředkované proteinem N se NusA váže na spacer mezi vazebným místem pro NusB/NusE faktory nazvaným BoxA na 5' konci RNA a vazebným místem pro protein N nazvaným BoxB (Greenblatt & Li, 1981; Prash *et al.*, 2009). NusA interaguje prostřednictvím své AR1 domény s N proteinem navázaným na vlásenkovou strukturu vzniklou v BoxB a svou doménou AR2 je v kontaktu s C koncovou doménou  $\alpha$  podjednotky RNA polymerázy (viz Obr. 4). Obdobná role NusA se předpokládá při terminaci závislé na  $\rho$  faktoru (Schmidt & Chamberlin, 1984).

## 2.8. Vazebné místo transkripčního faktoru NusA

V přítomnosti proteinu NusA dojde prakticky ke 100% rozpoznání signálu pro zastavení ve vedoucí sekvenci (Zengel & Lindahl, 1992). Atenuační vlásenka HE obsahuje veškeré potřebné informace, aby RNA polymeráza mohla zastavit; čtyři předcházející vlásenky lze odstranit bez jakýchkoliv následků pro tuto pauzu (Zengel & Lindahl, 1992). V tomto ohledu je nejdůležitější částí vlásenky HE její smyčka – RNA polymeráza ztratí schopnost reagovat na přítomnost NusA pokud dojde k narušení koncové smyčky, ať už prostřednictvím substituce bází nebo hybridizací s komplementárním oligonukleotidem.

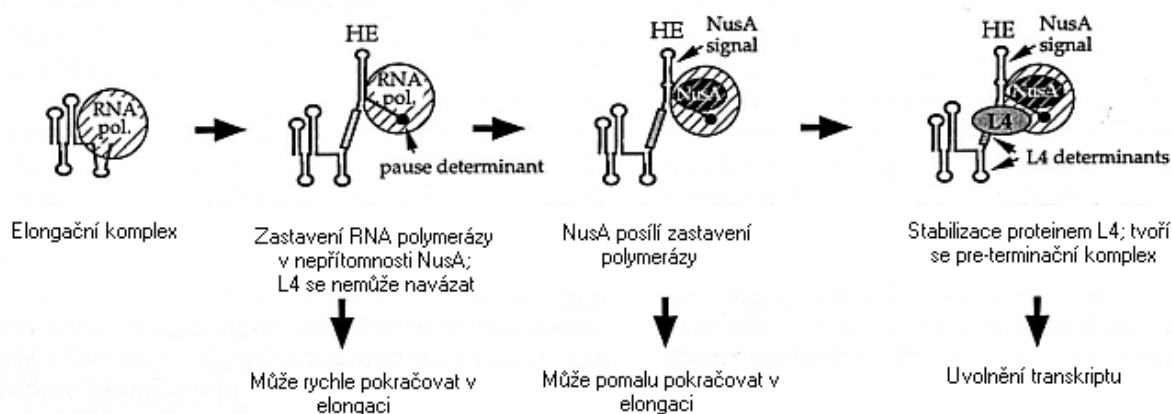
Struktura horní části HE vlásenky je schopna přesně kontrolovat místo, kde RNA polymeráza zastaví během transkripce. I u jiných operonů je tato struktura důležitá: například některé bodové mutace ovlivňující konformaci atenuační vlásenky v *thr* operonu způsobí, že RNA polymeráza zastaví o několik bází dále od svého obvyklého místa (Lynn, Kasper & Gardner, 1988). Sekvence v nižší vzestupné části atenuační vlásenky se zřejmě neúčastní interakce s NusA, protože oligomery komplementární s touto oblastí nemají žádný vliv na zastavení polymerázy v přítomnosti NusA proteinu. Ani sekvence přímo v místě, kde polymeráza zastavuje, není esenciální, jak vyplývá z pokusů, ve kterých byla tato sekvence odstraněna. I v tomto případě došlo ke stimulaci zastavení RNA polymerázy pomocí NusA ve stejné pozici jako u divokého typu atenuátoru. Výsledky těchto



delečních pokusů naznačují, že vlásenková struktura a možná i sekvence bezprostředně směrem 3' od ní poskytují v přítomnosti NusA signál pro RNA polymerázu k zastavení transkripce v oblasti přibližně 20 bází od horní smyčky vlásenky HE, bez ohledu na sekvenci v tomto místě. Přesto se RNA polymeráza nemůže na tomto místě stabilně zastavit v nepřítomnosti specifických signálů (jako například oblast bohatá na U). Podobné podmínky vyžadují i jiné atenuátory. Například k zastavení RNA polymerázy ve vedoucí sekvenci *his* a *trp* operonů *E. coli* přispívá jak struktura vlásenky, tak samotná oblast kolem tohoto místa (Lynn *et al.*, 1988), podobně je tomu také u atenuační vlásenky *his* operonu u *S. typhimurium* (Chan & Landick, 1989). Někteří autoři se domnívají, že zastavení RNA polymerázy závislé na okolní sekvenci by mohl být jeden z prvních kroků při terminaci transkripce (Chan & Landick, 1989; Lee *et al.*, 1990; Chan & Landick, 1993).

## 2.9. Charakteristiky mRNA rozpoznávané proteiny L4, NusA a RNA polymerázou při předčasné terminaci transkripce

Bylo experimentálně prokázáno, že atenuační vlásenka HE ve vedoucí sekvenci S10 operonu obsahuje dostatek informací pro navázání NusA proteinu, zatímco pro vazbu L4 je navíc důležitá i oblast ve směru 5' konce mRNA. Je zřejmé, že na kontrole atenuace pomocí L4 se podílejí alespoň tři odlišné odpovědi RNA polymerázy: vnitřně specifická pauza nezávislá na NusA, dále krok vyžadující protein NusA a nakonec stabilizace pozastaveného transkripčního komplexu proteinem L4. Těmto třem krokům odpovídají tři signály pro zastavení RNA polymerázy, které byly identifikovány v rámci vedoucí sekvence S10 operonu *E. coli* (Zengel & Lindahl, 1992; Sha, Lindahl & Zengel, 1995b). Třebaže signály pro tyto tři odpovědi jsou lokalizovány v tomtéž atenuačním místě, jejich charakteristiky se liší (viz Obr. 6).



Obr. 6: RNA determinanty vedoucí sekvence S10 operonu *E. coli* (Sha *et al.*, 1995b).

Signály na nepřekládané části S10 mRNA se zapojují do procesu přechasně terminace transkripce postupně, přičemž každý následující signál podporuje a zesiluje účinek přechodícího signálu.

### 2.9.1. Deteminanty rozhodující pro vazbu proteinu L4

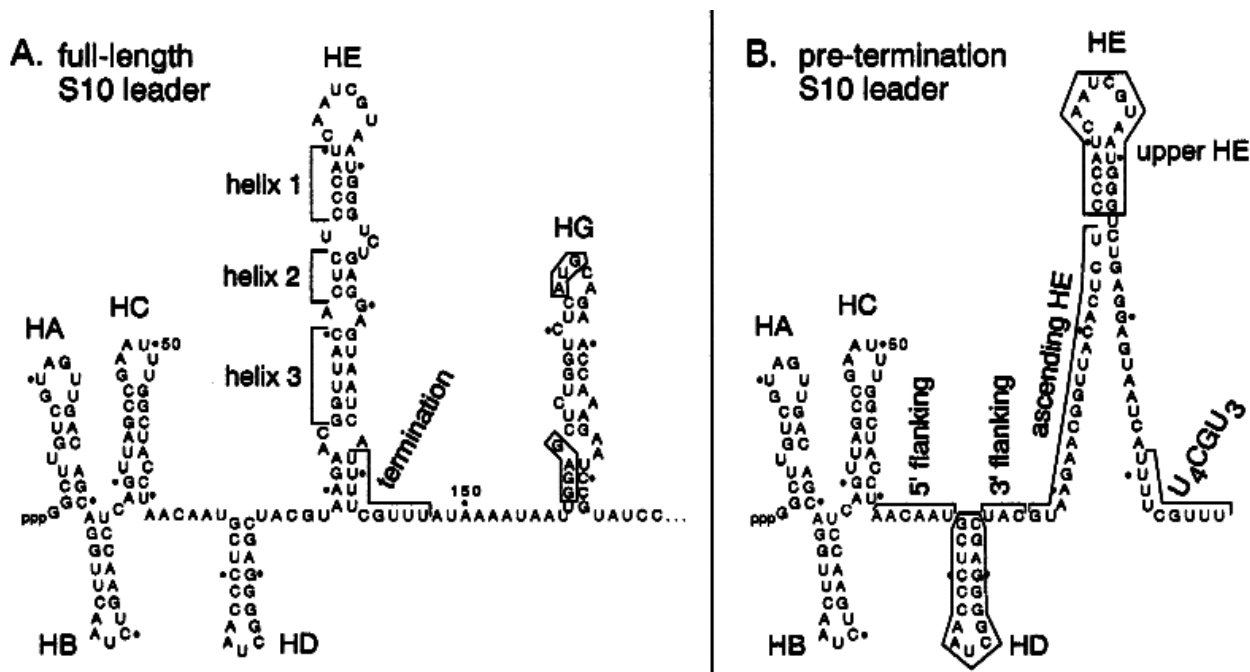
Z pokusů *in vivo* vyplynulo, že delece 72 bází na 5' konci S10 vedoucí sekvence zruší vliv L4 proteinu na inhibici transkripce (Freedman *et al.*, 1987). Série dalších experimentů v *in vitro* transkripčním systému objasnila, že delece eliminující čtyři proximální vlásenky nemá žádný vliv na vazbu NusA proteinu, ale zruší stabilizaci zastaveného ternárního komplexu pomocí L4 (Zengel & Lindahl, 1992). Z výsledků je patrné, že signály vedoucí k zastavení RNA polymerázy v přítomnosti NusA i bez něho leží v oblasti atenuační vlásenky HE. Z dalších genetických a fylogenetických studií vyplynulo, že tři proximální vlásenky ve vedoucí sekvenci jsou pro transkripční regulaci pomocí L4 postradatelné. Naopak přítomnost vlásenky HD je nutná pro funkci L4 (Zengel *et al.*, 1995). Pro uplatnění regulace transkripce pomocí proteinu L4 jsou tedy zapotřebí sekvence obsažené jak ve vlásence HD, konkrétně smyčka vlásenky, tak v atenuační vlásence HE.

Přestože regulace S10 operonu *E. coli* zprostředkovaná proteinem L4 je velmi specifická, genetická analýza S10 vedoucí sekvence ukázala, že existuje výrazná sekvenční flexibilita v 80 nukleotidové oblasti nutné pro tuto odpověď (Zengel, Sha & Lindahl, 2002). Prvky mRNA schopné zajistit optimální regulaci proteinem L4 zahrnují: 1) nepozměněnou horní část atenuační HE vlásenky se smyčkou 4-9 nukleotidů se zřejmě libovolnou sekvencí; 2) spojku 20-29 nukleotidů zahrnující nukleotidy downstream od vlásenky HD a vzestupnou část vlásenky HE s relativně flexibilní sekvencí, i když nahrazení 18 nukleotidového segmentu cizí sekvencí kontrolu zruší; 3) nezměněnou vlásenku HD s šesti páry bází dlouhým stonkem a smyčkou mezi 5-7 nukleotidy dlouhou s relativně flexibilní sekvencí; 4) a specifické nepárové sekvence obklopující HD vlásenku (viz Obr. 7). Tato data naznačují, že regulace S10 operonu by mohla záviset na redundatních vazebných místech s vícero interakcemi mezi mRNA, komplexem RNA polymerázy a proteinem L4 (Zengel *et al.*, 2002).

Zatímco oblasti ve vedoucí sekvenci mRNA nutné pro funkci L4 už známe, stále nelze s jistotou odlišit skutečné vazebné místo L4 od pomocných sekvencí či struktur.

Výsledky experimentů s oligonukleotidy komplementárními s určitými úseky mRNA ukázaly, že vazbu proteinu L4 je možné zablokovat oligomery vázajícími se na vzestupnou část atenuační vlásenky. Zatímco RNA polymeráza je zastavena v ATT/ATT' místě, 12 bází bezprostředně nad místem zastavení RNA polymerázy netvoří intramolekulární páry (Yager & von Hippel, 1987) a je tak umožněna tvorba pouze vrchní části vlásenky HE. Vedoucí sekvence S10 zaujme tzv. „pre-terminační“ konformaci, při které RNA polymeráza pozastaví v U-bohaté oblasti za vlásenkou HE a nukleotidy spodních dvou třetin vlásenky HE nejsou v nascentní vedoucí sekvenci spárovány (viz Obr. 7). Báze ve vzestupné části vlásenky musí tedy při zastavení polymerázy zaujmout přechodnou

alternativní strukturu (snad i jednořetězcovou). Tato struktura by mohla reprezentovat skutečný cíl pro vazbu L4. Alternativou je, že vzestupná část vlásenky se pouze nepřímo účastní kontroly transkripce prostřednictvím L4, zatímco vlastním cílem vazby L4 je vlásenka HD. Nicméně L4 může stále zprostředkovat kontrolu translace i když vlásenka HD chybí (Freedman *et al.*, 1987). Proto pokud je vlásenka HD cílem vazby L4 při kontrole transkripce atenuací, potom musí mít tento ribozomální protein další vazebné místo pro umožnění kontroly translace (viz Obr. 1).



Obr. 7: Sekundární struktury vedoucí sekvence S10 mRNA (Zengel *et al.*, 2002).

A: Základní konformace vedoucí sekvence S10 (Adaptováno od Shen, Zengel & Lindahl, 1988).

B: Předpokládaná pre-terminační konformace vedoucí sekvence S10 s RNA polymerázou navázanou na U4CGU3 oblast.

### 2.9.2. Oblast komplementárního párování bází a smyčky vlásenky HE v nepřítomnosti NusA

Při studiu jednotlivých funkčních komponent atenuační vlásenky pomocí substituce bází a následného narušení jejich párování se ukázalo, že mutace rušící párování ve stopce vlásenky HE eliminuje jak transkripční, tak translační kontrolu proteinem L4 (Freedman *et al.*, 1987), zatímco kompenzující mutace, které párování bází obnovují, umožní kontrolu translace, ale nikoliv transkripce.

Mutace lokalizované ve smyčce vlásenky HE naopak eliminují translační kontrolu, na kontrolu transkripce mají malý nebo vůbec žádný vliv. Narušené párování bází v komplementární části vlásenky způsobuje, že se RNA polymeráza za nepřítomnosti L4 ani nezastaví, popřípadě jen na velmi krátkou dobu. Kompenzační mutace obnovující párování bází obnoví také signál pro

zastavení RNA polymerázy. Z toho vyplývá, že oblast komplementárního párování bází v atenuační vlásence HE je rozhodujícím činitelem při zastavení RNA polymerázy. Smyčka atenuační vlásenky se naopak na zastavení RNA polymerázy nepodílí. Při pokusech se substitucemi bází ve smyčce vlásenky se RNA polymeráza zastaví – vedoucí sekvence S10 operonu se substitucí ve vlásence HE reaguje na přítomnost proteinu L4 normálně (Zengel *et al.*, 1995).

### **2.9.3. Komplementace úseků vedoucí sekvence syntetickými oligonukleotidy**

Celá řada pokusů o charakterizaci S10 operonu a jeho chování při transkripci byla založena na aplikaci oligonukleotidů komplementárních k různým oblastem vedoucí sekvence S10 operonu. Tyto oligonukleotidy se mohou párovat se vznikajícím řetězcem mRNA a tak narušit jeho sekundární strukturu. Tento přístup tak poskytuje vhodný způsob pro otestování významu specifických sekundárních struktur RNA (Zengel & Lindahl, 1996).

Oligomery použité v těchto pokusech byly navrženy tak, aby se párovaly s vlásenkou HD a se vzestupnou částí vlásenky HE (příčemž dojde pouze k narušení spodní části vlásenky, ale smyčka zůstane neporušená), další oligonukleotidy byly komplementární ke smyčce atenuační vlásenky HE. Jako negativní kontrola byl použit “univerzální primer” pro M13, který s vedoucí sekvencí S10 operonu neinteraguje.

Oligonukleotid, který narušil strukturu vlásenky HD, neměl prakticky žádný vliv na zastavení RNA polymerázy, ale znemožnil stabilizaci zastaveného komplexu pomocí L4. Oligonukleotidy narušující vzestupnou část vlásenky HE vykazovaly stejný efekt – nebránily zastavení polymerázy, ale vyřadily vliv L4. Je tedy zřejmé, že primární a sekundární struktura v oblasti vedoucí sekvence S10 operonu, která zahrnuje vlásenku HD a vzestupnou část atenuační vlásenky HE, není důležitá pro zastavení RNA polymerázy stabilizované proteinem NusA, ale je esenciální pro uplatnění stabilizačního vlivu proteinu L4.

Na druhou stranu, oligomery, které narušují strukturu smyčky vlásenky HE, zároveň ruší i stabilizační vliv proteinu NusA. Podle předpokladů, protein L4 nemá v tomto případě žádný vliv na transkripční komplex, protože stabilizace RNA polymerázy proteinem NusA je podmínkou pro funkci L4. Tyto pokusy tedy potvrdily závěr, že horní struktura atenuační vlásenky HE je rozhodujícím činitelem pro stabilizaci transkripčního komplexu v přítomnosti NusA proteinu. Její vliv na samotné zastavení RNA polymerázy je zanedbatelný.

Tyto experimenty dále ukázaly, že ke stabilizaci zastaveného transkripčního komplexu v přítomnosti NusA prostřednictvím proteinu L4 je nutná přítomnost vlásenky HD a vzestupné části vlásenky HE. Jak horní část vlásenky HE, tak oblast sousedící s místem zastavení polymerázy nemají na stabilizační efekt proteinu L4 vliv.

Je zajímavé, že ačkoliv interakce NusA s vedoucí sekvencí časově předchází akci proteinu L4, signál pro vazbu L4 je lokalizován proximálně od signálu pro NusA. Vazebné místo na mRNA rozpoznávané proteinem L4, které leží asi 50 bází ve směru 5' od místa, kde RNA polymeráza zastavuje, se může zapojit do atenuace pravděpodobně díky své trojrozměrné struktuře. Předpokládá se, že struktura mRNA se formuje takovým způsobem, aby signál pro vazbu L4 byl přístupný v okolí zastaveného transkripčního komplexu.

*In vitro* pokusy s transkripčním systémem S10 operonu naznačují, že tři typy zastavení polymerázy, které se účastní atenuace, tvoří hierarchii. První zastavení RNA polymerázy na specifickém místě vedoucí sekvence S10 operonu se může při nízké koncentraci UTP uskutečnit i bez NusA. Jakmile RNA polymeráza na okamžik zastaví svůj postup, může dojít k navázání NusA proteinu, který tak prodlužuje a stabilizuje zastavení transkripčního komplexu. Následně se k tomuto modifikovanému komplexu může přidat i protein L4 a tím dochází k další stabilizaci a k vytvoření pre-terminačního komplexu (Zengel & Lindahl, 1992).

I když je jasné, že zastavení polymerázy hraje v kontrole transkripce důležitou roli, stále nejsou k dispozici detaily rozhodující mezi zastavením a ireverzibilním terminačním krokem (Yager & von Hippel, 1987; Chan & Landick, 1993). Například signál, který vede k zastavení RNA polymerázy s velkou afinitou a pevností, v sobě ještě nemusí nést informaci pro terminaci transkripce (populární model terminace transkripce nezávislé na  $\rho$ -faktoru je založen na stabilní vlásence bohaté na G+C páry (signál pro zastavení) následované 3' koncem bohatým na U, který zprostředkuje uvolnění RNA z matrice (Farnham & Platt, 1980). Rozhodující roli při terminaci hrají interakce RNA s odpovídajícími vazebnými místy na RNA polymeráze (Surratt, Milan & Chamberlin, 1991; Altmann, Solow-Cordero & Chamberlin, 1994).

Zastavený ternární komplex obsahující L4 je velmi stabilní *in vitro*, ale vysoká koncentrace UTP může opět indukovat elongaci transkripce (Zengel & Lindahl, 1992). K terminaci transkripce dochází velmi zřídka, v tom se systém liší od atenuace probíhající *in vivo*. Pravděpodobně v *in vitro* systému chybí dosud neznámý dodatečný faktor, který by zprostředkoval uvolnění RNA z matrice. Nicméně výsledky měření transkripce *in vivo* před a po indukované nadprodukcí L4 potvrzují realizaci terminace transkripce a ne pouhého zastavení RNA polymerázy (Lindahl *et al.*, 1983; Freedman, Zengel & Lindahl, 1985). Třebaže se nepodařilo detekovat uvolnění atenuovaného transkriptu z DNA, pravděpodobnost tohoto kroku v *in vivo* systému je velmi vysoká.

Určité typy vlásenek bohaté na G+C páry mohou indukovat konformaci, která usnadňuje uvolnění RNA z templátu a mohla by tak být jedním z intermediátů terminace transkripce. V S10 operonu *E. coli* tuto roli hraje horní část atenuační vlásenky HE. V případech jiných transkripčních systémů se ale jedná o vlásenky nacházející se 6-10 bází proximálně od konce transkripce. Pokud je

vzdálenost vlásenky od 3' konce nascentní RNA větší než 10 nukleotidů, pak taková vlásenka představuje spíše signál pro zastavení transkripce než pro její ukončení, částečně proto, že nemá takový vliv na konfiguraci 3' oblasti nascentní RNA (Yager & von Hippel, 1987; Lee *et al.*, 1990; Chan & Landick, 1993). Jelikož vlásenka HE ve vedoucí sekvenci S10 operonu je lokalizována více než 15 nukleotidů od místa terminace, je třeba účasti dodatečných faktorů, jako například proteinu L4.

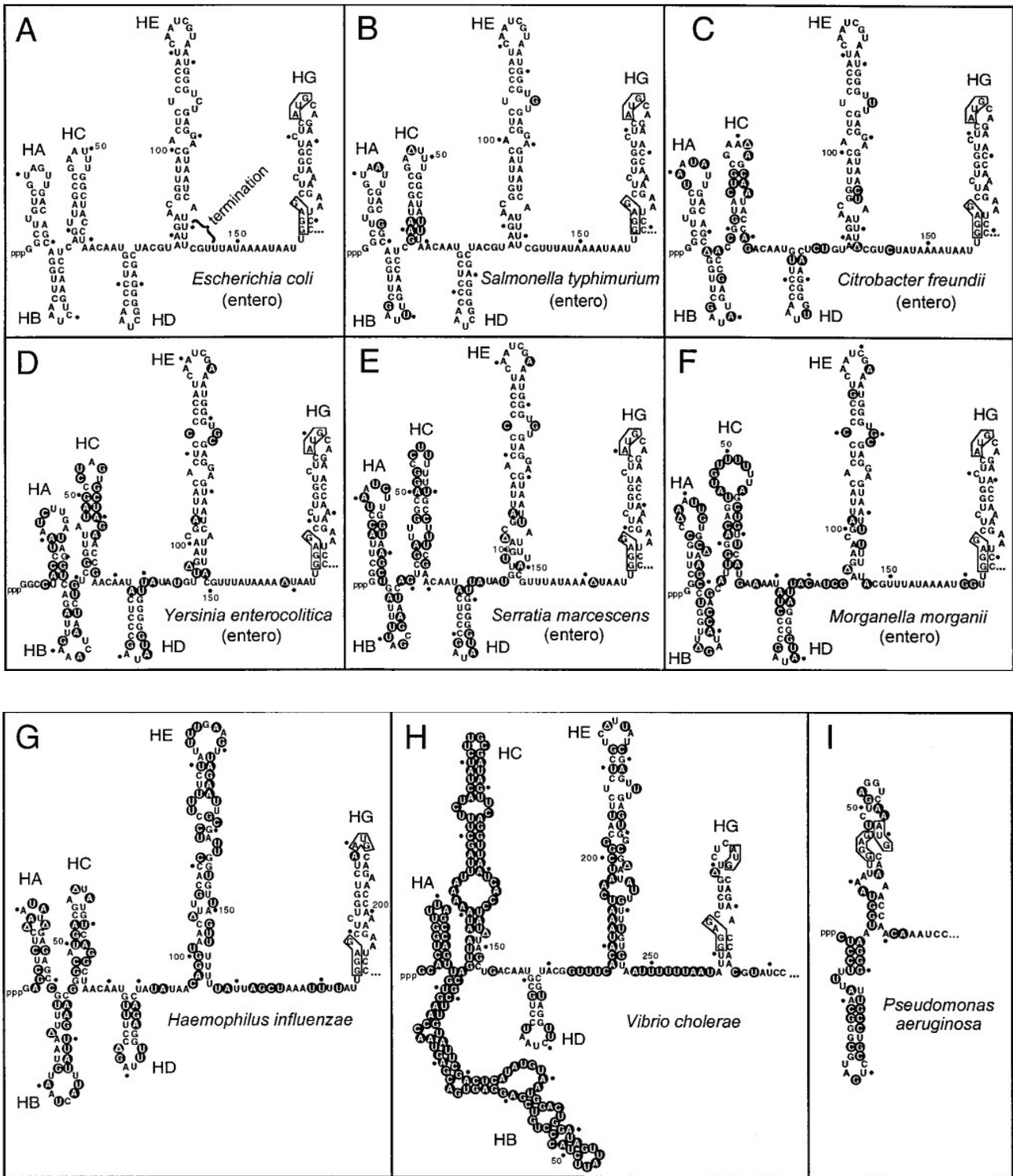
### 3. Role proteinu L4 z *E. coli* a jiných bakteriálních organizmů v regulaci operonu S10

Také pro jiné dosud studované bakteriální organizmy platí, že většina r-proteinů je seskupená do operonů podobným způsobem jako v *E. coli*. Jako příklad lze uvést *Bacillus subtilis* (Boylan, Suh, Thomas & Price, 1989), *Thermotoga maritima* (Sanangelantoni, Bocchetta, Cammarano & Tiboni, 1994), *Mycoplasma capricolum* (Ohkubo, Muto, Kawauchi, Yamao & Osawa, 1987) nebo *Methanococcus vannielii* (Auer, Spicker & Böck, 1989). V mnoha případech je dokonce pořadí genů identické. Přes svou podobnost se důležité prvky genetické organizace u různých druhů mezi sebou liší, což naznačuje, že u rozdílných druhů mikroorganismů se vyvinuly odlišné systémy pro regulaci syntézy r-proteinů. Například gen pro protein S4 u *B. subtilis*, který je autogenně kontrolován (Grundy & Henkin, 1991) není součástí  $\alpha$  operonu jako u *E. coli*, kde kontroluje translaci čtyř r-proteinů. Místo toho tvoří monocistronní operon a reguluje pouze svou vlastní syntézu.

I v případě, že je pořadí genů pro r-proteiny v genomu stejné, porovnání transkripční organizace mezi různými druhy ukazuje, že v průběhu evoluce došlo k rozrůznění regulačních mechanismů. Například i když genom *B. subtilis* obsahuje skupinu genů zahrnující většinu proteinů z *spc* lokusu *E. coli*, nebyl nalezen žádný promotor odpovídající promotorům  $\alpha$  a *spc* v *E. coli*. Zdá se, že transkripce celé skupiny genů probíhá z jedné regulační oblasti, která odpovídá promotoru S10 operonu v *E. coli*. Jelikož tento soubor genů je v *E. coli* přepisován do tří transkripčních jednotek (S10, *spc* a  $\alpha$ ) a translace každé z nich je regulována odlišným regulačním r-proteinem (L4, S8 a S4), je otázka jaký protein funguje jako regulátor v *B. subtilis*. Podobný případ nastal u L11-L1 a L10-L12 operonů. V *E. coli* jsou to dvě transkripční jednotky regulované proteinem L1, respektive L10. U archebakterií *Halobacterium cutirubrum* (Shimmin & Dennis, 1989) a *Methanococcus vannielii* (Hanner, Mayer, Köhrer, Golderer, Gröbner & Piendl, 1994) jsou geny pro L1, L10 a L12 kotranskribovány, a znovu tedy vyvstává otázka, který protein slouží v těchto mikroorganizmech jako regulátor. V případě *Methanococcus vannielii* je to pravděpodobně protein L1, který má funkci translačního regulátoru transkripční jednotky s geny pro L1-L10-L12 (Hanner *et al.*, 1994).

Za účelem porovnání regulace exprese genů odpovídajících operonu S10 v *E. coli* u vybraných odlišných mikroorganismů (konkrétně se jedná o enterobakterie *Yersenia pseudotuberculosis* a *Morganella morganii*, o non-enterickou Gram-negativní bakterii *Haemophilus influenzae* a o Gram-pozitivní *Bacillus subtilis*, *Thermotoga maritima* a *Methanococcus vannielii*) byly izolovány geny pro jejich proteiny L4 a ty byly exprimovány v *E. coli* (Zengel *et al.*, 1995). Ukázalo se, že tyto heterologické proteiny mohou sloužit jako regulátory operonu S10 i v *E. coli*. Indukce nadměrné syntézy cizího L4, jehož gen byl vnesen do organismu na P<sub>lac</sub> plazmidu, způsobí inhibici tvorby kolonií *E. coli*. Předpokládá se, že k inhibici růstu kolonií dochází mechanismem represe translace genu pro S10 a tím také genů pro ribozomální proteiny lokalizovaných distálně. Stejný efekt na buňku má nadprodukce vlastního proteinu L4 (Lindahl & Zengel, 1979). Tyto výsledky také naznačují, že by jednotlivé L4 proteiny mohly u těchto organismů účinkovat jako autoregulátory syntézy r-proteinů z vlastních operonů.

Vedoucí sekvence S10 mRNA řady zkoumaných enterobakterií jsou potenciálně schopné vytvořit vlásenkovou strukturu podobnou vlásenkám HA-HE u *E. coli*, i přesto, že se jejich primární sekvence mohou velmi lišit (Allen, Shen, Samsel, Liu, Lindahl & Zengel, 1999). Vlášenska HD, která je esenciální pro kontrolu transkripce (ne translace) a vlášenska HE, která je nezbytná pro kontrolu jak transkripce tak translace S10 operonu pomocí proteinu L4, jsou přítomné ve zkoumaných vedoucích sekvencích operonů S10, a pouze v případě *Pseudomonas aeruginosa* je vedoucí sekvence příliš krátká a chybí jí struktury schopné vytvořit vlásenky HA-HC (viz Obr. 8).



Obr. 8: Předpokládané sekundární struktury vedoucích sekvencí S10 operonu (Allen *et al.*, 1999).

Sekundární struktury byly určeny za použití programu mfold. Struktura vedoucí sekvence *E. coli* byla potvrzena chemickým a enzymatickým strukturním mapováním. Místo předčasné terminace transkripce zprostředkované proteinem L4 ve vedoucí sekvenci *E. coli* je vyznačeno. Nukleotidy odlišné v alignmentu sekvencí od *E. coli* jsou vyznačeny černým kruhem s bílým textem. Shine-Dalgarnova sekvence a iniciační translační kodón AUG jsou ohraničeny jednoduchou nepřerušovanou čarou.



### 3.1. Srovnání proteinů L4 z *E. coli* a dalších mikroorganismů

Byly porovnány nukleotidové a aminokyselinové sekvence L4 proteinů pocházejících z různých bakterií navzájem mezi sebou i vzhledem k proteinu L4 v *E. coli* (Allen *et al.*, 1999). Nukleotidové sekvence vykazují vysokou homologii s L4 v *E. coli* (92% u *M. morgani* a 83% u *H. influenzae*), přičemž většina změn je na třetí pozici, která tvoří synonymní kodóny. Jiné odlišnosti na první a druhé pozici jsou také buď součástí synonymních kodónů, anebo kódují aminokyseliny s postranním řetězcem podobného charakteru. Nejméně konzervované oblasti jsou N a C-konce proteinů, ale i zde je mnoho změn chemicky konzervativních.

Sekvence genů pro L4 u archebakterií obsahují „expanzivní“ oblast a jsou tedy delší než uvedené eubakteriální geny, ale jejich srovnání ukázalo, že jsou rovněž vysoce konzervované vzhledem k eubakteriálním proteinům.

### 3.2. Regulace S10 operonu *E. coli* cizími proteiny L4

Regulace operonu S10 v *E. coli* prostřednictvím r-proteinu L4 se realizuje kombinací inhibice transkripce a translace. Každý z těchto procesů snižuje expresi uvažovaných genů 4-5 krát, takže celkový regulační efekt je výsledkem obou (Freedman *et al.*, 1985; Zengel & Lindahl, 1990b). Protože represe syntézy proteinů S10 operonu je v *E. coli* často stejně silná při regulaci cizími proteiny jako vlastním L4, lze předpokládat, že tyto heterologické proteiny regulují rovněž jak translační, tak transkripční úroveň exprese. Rozdílně se chová protein L4 původem z *Thermotoga maritima*, který vykazuje pouze slabou kontrolu S10 operonu u *E. coli* (Worbs, Wahl, Lindahl & Zengel, 2002). Skutečnost, že proteiny L4 pocházející z jiných mikroorganismů mohou regulovat expresi genů v S10 operonu u *E. coli* naznačuje, že vlastnosti nutné pro autogenní kontrolu jsou velmi konzervované. Zatímco tento závěr není překvapující v případě vysoce homologických enterobakteriálních L4, nelze tak snadno odhadnout, zda například protein L4 pocházející z bakterie *Bacillus stearothermophilus* s pouhou 42% sekvenční homologií bude působit jako regulační protein v *E. coli*. Schopnost L4 z jiných bakteriálních druhů regulovat S10 operon *E. coli* naznačuje, že exprese odpovídajících genů u jiných organismů by také mohla být regulována vlastními proteiny L4. Na druhou stranu je možné, že struktury nutné pro regulaci jsou také nepostradatelné při interakcích typu RNA-protein a protein-protein ve funkčním ribozómu. Proto by mohly být za zachování tohoto regulačního potenciálu odpovědné i selektivní tlaky zaměřené na jiné charakteristiky L4 proteinů.

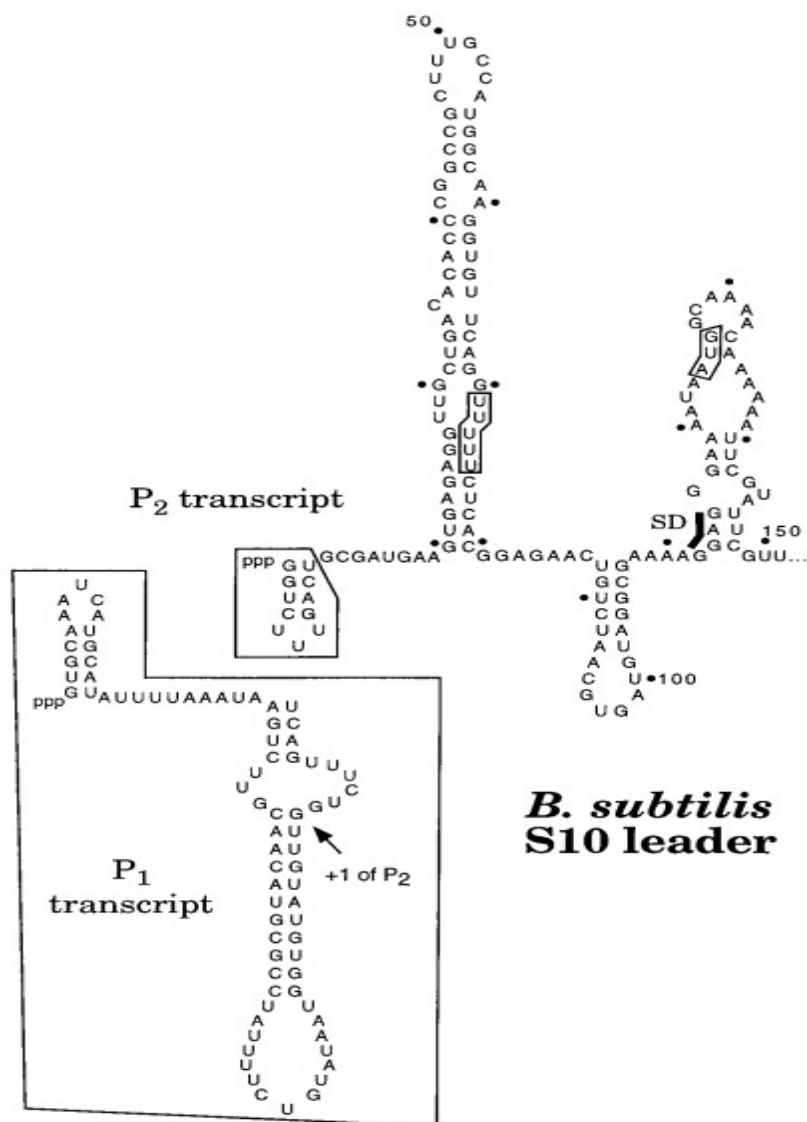
#### 4. Genové klastry operonu S10 z *E. coli* u vybraných druhů mikroorganismů

I když obecně neexistuje rozsáhlá kolinearita v pořadí genů v bakteriálních genomech, některé oblasti prokaryotických genomů mají pořadí genů v průběhu evoluce silně konzervováno. Jednou takovou oblastí je lokus nazvaný *str* obsahující operony S10, *spc* a  $\alpha$  (Coenye & Vandamme, 2005).

Různé studie navrhují, že mechanismus regulace S10 operonu prostřednictvím r-proteinu L4 u *E. coli* (Lindahl *et al.*, 1990; Yates & Nomura, 1980; Zengel & Lindahl, 1990b) je velmi rozšířený v poddivizi  $\gamma$ -proteobakterií (Allen *et al.*, 1999), i když ne zcela univerzální. Např. *Pseudomonas aeruginosa* nevykazuje kontrolu S10 operonu typu *E. coli* (Allen *et al.*, 1999). Data, která máme k dispozici, naznačují, že autogenní regulační mechanismy řídící expresi r-proteinů v *E. coli* se vyvinuly ze společného předka některých, ale ne všech,  $\gamma$ -proteobakterií (Allen, Watkins, Lindahl & Zengel, 2004). Kontrast mezi široce rozšířenou organizací hlavních skupin genů r-proteinů a relativně mnohem menším rozšířením regulačních mechanismů typu *E. coli* nabízí závěr, že některé regulační mechanismy se vyvinuly mnohem později než samotný operon, který regulují.

##### 4.1. Operon S10 u G+ bakterie *Bacillus subtilis*

V genomu *B. subtilis* byla nalezena skupina genů pro ribozomální proteiny, která obsahuje většinu genů nacházejících se v *E. coli* v operonech L11,  $\beta$ , S10, *spc* a  $\alpha$  (Li, Lindahl, Sha & Zengel, 1997). Všechny dosud identifikované geny kódující ribozomální proteiny jsou uspořádány ve stejném pořadí jako v *E. coli*, ale liší se v transkripčním uspořádání (viz Obr. 10). V *E. coli* byl na konci S10 operonu identifikován terminátor, na kterém je ukončeno asi 80% transkriptů přepisovaných z promotoru S10. V téže oblasti *B. subtilis* nebyla nalezena žádná podobná struktura. V oblasti zahrnující skupinu genů odpovídajících S10 operonu až po  $\alpha$  operon byly nalezeny pouze dva slabé doplňující promotory před geny *secY* a *map* (viz Obr. 10), jinak celá oblast S10-*spc*- $\alpha$  tvoří jedinou transkripční jednotku (Suh, Boylan, Oh & Price, 1996). Regulační mechanismus řídící genovou expresi se rovněž liší od *E. coli*. Regulační proteiny pro  $\alpha$  a *spc* operony v *B. subtilis* nejsou známy (Barloy-Hubler, Lelaure & Galibert, 2001).



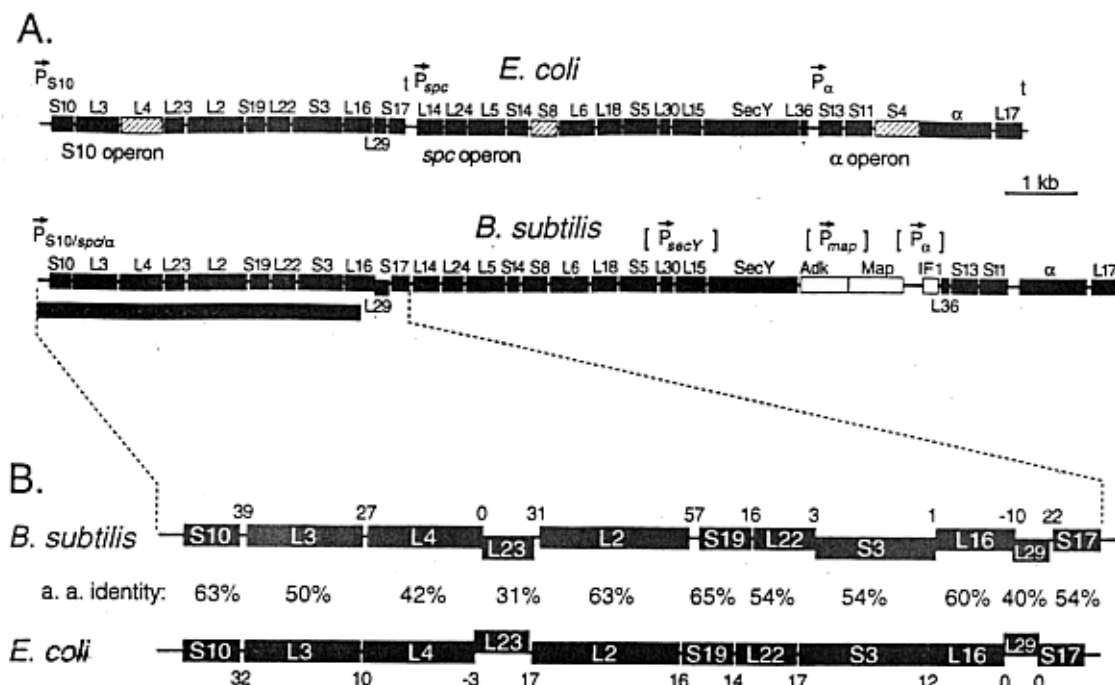
**Obr. 9: Vedoucí sekvence S10 operonu *B. Subtilis* (Li *et al.*, 1997).**

Sekundární struktura vedoucích sekvencí S10 operonu byla určena pomocí programu mfold. Transkripty označené P1 a P2 představují vedoucí sekvence z prvního nebo druhého promotoru. Shine-Dalgarnova sekvence je označena silným černým pruhem, sekvence pěti U připomínající místo přechasně terminace transkripce u *E. coli* zprostředkované proteinem L4 a iniciační kodón AUG jsou ohraničeny jednoduchou nepřerušovanou čarou.

Byly nalezeny dva promotory odpovědné za transkripci celé 15 kb oblasti S10-*spc-α*. Leží přibližně -200, respektive -140 bp proximálně od genu pro S10. -10 hexamery přímo odpovídají konsenzu 5' TATAAT 3', -35 sekvence se s kanonickou sekvencí 5' TTGACA 3' shodují v pěti a ve čtyřech nukleotidech. Oba promotory jsou funkční, i když distální promotor je o něco aktivnější (Li *et al.*, 1997).

I když má vedoucí sekvence podobnou délku jako vedoucí sekvence S10 operonu v *E. coli*, není mezi nimi žádná nápadná homologie, ať už v primární struktuře nebo v počítačem určené sekundární struktuře. Regulační funkce proteinu L4 nebyla u *B. subtilis* zatím prokázána. Regulace

expresi genů pro ribozomální bílkoviny vyžaduje specifické struktury ve vedoucí sekvenci mRNA, které jsou v *B. subtilis* jiné než v *E. coli*. Mechanismus řídicí expresi skupiny genů S10-*spc-α* v *B. subtilis* je prozatím neznámý (Li *et al.*, 1997).



Obr. 10: Operony S10, *spc* a  $\alpha$  *E. coli* a skupina genů S10-*spc-α* u *B. subtilis* (Li *et al.*, 1997).

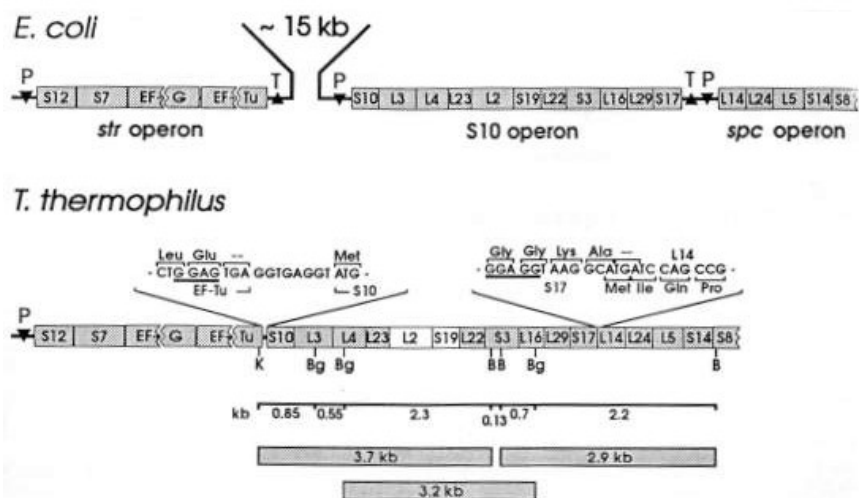
A: Promotory (P) a terminátory (t) jsou vyznačeny nad mapami organizace genů oblastí S10-*spc-α* u *E. coli* a *B. subtilis*. Slabé promotory identifikované v oblasti *spc-α* u *B. subtilis* jsou označeny hranatými závorkami. Geny pro r-proteiny u *E. coli* s autogenně regulační funkcí jsou vyšrafované. Geny nalezené u *B. subtilis* ale ne u *E. coli* jsou vyznačeny bílou barvou.

B: Srovnání aminokyselinových sekvencí r-proteinů kódovaných v S10 operonu u *E. coli* a *B. subtilis*. Mezigenové oblasti jsou vyznačeny nad mapou *B. subtilis* a pod mapou *E. coli*. Alignment aminokyselinových sekvencí byl proveden v programu LaserGene (DNASar, Inc.). Identita aminokyselinových sekvencí byla spočítána jako podíl aminokyselin v *E. coli* konzervovaných v *B. subtilis*.

#### 4.2. Operon S10 u G- bakterie *Thermus thermophilus*

Zajímavá organizace genů pro ribozomální proteiny byla objevena u termofilní bakterie *Thermus thermophilus*, jehož celý genom je kódován pouhými 1700 kb. Geny leží ve stejném pořadí jako jim odpovídající geny v *str*, S10 a *spc* operonech *E. coli*, avšak na rozdíl od *E. coli*, u které se úsek mezi operony *str*-S10 rozkládá na více než 15 kb, dělí poslední gen *str* operonu *tufA* (EF-Tu) od prvního genu S10 operonu *rpsJ* (S10) pouhých osm nukleotidů a operony S10 a *spc* se dokonce překrývají (viz Obr. 11). Je pravděpodobné, že celá skupina genů je přepisována z jediného promotoru umístěného před *str* operonem. Experimenty s příbuzným *Thermus aquaticus*

také prokázaly kotranskripci genů odpovídajících *str* a S10 operonům *E. coli* (Jahn, Hartmann & Erdmann, 1991).



**Obr. 11: Organizace genového klastru S10-*spc-a* u *E. coli* a *Thermus thermophilus* (Pfeiffer, Jorcke, Feltens & Hartmann, 1995)**

Promotory (P) a terminátory (T) jsou vyznačeny nad mapami organizace genů oblastí S10-*spc-a* u *E. coli* a *T. thermophilus*. U *E. coli* je vyznačena ~15 kb intergenová oblast oddělující operon *str* od operonu S10. U *T. thermophilus* jsou nad mapou genů vyznačeny detaily hranic mezi operony *str* a S10 (8 nukleotidů odděluje terminační kodón genu pro EF-Tu od iniciačního kodónu genu pro S10) a mezi operony S10 a *spc* (4 nukleotidový překryv kódujících sekvencí genu pro S17 a genu pro L14).

Z evolučního hlediska je zajímavé, že geny pro elongační faktor EF-Tu a r-protein S10 jsou rovněž transkripčně propojeny u některých Archaeobakterií (Creti, Citarella, Tiboni, Sanangelantoni, Palm & Cammarano, 1991) jako například *Methanococcus vannielii* (Lechner, Heller & Böck, 1989). Kotranskripce genů S10 a *spc* operonu se pravděpodobně odehrává také u G+ bakterie *Mycoplasma capricolum* a již dříve zmíněného *Bacillus subtilis* a u archeobakterie *Haloarcuula marismortui*.

#### 4.3. Operon S10 u G- *Vibrio cholerae*

Regulace genové exprese S10 operonu *Vibrio cholerae* je rovněž řízena autogenní kontrolou zprostředkovanou r-proteinem L4 (Allen *et al.*, 2004). I v tomto případě se ukazuje, že skupiny genů r-proteinů mají nápadně podobnou organizaci u organismů tak rozdílných jako jsou eubaktérie, archea, cyanely protist, chloroplasty a mitochondrie. Podrobná studie regulace S10 operonu *Vibrio cholerae* ukázala, že L4 kontroluje promotor podobně jako u *E. coli*, tj. na úrovni transkripce i translace, a přestože vedoucí sekvence je mnohem delší než u *E. coli*, počítačem vygenerovaná sekundární struktura ukazuje, že se pravděpodobně tvoří vlásenky analogické HA-

HE z *E. coli* (viz Obr. 8). Deleční analýza ukázala, že stejně jako u *E. coli*, první tři vlásenky nemají vliv na autogenní kontrolu promotoru S10 r-proteinem L4, delece vlásenky HD eliminovala strukturu esenciální pro vazbu L4 *in vitro* a významně snížila schopnost kontroly over-expressí L4 *in vivo*. Výsledky naznačují, že L4 reguluje expresi S10 operonu především indukcí předčasné terminace transkripce jako se děje u *E. coli*. Reziduální regulace u mutant s deletovanou HD vlásenkou by mohla odrážet kontrolu na úrovni translace.

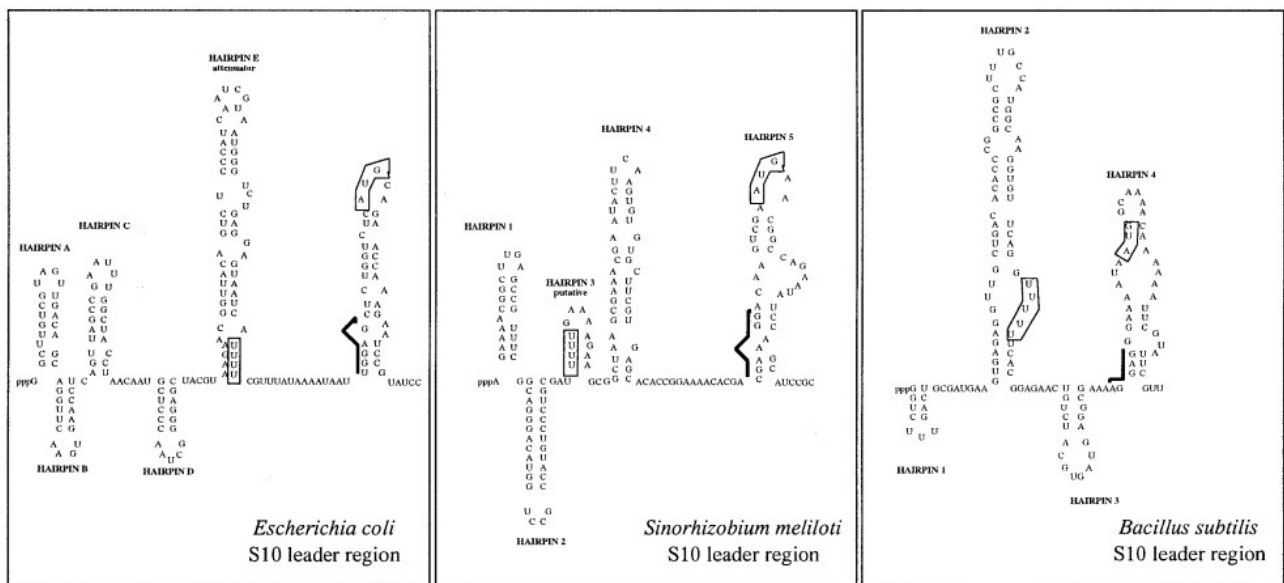
#### 4.4. Operon S10 u G- spirochety *Leptospira interrogans*

V případě spirochety *Leptospira interrogans* byly podobně jako u *B. subtilis* objeveny dva tandemové promotory. Na rozdíl od *B. subtilis*, tyto promotory nebyly lokalizovány před genem *rpsJ* (S10), ale před genem *fus* směrem upstream od *rpsJ*, a řídily tak transkripci celého genového klastru S10-*spc-α* o velikosti 17.5 kb (Zuerner, Hartskeerl, van de Kemp & Bal, 2000).

#### 4.5. Operon S10 u G- *Sinorhizobium meliloti*

Z porovnání genových klastrů *rpoBC*, *str*, S10, *spc* a  $\alpha$  byla odhalena podobnost mezi G- *S. meliloti* a G+ *B. subtilis*, která se ale považuje spíše za znak konvergence funkčních a genomových evolučních procesů než fylogenetické příbuznosti těchto dvou bakterií (Barloy-Hubler *et al.*, 2001).

Při zkoumání vedoucí sekvence S10 operonu *S. meliloti* se ukázalo, že tato je schopná tvořit 5 vlásenek, přičemž první tři vlásenky vykazují podobnost s vlásenkami HA-HC v *E. coli*, ale nebyly detekovány žádné struktury odpovídající vlásenkám HD nebo HE (viz Obr. 12). S vedoucí sekvencí *B. subtilis* nebyla pozorována žádná výrazná podobnost. Z uvedených údajů vyplývá, že i když u *S. meliloti* nelze vyloučit existenci regulačního mechanismu zprostředkovaného proteinem L4 jako je tomu u *E. coli*, struktura vedoucí sekvence S10 je dostatečně odlišná, že lze předpokládat existenci jiného regulačního mechanismu (Barloy-Hubler *et al.*, 2001).



Obr. 12: S10 vedoucí sekvence *E. coli*, *S. meliloti* a *B. subtilis* (Barloy-Hubler *et al.*, 2001).

Sekundární struktury vedoucích oblastí byly určeny pomocí programu mfold a struktura vedoucí sekvence *E. coli* byla navíc potvrzena chemickým a enzymatickým strukturním mapováním. Shine-Dalgarnova sekvence u *E. coli* a předpokládané Shine-Dalgarnovy sekvence u *S. meliloti* a *B. subtilis* jsou označeny silným černým pruhem, sekvence několika U a iniciační kodón AUG jsou ohraničeny jednoduchou nepřerušovanou čarou.

## 5. Závěr

Srovnání různých bakteriálních genomů ukazuje, že mezi vzdálenějšími taxony neexistuje významná kolinearita v pořadí genů: pravděpodobně v průběhu evoluce došlo dostatečně často k dynamickému přeskupení genetického materiálu a následně k narušení operonové struktury u různých taxonů. Přesto lze i u méně příbuzných genomů nalézt některé vysoce konzervované oblasti.

Organizace genů kódujících ribozomální proteiny je v genomech bakteriálních organismů obzvláště konzervována. Byla určena struktura S10 operonu *E. coli*, a následně byly zmapovány genové klastry odpovídající S10 operonu u řady dalších bakteriálních taxonů. Struktura této oblasti je do vysoké míry zachována u širokého spektra bakterií, i když byly také pozorovány delece nebo inserce dodatečných genů a/nebo genové translokace.

Operon S10 *E. coli* patří mezi jedny z nejvíce evolučně konzervovaných genových klastrů v genomu bakteriálních organismů, ale mechanismy řídicí exprese těchto genů se často liší, jak je zřejmé ze skutečnosti, že pozice regulačních prvků jako jsou promotory a terminátory transkripce jsou zachované mnohem méně často. Například u *E. coli* je organizováno 28 genů r-proteinů S10-*spc-α* klastru do 3 transkripčních jednotek, ale odpovídající geny v *B. subtilis* jsou součástí pouze jedné transkripční jednotky.

U *E. coli* je většina r-proteinů řízena autogenní zpětnou vazbou, kdy jeden r-protein operonu funguje nejen jako strukturní komponenta ribozómu, ale také jako regulační protein odpovědný za koordinaci exprese vlastního operonu s množstvím dostupných rRNA a dalších r-proteinů.

11-ti genový operon S10 *E. coli* je regulován proteinem L4, který je rovněž strukturní součástí větší ribozomální podjednotky 50S, a je kódován třetím genem operonu. Na rozdíl od mnoha jiných autogenně kontrolovaných operonů kódujících geny pro r-proteiny, které jsou regulovány pouze na úrovni translace, je operon S10 řízen na úrovni translace i transkripce. Tyto dva kontrolní mechanismy vyžadují rovněž dvě částečně překrývající se, ale přesto distinktní, determinanty v nepřekládané vedoucí sekvenci S10 mRNA.

Protein L4 samotný nestačí pro autogenní regulaci S10 operonu na úrovni transkripce, pro svou regulační funkci vyžaduje spolupráci s transkripčním faktorem NusA. Při zvýšené intracelulární koncentraci proteinu L4 dochází v přítomnosti NusA k regulaci na úrovni transkripce, prostřednictvím předčasné terminace transkripce S10 operonu, i na úrovni translace, kdy L4 reprimuje translaci S10 mRNA.

Studie ukázaly, že protein L4 i z bakterie tak evolučně vzdálené od *E. coli* jako je např. *Bacillus stearothermophilus* má zachované prvky nutné pro autogenní kontrolu S10 operonu v *E. coli*, ale samotný mechanismus autogenní kontroly už tak dobře mezi vzdálenými bakteriálními taxony konzervovaný není. Potenciální sekundární struktury vedoucích sekvencí mRNA S10 operonu u různých eubakteriálních druhů naznačují, že pouze určitá podmnožina druhů spadajících pod vývojovou větev  $\gamma$ -proteobakterií má zachované strukturní determinanty ve vedoucí sekvenci S10 mRNA, které jsou nezbytné pro autogenní kontrolu operonu zprostředkovanou proteinem L4 u *E. coli*.

## 6. Seznam použité literatury

Allen, T.; Shen, P.; Samsel, L.; Liu, R.; Lindahl, L. & Zengel, J. M. (1999): Phylogenetic analysis of L4-mediated autogenous control of the S10 ribosomal protein operon., *J Bacteriol* 181:6124-6132.

Allen, T. D.; Watkins, T.; Lindahl, L. & Zengel, J. M. (2004): Regulation of ribosomal protein synthesis in *Vibrio cholerae*., *J Bacteriol* 186:5933-5937.

Altmann, C. R.; Solow-Cordero, D. E. & Chamberlin, M. J. (1994): RNA cleavage and chain elongation by *Escherichia coli* DNA-dependent RNA polymerase in a binary enzyme.RNA complex., *Proc Natl Acad Sci U S A* 91:3784-3788.

Arnold, R. J. & Reilly, J. P. (1999): Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry., *Anal Biochem* 269:105-112.



- Artsimovitch, I. & Landick, R. (2000): Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals., *Proc Natl Acad Sci U S A* 97:7090-7095.
- Asato, Y. (2005): Control of ribosome synthesis during the cell division cycles of *E. coli* and *Synechococcus*., *Curr Issues Mol Biol* 7:109-117.
- Auer, J.; Spicker, G. & Böck, A. (1989): Organization and structure of the *Methanococcus* transcriptional unit homologous to the *Escherichia coli* "spectinomycin operon". Implications for the evolutionary relationship of 70 S and 80 S ribosomes., *J Mol Biol* 209:21-36.
- Barloy-Hubler, F.; Lelaure, V. & Galibert, F. (2001): Ribosomal protein gene cluster analysis in eubacterium genomics: homology between *Sinorhizobium meliloti* strain 1021 and *Bacillus subtilis*., *Nucleic Acids Res* 29:2747-2756.
- Berg, O. G. & von Hippel, P. H. (1987): Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters., *J Mol Biol* 193:723-750.
- Boylan, S. A.; Suh, J. W.; Thomas, S. M. & Price, C. W. (1989): Gene encoding the alpha core subunit of *Bacillus subtilis* RNA polymerase is cotranscribed with the genes for initiation factor 1 and ribosomal proteins B, S13, S11, and L17., *J Bacteriol* 171:2553-2562.
- Branlant, C.; Krol, A.; Machatt, A. & Ebel, J. P. (1981): The secondary structure of the protein L1 binding region of ribosomal 23S RNA. Homologies with putative secondary structures of the L11 mRNA and of a region of mitochondrial 16S rRNA., *Nucleic Acids Res* 9:293-307.
- Burd, C. G. & Dreyfuss, G. (1994): Conserved structures and diversity of functions of RNA-binding proteins., *Science* 265:615-621.
- Chan, C. L. & Landick, R. (1989): The *Salmonella typhimurium* his operon leader region contains an RNA hairpin-dependent transcription pause site. Mechanistic implications of the effect on pausing of altered RNA hairpins., *J Biol Chem* 264:20796-20804.
- Chan, C. L. & Landick, R. (1993): Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin., *J Mol Biol* 233:25-42.
- Coenye, T. & Vandamme, P. (2005): Organisation of the S10, spc and alpha ribosomal protein gene clusters in prokaryotic genomes., *FEMS Microbiol Lett* 242:117-126.
- Condon, C.; Squires, C. & Squires, C. L. (1995): Control of rRNA transcription in *Escherichia coli*., *Microbiol Rev* 59:623-645.
- Creti, R.; Citarella, F.; Tiboni, O.; Sanangelantoni, A.; Palm, P. & Cammarano, P. (1991): Nucleotide sequence of a DNA region comprising the gene for elongation factor 1 alpha (EF-1 alpha) from the ultrathermophilic archaeote *Pyrococcus woesei*: phylogenetic implications., *J Mol Evol* 33:332-342.
- Farnham, P. J.; Greenblatt, J. & Platt, T. (1982): Effects of NusA protein on transcription termination in the tryptophan operon of *Escherichia coli*., *Cell* 29:945-951.
- Farnham, P. J. & Platt, T. (1980): A model for transcription termination suggested by studies on the trp attenuator in vitro using base analogs., *Cell* 20:739-748.
- Feng, P.; Weagant, S. D. & Grant, M. A., (2002): *Bacteriological Analytical Manual*. FDA U.S. Food and Drug Administration.

- Freedman, L. P.; Zengel, J. M.; Archer, R. H. & Lindahl, L. (1987): Autogenous control of the S10 ribosomal protein operon of *Escherichia coli*: genetic dissection of transcriptional and posttranscriptional regulation., *Proc Natl Acad Sci U S A* 84:6516-6520.
- Freedman, L. P.; Zengel, J. M. & Lindahl, L. (1985): Genetic dissection of stringent control and nutritional shift-up response of the *Escherichia coli* S10 ribosomal protein operon., *J Mol Biol* 185:701-712.
- Gourse, R. L.; Thurlow, D. L.; Gerbi, S. A. & Zimmermann, R. A. (1981): Specific binding of a prokaryotic ribosomal protein to a eukaryotic ribosomal RNA: implications for evolution and autoregulation., *Proc Natl Acad Sci U S A* 78:2722-2726.
- Grayhack, E. J.; Yang, X. J.; Lau, L. F. & Roberts, J. W. (1985): Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site., *Cell* 42:259-269.
- Greenblatt, J. (1981): Regulation of transcription termination by the N gene protein of bacteriophage lambda., *Cell* 24:8-9.
- Greenblatt, J. & Li, J. (1981): The nusA gene protein of *Escherichia coli*. Its identification and a demonstration that it interacts with the gene N transcription anti-termination protein of bacteriophage lambda., *J Mol Biol* 147:11-23.
- Grundy, F. J. & Henkin, T. M. (1991): The rpsD gene, encoding ribosomal protein S4, is autogenously regulated in *Bacillus subtilis*., *J Bacteriol* 173:4595-4602.
- Hanner, M.; Mayer, C.; Köhrer, C.; Golderer, G.; Gröbner, P. & Piendl, W. (1994): Autogenous translational regulation of the ribosomal MvaL1 operon in the archaeobacterium *Methanococcus vannielii*., *J Bacteriol* 176:409-418.
- Herbert, K. M.; Porta, A. L.; Wong, B. J.; Mooney, R. A.; Neuman, K. C.; Landick, R. & Block, S. M. (2006): Sequence-resolved detection of pausing by single RNA polymerase molecules., *Cell* 125:1083-1094.
- Jahn, O.; Hartmann, R. K. & Erdmann, V. A. (1991): Analysis of the spc ribosomal protein operon of *Thermus aquaticus*., *Eur J Biochem* 197:733-740.
- Jin, D. J.; Burgess, R. R.; Richardson, J. P. & Gross, C. A. (1992): Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho., *Proc Natl Acad Sci U S A* 89:1453-1457.
- Kaczanowska, M. & Rydén-Aulin, M. (2007): Ribosome biogenesis and the translation process in *Escherichia coli*., *Microbiol Mol Biol Rev* 71:477-494.
- Kolsto, A. B. (1997): Dynamic bacterial genome organization., *Mol Microbiol* 24:241-248.
- Koonin, E. V.; Bork, P. & Sander, C. (1994): A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme?, *Nucleic Acids Res* 22:2166-2167.
- Landick, R. & Yanofsky, C. (1987): Isolation and structural analysis of the *Escherichia coli* trp leader paused transcription complex., *J Mol Biol* 196:363-377.
- Lechner, K.; Heller, G. & Böck, A. (1989): Organization and nucleotide sequence of a transcriptional unit of *Methanococcus vannielii* comprising genes for protein synthesis elongation factors and ribosomal proteins., *J Mol Evol* 29:20-27.

- Lee, D. N.; Phung, L.; Stewart, J. & Landick, R. (1990): Transcription pausing by *Escherichia coli* RNA polymerase is modulated by downstream DNA sequences., *J Biol Chem* 265:15145-15153.
- Levin, J. R. & Chamberlin, M. J. (1987): Mapping and characterization of transcriptional pause sites in the early genetic region of bacteriophage T7., *J Mol Biol* 196:61-84.
- Li, X.; Lindahl, L.; Sha, Y. & Zengel, J. M. (1997): Analysis of the *Bacillus subtilis* S10 ribosomal protein gene cluster identifies two promoters that may be responsible for transcription of the entire 15-kilobase S10-spc-alpha cluster., *J Bacteriol* 179:7046-7054.
- Li, X.; Lindahl, L. & Zengel, J. M. (1996): Ribosomal protein L4 from *Escherichia coli* utilizes nonidentical determinants for its structural and regulatory functions., *RNA* 2:24-37.
- Lindahl, L.; Archer, R. & Zengel, J. M. (1983): Transcription of the S10 ribosomal protein operon is regulated by an attenuator in the leader., *Cell* 33:241-248.
- Lindahl, L.; Archer, R. H.; McCormick, J. R.; Freedman, L. P. & Zengel, J. M. (1989): Translational coupling of the two proximal genes in the S10 ribosomal protein operon of *Escherichia coli*., *J Bacteriol* 171:2639-2645.
- Lindahl, L.; Sor, F.; Archer, R. H.; Nomura, M. & Zengel, J. M. (1990): Transcriptional organization of the S10, spc and alpha operons of *Escherichia coli*., *Biochim Biophys Acta* 1050:337-342.
- Lindahl, L. & Zengel, J. M. (1979): Operon-specific regulation of ribosomal protein synthesis in *Escherichia coli*., *Proc Natl Acad Sci U S A* 76:6542-6546.
- Lindahl, L. & Zengel, J. M. (1986): Ribosomal genes in *Escherichia coli*., *Annu Rev Genet* 20:297-326.
- Lindahl, L. & Zengel, J. M. (1990): Autogenous control is not sufficient to ensure steady-state growth rate-dependent regulation of the S10 ribosomal protein operon of *Escherichia coli*., *J Bacteriol* 172:305-309.
- Lodish, H.; Berk, A.; Zipursky, S. Lawrenceand Matsudaira, P.; Baltimore, D. & Darnell, J. E., 1999. *Molecular Cell Biology*. W. H. Freeman & Co.
- Lynn, S. P.; Kasper, L. M. & Gardner, J. F. (1988): Contributions of RNA secondary structure and length of the thymidine tract to transcription termination at the thr operon attenuator., *J Biol Chem* 263:472-479.
- Nomura, M. (1999): Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: diversity and common principles., *J Bacteriol* 181:6857-6864.
- Nomura, M.; Gourse, R. & Baughman, G. (1984): Regulation of the synthesis of ribosomes and ribosomal components., *Annu Rev Biochem* 53:75-117.
- Nomura, M.; Yates, J. L.; Dean, D. & Post, L. E. (1980): Feedback regulation of ribosomal protein gene expression in *Escherichia coli*: structural homology of ribosomal RNA and ribosomal protein mRNA., *Proc Natl Acad Sci U S A* 77:7084-7088.
- Ohkubo, S.; Muto, A.; Kawauchi, Y.; Yamao, F. & Osawa, S. (1987): The ribosomal protein gene cluster of *Mycoplasma capricolum*., *Mol Gen Genet* 210:314-322.
- Olins, P. O. & Nomura, M. (1981): Regulation of the S10 ribosomal protein operon in *E. coli*: nucleotide sequence at the start of the operon., *Cell* 26:205-211.

- Oppenheim, D. S. & Yanofsky, C. (1980): Translational coupling during expression of the tryptophan operon of *Escherichia coli*., *Genetics* 95:785-795.
- Pfeiffer, T.; Jorcke, D.; Feltens, R. & Hartmann, R. K. (1995): Direct linkage of *str*-, *S10*- and *spc*-related gene clusters in *Thermus thermophilus* HB8, and sequences of ribosomal proteins L4 and *S10*., *Gene* 167:141-145.
- Platt, T. (1994): Rho and RNA: models for recognition and response., *Mol Microbiol* 11:983-990.
- Prasch, S.; Jurk, M.; Washburn, R. S.; Gottesman, M. E.; Wöhrl, B. M. & Rösch, P. (2009): RNA-binding specificity of *E. coli* NusA., *Nucleic Acids Res* 37:4736-4742.
- Sanangelantoni, A. M.; Bocchetta, M.; Cammarano, P. & Tiboni, O. (1994): Phylogenetic depth of *S10* and *spc* operons: cloning and sequencing of a ribosomal protein gene cluster from the extremely thermophilic bacterium *Thermotoga maritima*., *J Bacteriol* 176:7703-7710.
- Schmidt, M. C. & Chamberlin, M. J. (1984): Amplification and isolation of *Escherichia coli* nusA protein and studies of its effects on in vitro RNA chain elongation., *Biochemistry* 23:197-203.
- Schmidt, M. C. & Chamberlin, M. J. (1987): nusA protein of *Escherichia coli* is an efficient transcription termination factor for certain terminator sites., *J Mol Biol* 195:809-818.
- Sha, Y.; Lindahl, L. & Zengel, J. M. (1995a): Role of NusA in L4-mediated attenuation control of the *S10* r-protein operon of *Escherichia coli*., *J Mol Biol* 245:474-485.
- Sha, Y.; Lindahl, L. & Zengel, J. M. (1995b): RNA determinants required for L4-mediated attenuation control of the *S10* r-protein operon of *Escherichia coli*., *J Mol Biol* 245:486-498.
- Shen, P.; Zengel, J. M. & Lindahl, L. (1988): Secondary structure of the leader transcript from the *Escherichia coli* *S10* ribosomal protein operon., *Nucleic Acids Res* 16:8905-8924.
- Shimmin, L. C. & Dennis, P. P. (1989): Characterization of the L11, L1, L10 and L12 equivalent ribosomal protein gene cluster of the halophilic archaeobacterium *Halobacterium cutirubrum*., *EMBO J* 8:1225-1235.
- Singh, D.; Chang, S.-J.; Lin, P.-H.; Averina, O. V.; Kaberdin, V. R. & Lin-Chao, S. (2009): Regulation of ribonuclease E activity by the L4 ribosomal protein of *Escherichia coli*., *Proc Natl Acad Sci U S A* 106:864-869.
- Stelzl, U.; Zengel, J. M.; Tovbina, M.; Walker, M.; Nierhaus, K. H.; Lindahl, L. & Patel, D. J. (2003): RNA-structural mimicry in *Escherichia coli* ribosomal protein L4-dependent regulation of the *S10* operon., *J Biol Chem* 278:28237-28245.
- Suh, J. W.; Boylan, S. A.; Oh, S. H. & Price, C. W. (1996): Genetic and transcriptional organization of the *Bacillus subtilis* *spc*-alpha region., *Gene* 169:17-23.
- Surratt, C. K.; Milan, S. C. & Chamberlin, M. J. (1991): Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription., *Proc Natl Acad Sci U S A* 88:7983-7987.
- Wikipedie (2010). *Escherichia coli* --- Wikipedie: Otevřená encyklopedie.
- Wong, B. (2004). *The Study of RNA Polymerase Pausing by Optical Traps*.
- Worbs, M.; Wahl, M. C.; Lindahl, L. & Zengel, J. M. (2002): Comparative anatomy of a regulatory ribosomal protein., *Biochimie* 84:731-743.

- Yager, T. & von Hippel (1987): Transcript elongation and termination in *E. coli*. In "Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology", American Society of Microbiology .
- Yang, M. T. & Gardner, J. F. (1991): Isolation and footprint analysis of the Escherichia coli thr leader paused transcription complex., *Nucleic Acids Res* 19:1671-1680.
- Yang, X. J. & Roberts, J. W. (1989): Gene Q antiterminator proteins of Escherichia coli phages 82 and lambda suppress pausing by RNA polymerase at a rho-dependent terminator and at other sites., *Proc Natl Acad Sci U S A* 86:5301-5305.
- Yates, J. L. & Nomura, M. (1980): *E. coli* ribosomal protein L4 is a feedback regulatory protein., *Cell* 21:517-522.
- Zengel, J. M.; Archer, R. H.; Freedman, L. P. & Lindahl, L. (1984): Role of attenuation in growth rate-dependent regulation of the S10 r-protein operon of *E. coli*., *EMBO J* 3:1561-1565.
- Zengel, J. M. & Lindahl, L. (1990a): Ribosomal protein L4 stimulates in vitro termination of transcription at a NusA-dependent terminator in the S10 operon leader., *Proc Natl Acad Sci U S A* 87:2675-2679.
- Zengel, J. M. & Lindahl, L. (1990b): Escherichia coli ribosomal protein L4 stimulates transcription termination at a specific site in the leader of the S10 operon independent of L4-mediated inhibition of translation., *J Mol Biol* 213:67-78.
- Zengel, J. M. & Lindahl, L. (1992): Ribosomal protein L4 and transcription factor NusA have separable roles in mediating terminating of transcription within the leader of the S10 operon of Escherichia coli., *Genes Dev* 6:2655-2662.
- Zengel, J. M. & Lindahl, L. (1993): Domain I of 23S rRNA competes with a paused transcription complex for ribosomal protein L4 of Escherichia coli., *Nucleic Acids Res* 21:2429-2435.
- Zengel, J. M. & Lindahl, L. (1994): Diverse mechanisms for regulating ribosomal protein synthesis in Escherichia coli., *Prog Nucleic Acid Res Mol Biol* 47:331-370.
- Zengel, J. M. & Lindahl, L. (1996): A hairpin structure upstream of the terminator hairpin required for ribosomal protein L4-mediated attenuation control of the S10 operon of Escherichia coli., *J Bacteriol* 178:2383-2387.
- Zengel, J. M.; Sha, Y. & Lindahl, L. (2002): Surprising flexibility of leader RNA determinants for r-protein L4-mediated transcription termination in the Escherichia coli S10 operon., *RNA* 8:572-578.
- Zengel, J. M.; Vorozheikina, D.; Li, X. & Lindahl, L. (1995): Regulation of the Escherichia coli S10 ribosomal protein operon by heterologous L4 ribosomal proteins., *Biochem Cell Biol* 73:1105-1112.
- Zhang, Y. & Hanna, M. M. (1994): NusA changes the conformation of Escherichia coli RNA polymerase at the binding site for the 3' end of the nascent RNA., *J Bacteriol* 176:1787-1789.
- Zuerner, R. L.; Hartskeerl, R. A.; van de Kemp, H. & Bal, A. E. (2000): Characterization of the *Leptospira interrogans* S10-spc-alpha operon., *FEMS Microbiol Lett* 182:303-308.