

BIOINFORMATICS

Phylogenetic analysis of structural preconditions of autogenous control of gene expression of *rplJL* operon at the translation level in α -subdivision.

I. V. Kroupskaya, E. B. Paton

Institute of Molecular Biology and Genetics NAS of Ukraine
150, Academician Zabolotny Str., Kyiv, 03143, Ukraine
Institute of Cell Biology and Genetic Engineering, NAS of Ukraine
e-mail: i.v.kroupskaya@imbg.org.ua

*Phylogenetic analysis of structural requirements of autogenous control of gene expression of *rplJL* operon at the translation level in α -subdivision has been performed. The structural requirements of both L10 ribosomal protein and its RNA targets, which determine the possibility and the specificity of L10-RNA interaction have been studied. Phylogenetic comparison provides evidence for similarity of structural requirements of the feedback regulation mechanism in α -subdivision and *E.coli*, except for the endosymbionts.*

Keywords: *autogenous control, gene expression, protein biosynthesis, L10 ribosomal protein, mRNA, *rplJL* operon.*

Introduction. Autogenous control of expression of protein-coding gene is common for prokaryotes [1–3]. Autoregulation of *rplJL* genes, which code L10 and L12, is performed at the translation level by regulatory protein L10 or by L8 complex (L10:4L12) [4, 5]. Regulatory protein can bind competitively with the target sequence on both mRNA and rRNA. In case of redundant synthesis of the mentioned protein in regards to rRNA synthesis, either L10 or L10:4L12 complex binds with mRNA target site, blocks translation of both its own and adjacent cistrons [2, 3]. Simultaneous

blocking of transcription initiation of both cistrons (L10 and L12) takes place both due to L10 binding mRNA-target, located in the untranslated 5'-region of the former, and as a result of firm binding between 5'-terminal regions (500-nucleotide apart from L10 translation initiation site) of these cistrons on one mRNA [4, 5]. Supposedly, *Escherichia coli* leader area before *rplJ* gene exists in thermodynamically equilibrated conformations, *i.e.* translationally active (“open”) and transformationally blocked (“closed”) [6]. It is assumed that the binding of L10 protein is possible with the “closed” form of *rplJ*-leader only.

Taking into account the data in [7], the goal of our work was to search and investigate structural

prerequisites for implementing autogenous control mechanism of *rplJ* expression in r-subdivision. The structural analysis consisted in searching and characterization of L10 mRNA-target structure; constructing the model of secondary leader sequence of *rplJ*, as well as the search for structural requirements for long-range translation coupling of L10 and L12 cistrons at the expense of binding of 5'-regions of these cistrons.

The data on r-subdivision nucleotide sequences were obtained from NCBI, TIGR Microbial Data-base, Sanger Centre, Joint Genome Institute, University of Washington, and University of Wisconsin. To search for *rplJL* genes with incomplete genome sequencing BLAST 2.0 (NCBI BLAST) was used.

The following representatives of r-subdivision have been investigated:

Enterobacteriales: *Buchnera aphidicola* str. APS NC_002528; *Candidatus Blochmannia floridanus* NC_005061; *Erwinia carotovora* subsp. *atroseptica* SCRI1043; *Erwinia chrysanthemi* str. 3937; *Escherichia coli* K12 NC_000913; *Klebsiella pneumoniae*; *Salmonella enterica* subsp. *enterica* serovar Typhi NC_0031; *Salmonella enteritidis* NC_003455; *Salmonella paratyphi*; *Salmonella typhimurium* DT104 NC_002056; *Shigella dysenteriae* M131649; *Shigella flexneri* 2a str. 301 NC_002773; *Shigella sonnei* 53G NC_002809; *Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis* NC_005139; *Yersinia enterocolitica* (type 0:8) NC_002120; *Yersinia pestis* CO92 NC_003143;

Pasteurellaceae: *Actinobacillus actinomycetemcomitans* NK_1651; *Haemophilus ducreyi* 35000HP NC_002940; *Haemophilus influenzae* Rd KW20 NC_000907; *Haemophilus somnus* 129PT NC_002664; *Mannheimia succiniciproducens* MBEL55E NC_006300
Pasteurella multocida NC_001774;

Pseudomonadaceae: *Pseudomonas aeruginosa* PAO1 NC_002516; *Pseudomonas fluorescens* SBW25; *Pseudomonas putida* KT2440 NC_002947; *Pseudomonas syringae* pv. *syringae* B728a NC_002759;

Vibrionaceae: *Photobacterium profundum* SS9 NC_006370; *Vibrio cholerae* O1 biovar eltor str. N16961 NC_004982; *Vibrio parahaemolyticus* RIMD 2210633 NC_002088; *Vibrio vulnificus* CMCP6 NC_004459; *Vibrio vulnificus* YJ016 NC_005139;

Xanthomonadaceae: *Xanthomonas campestris* pv. *campestris* str. ATCC 33913 NC_003902; *Xylella fastidiosa* 9a5c NC_002490; *Xylella fastidiosa* Temecula1 NC_002144;

Others: *Acidithiobacillus ferrooxidans* ATCC 23270; *Colwellia psychrerythraea* 34H; *Coxiella burnetii* RSA 493; *Legionella pneumophila*; *Methylococcus capsulatus* str.; *Shewanella oneidensis* MR-1 NC_004347.

Primary DNA sequences were compared using CLUSTAL W software [8].

rplJ leader secondary structure models were constructed with RNAFOLD [9] (PC GENE software pack) and RNADRAW software [10].

The following seven enterobacteria representatives were reported to have high conservatism degree of binding site structure of L10 with mRNA: *S. marcescens*, *P. vulgaris*, *E. coli*, *S. typhimurium*, *C. freundii*, *E. cloacae*, and *K. pneumoniae* [7]. Two conditions were kept in mind while searching for binding site of L10 on mRNA.

First, all known m- and rRNA targets of prokaryotic L10 protein contain AGA elements and C-G pairs, neighboring with unpaired A and UAAA element (Fig.1), specific for mRNA binding regions of L10 of so called enterobacterial kind [11-14]. L10-mRNA interaction sites of r-subdivision possess a series of conservative structural elements and clearly distinguished structural likeness with the region of L10 binding on 23SRNA (Fig.1). Functional significance of these elements has been earlier described [7].

Second, all known mRNA-targets of L10 are removed significantly from L10-cistron translation initiation site (~150 b.p.) and reveal evident structural homology with the binding area of 23S rRNA. On the basis of comparative analysis performed, consensual model of L10 binding region secondary structure on mRNA has been constructed. As it is seen in Fig.1, all mentioned above specific elements of mRNA binding site of L10 (Fig.1, sequences marked bold) are preserved in r-subdivision representatives.

From the standpoint of definite initial mRNA sequence of *rplJ* leaders of r-subdivision it is evident that their general structural organization is conservative (Fig.2, a). At the same time five double-stranded (ds) and five single-stranded (ss) regions remain which is typical for *E. coli*. Regions dsIV and dsIII are the most conservative ones which correlates well with their key role in autoregulation of *rplJL* gene expression [7]. DsIV contains binding site for regulatory protein L10(L8). DsIII is supposed to exist in two alternative forms corresponding to translationally active and translationally blocked state of L10-L12 mRNA. Sequences 5'-CCRUC-3' and 5'-GRUGG-3', which provide the possibility for rearrangement of mRNA L10-L12 leader due to alternative binding, turned out to be conservative in all investigated representatives of r-subdivision. Therefore, the assumption on conformational transfer being the basis for translation regulation of L10-L12 mRNA reasonable for both *E. coli* [4] and six kinds of Enterobacteriales [15] is reasonable for representatives of r-subdivision under investigation.

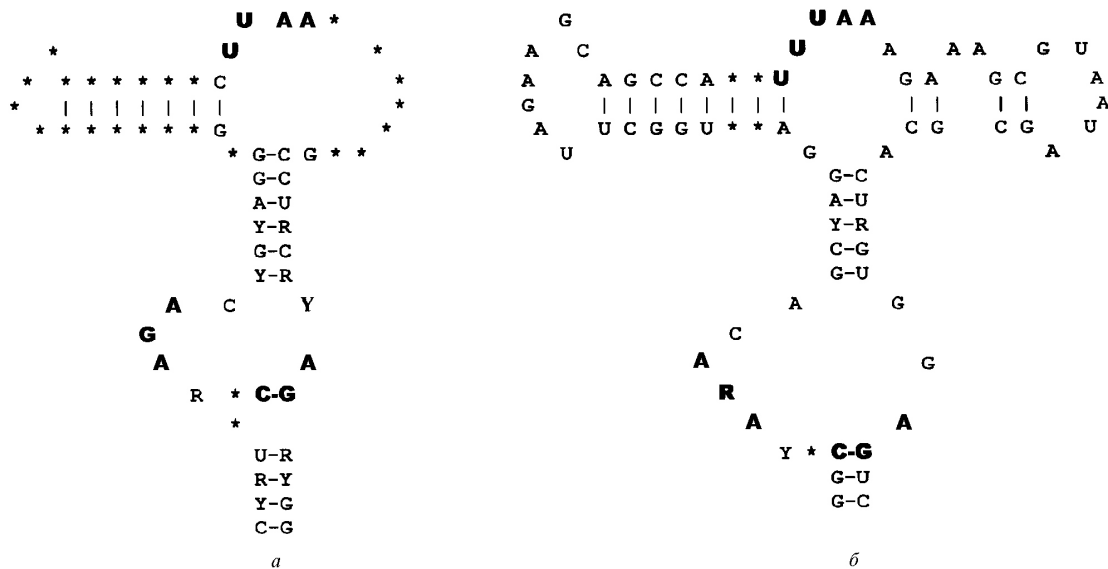


Fig.1 Secondary structure of regulatory protein L10 binding site on mRNA and on 23S rRNA: general consensus, introduced for L10 binding area on mRNA (a) and 23S rRNA (b). letters – conservative nucleotides, * – variable nucleotides. Y and R respectively C, U and G, A

Autoregulation mechanism of gene translation of *rplJL* operon presupposes that L10 binding the leader region of mRNA L10-12 blocks the translation of both cistrons. Structural requirement of associated translation of L10 and L12 cistrons in *E. coli* is the coupling of distant regions 1738–1773–L10 cistron and 2256–2290–L12 cistron of mRNA of L10-L12 [16]. Similar regions in all representatives of γ -subdivision are conservative, which provides the possibility for such coupling (Fig.2, b). The issues mentioned above allow supposing the existence of the same mechanism of associated translation of the mentioned cistrons of the investigated representatives.

It is noteworthy that localization of regions of *Pseudomonadaceae* representatives involved in coupling and resulting in blocking of associated cistrons is different from that of other representatives of γ -subdivision. This fact can be explained by phylogenetic remoteness from other groups of γ -subdivision. Some representatives of γ -subdivision did not reveal structural prerequisites for L10 translation regulation, e.g. *Buchnera sp.*, *C. ruddii*, *X. fastidiosa*, belonging to parasite organisms. The absence of prerequisites may be explained by endosymbiotic life style of these organisms [17, 18]. One more feature is specific to endosymbionts – the absence of a binding site for L1 protein before *rplK* gene (unpublished data). The latter fact gives grounds to

suppose that *rplKALJ* cluster is not regulated by protein L1. Cluster *rplKALJ* of all endosymbionts contains *rplJL* genes and, therefore, it is reasonable to assume the participation of *NusG* gene in the regulation of *rplJL* genes expression. Owing to the possibilities of obtaining the RNA crystals and defining their atomic structures it became possible to identify earlier unknown RNA motifs. Thus, so called flexible molecular motifs – kink-turns (K-turn or GA), which lead to bending of double-stranded RNA and are considered to be the important region for protein recognition have been identified [19-21]. K-turn is asymmetric internal loop, flanked with C-G pair from one side and with G-A pair from the other side. This motif is recognized by 9 of 31 of protein large 23S RNA subunit: L4, L7Ae, L10, L15e, L19e, L24, L29, L32e and L37Ae. L10 binding site on mRNA is similar to KT-42 in the 78th loop of the 4th domain of 23S of *E. coli* rRNA [16]. We have constructed a consensus model of K-turn motif structure for mRNA of L10 of γ -subdivision (Fig.3). On the basis of this model the conclusion on the presence of additional structural elements which participate in autogenous control of gene expression of *rplJL* operon can be made.

Therefore, phylogenetic analysis of autogenous control of gene expression of *rplJL* in representatives of γ -subdivision proved that these genes are regulated in accordance with classical feedback principle specific to prokaryotes and

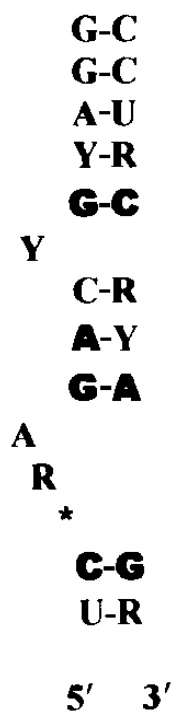


Fig.3 Consensus structure of K-turn motif for L10 mRNA

possess all structural requirements, namely, binding site for ribosomal protein L10 on mRNA; secondary structure of the leader region providing conformational transition, which is the basis of regulation of L10-L12 mRNA translation, as well as the conditions for associated translation of L10-L12 cistrons.

И. В. Крупская, Е. Б. Патон

Филогенетический анализ структурных предпосылок аутогенного контроля экспрессии генов rplJL оперона на уровне трансляции представителей -протеобактерий

Резюме

Проведен филогенетический анализ структурных предпосылок аутогенной регуляции экспрессии генов rplJL оперона у представителей -протеобактерий на уровне трансляции. Исследованы структурные предпосылки взаимодействия рибосомного белка L10 с мРНК. Показано, что у представителей -протеобактерий аутогенный контроль экспрессии генов L10 белка осуществляется по тому же механизму, что и у *Escherichia coli*, за исключением эндосимбионтов.

Ключевые слова: аутогенный контроль, экспрессия генов, рибосомный белок L10, мРНК, сайт связывания.

REFERENCES:

1. Post L. E., Strycharz G. D., Nomura M., Lewis H., Dennis P. P. Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene for RNA polymerase subunit B in *E. coli* // Proc. Nat. Acad. Sci. USA.-1979.-76.-P. 1697-1701.
2. Zengel J. M., Lindahl L. Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli* // Prog. Nucl. Acids Res. and Mol. Biol.-1994.-47.-P. 331-370.
3. Nomura M. // Symp. of the Soc. for Gen. Microbiol. Regulation of Gene Expression / Eds I. Both, C. Higgins.-Cambridge: Univ. press, 1986.-P. 199-220.
4. Sor F., Nomura M. Cloning and DNA sequence determination of the L11 ribosomal protein of *Serratia marcescens* and *Proteus vulgaris* L11 operon by heterologous L1 protein // Mol. and Gen. Genet.-1987.-210.-P. 52-59.
5. Climie S. C., Friesen J. D. Feedback regulation of the rplJL-rpoBC ribosomal protein operon of *Escherichia coli* requires a region of mRNA secondary structure // J. Mol. Biol.-1987.-198.-P. 371-381.
6. Friesen J. D., Tropak M., An G. Mutations in the leader of *Escherichia coli* that abolish feedback regulation // Cell.-1983.-32.-P. 361-39.
7. Патон Е. Б. Особенности регуляции экспрессии генов кластера rplKAJL // Биополимеры и клетка.-1990.-6, № 5.-С. 5-23.
8. Thomson J. D., Higgins D. G., Gibson T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice // Nucl. Acids Res.-1994.22.-P. 4673-4680.
9. Zuker M., Stiegler P. Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information // Nucl. Acid Res.-1981.-9.-P. 133-148.
10. Matzura O., Wennborg A. RNA draw: an integrated program for RNA secondary structure calculation and analysis under 32-bit Microsoft Windows // Comp. Appl. Biosci. (CABIOS).-1996.-12.-P. 247-249.
11. Christensen T., Johnson M., Fill N. P., Frisen J. D. RNA secondary structure and translation translation inhibition: analysis of mutants in the rplJ leader // EMBO J.-984.-3.-P. 1609-1612.
12. Живолуп А. Н., Патон Е. Б. Общий механизм регуляции экспрессии в rplJL опероне

- энтеробактерий подтверждается консервативностью структурной организации лидерной области мРНК. Регуляторный консенсус лидерной области мРНК L10-L12 энтеробактерий // Докл. АН Украины.-1995.-340, № 1.-С. 104-107.
13. Патон Е. Б., Живолуп А. Н. Наличие сайта связывания с рибосомным белком L10 в нетранслируемой последовательности перед геном *rplJ* *Thermotoga maritima* указывает на аутогенный контроль экспрессии этого гена // Генетика.-1996.-32, № 1.-С. 139-143.
14. Патон Е. Б., Живолуп А. Н. Доказательство аутогенной регуляции экспрессии гена *rplJ* у *Thermotoga maritima* и возможности перекрестной регуляции экспрессии по этому принципу между *T. Maritima* и энтеробактериями // Молекуляр. генетика.-1997.-33, № 10.-С. 1341-1344.
15. Живолуп А. Н., Патон Е. Б. Общий механизм регуляции экспрессии в *rplJL* опероне энтеробактерий подтверждается консервативностью структурной организации лидерной области мРНК. Регуляторный консенсус лидерной области мРНК L10-L12 энтеробактерий // Докл. АН Украины.-1995.-340, № 1.-С. 104-107.
16. Petersen C. Long-rang translational coupling in the *rplJL-rpoBC* operon of *Escherichia coli* // J. Mol. Biol.-1989.-206.-P. 323-332.
17. Charles H., Mouchiroud D., Lobry J., Goncaves I., Rahbe Y. Gene size reduction in the bacterial aphid endosymbiont *Buchnera* // Mol. Biol. Evol.-1999.-16.-P. 1820-1822.
18. Clark H., Baumann L., Thao M. L., Moran N. A., Baumann P. Degenerative minimalism in the genome of a psyllid endosymbiont // J. Bacteriol.-2001.-183.-P. 1853-1861.
19. Klein D. J., Schmeing T. M., Moore P. B., Streitz T. A. The kink-tern: a new RNA secondary structure motif // The EMBO J.-2001.-20.-P. 4214-4221.
20. Matsumura S., Ikawa Y., Inoue T. Biochemical characterization of the kink-turn RNA motif // Nucl. Acids Res.-2003.-31.-P. 5544-5551.
21. Razga F., Spackova N., Reblova K. et al. Ribosomal RNA kink-turn motif-a flexible molecular hinge // J. Biomol. Struct. and Dyn.-2004.-22.-P. 183-194.

УДК 577.21
Надійшла до редакції 22.06.06