УДК 577.21

Eugenia B. Paton

THE POTENTIAL mRNA TARGET SITES FOR THE RIBOSOMAL PROTEIN L10 FOUND IN STREPTOMYCES GRISEUS AND S. COELICOLOR GIVE EVIDENCE FOR THE AUTOGENOUS REGULATION OF EXPRESSION OF THE rplJL GENES

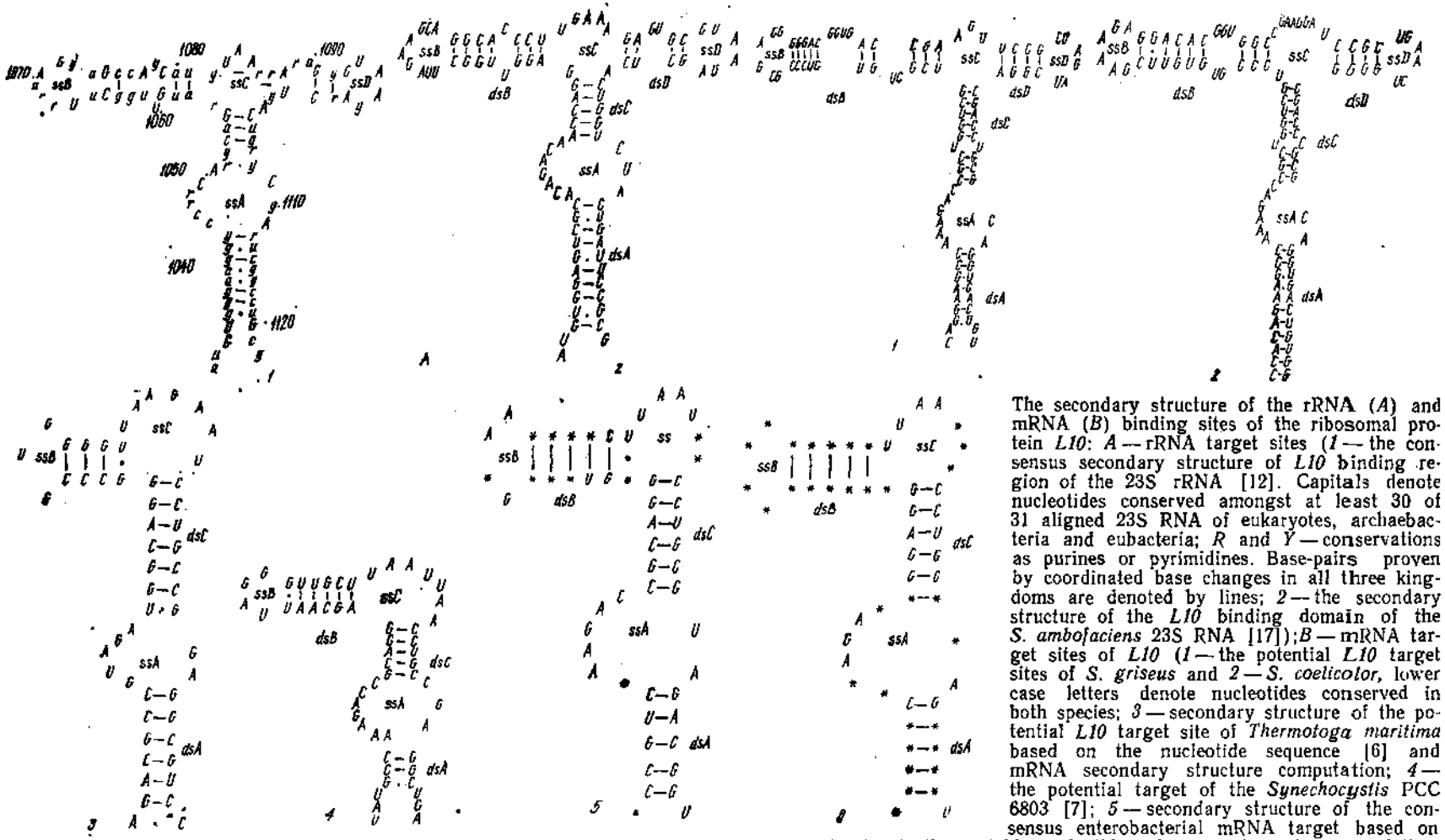
On the basis of structural similarity between the m- and rRNA binding sites which underlies the molecular mechanism of the autogenous control of expression of the rplJL genes, mediated in prokaryotic organisms by ribosomal protein L10, we attempted to find the potential mRNA target sites in S. griseus and S. coelicolor. The potential targets found in both species of Streptomyces are of highly conserved structure and contain elements similar with the L10 binding domain of the 23S RNA. Compared to the L10 mRNA targets of Enterobacteria, Thermotoga maritima and Synechocystis, the target sites of Streptomyces contain a larger number of conserved and specific structural elements, similar in both analyzed species. Comparison of the structural organization of the «Streptomyces» and «Enterobacterial» types of the L10 mRNA targets with the structure of the L10 binding domain in the 23S RNA reveals that the elements, conserved in all the mentioned targets, are located in the ssA and dsA regions. The structure of the potential L10 mRNA target sites, which mimic the rRNA target of the regulatory ribosomal protein L10, gives evidence for the autogenous L10-mediated control of expression of the rplJL genes in Streptomyces.

The expression of genes in the majority of ribosomal protein operons in prokaryotic organisms is regulated autogenously at the translational level [1]. This mode control is achieved due to the possibility of the regulatory ribosomal proteins, encoded by the mRNA of the controlled operons, to bind in a competitive manner the target sites in the m- and rRNA. The structural organization of these sites is highly homologous. In case of the *rplJL* operon, whose genes are regulated by r-protein *L10* (or its 1 : 4 complex with *L12*) [2] we observed the possibility for heterologous translational feedback control in *Enterobacteria* [3, 4], provided by the highly conserved structure of both, the regulatory *L10* proteins [5] and their mRNA target sites [6]. Of particular interest is the fact, that in case of the *L10* target sites, whose structure is proved for *Enterobacteria* [6] and putative — for *Synechocystis* PCC 6803 [7] and *Thermotoga maritima* ([8]; Paton and Zhyvoloup, unpublished data), the degree of homology between the mRNA target sites is higher than between the respective rRNA targets. Moreover, additional structural similarity between the m- and rRNA targets of the *L10* protein is characteristic of *Enterobacteria, Synechocystis* and *Thermotoga*. This makes it reasonable to assume that the *L10*-RNA interaction, which mediates the autogenous regulation of the *rplJL* genes expression is species- and organism-specific. Existence of the «specifying» complementary structural elements, is to be found in both, the regulatory proteins [9] and their mRNA binding sites [6—8]. The latter might be of importance to provide the efficient recognition and tight binding the *L10* protein to the mRNA, which in the case of the rRNA target would be achieved *via* the cooperative binding of the neighbouring r-proteins. Comparison of the *L10* binding regions in the *rplJ* leaders of seven enterobacterial species [6] reveals that the AA1571—1572 (ennumeration as in *E. coli* [10, 11]), whose interaction with the *L10* protein had been proven by chemical analysis [11], are con-

served. The conservation of AA1571—1572 in the potential *L10* mRNA targets of *Synechocystis* and *Thermotoga* (see Fig.) allows an assumption that AA1571—1572 in all of these mRNA targets, are involved in binding to a complementary conserved structural elements of the regulatory protein *L10*. It is notable that the structural homologs of AA1571—1572 in the 23S rRNA belong to binding domain of r-protein *L11* [12]. We reasoned that in other prokaryotic organisms whose *rplJL* genes were feedback-regulated by *L10*, the regulatory *L10*-mRNA interaction might also involve additional structural elements, however, possibly different from AA1571—1572. The present paper is devoted to the analysis of the *rplJ* leader region of *S. griseus* and *S. coelicolor*, whose nucleotide sequence was recently determined [13, 14]. The special attempt was made to find the potential *L10* binding sites and, if it were the case, to compare their structural organization in the both species to the structure of the so far determined mRNA targets of *L10*. The following logic underlied the search for the potential *L10* mRNA target sites.

The mRNA target site of the regulatory protein *L10* is to be homologous to the rRNA target in as much as to provide the competitive binding of the *L10* protein. In all the cases documented so far [2, 6—8, 11] the mRNA target sites for the *L10* protein are located in the untranslated *rplJ* leader region. In both species of *Streptomyces*, the untranslated leader precedes the start of the *L10* cistron. We further analysed the structural homology of the *rplJ* leaders of *Streptomyces* with the *L10* mRNA target sites of *Enterobacteria, Synechocystis* PCC 6803 and *Thermotoga maritima*, as well as with the *L10* binding domain of the 23S RNA of *Streptomyces ambofaciens* (see Fig., *A*). The alignment of the nucleotide sequences was made using the ALIGN program (PCGENE software package). The secondary structure prediction was carried out using the RNAFOLD program (PCGENE). Motives of primary structure identical to the *Synechocystis* and the *Enterobacterial* consensus *L10* mRNA target sites were found in the *rplJ* leaders of both species of *Streptomyces*. Further on, the secondary structure of the potential mRNA *L10* binding sites was calculated (results shown in Figure).

One can notice the pronounced similarity of structure of both *S. griseus* and *S. coelicolor* putative *L10* binding sites with the consensus *L10* binding domain of the 23S RNA [12]. The most striking similarity with the 23S RNA is accounted for by the common general organization, where of particular interest is the presence in the mRNA targets of *Streptomyces* of the two (dsB and dsD) double-stranded regions, rather than a single (dsB) region, which is typical for all of the so far known *L10* mRNA binding sites. The structural organization of the potential *L10* binding sites is highly similar in *S. griseus* and *S. coelicolor*. The single-stranded region C (ssC) is different from the relevant region of the 23S RNA, as a well as from the other prokaryotic mRNA targets. The latter would contain the typical conserved UUAA motif (see Fig., *B*). It is noteworthy that the dsC regions are conserved in both types of mRNA targets «enterobacterial» and «*Streptomyces*» like. However identical within the two types of targets are only the two pairs, G-C and C-G in the dsC region. The highest degree of homology of the mRNA targets in *S. griseus* and *S. coelicolor* is found in the dsC, ssA and sdA regions (see Fig., *B*). The structure in dsC region is completely conserved, except the internal loop, which in *S. griseus* is formed of the U, U, while in *S. coelicolor* — of U, C. The bulged nucleotides are typical structural elements for the protein-nucleic acid interaction. Therefore, we should not exclude the possible species-specifying role of the loop in the dsC region. The dsB, dsC and dsD regions in both pro- and eukaryotic 23S RNA belong to the *L11* binding domain [12], which is interchangable for *E. coli* and yeasts [15]. As already noticed, the mRNA targets of *Enterobacteria, Synechocystis* and *Thermotoga*, although lacking the dsD region, do contain the conserved UUAA motif in the dsC (Fig.). Therefore the structural similarity to the *L11* binding domain of the 23S RNA is present in both



The secondary structure of the rRNA (*A*) and mRNA (*B*) binding sites of the ribosomal protein *L10*: *A* — rRNA target sites (*1* — the consensus secondary structure of *L10* binding region of the 23S rRNA [12]. Capitals denote nucleotides conserved amongst at least 30 of 31 aligned 23S RNA of eukaryotes, archaebacteria and eubacteria; *R* and *Y* — conservations as purines or pyrimidines. Base-pairs proven by coordinated base changes in all three kingdoms are denoted by lines; *2* — the secondary structure of the *L10* binding domain of the *S. ambofaciens* 23S RNA [17]); *B* — mRNA target sites of *L10* (*1* — the potential *L10* target sites of *S. griseus* and *2* — *S. coelicolor*, lower case letters denote nucleotides conserved in both species; *3* — secondary structure of the potential *L10* target site of *Thermotoga maritima* based on the nucleotide sequence [6] and mRNA secondary structure computation; *4* — the potential target of the *Synechocystis* PCC 6803 [7]; *5* — secondary structure of the consensus enterobacterial mRNA target based on comparison of the nucleotide sequences [2, 4] and secondary structure prediction. Asterisks denote the variable nucleotides; *6* — secondary structure of the *L10* mRNA target, consensus for *Enterobacteria*, *Th. maritima* and *Synechocystis*. Asterisks denote the variable nucleotides. The conserved base pairs are denoted by lines

ISSN 0233-7657. БИОПОЛИМЕРЫ И КЛЕТКА. 1995. Т. 11. № 1

types of the mRNA *L10* targets («*Enterobacterial*» and «*Streptomyces*»). In *Streptomyces*, however, this similarity is much higher due to the presence of the dsD region and therefore makes it tempting to assume that *L11* protein, in addition to *L10*, might be involved in *Streptomyces* in the control of expression of the *rplJL* genes. The functional role of *L11* here might consist in co-regulation of expression of the *rplA* and the *rplJL* genes.

The structural elements which comprise the most important similarity between the two types of mRNA targets, the «*Enterobacterial*» and «*Streptomyces*» like, are to be found in regions ssA and dsA. The conserved elements in the ssA region are: the AGA motif on the left side and the A, adjacent to the first C-G pair of the dsA region, on the right side of region ssA. The C-G pair is the most essential conservation in the dsA region. The formation — of this C-G pair provides the most complete structural similarity between the m- and rRNA targets of *L10* [16]. This feature is exploited in the models of the conformational switches of the *rplJ* leaders of *E. coli* [2] and the six other species of *Enterobacteria* [6], as well as those of *Thermotoga maritima* and *Synechocystis* (Paton and Zhyvoloup, unpublished) to explain the molecular mechanism of the *L10*-mediated feedback regulation. The conservation of the C-G pair (dsA), and the adjacent A (ssA) in the putative mRNA target sites of the *L10* protein of *Streptomyces* proves the functional importance of these elements for *L10* binding. The identical elements of the mRNA and rRNA targets (ssA region and the C-G pair of dsA) suggest a complementary structural conservation in the regulatory protein. We reason that the -RNTLL- motif, which is the only one conserved among the 17 analyzed *L10* protein sequences, is complementary to the conserved structural elements in dsA and ssA regions of the RNA target sites [9].

It may be noted in conclusion that the potential *L10* mRNA target sites found in the two species of *Streptomyces, S. griseus* and *S. coelicolor*, prove once again the principle of feedback regulation of r-protein synthesis in prokaryotic organisms being based on the structural similarity of the m- and rRNA targets, competitive for binding of the regulatory proteins. Similarly to other prokaryotes, the mRNA target of *L10* is located in *Streptomyces* in the untranslated *rplJ* leader region more than a 100 nucleotides upstream the *L10* start codon and therefore points to the important functional role of the *rplJ* leader. The conserved structure of the found potential *L10* mRNA target is an evidence of its functional role for the autogenous control of expression of the *rplJL* genes in *Streptomyces*. The structural specificity of the *L10* mRNA targets found in *Streptomyces* might reflect the degree of phylogenetic remoteness. Given the peculiarities of the structural organization of the *L10* target, as well as of the whole *rplJ* leader region (Paton and Zhyvoloup, in preparation), one might expect some other specific features pertinent to the molecular mechanism of the *L10*-mediated autogenous control of the *rplJL* genes expression in *Streptomyces*.

Є. Б. Патон

ПОТЕНЦІЙНІ мРНКові САЙТИ-МІШЕНІ ДЛЯ РИБОСОМНОГО БІЛКА L10, ЯКІ ВИЯВЛЕНО У STREPTOMYCES GRISEUS І S. COELICOLOR, ВКАЗУЮТЬ НА АУТОГЕННИЙ КОНТРОЛЬ ЕКСПРЕСІЇ ГЕНІВ rplJL

Резюме

Описано структурові особливості виявленого у представників *Streptomyces* постійного сайта зв'язування рибосомного білка *L10*, який розміщується в лідерній послідовності, що передує генам *rplJL*. Консервативність структурової організації сайта у двох видів *Streptomyces* і значна подібність структури до 23S мРНК свідчать про аутогенну регуляцію генів *rplJL* стрептоміцетів рибосомним білком *L10*.

REFERENCES

1. *Lindahl L., Zengel J. M.* Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli* // Progr. Nucl. Acids Res. and Mol. Biol.— 1993.— N 1.— P. 3—57.
2. *Christensen T., Johnsen M., Fiil N. P., Friesen J. D.* RNA secondary structure and translational inhibition: analysis of the mutants in the *rplJ* leader // EMBO J.— 1984.— 3, N 7.— P. 1609—1612.
3. *Paton E. B., Woodmaska M. I., Kroupskaya I. V. et al.* Evidence for the ability of the *L10* ribosomal proteins of *Salmonella typhimurium* and *Kilebsiella pneumoniae* to regulate *rplJL* genes expression in *E. coli* // FEBS Let.— 1990.— **265**, N 1, 2.— P. 129—132.
4. *Zhyvoloup A. N., Paton E. B.* The *E. coli* ribosomal protein *L10* of *E. coli* is capable to regulate expression of the *rplJL* genes in *Citrobacter freundii* // Biopolymers and Cell.— 1993.— **9**, N 6.— P. 88—100.
5. *Paton E. B., Zhyvoloup A. N., Woodmaska M. I., Kroupskaya I. V.* Nucleotide sequence of the *rplJL* operon and the deduced primary structure of the encoded *L10* and *L7/L12*, proteins of *S. typhimurium* compared to that of *E. coli* // Nucl. Acids Res.— 1990.— **18**.— P. 4620.
6. *Zhyvoloup A. N., Kroupskaya I. V., Lyaskovsky T. M., Paton E. B.* Comparison of nucleotide sequences of the *rplJ* leader in *Enterobacteria* // Biopolymers and Cell.— 1994.— **10**, N 1.— P. 37—40.
7. *Schmidt J., Bubenko M., Subramanian A.* A novel operon organization involving the genes for the chorismate synthase aromatic biosynthesis pathway and ribosomal GTPase center proteins *L11, L1, L10, L12: rplKJL* in *Cyanobacterium synechocystis* PCC 6803 // J. Biol. Chem.— 1993.— **268**, N 36.— P. 27447—27457.
8. *Liao D., Dennis P. P.* The organization and expression of essential transcription translation component genes in the extremely thermophilic eubacterium *Thermotoga maritima* // Ibid.— 1992.— **267**, N 32.— P. 22787—22797.
9. *Tereschcenko F. M., Paton E. B.* Identification of the conserved structural elements in the ribosomal protein *L10* which are essential for the autogenous regulation of gene expression in the *rplJL* operon of *Enterobacteria* // Biolopymers and Cell.— 1995.— **11**, N 1.— P. 00—00.
10. *Post L. E., Strycharz C. D., Nomura M. et al.* Nucleotide sequence of the ribosomal protein cluster adjacent to the gene for RNA polymerase subunit *b* in *Escherichia coli* // Proc. Nat. Acad. Sci. USA.— 1979.— **76**, N 4.— P. 1697—1701.
11. *Climie S. C., Friesen J. D. In vivo* and *in vitro* structural analysis of the *rplJ* mRNA leader in *Escherichia coli* // J. Biol. Chem.— 1988.— **263**, N 29.— P. 16166—16175.
12. *Egebjerg J., Douthwaite S. R., Liljas A., Garrett R. A.* Characterization of the binding sites of protein *L11* and the *L10 (L12)* pentameric complex in the GTPase domain of 23S ribosomal RNA from *Escherichia coli* // J. Mol. Biol.— 1990.— **213**, N 1.— P. 275—289.
13. *Kuester K., Kuberski S., Piepersberg W., Distler J.* Cloning and nucleotide sequence analysis of the nusg-*rplK, rplA, rplJ, rplL* gene cluster of *S. griseus* // EMBL accession N X72787.
14. *Bianco G., Rodicio R., Puglia A. M. et al.* Synthesis of ribosomal proteins during growth of *Streptomyces coelicolor* // EMBL accession N L24552.
15. *Thompson J., Musters W., Cundliffe E., Dahlberg A. E.* Replacement of the *L11* binding region within *E. coli* 23S ribosomal RNA with its homologue from yeast: *in vivo* and *in vitro* analysis of hybrid ribosomes altered in the GTPase centre // EMBO J.— 1993.— **12**, N 4.— P. 1499—1504.
16. *Draper D. E.* How do proteins recognise specific RNA sites? New clews from autogenously regulated ribosomal proteins // Trends in Biochem. Sci.— 1989.— **14**, N 8.— P. 335—338.
17. *Gutell R. R., Schnare M. N., Gray M. W.* A compilation of large subunit (23S-like) ribosomal RNA sequences presented in a secondary structure format // Nucl. Acids Res.— 1991.— **18**, Suppl.— P. 2319.

Inst. Cell. Biol. and Genet. Engineer.
the Nat. Acad. Sci. of Ukraine, Kiev

15.09.94