Volume 219, number 2, 269–273 FEB 04924 July 1987

Hypothesis

A putative internal promoter in the 16 S/23 S intergenic spacer of the rRNA operon of archaebacteria and eubacteria

A.S. Mankin, E.A. Skripkin and V.K. Kagramanova

Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry, Moscow State University, Moscow 119899, USSR

Received 15 May 1987

The existence of the internal promoter Pi in the 16 S/23 S intergenic spacers of the rRNA operons of an eubacterium *Escherichia coli* and archaebacterium *Halobacterium halobium* is proposed. The possible functional significance of these promoters is discussed.

Internal promoter; rRNA expression; Archaebacteria; rRNA operon; (*Halobacterium halobium, Escherichia coli*)

1. INTRODUCTION

The genes of ribosomal RNAs of eubacteria are organized in operons in the order 16 S-23 S-5 S. It is widely assumed that transcription starts from promoters located upstream from the first structural gene and spans through the entire operon (review [1]). It seems that the same strategy of rRNA biosynthesis holds true for some archaebacteria, at least for those harbouring closely linked rRNA genes. However, while examining and comparing the primary structures of the *rrn*B operon of eubacterium *Escherichia coli* [2] and of the recently sequenced single rRNA operon of the archaebacterium *Halobacterium halobium* [3], we came across promoter-like sequences inside the operon bodies. In both cases these structures were localized in the 16 S/23 S intergenic spacer immediately upstream from the 23 S pre-rRNA processing signals (fig.1A). Being nonhomologous to each other, these *E. coli* and *H. halobium* sequences closely resembled the consensus promoter structures of eubacteria and archaebacteria, respectively. This provocative finding raised the possibility of utilization of the internal promoters for the rRNA biosynthesis in some procaryotes.

Correspondence address: A.S. Mankin, Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry, Moscow State University, Moscow 119899, USSR

2. RESULTS OF THE NUCLEOTIDE SEQUENCE ANALYSIS

2.1. *A putative internal promoter in the* rrn*B operon of the* E. coli

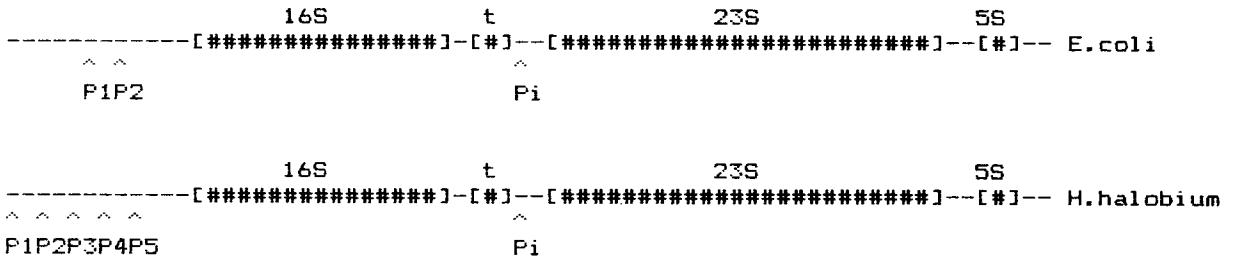
The canonical structure of the eubacterial promoter is represented by the two consensus hexanucleotide sequences, TTGACA and TATAAT, separated by 16–18 bp and located at approx. 35 and 10 bp, respectively, upstream from the transcription starting point [4]. The transcription of the ribosomal RNA operons of *E. coli* was shown to be regulated by the two tandem promoters, P1 and P2, closely resembling the 'ideal' promoter sequence. However, a scan of the complete nucleotide sequence of the *rrn*B operon

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50 © 1987 Federation of European Biochemical Societies

of *E. coli* revealed, in the spacer region separating the 16 S and 23 S rRNA genes, the presence of a structure that fits these criteria well (fig.1B). The '−35' and '−10' boxes of the putative internal promoter (Pi) are represented by the sequences TTGTGA and TAATAT positioned 17 bp from each other. Although the homology score for the Pi (reflecting the degree of homology to the consensus eubacterial promoter) is not very high, 41.2% if calculated according to [5], it falls within a range of homology scores for the moderate *E. coli* promoters [4].

It is well established that the authentic promoters of the *E. coli* rRNA operons confer the antitermination properties on the RNA polymerase, which enables it to ignore potential cryptic terminators inside the nontranslatable operon [6–8]. The DNA sequences located in close proximity to the promoters play an important role in this mechanism. Among them is a *nus A* utilization

A.



B.

```
TTGACTCTGAAAGAGGAA AGCGTAATATA -  37bp  - TGCTCTTTA : P2 rrnE
***    ******  *   * * ******              *********
TTGTGAGTGAAAGTCGCCGACCTTAATATC -  34bp  - TGCTCTTTA : Pi rrnB

"-35"                  "-10"                "box A"
```

C.

```
AGCGTGATTCGATGCCCTTAAGTAATA-ACGGGGCGTTCGGGGAAATGCGAACGTC    P5
* * **  ***   *********   * *******   **  ***  ****
ATCTTGGATCGTGTCCCTTAAGTGGGAGACGGGGCAA-CGATGAATCGCGACGAAG    Pi
```

Fig.1. (A) Physical maps of the *rrn*B operon of *E. coli* and of the single rRNA operon of *H. halobium*. Structural rRNA and tRNA genes and location of outer and putative inner promoters are indicated. (B) The homology of the P2 promoter of the *rrn*E operon and the Pi promoter of the *rrn*B operon of *E. coli*. Asterisks indicate the homologous nucleotides. The −35 and −10 boxes and *box A* are underlined. (C) The homology of the 16 S rRNA gene proximal P5 promoter and Pi promoter of the rRNA operon of *H. halobium*. Asterisks show homologous nucleotides; consensus boxes are underlined.

sequence (*box A*) which is present in the leader regions of all seven *rrn* operons at 37–38 bp downstream from the Pribnow box of the P2 promoters [8]. *box A* sequences were also found by M. Cashel prior to the 23 S rRNA genes (see [8]). Having identified the putative internal promoter in the 16 S/23 S intergenic spacer, we were pleased to see that the 23 S gene proximal *nus A* utilization sequence is localized at almost precisely the same distance from the internal promoter −10 box as the leader *box A* from that of P2 promoter (35 and 37 bp, respectively). This observation suggests that the internal promoter (if active), like the authentic *rrn* P1,P2 promoters, might confer the ability to transcribe through cryptic termination signals.

2.2. *A putative internal promoter in the rRNA operon of the* H. halobium

Recently we determined a complete primary structure of the single rRNA operon of the halophilic archaebacterium *H. halobium*. It was the first rRNA operon from the archaebacterial kingdom that was sequenced in its entirety [3]. In some respects, viz., the order and sizes of structural rRNA genes, processing signals, the sizes of the spacer sequences, the presence of the tRNA gene in the 16 S/23 S spacer, it closely resembles the 'typical' eubacterial rRNA operon. According to some experimental results, it could be transcribed as a single transcriptional unit [9,10]. The 900 bp region preceding the 16 S rRNA gene contains five imperfect direct repeat sequences over 80 bp length. It has been suggested that these repeats represented tandem promoters (P1–P5) of the operon [11,12]. The results of the nuclease S1 mapping of the transcription starting points supported this proposition [13].

Though the precise structure of the archaebacterial promoter is still obscure, a comparison of the rRNA operon promoter repeats of *H. halobium* provided some insight into certain features of the stable RNA gene transcription signals in the halophilic archaebacteria. Two particular nucleotide stretches, more conservative than the other parts of the repeats, were outlined [11,12]. One of them, CCCTTAAGT, was present in all but one repeat. The second one, TGCGAAC, located at 25–30 bp downstream from the first box, was identified in all five repeats and was associated with the transcription starting point ([13], our unpublished results). Our analysis [14] also revealed the presence of the first box and some reminiscences of the second one in the promoter regions of another two stable RNA genes of Halobacteria: that of the 7 S RNA gene of *H. halobium* [15] and of the Trp tRNA gene of *H. volcanii* [16]. These two boxes were proposed to be significant for promoter activity [11,12,17].

Again, as in the case of eubacteria, we were able to identify a promoter motif in the 16 S/23 S intergenic spacer of the rRNA operon of *H. halobium* (fig.1A). In the archaebacterium the homology of the internal promoter structure to the 'outer' operon promoters was even more pronounced (fig.1C). It covered not only the defined boxes, but also the sequences between them, so that the degree of homology with the 16 S rRNA gene proximal promoter, P5, raised up to 68.5%. It may be relevant, as the *nus A* utilizing sequence (*box A*) is located downstream from the P2 and Pi promoters of the *rrn* operons in *E. coli*, that a conserved stretch of nucleotides (GACTGAT) is found at a short distance downstream from the transcription start box of the P5 promoter and from that of internal Pi promoter of *H. halobium*. If archaebacteria, like eubacteria, possess an antitermination mechanism (that seems reasonable), then one can speculate that this sequence could be a candidate for the halobacterial functional analog of the *box A*.

3. DISCUSSION

In the previous sections we have demonstrated the presence of internal promoter-like structures in the 16 S/23 S intergenic spacers of *rrn* operons of an eubacterium *E. coli* and an archaebacterium *H. halobium*. Some arguments could be advanced for the functional significance of these internal promoters.

It should be emphasized that the putative spacer promoters in *E. coli* and *H. halobium* are encoded by the nonhomologous nucleotide stretches. Their location in the highly variable noncoding sequences suggests that the presence of the putative promoters in the spacers of the rRNA operon of both organisms is not related to any primary structure conservation constraints.

In this respect, the here described putative inter-

nal promoters should be distinguished from the promoter proposed by Amemiya et al. [18] inside the 16 S rRNA structural gene of an eubacterium *Caulobacter crescentus*. The sequence identified by these authors is located in one of the most conserved segments of the 16 S rRNA and is present not only in eubacterial, but also in archaebacterial and eucaryotic 16 S-like rRNA genes. Assuming that the consensus promoter motives in the three primary kingdoms differ significantly, it is unlikely that this sequence bears a promoter function also in archaebacteria and eucaryotes. Therefore the proposition of Amemiya et al. [18] that this conserved region in the 16 S rRNA gene serves a purpose other than that of a promoter seems reasonable. Thus, the transcription initiation from this sequence in *C. crescentus* might reflect some aberrant, rather than common biosynthesis pathway.

The internal promoter activity (if any) inside the ribosomal operon should not interfere with the major posttranscriptional processing events that lead to the appearance of the ultimate rRNAs in the cell. It is noteworthy that spacer Pi promoters in *E. coli* and *H. halobium* are found just upstream from the nucleotide stretches which participate in the formation of a long stable hairpin, flanking the 23 S rRNA sequence. This hairpin is recognized by the processing nuclease in *E. coli* and presumably by an analogous enzyme in Halobacteria [3,12,19]. Thus, the transcripts initiated from the internal promoters Pi would be normally processed to mature 23 S and 5 S rRNAs.

At first glance the activity of the Pi promoters might seem to lead to the overproduction of the 23 S and 5 S rRNAs over the 16 S rRNA. However, this contradiction could be easily withdrawn if one considers the possibility of the transcription termination on the cryptic terminators inside the ribosomal operons. In spite of the existence of the mechanism of antitermination provided by the rRNA operon promoters, it is not sufficient to overcome completely premature termination [8]. Such aberrant termination would lead to the slight increase in the level of the promoter proximal gene product (16 S rRNA) over the promoter distal rRNAs (23 S and 5 S). The internal promoter located upstream from the 23 S rRNA gene could compensate for such an imbalance and adjust the cellular level of the ribosomal RNAs.

It still remains to be determined whether RNA polymerases in the bacterial cells can initiate in vivo from the internal Pi promoters under certain physiological conditions. We were able to detect in exponentially growing *E. coli*, as well as in *H. halobium* cells, the presence of RNA molecules with 5′-ends mapping exactly to the internal promoter transcription starting point in the rRNA operons of these organisms (unpublished results). However, the triphosphate nature of the 5′-RNA termini has not yet been demonstrated and therefore these molecules may be attributed not only to the initiation of transcription from the internal promoters, but also to the processing of the transcripts initiated somewhere upstream. This work is now in progress.

ACKNOWLEDGEMENTS

We are grateful to Professor Bogdanov and Dr Baratova for stimulating discussions.

REFERENCES

[1] Lamond, A.I. (1985) Trends Biochem. Sci. 10, 271–274.
[2] Brosius, J., Dull, T.J., Sleeter, D.D. and Noller, H.F. (1981) J. Mol. Biol. 148, 107–127.
[3] Mankin, A.S. and Kagramanova, V.K. (1986) Mol. Gen. Genet. 202, 152–161.
[4] Hawley, D.K. and McClure, W.R. (1983) Nucleic Acids Res. 11, 2237–2255.
[5] Mulligan, M.E., Hawley, D.K., Entriken, R. and McClure, W.R. (1984) Nucleic Acids Res. 12, 789–800.
[6] Holben, W.E. and Morgan, E.A. (1984) Proc. Natl. Acad. Sci. USA 81, 6789–6793.
[7] Li, S.C., Squires, C.L. and Squires, C. (1984) Cell 38, 851–860.
[8] Holben, W.E., Prazad, S.M. and Morgan, E.A. (1985) Proc. Natl. Acad. Sci. USA 82, 5073–5077.
[9] Hofman, J.D., Lau, R.H. and Doolittle, W.F. (1979) Nucleic Acids Res. 7, 1321–1333.
[10] Chant, J. and Dennis, P. (1986) EMBO J. 5, 1091–1097.
[11] Mankin, A.S., Teterina, N.L., Rubtsov, P.M., Baratova, L.A. and Kagramanova, V.K. (1984) Nucleic Acids Res. 12, 6537–6546.
[12] Hui, I. and Dennis, P.P. (1985) J. Biol. Chem. 260, 899–906.

[13] Dennis, P.P. (1986) J. Mol. Biol. 186, 457–461.
[14] Mankin, A.S., Skripkin, E.A. and Kagramanova, V.K. (1987) Biol. Nauki, in press.
[15] Moritz, A., Lankat-Buttgereit, B., Gross, H.J. and Goebel, W. (1985) Nucleic Acids Res. 13, 31–43.
[16] Daniels, C.J., Gupta, R. and Dolittle, W.F. (1985) J. Biol. Chem. 260, 3132–3134.
[17] Wich, G., Hummel, H., Jarsch, M., Bar, U. and Bock, A. (1986) Nucleic Acids Res. 14, 2459–2479.
[18] Amemiya, K., Belofatto, V., Shapiro, L. and Feingold, J. (1986) J. Mol. Biol. 187, 1–14.
[19] Young, R. and Steitz, J. (1978) Proc. Natl. Acad. Sci. USA 75, 3593–3597.