Volume 297, number 3, 209–211 FEBS 10680 February 1992

Hypothesis

'*box*A'-like sequence between the 16 S/23 S spacer in rRNA operon of mycoplasmas

R. Harasawa[a], T. Uemori[b], K. Asada[b], I. Kato[b] and N. Shiragami[c]

[a]*Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan,* [b]*Biotechnology Research Laboratories, Takara Shuzo Co., Otsu-shi, Shiga 520-21, Japan* and [c]*Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama 227, Japan*

Received 2 December 1991

We have found that a *boxA*-like sequence is conserved in the 16 S and 23 S rRNA intergenic spacer regions of mycoplasmas, and that it always locates on loop regions of the hypothetical secondary stem-loop structures. A nucleotide sequence similar to the '−10' box of prokaryotic promoters was identified at upstream sites of the *boxA*-like sequence in the 16 S/23 S spacer regions. These structures may represent an internal promoter between the 16 S and 23 S rRNA genes in mycoplasmas.

boxA; rRNA; Antitermination; Mycoplasma; Promoter

The so-called *boxA* sequence, originally found at upstream regions of the *nut* site of the lambda phage genome, is considered to be the recognition site for the *Escherichia coli* host NusA protein, a transcription termination factor [1]. A sequence homologous to the *boxA* of phage lambda has also been identified in the tryptophanase operon [2], and in the leader sequence of glutamate synthase structural genes [3] of *E. coli*. A large stem-loop structure with the *boxA* sequence, which causes strong reduction of gene expression, has been demonstrated between *melA* and *melB* of the melibiose operon of *E. coli* [4]. Although the *boxA* sequence in lambda phage is implicated in trascription antitermination [5], function of the *boxA*-like sequences located in the structural genes of *E. coli* has not been well understood. The *boxA* sequence has been found in a wheat tRNA gene of *Triticum vulgare* var. *aria* [6]. The *boxA*-like sequences have also been reported in the 16 S/23 S rRNA intergenic spacer of *rrn*B of *E. coli* [7] as well as in the promoter regions of *rrn*B and *rrn*E of *E. coli* [8,9] and of *rrn*B and *rrn*O of *Bacillus subtilis* [10,11].

We have found the *boxA*-like sequence in the 16 S/23 S rRNA intergenic spacer regions of mycoplasmas, the smallest and simplest self-replicating prokaryotes. The mycoplasma genome is also the smallest in size among free-living cells. Therefore, mycoplasma cells are considered to represent the minimum living system and to have minimal numbers of genes indispensable for growth [12]. Mycoplasmas are known to carry only one or two sets of rRNA genes in their genome [13].

Ribosomal RNA genes of mycoplasmas are organized in rRNA operons and transcribed from an upstream promoter of the 16 S rRNA gene in the arrangement of 5′-16 S-23 S-5 S-3′ [14]. In one minor variation, *Mycoplasma hyopneumoniae*, the 16 S and 23 S rRNA genes are close, but the 5 S rRNA gene is separated by about 4 kb [15]. A different structural organization has recently reported for *M. gallisepticum* in which one locus contains 16 S, 23 S and 5 S rRNA genes; a second contains 23 S and presumably 5 S; and a third appears to have only the 16 S rRNA gene [16]. In *E. coli* and *B. subtilis*, some tRNA genes are known to be located in the spacer region between the 16 S and 23 S rRNA genes and are co-transcribed with them. However no tRNA genes have been evident so far between the 16 S and 23 S rRNA genes of mycoplasmas like eukaryotes [15,17,18]. The spacer regions between the 16 S and 23 S rRNA genes of *M. hyopneumoniae* and *M. capricolum* contain sequences complementary to the 5′- and 3′-flanking sequences of both 16 S and 23 S rRNA genes [17,19], suggesting that the rRNA transcript can generate large stem-loop structures by pairing between the spacer regions and the 5′ or 3′-flanking sequences of 16 S and 23 S rRNA. These stem-loop structures in primary rRNA transcripts are known as possible substrates for processing enzymes during rRNA maturation [9,11] in other eubacteria such as *E. coli* [20] and *B. subtilis* [10]. This processing site between the 16 S/23 S rRNA spacer regions has also been demonstrated in other *Mycoplasma* species [18].

Correspondence address: R. Harasawa, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan. Fax: (81) (3) 3816 5680.

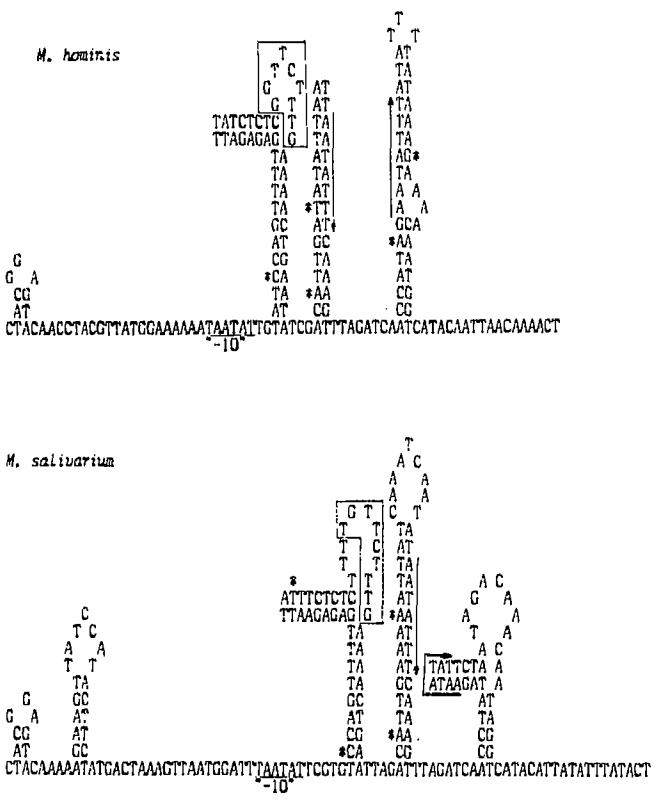

Fig. 1. Hypothetical secondary structure of the spacer region between 16 S and 23 S rRNA genes of *M. hominis* and *M. salivarium*. The '*boxA*'-like sequence is boxed. Direct repeats are indicated by arrows. Mismatch pairing is shown by an asterisk.

We have previously sequenced the spacer regions between the 16 S and 23 S rRNA genes of the following mycoplasmas and deposited their sequences to the EMBL Data Library under the accession numbers given in parentheses [18]: *M. orale* CH19299 (X58556), *M. salivarium* PG20 (X58558), *M. hominis* PG21 (X58559), *M. fermentans* PG18 (X58553), *M. arginini* G230 (X58560), *M. hyorhinis* BTS7 (X58555), *M. arthritidis* PG6 (X58557), *M. pulmonis* m53 (X58554), *M. neurolyticum* PG28 (X58552), *M. hyopneumoniae* VPP11 (X58551), and *Ureaplasma urealyticum* T960 (X58561). All of these sequences were investigated for direct repeats, inverted repeats and specific motifs by using a SEQA program (Shiragami, unpublished). Regions of significant sequence similarity were aligned by a combination of visual inspection. Hypothetical secondary structures were edited and plotted from the results deduced from the SEQA program.

Several palindromic sequences, which may be responsible for pausing, were identified between the 16 S/23 S rRNA intergenic spacer regions of the 11 *Mycoplasma* species examined (Fig. 1). However neither tRNA genes nor their pseudogenes were found between the 16 S/23 S rRNA intergenic spacer regions of mycoplasmas, and this is similar to eukaryotic spacers. A *boxA*-like sequence was seen within the rRNA processing sites of all the *Mycoplasma* species examined and it always locates on the loop of the possible secondary stem-loop structures (Fig. 1). A short consensus motif 5'-CTTT(G/A)-3' of the *boxA* is similar to a direct repeat in the controlling region of the *E. coli lac* operon. A direct repeat sequence, 5'-AATATTT-3', is conserved among several mycoplasmas examined, but its logical function is unknown. A putative promoter sequence similar to the '−10 (TAATAT)' box of the P2 promoter of the *E. coli* rRNA operon E was found between 19 and 35 bp upstream from the *boxA*-like sequence in *M. salivarium*, *M. hominis* and *U. urealyticum* (Fig. 2). These common sequences may represent an internal promoter between the 16 S/23 S rRNA genes though its promoter activity has not been examined yet. If it is active, *M. hyopneumoniae* and *M. gallisepticum* may use this internal promoter for the transcription of their split rRNA genes which lack upstream leader promoters of the 16 S rRNA gene. It has been reported that some eubacteria and archebacteria have also putative internal promoters between the 16 S and 23 S rRNA genes [21,22]. Although the presence of a promoter-like sequence in the 16 S/23 S rRNA intergenic spacer region has been considered to be responsible for a different function [8], it may not be easy to interpret this common structure by chance because it is highly conserved among such distant microorganisms. Our observation demonstrates that the *boxA*-like sequences are conserved in mycoplasmas, the smallest self-replicating procaryotes. Mycoplasmas may provide a simple model to examine the function of the '*boxA*' sequences in vivo.

"−35"	"−10"	"boxA"	
TTGACT.....	TAATAT..... 37bp.....	TGCTCTTTA:	*E. coli* (P2 rrnE)
TTGACC.....	TACTAT..... 28bp.....	AGTTCTTTG:	*B. subtilis* (rrnO)
TTGACA.....	TATATT..... 31bp.....	TGATCTTTG:	*B. subtilis* (rrnB)
ATGACT.....	TAATAT..... 35bp.....	TGTTCTTTG:	*M. salivarium*
TAGACC.....	TATATC..... 34bp.....	TGTTCTTTG:	*M. fermentans*
GAGACT.....	TCGTAT..... 31bp.....	TGTTCTTTG:	*M. arthritidis*
(?)	ATTATT..... 33bp.....	TGTTCTTTG:	*M. arginini*
(?)	CAATAT..... 34bp.....	TATTCTTTG:	*M. orale*
(?)	TCTAAT..... 17bp.....	TATTCTTTG:	*M. neurolyticum*
(?)	TCTAAA..... 36bp.....	AGCTCTTTC:	*M. hyopneumoniae*
TTGACT.....	CATAAT..... 41bp.....	AGTTCTTTG:	*M. hyorhinis*
(?)	TACTAT..... 33bp.....	AGTTCTTTA:	*M. pulmonis*
(?)	TAATAT..... 29bp.....	GGTTCTTTG:	*M. hominis*
(?)	AATATT..... 19bp.....	CGATCTTTG:	*U. urealyticum*

Fig. 2. Comparison between the authentic *boxA* and promoter sequences in *E. coli* and *B. subtilis* and the *boxA*-like sequences and putative internal promoters in mycoplasmas.

REFERENCES

[1] Olson, E.R., Flamm, E.L. and Friedman, D.I. (1982) Cell 31, 61–70.
[2] Stewart, V. and Yanofsky, C. (1985) J. Bacteriol. 164, 731–740.
[3] Oliver, G., Gosset, G., Sanchez-Pescador, R., Lozoya, E., Ku, L.M., Flores, N., Becerril, B., Valle, F. and Bolivar, F. (1987) Gene 60, 1–11.
[4] Shimamoto, T., Noguchi, K., Kuroda, M. and Tsuchiya, T. (1988) Nucleic Acids Symp. Ser. 19, 171–173.
[5] Olson, E.R., Tomich, C.-S. and Friedman, D.L. (1984) J. Mol. Biol. 180, 1053–1063.
[6] Szweykowska-Kulinska, Z., Jarmolowski, A. and Augustyniak, J. (1989) Gene 77, 163–167.
[7] Mankin, A.S., Skripkin, E.A. and Kagramanova, V.K. (1987) FEBS Lett. 219, 269–273.
[8] Zacharias, M. and Wagner, R. (1989) Mol. Microbiol. 3, 405–410.
[9] Young, R.A. and Steitz, J.A. (1978) Proc. Natl. Acad. Sci. USA 75, 3593–3597.
[10] Stewart, G.C. and Bott, K.F. (1983) Nucleic Acids Res. 11, 6289–6300.
[11] Ogasawara, N., Moriya, S. and Yoshikawa, H. (1983) Nucleic Acids Res. 11, 6301–6318.
[12] Muto, A., Yamao, F. and Osawa, S. (1987) Prog. Nucleic Acid Res. Mol. Biol. 34, 29–58.
[13] Amikam, D., Glaser, G. and Razin, S. (1984) J. Bacteriol. 158, 376–378.
[14] Razin, S. (1985) Microbiol. Rev. 49, 419–455.
[15] Taschke, C., Klinkert, M.-Q., Wolters, J. and Herrmann, R. (1986) Mol. Gen. Genet. 205, 42282–433.
[16] Chen, X. and Finch, L.R. (1989) J. Bacteriol. 171, 2876–2878.
[17] Iwami, M., Muto, A., Yamao, F. and Osawa, S. (1984) Mol. Gen. Genet. 196, 311–316.
[18] Uemori, T., Asada, K., Kato, I. and Harasawa, R. (1992) Syst. Appl. Microbiol. (in press).
[19] Taschke, C. and Herrmann, R. (1986) Mol. Gen. Genet. 205, 434–441.
[20] Young, R.A., Macklis, R. and Steitz, J.A. (1979) J. Biol. Chem. 254, 3264–3271.
[21] Mankin, A.S. and Kagramanova, V.K. (1988) Nucleic Acids Res. 16, 4679–4692.
[22] Kjems, J. and Garrett, R.A. (1990) J. Mol. Evol. 31, 25–32.