Function of the repeated sequence in the 3′ flanking region of the *Escherichia coli rnpB* gene on transcription termination and RNA processing

Semi Kim, Younghoon Lee*

Department of Chemistry, Korea Advanced Institute of Science and Technology, Taejon 305-701, South Korea

Received 12 February 1997; revised version received 25 March 1997

Abstract The 3′ flanking region of the *Escherichia coli rnpB* gene-encoding M1 RNA, the RNA component of RNase P, contains a 113 bp repeated sequence. This sequence, successively reiterating 3.5 times, includes the region for intrinsic termination. In vivo termination of transcription occurs mostly at the first terminator (T1). Analysis of deletions at the 3′ flanking region revealed that the second terminator (T2) and third (T3) are functional in vivo and that the sequences preceding the region coding for an RNA-terminator hairpin and U-rich 3′ tail are essential for efficient termination. Transcripts terminating at T2 and T3 were also processed at the 3′ end in a manner similar to those terminating at T1.

Key words: M1 RNA; RNA processing; RNase P; *rnpB*; Termination; *Escherichia coli*

1. Introduction

M1 RNA is the catalytic component of *Escherichia coli* RNase P, which is a tRNA-processing enzyme [1,2]. The 377 nucleotide M1 RNA is synthesized by transcription of the *rnpB* gene [3,4] and subsequent processing of the resulting transcripts [5–8]. The 3′ flanking region of the *rnpB* gene contains a repeated unit of 113 bp [9] which includes both the sequence coding for the 3′ terminal 24 nucleotides of M1 RNA and the region for intrinsic transcription termination [10–12]. This unit starts at position +354 and successively reiterates almost 3.5 times. The overlapping coding regions for two small peptides extend from one repeat to the next repeat [2]. In vivo transcription terminates almost exclusively at the first terminator, T1, in the first repeating unit [8,13]. The RNA transcripts terminating at T1 have an extra 36 nucleotides at the 3′ end of M1 RNA, which should be processed to generate the mature 3′ end [5,8]. This 3′ processing event does not require the intact M1 RNA sequence but depends on the *rne*-dependent site, immediately 3′ to the processing site [8], which is coded by the corresponding sequence within the repeating unit. Although the major termination occurs at T1 with a termination efficiency of 94% in a GalK reporter system, the repeated sequences following T1 are required for complete termination [13]. This suggests that the second and third terminators, T2 and T3, respectively, in the following repeating units are functional. However, there is no direct evidence whether T2 and T3 are functional independently. If T2 and T3 are functional, the three terminators all together would work for securing no run-through transcription into genes downstream of *rnpB* [14]. Furthermore, translation of the coding regions for the putative small peptides could be interrupted by termination at T2 and/or T3 [9].

In this report, we carried out a deletion experiment at the 3′ flanking region to examine in vivo termination at T2 or T3. We show that T2 and T3 are functional. From this study, in addition, we find that the RNA-terminator hairpin and U-rich 3′ tail which is known as an essential element of intrinsic terminators [10–12] is not sufficient for efficient termination of *rnpB* transcription in vivo and that 3′ processing of transcripts terminating at T2 and T3 occurs like that of transcripts terminating at T1.

2. Materials and methods

2.1. Bacterial strain and plasmids

Plasmid pDBSK [15], a derivative of the pGEM3 vector, was a parental plasmid containing the *rnpB* sequence of positions −270 to +1286, from which the deletion derivatives originated. Plasmids pLMd23, pLMdd5-9, pLMdd5-23 and pLMdd5-50 are some of the derivatives with the internal deletion of the sequences between positions +57 and +330, +139 and +312, +139 and +330, and +139 and +363, respectively [8,15]. *E. coli* strain JM109 [16] was used for the construction and propagation of plasmids.

2.2. Construction of deletion plasmids

Plasmids with internal deletions in the *rnpB* gene were constructed using *Bal*31 as described [8]. The deletion was bidirectionally made from position +290 of the *rnpB* gene and the deletion ends contained an *Eco*RI linker (GGAATTCC). The deletion end points of some clones were determined by dideoxy sequencing [17]. The final deletion plasmids were constructed by replacing the *rnpB*-containing *Hin*dIII-*Eco*RI DNA fragments of the primary deletion plasmids with the corresponding fragment of pLMdd5-9 [8] so that all the deletions should extend from position +138 into the 3′ flanking sequence.

2.3. Northern blot analysis

E. coli cells containing plasmids were grown to an A_{600} of ≈0.6 at 37°C. Total cellular RNA was prepared by directly extracting the culture with a phenol mixture at 65°C, which contained 0.1 volume of 10×RNA extraction buffer (RNA extraction buffer: 20 mM sodium acetate, 0.5% SDS, 1 mM EDTA, pH 5.2) and then with the same volume of phenol saturated with the RNA extraction buffer as described previously [15]. RNA fractionated in a 5% polyacrylamide gel containing 7 M urea was electrotransferred to a nylon membrane (Hybond N+; Amersham) with a Hoefer Semi-Phor Semi-dry transfer unit. Hybridization was performed according to the manufacturer's instructions. The probe used was the ^{32}P-labeled antisense RNA prepared by in vitro transcription with T7 RNA polymerase and the *Hin*dIII-treated pLMd23 as a template DNA [18]. Relative amounts of RNA species were estimated by analyzing the filter with a Molecular Dynamics PhosphorImager, Image Quant Version 3.3.

*Corresponding author. Fax: (82) 42-869-2810.
E-mail: ylee@sorak.kaist.ac.kr

PII S0014-5793(97)00377-3

3. Results and discussion

3.1. T2 and T3 are functional in vivo

The region for transcription termination in the 3′ flanking region of the *E. coli rnpB* gene, which encodes M1 RNA, is repeated three times [9]. A GalK fusion study [13] has shown that most transcription is terminated at the first terminator (T1) in the first repeat while complete termination is due to the presence of T2 and T3 in the following repeats. Therefore, T2 and T3 seemed to be required for efficient termination of the *rnpB* gene. Transcripts terminating at T1 have been observed in vivo [8] as well as in vitro [6,9] and the sequences identified as T2 or T3 are nearly the same as the corresponding region of T1 [9]. Therefore, T2 and T3 were expected to function with a similar termination efficiency. However, any transcripts terminating at T2 and T3 have not yet been identified, possibly due to exclusive termination at T1 [13] which results in little run-through transcription over T1. To determine the presence of transcripts terminating at T2 or T3, we constructed plasmids carrying various deletions at the 3′ flanking region of the *rnpB* gene by extending the deletion from position +138 in the M1 RNA structural gene into the 3′ flanking sequences, with the expectation that the deletion of a terminator could cause transcription termination at the next terminator (Fig. 1). The deletion plasmids generated discrete truncated M1 RNA transcripts in vivo, with the size dependent both on the extent of the deletion and on the termination site (Fig. 2A). Since most transcripts terminating at each terminator had the 3′ processing site, the corresponding processed RNAs lacking 36 nucleotides at the 3′ end [5,6] had been expected. The primary and processed transcripts generated by pLMdd5-9, pLMdd5-23 and pLMdd5-50 have previously been analyzed [8]. On the basis of their estimated sizes, it was determined at which terminator each transcript formed through transcription termination, or from which primary transcript each transcript formed through RNA processing (Fig. 2B). In some cases where expected sizes of RNA were similar to the size of M1 RNA, we could not assign the corresponding bands due to the endogenous M1 RNA band. As shown in Fig. 2, transcripts which run through a defective preceding terminator terminate at the next terminator. Therefore, these results firmly show that T2 and T3 are functional in vivo and should contribute to the complete termination of *rnpB* transcription in vivo.

Transcripts terminating at T2 or T3 were not observed when T1 is fully functional (plasmids pLMdd5-9, pLMdd5-

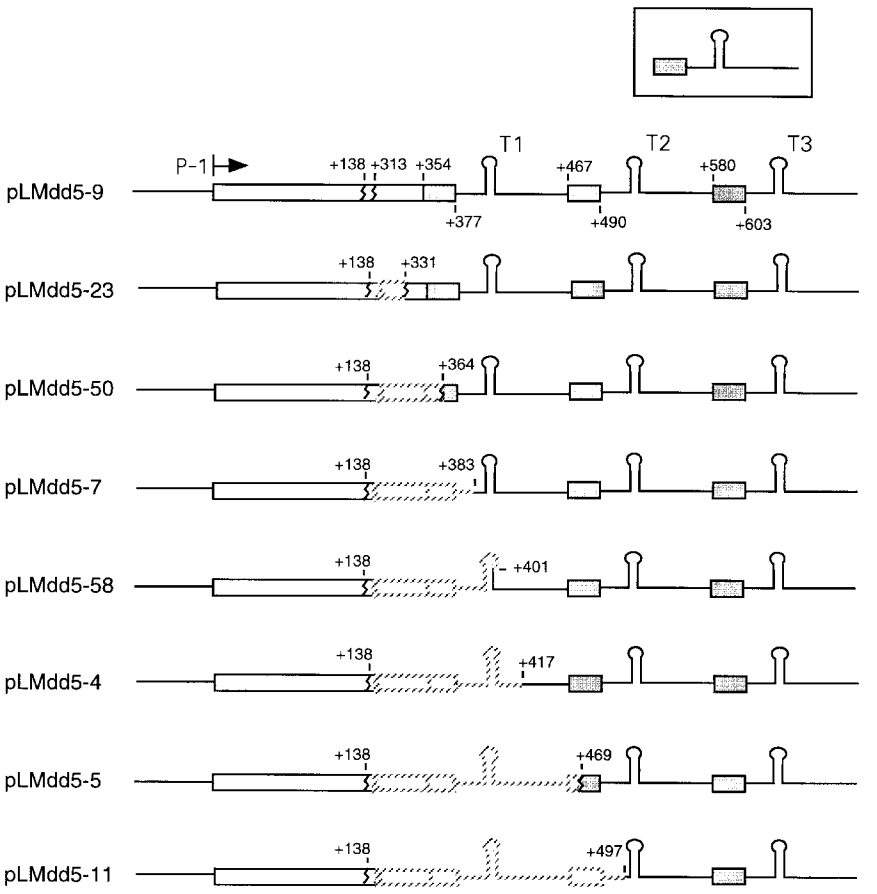

Fig. 1. Schematic representation of the deletion plasmids. The rectangle indicates the structural sequence of M1 RNA, and the numbers refer to positions in the *rnpB* gene. P-1 stands for the major *rnpB* promoter. T1, T2 and T3 refer to the sites of terminators. Deletion points are indicated by wavy lines and numbers. An *Eco*RI linker (GGAATTCC) was inserted in the deletion points of plasmids. pLMdd5-9 carries a deletion of nucleotides 139 to 312 in the M1 RNA sequence. Other plasmids are represented by showing further deletions as broken lines in the schematic representation. The deletion ends in pLMdd5-9, pLMdd5-23 and pLMdd5-50 were previously identified [8]. The repeated sequence of 113 bp starts with position +354 and includes the 3′ terminal 24 nucleotides of the M1 RNA coding sequence as indicated by the rectangle with grey-shading, and the region for transcription termination is indicated by the stem and loop structure. The repeated sequence is also schematically represented in the inset.

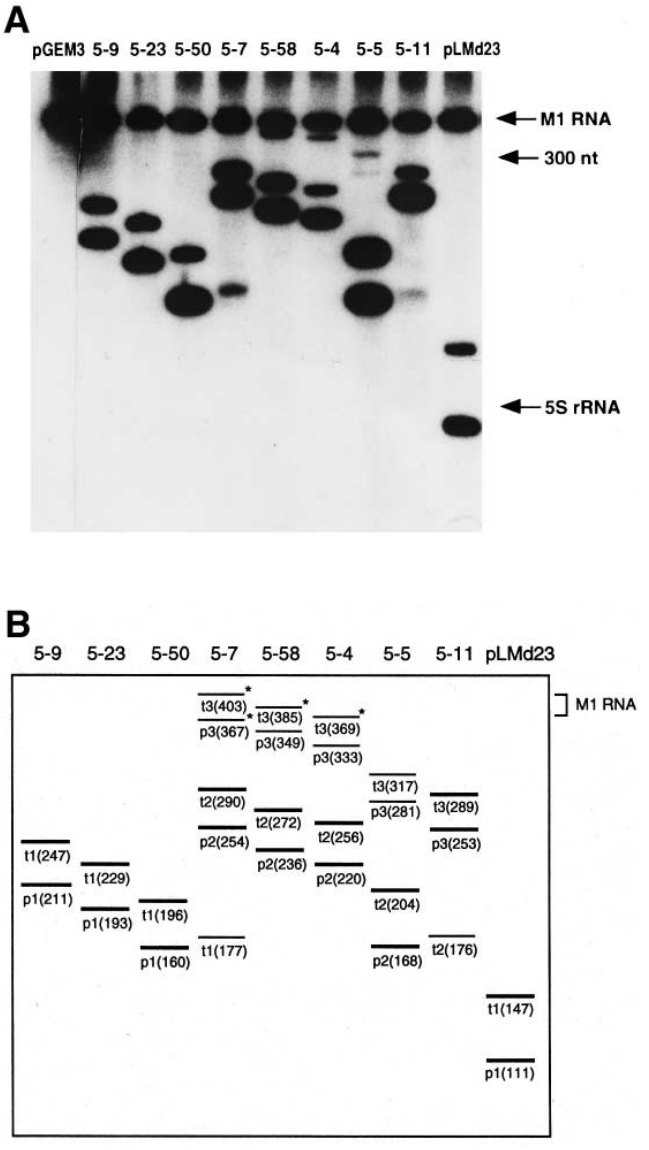

Fig. 2. Northern analysis of total cellular RNA. Total cellular RNAs prepared from 2×10^7 cells were analyzed in a 5% polyacrylamide gel containing 7 M urea. An RNA molecular weight marker from Boehringer Mannheim was used. For RNA preparation, JM109 cells containing the indicated plasmids were used. Plasmids of the pLMdd series were indicated only by numbers. A, Autoradiogram. B, Schematic representation of the Northern filter. Primary transcripts are indicated by t1, t2 and t3 according to where they terminated. Processed RNAs from primary transcripts terminated at T1, T2 and T3 are indicated by p1, p2 and p3, respectively. The expected sizes of RNAs are given in parentheses. Undetectable RNA bands due to overlapping with the M1 RNA band are shown with asterisks.

23 and pLMdd5-50 in Fig. 2), while transcripts terminating at T3 were observed even when T2 is intact (pLMdd5-58, pLMdd5-4 and pLMdd5-5 in Fig. 2). This suggests that T1 is more efficient than T2. This difference may be ascribed to the base differences between the repeated sequences comprising T1 and T2. The difference bases are located at positions +354, +379 and +450 of the T1 frame, although the two bases at positions +379 and +450 appear to be accountable for the difference of termination efficiency because this termination difference was not affected by the absence of the base at position +354 as shown by plasmids pLMdd5-50 and pLMdd5-5 in Fig. 2. The two bases are not within the sequence coding for the RNA-terminator hairpin and U-rich tail (Fig. 3), but they may be in contact with RNA polymerase as it transcribes the terminator region. It remains to be determined how this difference would affect the termination efficiency.

The repeated sequences in the 3′ flanking region of *rnpB* have a potential to code for small peptides [2]. Coding regions for small peptides in 3′ terminal repeated sequences have been found also in several tRNA genes [19–22], although their function is not yet known. For translation of the first overlapping coding regions, in the case of *rnpB*, transcription should run through both T1 and T2 (Fig. 3). T1 is not complete for termination [13] and efficiency of termination by T2 is less than that by T1 (Fig. 2). Therefore, read-through transcripts over both T1 and T2 can be produced even if the amount of the transcripts is very small and consequently their translation into the small peptides is possible.

3.2. Sequences required for transcription termination by the *rnpB* terminators

The boundary of the *rnpB* terminator region has not been defined, although it was expected that the sequence coding for the RNA hairpin followed by the U-rich sequence is responsible [3,5,9]. As expected, the deletion of the stem and loop structure abolished the function of the *rnpB* terminators. Besides this structural element, however, termination requires other sequences because most (more than 95%) of transcription was not terminated by this element when the preceding sequence between positions +364 and +382 or between positions +469 and +496 (equivalent to positions +356 and +383 of the T1 frame, respectively) was deleted (plasmids pLMdd5-

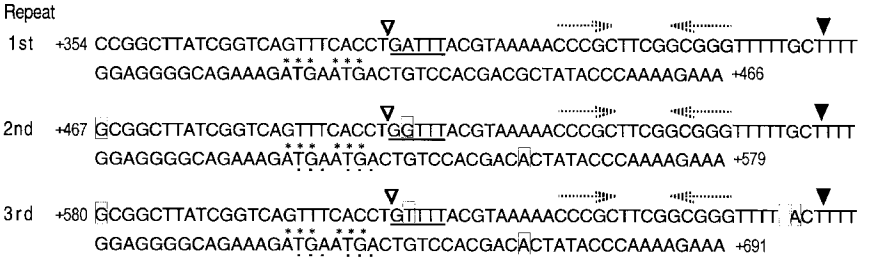

Fig. 3. Sequence comparison of the repeated sequences containing T1, T2 and T3. The different bases from those in the first repeat are boxed in the second and third repeats. The initiation and stop codons in the overlapping coding regions for the putative small peptides are asterisked above and dotted below the sequences, respectively. The sequence able to form a terminator hairpin is represented by the shaded arrows. Termination sites and processing sites are indicated by the closed and open arrowheads, respectively. The *rne*-dependent site is underlined.

7 and pLMdd5-11 in Fig. 2; see also Figs. 1 and 3). The requirement of preceding sequences for transcription termination has been reported in other intrinsic terminators [23–25]. Replacement of sequences upstream from an RNA-terminator hairpin of the *trp* attenuator with the comparable sequences of the phage T7 early terminator reduces the termination efficiency, but the exact nucleotides that are critical for the alteration of termination efficiency are not known [23]. In addition, the decrease of termination efficiency by the exchange of the preceding sequence in the *trp* attenuator is not dramatic, in contrast to the decrease by the deletion of the preceding sequence in the *rnpB* terminators. The early sequences transcribed from some promoters reduce the termination efficiency by acting like 'intrinsic antiterminators' both in vitro and in vivo [24,25]. In the case of *rnpB* transcription, on the other hand, it seems likely that there might be a 'termination enhancer' within the preceding sequence between positions +364 and +382, which is more delimited than the sequence between positions +469 and +496 as viewed in the T1 frame. A previous in vivo study [8] has shown that the mutations at the sequence coding for the *rne*-dependent site (positions +378 to +382) or at the sequence between positions +367 to +373 do not significantly affect the termination efficiency at T1, suggesting that the remaining sequence such as at positions +364 to +366 or at positions +374 to +377 may be responsible for the 'termination enhancer'. Alternatively, efficient termination may need some sequence contexts or RNA structures rather than specific nucleotides.

3.3. *Processing at the 3′ ends of transcripts terminating at T2 and T3*

M1 RNA processing at the 3′ end is dependent on the *rne*-dependent site immediately 3′ to the processing site [8]. This site is included in each repeated sequence containing the terminator region [9], although one nucleotide in the pentanucleotide known as the consensus sequence [26] for the *rne*-dependent site is different among the three repeated sequences (Fig. 3). Plasmids pLMdd5-7 and pLMdd5-11, having deletions of the pentanucleotide in the first and second repeats, respectively, did not generate any processed products from transcripts with the deletions (Fig. 2). On the other hand, plasmids generating transcripts with the pentanucleotide gave the corresponding processed RNA irrespective of where transcription terminates, suggesting that the *rne*-dependent sites in the second and third repeats are functional and the 3′ processing is tolerant of one nucleotide variation in the pentanucleotide. The fact that this variation does not reduce the number of U nucleotides is in good agreement with the mutation analysis of the site showing that U nucleotides are preferred for processing [8]. These results, as well as the previous study [8], accentuate the importance of the *rne*-dependent site on the processing reaction.

Acknowledgements: This work was supported by a grant from the Genetic Engineering Research Program of the Korean Ministry of Education (1995).

References

[1] S. Altman, Adv. Enzymol. Relat. Areas Mol. Biol. 62 (1989) 1–36.
[2] R.E. Reed, M.F. Baer, C. Guerrier-Takada, H. Donis-Keller, S. Altman, Cell 30 (1982) 627–636.
[3] H. Motamedi, Y. Lee, F.J. Schmidt, Proc. Natl. Acad. Sci. USA 81 (1984) 3959–3963.
[4] Y. Lee, R. Ramamoorth, C.-U. Park, F.J. Schmidt, J. Biol. Chem. 264 (1989) 5098–5103.
[5] H. Sakamoto, N. Kimura, Y. Shimura, Proc. Natl. Acad. Sci. USA 80 (1983) 6187–6191.
[6] M. Gurevitz, K.J. Swatantra, D. Apirion, Proc. Natl. Acad. Sci. USA 80 (1983) 4450–4454.
[7] U. Lundberg, S. Altman, RNA 1 (1995) 327–334.
[8] S. Kim, H. Kim, I. Park, Y. Lee, J. Biol. Chem. 271 (1996) 19330–19337.
[9] R.E. Reed, S. Altman, Proc. Natl. Acad. Sci. USA 80 (1983) 5359–5363.
[10] S. Adhya, M. Gottesman, Annu. Rev. Biochem. 47 (1978) 967–996.
[11] T. Platt, Annu. Rev. Biochem. 55 (1986) 339–372.
[12] Yager, T.D. and Von Hippel, P.H. (1987) in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology (Neidhardt, F.C., Ed.), pp. 1241–1275, American Society for Microbiology, Washington, DC.
[13] Y.M. Lee, Y. Lee, C.-U. Park, Korean Biochem. J. 22 (1989) 276–281.
[14] Y. Komine, H. Inokuchi, J. Bacteriol. 173 (1991) 1813–1816.
[15] S.J. Lee, Y.H. Jung, C.-U. Park, Y. Lee, Mol. Cells 1 (1991) 415–420.
[16] C. Yanisch-Perron, J. Vieira, J. Messing, Gene 33 (1985) 103–119.
[17] F. Sanger, S. Nicklen, A.R. Coulson, Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.
[18] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[19] J. Rossi, J. Egan, L. Hudson, A. Landy, Cell 26 (1981) 305–314.
[20] Y. Kuchino, F. Mori, S. Nishimura, Nucleic Acids Res. 13 (1985) 3213–3220.
[21] M. Yoshimura, M. Kimura, M. Ohno, H. Inokuchi, H. Ozeki, J. Mol. Biol. 177 (1984) 609–625.
[22] R.L. Russel, J.N. Abelson, A. Landy, M.L. Gefter, S. Brenner, J.D. Smith, J. Mol. Biol. 47 (1970) 1–13.
[23] R. Reynolds, M.J. Chamberlin, J. Mol. Biol. 224 (1992) 53–63.
[24] A.P.W. Telesnitsky, M.J. Chamberlin, J. Mol. Biol. 205 (1989) 315–330.
[25] J.A. Goliger, X. Yang, H.-C. Guo, J.W. Roberts, J. Mol. Biol. 205 (1989) 331–341.
[26] C.P. Ehretsmann, A.J. Carpousis, H.M. Krisch, Genes Dev. 6 (1992) 149–159.