Analysis of the complex transcription termination region of the Escherichia coli rrnB gene

András OROSZ, Imre BOROS and Pál VENETIANER

Institute of Biochemistry, Biological Research Center, Szeged, Hungary

(Received April 16/June 20, 1991) — EJB 91 0504

The complex terminator region of the Escherichia coli rrnB gene was analyzed by subcloning the terminators T1 and T2 and the inverted repeats IR1 and IR2 individually, or in various combinations, in a normal or inverted orientation into a terminator probe vector. The in vivo terminating efficiency was assayed by measuring the galactokinase activity encoded by the downstream galK gene. Termination efficiencies of all fragments were compared in two constructs, differing in the presence or absence of readthrough translation over the investigated terminator signal. The following main conclusions were drawn. (a) T1 and T2 are both efficient terminators in isolated forms. (b) IR1 and IR2 have some terminating effect (much lower than the proper terminators), especially in the inverted orientation. Their presence modifies the effect of the proper terminators in a quite unpredictable way, especially if these regions are translated. (c) The terminators are not symmetrical; in the inverted orientation T1 is practically inactive and T2 termination is reduced. (d) Translation radically decreases the efficiency of the terminators. (e) Several sequences in the rrnB gene upstream of the terminator region (one in the 16S RNA and one in the SS RNA coding region), are very efficient in vivo terminators in the inverted orientation.

The seven genes coding for ribosomal RNA in Escherichia coli have basically similar structures. The ends of six genes have been sequenced [1—6] and are all factor-independent (type I) [7] transcriptional terminators. A characteristic feature of these very large untranslated transcription units is that one or more antitermination regions ensure the complete uninterrupted transcription of the whole operon [8, 9]. Several lines of evidence indicates that antiterminated transcriptional complexes of rrn genes are able to transcribe some type I terminator signals [10]. The experimental evidence is conflicting on the problem as to how the rrn gene terminator regions efficiently stop antiterminated transcription [11].

The rrnB gene is probably unique among the rrn genes because its termination region is large and complex. It contains two regions (T1 and T2) with typical factor-independent terminator-like sequences, two additional inverted repeats (IR1 and IR2) and a pair of direct repeats [1]. The whole region is frequently employed as an efficient terminator of transcription in several cloning vectors [12]. In this paper we report the detailed molecular analysis of the terminating properties of this region. In particular, we attempted to answer the following questions. (a) What are the terminating efficiencies of the individual subregions of the complex rrnB terminator region? (b) How efficient are the terminators when present in an inverted orientation? (c) What is the effect of translation on the terminating efficiencies of the terminators?

Materials and Methods

Bacterial strains

All experiments were performed and plasmid constructions maintained in the Escherichia coli C600 galK strain (F−, galK, lacYI, thrI, leuB6, thiI, tonA21, supE44) [13].

Plasmids

Starting plasmids. The terminator selection vector, plasmid pKL300, is a derivative of the KO vector series [14]. It was constructed from pKO-1 by inserting a SalI linker into the EcoRI site and a PvuII polylinker fragment from M13mp9 into the Smal site (McKenney, personal communication). Plasmid p418-23 [15], carrying the complex terminator region of the rrnB gene was the source of the constructions described below. The polylinker region was taken from the miniplasmid πVX [16].

Construction of the terminator probe vector pKL300ΔH. As we intended to clone the terminator fragments into the HindIII site downstream of the lac promoter of pKL300, the second upstream HindIII site had to be eliminated. pKL300 was partially digested with HindIII, the 5’-overhanging ends were filled with Klenow polymerase and the plasmid was recircularized. Plasmids with only the downstream HindIII site intact were selected and their structure was verified by sequencing (Fig. 1).

Construction of the terminator probe vector pKLO2. In plasmid pKL300ΔH, the terminator to be tested was preceded by an intact Shine-Dalgarno sequence and a functional translation initiation codon, therefore transcriptional termination could possibly be influenced by translation through the terminators to be studied [17]. In order to eliminate this possi-
Fig. 1. Structure of the terminator-probe vector pKL300.4H. Construction is described in Materials and Methods. ori, origin of replication; lac pr, wild-type promoter; lac op, lac operator; ApR, ampicillin-resistance gene

Fig. 2. Comparison of the promoter regions of vectors pKL300AH and pKL02. Promoter — 35 regions are boxed, — 10 regions are underlined and overlined, translation initiation codons underlined. (○) Beginning of transcription; (---) ribosome-binding site; (------) operator. The polylinker sequence is represented by lower case letters

RESULTS

Galactokinase assay

E. coli C600 galK cells, harboring the terminator probe plasmids, were grown in 10 g/l bacto-tryptone, 5 g/l yeast extract and 5 g/l NaCl pH 7.5, supplemented with 100 μg/ml ampicillin and 1 mM isopropyl-β-D-thiogalactopyranoside to an A₆₅₀ value of 0.6. Toluenized extracts were prepared in duplicate aliquots and galactokinase activities were determined according to McKenney et al. [14]. The values shown in the figures represent the average of at least two parallel experiments. The scatter of these experiments was within 10%.

Recombinant DNA methods

Standard techniques were used according to Maniatis et al. [19]. Bal31 deletions, the fidelity of the filling-in reactions and the linker attachment points were verified by Maxam-Gilbert sequencing.

Enzymes and chemicals

Restriction endonucleases were prepared in this laboratory or purchased from New England Biolabs. Bal31 nuclease and oligonucleotide linkers were also from New England Biolabs. The DNA polymerase large fragment (Klenow enzyme) and T4-ligase were from Boehringer (Mannheim, FRG), [α-³²P]dextaxyribonucleotide triphosphates were from Amersham or Izinta (Budapest), U-[¹⁴C]galactose from Chemapol (Czechoslovakia). All other chemicals were reagent grade commercial products.

RESULTS

Elimination of the promoter following the rrnB terminators

Our experimental approach was to clone the entire rrnB terminator region, or various deletion derivatives representing individual components of this region between the lac promoter and the galK gene, into a terminator probe vector and estimate the in vivo terminator efficiency by measuring the activity of the galactokinase enzyme. As several authors demonstrated that promoters can influence the efficiency of terminators [20, 21], we used the same promoter in all experiments.
Fig. 3. **Origin of the terminator fragment used throughout this work.** Upper line shows the schematic structure of the **rrnB** gene (not drawn to scale). Lower line, terminator fragment, originating from p418-23 [15] and further modified as described in Materials and Methods. Black boxes represent regions coding for mature rRNA or tRNA, open boxes are the terminators, black dots the promoters (P₁, P₂ and P₃).

Fig. 4. **Sequence of the terminator fragment.** The numbering system adopted is taken from [1]. The restriction sites used in the subcloning experiments, the inverted repeats forming the T₁ and T₂ terminators and the two smaller repeats (IR₁, IR₂) in between are underlined.

**Table 1. Properties of deletion plasmids of the **pKLT** series**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Start point</th>
<th>End point</th>
<th>Fragment length</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>pKLT1</td>
<td>1598</td>
<td>6686</td>
<td>392 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLT2</td>
<td>1598</td>
<td>6686</td>
<td>392 reverse</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLT3</td>
<td>1598</td>
<td>6775</td>
<td>481 normal</td>
<td>normal</td>
</tr>
<tr>
<td>pKLT9</td>
<td>1598</td>
<td>6588</td>
<td>294 normal</td>
<td>normal</td>
</tr>
<tr>
<td>pKLT10</td>
<td>1598</td>
<td>6588</td>
<td>294 reverse</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLT12</td>
<td>1598</td>
<td>6741</td>
<td>447 normal</td>
<td>normal</td>
</tr>
<tr>
<td>pKLT19</td>
<td>1598</td>
<td>6653</td>
<td>359 normal</td>
<td>normal</td>
</tr>
<tr>
<td>pKLT21</td>
<td>1598</td>
<td>6653</td>
<td>359 reverse</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLT23</td>
<td>1598</td>
<td>6712</td>
<td>418 normal</td>
<td>normal</td>
</tr>
<tr>
<td>pKLT24</td>
<td>1598</td>
<td>6712</td>
<td>418 reverse</td>
<td>reverse</td>
</tr>
</tbody>
</table>

**Brosius et al.** [1] found that the **rrnB** gene is followed by an open reading frame (termed ORFII) that is preceded by a ribosomal binding site and a promoter-like sequence downstream of the **rrnB** terminator. Thus, this region was eliminated by Bal31 deletion (to nucleotide 6837) because the presence of this promoter would have interfered with the galactokinase assay. The sequence of this truncated fragment is shown in Fig. 4. This fragment represents the insert in plasmids pKLT1 and pKLT6.

**Attempts to delineate the minimal terminating structure in T₁**

The largest structure similar to the rho-independent terminators within the **rrnB** terminator region is the T₁ terminator. Using deletion mutants, we wanted to answer two questions. (a) Is T₁ itself sufficient for termination? (b) What is the minimum sequence required for this termination? Several deletion mutants were generated starting from the 3'-end of the fragment cloned in pKLT1 (Table 1 and Fig. 5), and the truncated fragments were recloned in the terminator-probe vector pKL300.4H in either orientation, for measurement of the efficiency of termination. Fig. 6 shows the results. It can be seen that, in clones pKLT4 (deletion endpoint at nucleotide 6686) and pKLT23 (deletion endpoint at nucleotide 6712) the termination is nearly as efficient as in the original pKLT1, indicating that the T₁ region alone is sufficient for almost complete termination. In contrast, pKLT19 (deletion endpoint at nucleotide 6653) has a much lower efficiency of termination, suggesting that loss of the T-rich region following the G+C-
rich hairpin in the typical terminator structure, leads to a significant (but not complete) loss of function. The fact that this decreased, but measurable, terminating activity is due to the presence of the truncated T1 region, is supported by the results of the assay with pKLT9 (deletion endpoint at nucleotide 6588). In this mutant, where T1 was completely eliminated, the galactokinase activity was not decreased, but rather enhanced relative to the pKL300dH control. The reason of this increased activity is unexplained at present.

Are the terminators functionally symmetrical?

As there were some suggestions in the literature that rho-independent terminators might function symmetrically, i.e. they are able to terminate transcription in either orientation [12, 22], we assayed the efficiency of our deletion mutants to function in the inverted orientation. As Fig. 6 shows, surprisingly, all the analyzed fragments exhibited efficient termination in this reverse orientation. However, as this efficient termination can be observed in the absence of any of the putative terminators (pKLT10), the ability to terminate in the reverse orientation must be attributed to sequences located upstream of the terminators, either within the 16S coding region (positions 1598 – 1787) or within the region coding for 5S rRNA (6496 – 6615).

This question was resolved by cutting this region at nucleotide 6498 with BstNI and subcloning. The results of Fig. 7 show that both subregions, i.e. the 16S sequences (nucleotides 1598 – 1787 plus 6483 – 6497) and the 5S rRNA sequences (nucleotides 6496 – 6615) are efficient terminators in the reverse orientation.

Efficiency of the individual components of the complex rRNA terminator region

In the next series of experiments, we wanted to determine the in vivo terminating efficiencies of the individual components of the complex rRNA terminator region and their various combinations in both possible orientations. As the previous experiments revealed, the presence of cryptic reverse terminators in both the 16S and 5S rRNAs regions present in the insert of pKLT1, made it necessary to delete the 5'-half of the sequence shown in Fig. 4. Therefore, the insert of pKLT1 was partially digested with BstNI, filled in with Klenow enzyme, ligated with HindIII linkers and inserted into pKL300dH as described in Materials and Methods. Plasmids of the expected size were recovered (pKL1 and pKL2, containing 228-bp inserts), but their sequence analysis revealed that these inserts were not the original terminator regions. It appears that, in these clones, the 6498 – 6776 BstNI fragment was inverted and attached to the normally oriented T2-containing fragment 6776 – 6837. Thus, in pKL1, the sequence of elements was IR2 (inverted), IR1 (inverted), T1 (inverted) and T2; in pKL2 it was T2 (inverted), T1, IR1 and IR2. We were unable to find plasmids containing this region in the original structure. This failure is in agreement with the experience of Albrechtsen et al. [10]. They were also unable to clone the terminator region in its entirety. The structures of the other subclones are shown in Table 2 and Fig. 5 and the results of the galactokinase assays are summarized in Fig. 7.

Effect of translation on the efficiency of the terminators

In the pKLA plasmid series, the terminators were cloned in vector pKL300dH just after a functional translation initiation
Table 2. Properties of the pKLA and pKLQ plasmids

In pKLA1 and pKLQ1 the 6610–6678 BstNI fragment was in an inverted position and this was followed by nucleotides 6778–6838, thus the order of the functional elements was IR₂ (reverse)-IR₁ (reverse)-T₁ (reverse)-T (normal). In pKLA2 and pKLQ2 it was T₂ (reverse)-T (normal)-IR (normal)-IR₂ (normal)

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Start point</th>
<th>End point</th>
<th>Fragment length</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>pKLA1, pKLQ1</td>
<td>6610 (BstNI)</td>
<td>6838 (pKLT1)</td>
<td>228 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA2, pKLQ2</td>
<td>6610 (BstNI)</td>
<td>6838 (pKLT1)</td>
<td>228 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA3, pKLQ3</td>
<td>6610 (BstNI)</td>
<td>6777 (pKLT3)</td>
<td>167 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA4, pKLQ4</td>
<td>6610 (BstNI)</td>
<td>6777 (pKLT3)</td>
<td>167 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA5, pKLQ5</td>
<td>6610 (BstNI)</td>
<td>6741 (pKLT12)</td>
<td>131 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA6, pKLQ6</td>
<td>6610 (BstNI)</td>
<td>6741 (pKLT12)</td>
<td>131 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA7, pKLQ7</td>
<td>6610 (BstNI)</td>
<td>6686 (pKLT4)</td>
<td>76 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA8, pKLQ8</td>
<td>6610 (BstNI)</td>
<td>6686 (pKLT4)</td>
<td>76 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA9, pKLQ9</td>
<td>6689 (DdeI)</td>
<td>6838 (pKLT1)</td>
<td>149 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA10, pKLQ10</td>
<td>6689 (DdeI)</td>
<td>6838 (pKLT1)</td>
<td>149 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA11, pKLQ11</td>
<td>6742 (HpaII)</td>
<td>6838 (pKLT1)</td>
<td>96 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA12, pKLQ12</td>
<td>6742 (HpaII)</td>
<td>6838 (pKLT1)</td>
<td>96 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA13, pKLQ13</td>
<td>6778 (BstNI)</td>
<td>6838 (pKLT1)</td>
<td>60 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA14, pKLQ14</td>
<td>6778 (BstNI)</td>
<td>6838 (pKLT1)</td>
<td>60 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA15, pKLQ15</td>
<td>6689 (DdeI)</td>
<td>6777 (pKLT3)</td>
<td>88 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA16, pKLQ16</td>
<td>6689 (DdeI)</td>
<td>6777 (pKLT3)</td>
<td>88 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA17, pKLQ17</td>
<td>6689 (DdeI)</td>
<td>6741 (pKLT12)</td>
<td>52 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA18, pKLQ18</td>
<td>6689 (DdeI)</td>
<td>6741 (pKLT12)</td>
<td>52 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA19, pKLQ19</td>
<td>6739 (HaeIII)</td>
<td>6777 (pKLT3)</td>
<td>38 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA20, pKLQ20</td>
<td>6739 (HaeIII)</td>
<td>6777 (pKLT3)</td>
<td>38 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA21, pKLQ21</td>
<td>1598 (pKLT1)</td>
<td>6500 (BstNI)</td>
<td>207 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA22, pKLQ22</td>
<td>1598 (pKLT1)</td>
<td>6500 (BstNI)</td>
<td>207 bp</td>
<td>reverse</td>
</tr>
<tr>
<td>pKLA23, pKLQ23</td>
<td>6500 (BstNI)</td>
<td>6610 (BstNI)</td>
<td>110 bp</td>
<td>normal</td>
</tr>
<tr>
<td>pKLA24, pKLQ24</td>
<td>6500 (BstNI)</td>
<td>6610 (BstNI)</td>
<td>110 bp</td>
<td>reverse</td>
</tr>
</tbody>
</table>

DISCUSSION

Functional terminators outside the terminator region

The DNA fragment originally used in these experiments contained sequences from the 16S rRNA coding region and the total 5S rRNA coding sequence. It is obvious from the results shown in Figs 6 and 7 that, in the in vivo galactokinase assay employed throughout this work, both of these regions proved to be efficient terminators in the reverse orientation. As neither of these sequences exhibit any similarity to rho-independent type I terminators it seems likely that these are rho-dependent terminators. The presence of such a rho-dependent terminator within the 16S region was also reported by Li et al. [23]. Whereas these authors localised this terminator between nucleotides 1598–2166, the present work determines its position more precisely (nucleotides 1598–1787). The position of the inverted rho-dependent terminator within the 5S gene was localised by S1 mapping to nucleotides 6572–6580, just after one G+C-rich stem and loop structure (A. Orosz, unpublished results).

As both of these terminators work efficiently only in the reverse orientation, it seems to be unlikely that they have any
functional significance. Thus, we did not attempt to analyze them further.

Efficiency of individual terminator elements

The T<sub>2</sub> terminator appears to be fully efficient in the isolated form, and normal orientation under non-translated conditions (pKLQ13, 100%). The larger T<sub>1</sub> terminator in itself is somewhat less efficient (pKLQ7, 87%). The inverted repeats, either individually or together, display only slight terminating efficiencies (15—33%). The higher efficiency of T<sub>1</sub> is somewhat surprising because T<sub>1</sub> is larger and it is homologous to the single terminator of rrd<sub>D</sub> [3].

Are the terminators functionally symmetrical?

The results clearly indicate that, in the isolated form, the analyzed real terminators are not functionally symmetrical. Reversing the orientation decreased the efficiency of T<sub>2</sub> over 100—66% (pKLQ14) and that of T<sub>1</sub> over 87—11% (pKLQ8). In view of the non-symmetrical sequence of these elements, this is not surprising. It is interesting to note, however, that the reversal of the inverted repeats markedly increases their termination efficiency (54—81%).
5S and 16S rRNA regions. Nevertheless, according to S<sub>1</sub> mapping experiments (data not shown), some termination occurs in T<sub>2</sub> at the same positions as that in normal orientation [25], i.e. at several points in the region of nucleotides 6629 – 6610.

Why two terminators?

In view of the fact that (a) the *rrnB* T<sub>1</sub> and T<sub>2</sub> terminators appear to be fully effective in the isolated form, (b) the other *rrn* genes have only single terminators, (c) the *rrnB* inverted repeats are not efficient terminators, it seems likely that the complexity of the *rrnB* terminator region has no particular function in termination. Considering the near identity of the T<sub>2</sub> sequence with the *rrnG* terminator and the very strong similarity of T<sub>1</sub> and the *rrnD* terminator (Fig. 8), it might be assumed that *rrnG* and an ancestral *rrnB* (with only the T<sub>2</sub> terminator) were the results of gene duplication. The insertion of T<sub>1</sub> might have been the result of a later transposition, as suggested by Brosius et al. [1].

Specific features of rRNA terminators

Although the experimental evidence is conflicting about the specific ability of *rrn* terminators to terminate antiterminated transcripts [11], it seems to be very likely that this distinctive property is necessary. Comparing the sequences of all seven known *rrn* terminators (Fig. 8), the following common features (each shared by some, but not all, known non-ribosomal rho-independent terminators) can be found. (a) All are preceded by sequences coding for an untranslated small RNA, having strong and compact higher structure (in four cases 5S rRNA, in two cases tRNA). (b) The characteristic stem/loop structure of these terminators is unusually complex. The stem always contains a lower, 4 – 11-bp A + T-rich region (with predominantly A on the ascending and T on the descending strand), and an upper, 7 – 9-bp G + C-rich region. (c) The loop of these regions is four nucleotides long and is purine-rich, where the third position is always A. (d) A TYTG tetranucleotide motif follows the A + T-rich part of the stem. Further experimental work is required to clarify whether the simultaneous occurrence of these features endows the *rrn* terminators with distinctive functional properties.

Thanks are due to Drs C. Kari and A. Simoncsits for providing the original terminator probe plasmid and its host. The skilful technical assistance of Mrs I. Anton is gratefully acknowledged.

REFERENCES