Antitermination of transcription from an *Escherichia coli* ribosomal RNA promoter

(operon fusion/*rrnC* promoter-leader region/*lacZYA* genes/insertional polarity)

WILLIAM E. HOLBEN AND EDWARD A. MORGAN*

Department of Biology, State University of New York at Buffalo, Buffalo, NY 14260

Communicated by H. E. Umbarger, July 9, 1984

ABSTRACT The *Escherichia coli* lac and ara promoters and *rrnC* ribosomal RNA promoter-leader region were fused to *lacZYA*. Transcription termination signals were introduced into the *lac* genes of these fusions by *Tn*9 and ISI insertions. Measurement of *lac* enzymes from upstream and downstream of the insertions showed that termination signals resulting from these insertions are very efficient when transcription begins at *lac* or *ara* promoters but are very inefficient when transcription begins at the *rrnC* promoter-leader region. The *rrnC* promoter-leader region must, therefore, modify RNA polymerase to enable it to read through transcription termination signals.

Transcription and translation are functionally coupled in the *lac* operon and in most bacterial operons that code for translated RNAs. The essence of coupling is that ribosomes translating mRNA immediately behind transcribing RNA polymerase molecules enable RNA polymerase to read through sequences that would terminate transcription in the absence of ribosomes. Coupling therefore confers an advantage to the cell because it allows coding regions to evolve without the need to eliminate certain types of transcription termination signals (and perhaps also signals that cause termination-related events such as pausing). Coupling also may have evolved to allow coupling-dependent processes such as attenuation.

The disruption of coupling by premature transcription termination unmasks transcription termination sequences frequently present in protein-coding regions and results in premature transcription termination (1–10). The uncoupling of transcription and translation is therefore responsible for some forms of polarity caused by premature nonsense codons in promoter-proximal genes of multicistronic protein-coding operons (1, 2, 4–10). The polar effects resulting from disruption of coupling by premature nonsense codons are dependent on the transcription termination factor *p* and are relieved by mutations in the gene for *p* (1, 6–9).

Nontranslated operons that are transcribed with negligible natural polarity, such as ribosomal RNA operons (11–15), must not require coupling. Coupling would not be required if RNA polymerase initiating at these promoters is modified to prevent termination. Detailed models for the reduction of termination by modification of RNA polymerase are provided by regulatory events in the bacteriophage λ (1, 2, 10, 16–19). Coupling would also not be required if the transcript has a primary sequence devoid of termination signals, a secondary or tertiary structure unsuitable for the function of termination signals or termination factors, or if the binding of ribosomal proteins to the nascent transcript prevents premature termination.

Previous experiments in which transcription termination signals were introduced into ribosomal RNA operons by insertions of *Tn*10, *Tn*9, or ISI (insertion sequence 1) showed that transcription termination is suppressed in ribosomal RNA operons, suggesting that modification of the RNA polymerase at or near the ribosomal RNA promoters is likely (13–15). However, definitive proof that polymerase modification is involved requires elimination of any possible contribution by the sequences surrounding the termination signals.

In this paper we report the use of operon fusions to show that the ability of RNA polymerase to read through termination signals is a special property conferred by the ribosomal RNA promoter-leader region and is independent of the sequences immediately surrounding the termination signals. These observations suggest that RNA polymerase is modified at the ribosomal RNA promoter-leader region to enable read-through of transcription termination signals.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. EM343 (ΔlaczZYA melB ara *λkimM* Str') is an ara derivative of RE16 (20) obtained by ethylmethanesulfonate mutagenesis and ampicillin enrichment. EM350 is a recA derivative of EM343 obtained by trimethoprim selection of a thyA derivative followed by introduction of thyA*′* recA after conjugation with KL16-99. EM388 [Lac*ZU*A-*thyA*36* polA1* deoC2* 0* (rrnD-rrnE*1) *λ* F'] was obtained by transposition of *Tn*9 into lac*Y* of P3478 (21) and subsequent ampicillin enrichment of chloramphenicol-sensitive bacteria followed by genetic and enzymatic screening for a chloramphenicol-sensitive LaciZYA*−* phenotype.

λ*Tn*9::4.33 is a defective specialized transducing phage that has a *Tn*9 insertion in lambda DNA near or in *int* (14). pMC306 (22) is described in Fig. 1. pBH16 and pBH30 are described in Fig. 3 and contain a replication origin of *F* on an *Hpa* I fragment derived from pDF42 (23) and a replication origin derived from pUC9 (24). *Tn*9 consists of a chloramphenicol acetyltransferase gene between direct repeats of *IS* (25).

Insertions in *lac*. EM350 containing pMC306 was infected with a helper-depleted preparation of λ*Tn*9::4.33. Lac*−* survivors were detected on lactose tetrazolium agar containing 0.1% arabinose, ampicillin, and chloramphenicol. Transposition is selected for by this procedure because integration of *λ* by site-specific or general recombination is prevented by the *recA* mutation and by the *λ* repressor present in the cells prior to infection. The location of each *Tn*9 insertion was determined by use of melibiose and lactose minimal medium and by melibiose, lactose, and 5-bromo-4-chloro-3-indolyl β-D-thiogalactoside (X-Gal) indicator agar. Purified plasmid DNA was also analyzed by agarose gel electrophoresis after

*Abbreviation: IS, insertion sequence.*

*To whom reprint requests should be addressed.*
RESULTS

Experiments designed to prove that RNA polymerase transcribing from the ribosomal RNA promoter-leader regions is resistant to termination require that a termination signal in a single sequence context be shown to cause termination with different efficiencies when rrrn and other promoters are compared. It is important to measure the efficiency of termination, not just the level of transcription downstream of a termination signal, because the level of transcription downstream of a leaky termination signal is also dependent on promoter strength. The operon fusions we have constructed enable termination efficiency to be measured independently of promoter strength.

Insertions in ara-lac. Twenty-nine derivatives of pMC306 were isolated with insertions of Tn9 in the lacZ and lacY genes of the ara-lac fusion operon (Figs. 1 and 2). An IS1 insertion was then generated at the site of each Tn9 insertion by homologous recombination between the directly repeated IS1 elements of Tn9. The location and orientation of each insertion was determined by genetic and restriction nuclease analysis (see Materials and Methods and Fig. 2). The regional preference for insertions in or near lacY has been observed previously (28).

All Insertions in ara-lac Cause Efficient Termination. Enzymatic assays of lacZ and lacA gene products reveal that all insertions in lacZ and lacY reduce lacA expression at least 95% (Table 1, some data not shown). Since lacZ expression is not substantially affected by insertions in lacY, transcription rates and plasmid copy numbers are not significantly affected by the insertions. It is therefore likely that insertions in lacZ also do not substantially affect transcription initiation rates or plasmid copy numbers. It can be concluded that Tn9 and IS1 insertions in many locations and both orientations are at least 95% polar on lacA. Since IS1 insertions cause polarity that is relieved by mutations in p (29–31), it follows that Tn9 and IS1 insertions contain or unmask strong p-dependent transcription termination signals. For the purposes of this paper, it is unnecessary to know whether the termination signals are within the insertions or unmasked by the insertions. However, Tn9 and IS1 insertions are very polar even when located very near the 3' end of lacZ or lacY (Table 1). The strong and universal polarity caused by these insertions is unlike the uneven gradient of generally weaker polarity that is observed when p-dependent termination signals in lac are unmasked by premature nonsense codons (32). It is therefore very likely that Tn9 and IS1 contain strong p-dependent termination signals that function in both orientations.

Insertions in rrrrn-lac Do Not Cause Efficient Termination. HindIII/BamHI fragments carrying lac genes and lac insertion derivatives were cloned from pMC306 into the HindIII/
Table 1. Synthesis of β-galactosidase (lac\(Z\)) and thiogalactosidase transacetylase (lac\(A\)) from ara-lac, rrn-lac, and lac-lac fusion operons

<table>
<thead>
<tr>
<th>Insert</th>
<th>Z</th>
<th>A</th>
<th>A/Z</th>
<th>Z</th>
<th>A</th>
<th>A/Z</th>
<th>Z</th>
<th>A</th>
<th>A/Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>14</td>
<td>6.8</td>
<td>0.48</td>
<td>0.10</td>
<td>2.0</td>
<td>0.20</td>
<td>0.16</td>
<td>0.33</td>
<td>2.1</td>
</tr>
<tr>
<td>IS1</td>
<td>14L</td>
<td>12</td>
<td>0.0</td>
<td>0.0</td>
<td>0.09</td>
<td>1.6</td>
<td>17</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>15L</td>
<td>12</td>
<td>0.0</td>
<td>0.0</td>
<td>0.17</td>
<td>1.0</td>
<td>6.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16L</td>
<td>8.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.17</td>
<td>1.2</td>
<td>7.3</td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>17R</td>
<td>9.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.16</td>
<td>0.57</td>
<td>3.6</td>
<td>0.16</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>19R</td>
<td>16</td>
<td>0.0</td>
<td>0.0</td>
<td>0.11</td>
<td>0.63</td>
<td>5.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20L</td>
<td>17</td>
<td>0.0</td>
<td>0.0</td>
<td>0.12</td>
<td>1.4</td>
<td>11</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>22R</td>
<td>10</td>
<td>0.0</td>
<td>0.0</td>
<td>0.10</td>
<td>0.71</td>
<td>7.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23R</td>
<td>19</td>
<td>0.0</td>
<td>0.0</td>
<td>0.10</td>
<td>0.64</td>
<td>6.4</td>
<td>0.18</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>24L</td>
<td>13</td>
<td>0.0</td>
<td>0.0</td>
<td>0.10</td>
<td>1.1</td>
<td>31</td>
<td>0.11</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>26L</td>
<td>12</td>
<td>0.0</td>
<td>0.0</td>
<td>0.10</td>
<td>1.3</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27R</td>
<td>13</td>
<td>0.0</td>
<td>0.0</td>
<td>0.06</td>
<td>0.69</td>
<td>12</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

One unit of β-galactosidase is defined as 1 µmol of o-nitrophenyl β-D-thiogalactoside hydrolyzed per min per mg of protein. One unit of thiogalactosidase as defined is 1 µmol of thionitrobenzoic acid produced per mg of protein. A blank in the table indicates that the corresponding plasmid construct was not made. All insertions in lac\(Z\) reduced lac\(A\) activity by >95% in ara-lac fusions (they were not examined by lac-lac fusions). In rrn-lac fusions, lac\(Z\) insertions 3R and 4L reduced lac\(A\) expression to 43% and 17% of that of an rrn-lac fusion without insertions. The only significant variation in the assays arises from culture variability and is an average of 10% of the mean value. The values given are the average of assays performed on at least two independent cultures of cells containing each plasmid construct.

BamHI sites of pBH16 (Fig. 3) and transformed into the polA strain EM388, in which pBH16 replicates from the origin of F. These constructions fuse the rrnC promoter-leader region to the trpB portion of the ara-Mu-trp-lac fusion operon from pMC306. Since trpB is translationally coupled to trpA and a trpA-lacZ fusion protein is produced, this fusion junction results in a low level of synthesis of the trpA-lacZ fusion protein (ref. 33 and Table 1). The low level of translation of the trpB-trpA-lacZ portion of the transcript results in only a 2-fold transcriptional polarity on lacA when transcription begins from the ara promoter (33).

In the polA bacterial strain EM388, the rrnC-lac fusion plasmids replicate only from the origin of F and therefore probably exist in a single copy per cell. The rrnC-lac fusion plasmids do not severely affect the growth of EM388 and are very stable. In polA + strains, the same fusion plasmids exist in multicopy form, are unstable, and severely affect cell growth (unpublished observations). For this reason EM388 was used to assay rrnC-lac fusions, whereas ara-lac fusions on pMC306 were assayed using strain EM350. In EM350, pMC306 is a multicopy plasmid that replicates from the CoE1 origin.

Enzymatic assays using extracts of EM388 containing rrnC-lac fusion plasmids reveal that insertions in lac\(Y\) do not significantly affect lac\(Z\) expression (Table 1). Therefore, insertions in lac\(Z\) or lac\(Y\) probably do not affect plasmid copy number or transcription initiation. As determined by lac\(A\) measurement and lac\(A\)/lac\(Z\) ratios (Table 1), all insertions tested are substantially less polar when synthesis is directed by the rrnC promoter than when directed by the ara promoter.

When lac\(A\) expression downstream of insertions is compared with expression from uninterrupted rrnC-lac fusion operons, it is apparent that IS1 insertions very near the carboxyl terminus of lac\(Y\) or between lac\(Z\) and lac\(Y\) cause little termination in rrnC-lac fusions (Table 1). From this observation it can be concluded that RNA polymerase initiating at the rrnC promoter-leader region efficiently reads through termination signals that cause >95% termination when transcription is initiated at the ara promoter. A gradient of polarity on lac\(A\) expression is observed when insertions in a single orientation in the lac\(Y\) gene are examined (Table 1). This gradient probably occurs because insertions nearer the 5' end of lac\(Y\) prevent coupling over larger regions and probably unmask more transcription termination signals in lac\(Y\) coding regions between the insertion and lac\(A\). A gradient of polarity of insertions would then be expected if RNA polymerase initiating at ribosomal promoters is not 100% efficient in reading through termination sequences. Partial termination would also be consistent with the greater polarity of Ts9 insertions (which have two IS1 sequences and may have other termination signals between the IS1 elements) than polarity resulting from a single IS1 sequence in the same orientation at the same site (Table 1).

Insertions in lac-lac Cause Efficient Termination. Differences in polarity observed using the ara and rrnC promoter regions could be due to disruption of translational coupling resulting from use of the trpB HindIII site to construct rrnC-lac fusions or use of the strain EM388 to assay rrnC-lac fusions. To rule out these possibilities the HindIII/BamHI lac

![FIG. 3. The structures of pBH16 and pBH30. Both plasmids contain origins of replication derived from F and pUC9. Insertions in the polylinker of pBH30 are transcribed from the lac promoter. Insertions in the polylinker of pBH16 are transcribed from the rrnC promoter. lac sequences in pBH16 are represented as thick lines, and flanking non-rrn chromosomal DNA as thin lines. The tandem rrnC promoters are followed by the entire leader region, 16S rRNA gene sequences from positions 1–16 and 613–650, a polylinker with HindIII and EcoRI ends, the rrnC trnA + trp and trnA + trp genes, and the rrnC termination sequence. A SmaI site was generated within the 16S rRNA sequences between positions 16 and 63 by ligating blunt ends resulting from Bcl I and Xma I termini that had been filled in with the Klenow fragment of DNA polymerase 1. pBH30 contains E. coli chromosomal DNA extending approximately 350 bases upstream of the rrnC PI promoter and 150 bases downstream of the rrnC termination signal. Upstream Hpa I and downstream Pvu II sites flanking rrnC in pCL22–26 (13) are the boundaries of rrnC and flanking E. coli chromosomal DNA of pBH16. Nucleotide sequence determination has verified that the rrnC promoter-leader region on pBH16 is intact. The origin of F is an Hpa I fragment from pDF42 (22) inserted in a Pvu II site located downstream of lac in pUC9 (24). The lac promoter of pUC9 was removed during construction of pBH16. kb, Kilobases.](image-url)
fragments of pMC306 and several insertion derivatives were cloned under control of the pBH30 lac promoter (Fig. 3). These constructions result in lacZ'-trpB-trpA'-lacZYA fusion operons with a frameshift in the lacZ'-trpB gene that introduces many translation termination codons beginning in trpB very near the HindIII site (24, 41). Therefore, trpB-trpA translational coupling is prevented in these lac-lac constructions as it is in the rrn-lac constructions. Synthesis of lac enzymes from the lac-lac fusion operons was assayed using strain EM38, where the lac-lac fusion plasmids replicate from the origin of F. Insertions in lac-lac fusion operons are extremely polar (Table 1).

Relief of Natural Termination in lac. When rrnC promoters on pBH16 are fused to lac, the ratio of thiogalactoside acetylase to β-galactosidase is 9.6 times greater than that from similarly constructed lac-lac fusions made with pBH30 (Table 1). This observation suggests that substantial termination may occur between lacZ and lacA in lac-lac fusions but not in rrnC-lac fusions.

DISCUSSION

RNA polymerase initiating at the rrnC promoter-leader region can efficiently read through β-dependent transcription termination signals in lac. It is therefore likely that RNA polymerase is modified at the rrnC promoter-leader region to allow read-through of β-dependent transcription termination signals. Modification of polymerase presumably compensates for the lack of transcriptional-translational coupling in rrn operons and allows transcription of rrn operons with little or no polarity. The modification of polymerase may also play an as yet unknown role in regulation of rrn operon transcription.

RNA polymerase could be modified during transcription initiation at the rrn promoters or while polymerase traverses the rrn leader region. The λ phage antitermination proteins N and Q (16, 19) provide models that suggest how RNA polymerase modification at the rrn promoter-leader region might occur. Modification of polymerase by N and Q proteins requires E. coli nusA protein (16-18). It is therefore significant that a nusA utilization sequence is present in the leader region of all seven rrn operons of E. coli, including the rrnC operon (ref. 34, unpublished data). There are additional nusA utilization sequences just prior to the 23S rRNA genes. These observations suggest that the nusA protein might be involved in antitermination in rrn operons either alone or in cooperation with unknown E. coli proteins.

In vitro, the nusA protein causes RNA polymerase to pause in the rrn leader for the unphysiologically long period of 20 sec (34). Perhaps polymerase pauses at this site in vivo until modification by a host antitermination protein allows transcription to continue. This mechanism would ensure that polymerase molecules will not prematurely terminate after entering rRNA genes. Alternatively, the antitermination mechanism may modify polymerase before the pause site to prevent significant pausing from occurring in vivo. It is also possible that the antitermination mechanism and pause site cooperate in a regulatory response that determines transcriptional rates of rRNA genes.

rrn operon termination signals might need special structures to terminate rrn transcripts. Interestingly, two of the four sequenced rrn termination signals have tandem stem-loop structures (35, 36). However, two operons, including the rrnC operon used here, have a single stem-loop structure (37, 38). Therefore, the tandem arrangement is not critical for termination of rrn transcripts. In four rrn termination regions for which the nucleotide sequence is known, the first (or only) rrn termination signals encountered by polymerase have conserved elements of structure and sequence not found in other bacterial termination signals. It has been suggested that these conserved elements are important to rrn transcript termination (39).

The termination signal of rrnC functions in vitro without ρ when purified RNA polymerase initiates at nuclelease-generated ends of DNA molecules (38). It remains to be determined which proteins are required for termination at this signal in vivo when transcription initiates at rrn promoter-leader regions or whether all ρ-independent termination signals can terminate rrn transcripts.

Transcription of lac from the rrnC promoter-leader region results in a 9.6-fold greater ratio of thiogalactoside transacylase to β-galactosidase compared with transcription from the lac promoter. This observation suggests that 89% of lac transcripts may terminate between lacZ and lacA when polymerase is not modified to prevent termination. Termination in this region may be due to a termination signal located between lacZ and lacY (†). Therefore, the 7:1 molar ratio of β-galactosidase to thiogalactoside transacylase synthesized from the wild-type lac operon (40) may be largely due to termination between lacZ and lacA.


We thank Michael Cashel for drawing our attention to the fact that there are additional nusA utilization sequences just prior to the 23S rRNA genes.