Transcriptional termination sequence at the end of the *Escherichia coli* ribosomal RNA *G* operon: complex terminators and antitermination

Bjarne Albrechtsen, Barbara M.Ross, Craig Squires and Catherine L.Squires*
Department of Biological Sciences, Columbia University, New York, NY 10027, USA

Received December 4, 1990; Revised and Accepted February 18, 1991 EMBL accession no. X56780

ABSTRACT

We have examined the termination region sequence of the *rrnG* operon and have observed its properties *in vivo* using a fusion plasmid test system. Transcription of *rrnG* terminator fragments was also studied *in vitro*. We found that termination of *rrnG* transcription is a complex process controlled by a tandem Rho-independent and Rho-dependent terminator arrangement which we designate *rrnG-tt'*. Together, these two elements were 98% efficient at terminating transcription initiated at the *rrnG-P2* promoter. When the two elements were separated, however, we found that the Rho-independent structure was only 59% efficient while the Rho-dependent fragment alone could account for total transcriptional termination of the tandem arrangement. The *rrnG* termination region was resistant to *rrn* antitermination and, therefore, possesses some means of stopping antiterminated transcription. The distal *rrnG* sequence contains several additional noteworthy features; the *rrnGt'* fragment contains a REP (repetitive extragenic palindromic) sequence and homology with a small unidentified reading frame following *rrnE*. This sequence is followed by *witA*, which is homologous to a citrate transport gene, *citB*. Finally, our sequence, obtained from plasmid pLC23-30, contains a Tn*1000* insertion that is absent from the *E. coli* chromosome. This insertion lies 975 bp beyond the 5S gene and is not involved in the termination events examined in this study.

INTRODUCTION

The transcriptional termination process of *Escherichia coli* ribosomal RNA (*rrn*) operons is exceptional in two regards. First, it must terminate transcription that has originated from the very active tandem *rrn* promoters. Second, it must stop transcription that is altered by an *rrn* antitermination system (1) which renders transcribing RNA polymerase less sensitive to many termination signals (2, 3). We report here the characterization of the *E. coli* *rrnG* termination region with respect to termination efficiency, termination factor Rho dependency, and response to the *rrnG* operon antiterminator (*rrnG-AT*). We found that the *rrnG* operon is followed by a complex termination region consisting of type I (Rho-independent) and type II (Rho-dependent) terminators which, together, caused highly efficient transcriptional termination. We also found that termination by this region was resistant to the *rrn* antitermination mechanism.

Type I terminator sequences are characterized by a G-C rich region capable of forming an RNA stem-loop structure when transcribed. This region is followed by four to eight thymidine residues in the DNA (4). These terminators function *in vitro* without added factors, stopping transcription abruptly after the stem structure at one or more uridines specified by the poly-thymidine run. Type II terminators require factors, such as Rho, for *in vitro* termination and do not have extensive sequence or structural similarities. Rho utilization (*rut*) sites have, however, been identified for the λ*t*R1 terminator (5). Termination usually occurs at multiple sites 80–100 nucleotides downstream of the Rho utilization site, typically spread over a 20 or 30 nucleotide region (6).

Termination region sequences have been published for all of the *E. coli* ribosomal RNA operons except *rrnA* (7–12). The *rrnB* and *rrnD* operons are essentially identical for 213 bp following the final 5S gene, and possess two type I terminators designated *t1* and *t2* (7, 9; M.Holmes, personal communication). The single *rrnG* type I terminator sequence is identical to the *rrnBt2* (11), and *rrnDt2* sequences. The remaining sequenced *rrn* operons possess only one type I terminator structure following their 5S genes and have no primary sequence similarities among them. At first it appeared that the type I terminators were sufficient to explain *rrn* termination. Several studies showed that transcription stops following type I terminators; 3'-ends of transcripts were identified for *rrnB t1* and *t2* (13), *rrnCt* (8), *rrnDt1* (14) and *rrnHt* (12). However, when measured quantitatively, *rrn* terminator fragments gave varied termination strengths and some of the type I-containing fragments were not very efficient (2). In one example, termination strength depended on the size of the fragment. A 351 base pair (bp) fragment containing *rrnCt* was 99% efficient at terminating transcripts from the *rrnG-P2* promoter. However, a smaller (228 bp) fragment

* To whom correspondence should be addressed

possessing less *rrnCt* 3′ sequence was only 87% efficient. The finding that sequences downstream of the *rrnC* type I terminator improved termination efficiency first suggested to us that all *rrn* operons might require sequences in addition to their type I terminators for full termination efficiency (2). This paper, and other work from this laboratory (15) confirm the presence of complex terminators at the ends of other *rrn* operons.

The *rrnG-AT* system consists of boxB and boxA sequences occurring early in the *rrnG-P2* transcribed region (1). These elements resemble the *nut* boxA and boxB features which mediate phage λ antitermination, except that in the *rrn* operon leaders, boxB precedes boxA. Although a short boxA sequence is sufficient for *rrn* antitermination, a mutational study has shown that boxB also influences this process (16). The strong similarities between the two antitermination sequences suggest that *rrn* and λ antitermination proceed by similar mechanisms. In λ, RNA polymerase interacts with *nut* sequences, λ N protein, and several Nus factor proteins to form an antiterminated transcription system that is resistant to transcriptional termination signals (17). The *rrnG-AT* causes transcription to proceed through type II Rho-dependent terminators that are present within *rrn* genes (3), and through Rho-dependent terminators and several type I terminators isolated from other *E. coli* operons (2). However, large fragments containing type I terminators of some of the *rrn* operons and the *rpoBC* operon can stop antiterminated transcription (2).

In this paper, we have used a previously developed plasmid test system to examine *rrnG* termination region fragments for *in vivo* termination efficiency and response to *rrn* antitermination (Figure 1). We have shown that the type I terminator is followed by extensive 3′ sequences that include a Rho-dependent terminator. This Rho-dependent terminator was required for efficient *rrnG* operon termination. Terminators were characterized as Rho-dependent by *in vivo* tests in a *rho115* mutant strain, and by *in vitro* transcription in the presence of purified Rho protein. The *in vivo* plasmid test system also showed that *rrnGtt′* stops antiterminated transcription. We have sequenced 1828 bp at the end of the *rrnG* operon, and present some of its noteworthy characteristics, augmenting information already published for that region (11).

MATERIALS AND METHODS

Bacterial strains, plasmids and phage

Strain MC1009 [Δ(*lacIPOZY*)*galU galK* Δ(*ara-leu*) *rpsL srl::Tl0 recA56 spotT relA*] (18) was provided by M.Casadaban. DW319 [F⁻ *ilv lacZ*::ISI-MS319] and M41 [DW319 (*rho115*)] (19) were obtained from D.Calhoun and made *recA56* by P1 transduction. Strain C600 (F⁻*supE tonA thr leu lacY*) was used to grow λ and was obtained from Lynn Enquist. The *E. coli* gene bank clone, pLC23−30 (20), was given to us by Masayasu Nomura. Plasmids pSL102, pSL114 (1), pHBA1, pHBA17 (2), and pHBA2 were made in this laboratory. Bacteriophages M13mp8, M13mp10 and their host strain 71.18 [Δ(*lac-proAB*) *supE thi/F′lacI*^{q}Z Δ*lacZM15 proA*$^+$B$^+$] were obtained from J. Messing (21). The *E. coli* λ gene bank of Kohara (22), containing λ436, was the very kind gift of Akira Ishihama.

Enzymes and reagents

Rho protein preparations were generous gifts from Barbara Stitt and Terry Platt. RNA polymerase was purchased from Sigma Chemical Company. RNasin was from Promega, Inc. Bal-31 was purchased from Bethesda Research Laboratories. Deoxyadenosine 5′-[α-^{32}P] triphosphate (>400 Ci/mmol) and guanosine 5′-[α-^{32}P] triphosphate (>400 Ci/mmol) were purchased from Amersham Corp. S-Acetyl CoA, 5,5′-dithiobis-(2-nitrobenzoic acid), cephaloridine, chloramphenicol, ampicillin, and isopropyl-β-2-β-D-thiogalactopyranoside were obtained from Sigma Chemical Company. 5-bromo-4-chloro-3-indolyl-β-2-D-galactopyranoside, calf intestinal alkaline phosphatase, and HincII were purchased from Boehringer Mannheim. Other restriction enzymes were from New England Biolabs or made in this laboratory. DNA polymerase−Klenow fragment was obtained from Bethesda Research Laboratories, and Sequenase™ was obtained from US Biochemicals. NA-45 diethylaminoethyl cellulose membranes and DNA purification columns (Elutip™) were purchased from Schleicher & Schuell, Inc.

DNA manipulations

DNA was prepared according to Birnboim & Doly (23) except that RNaseA was added to solution I and the cleared lysate was phenol−chloroform (1:1) extracted. Some DNA preparations were purified over Elutip™ columns. DNA fragments were isolated using NA-45 membranes and cloned as described (24) or used as templates in the *in vitro* reactions. Fragment DNA concentrations were estimated by spotting dilutions onto ethidium bromide agarose plates (25). When appropriate, vector DNA was treated with 5−10 units of alkaline phosphatase for 20 minutes at 37°C to prevent self-legation. The enzyme was inactivated by 0.1% diethyl pyrocarbonate treatment at 65°C for 15 minutes. Hybridization screening of single-stranded M13 DNA was done as described (26). Single-stranded DNA was sequenced by dideoxy chain terminating methods (26, 27). Double stranded DNA sequencing was done by alkaline denaturation of template in the presence of excess primer, followed by a standard chain termination sequencing method (28).

Construction of pHBA2

To make a vector with a blunt-end cloning-site between *rrnG-P2-AT* and the *cat* gene, we cleaved pSL114 (1) with SalI, treated with Klenow fragment and ligated it together with a HpaI hexamer linker. Insertion of the HpaI site was verified by sequencing, revealing that the last three base pairs of the 5′ TCGA were not restored by the Klenow step.

Sequencing and subcloning the *rrnG* termination region

An 1828 bp AvaI DNA fragment containing the *rrnG* termination region was isolated from pLC23−30, treated with Klenow, and cloned in both orientations into the HincII site of the M13 vector, mp8. This gave rise to #17(−) and #6(+). Three different sequencing strategies were employed to obtain the termination region sequence. In the first method, families of 3′ and 5′ deletions were generated by cutting #17 and #6 DNAs with a restriction endonuclease at appropriate sites and treating with Bal-31 exonuclease. After varied times of digestion, the insert was cut out of mp8 and recloned into a new M13 vector DNA. Subclones of progressively smaller sizes were identified and ordered by M13 DNA hybridization followed by S1 nuclease digestion. Single-stranded DNAs from these subclones were sequenced by the dideoxy chain terminating method using M13-specific primer. (LM49, used to isolate and subclone the *rrnG* terminators, was one such characterized clone. See below.) The second sequencing method used synthetic primers and the #17(−) and #6(+) single-stranded M13 DNA templates. The final sequencing method used double-stranded DNA terminator

fragments templates. Purified EcoRI fragments from λ436 and synthetic primers were used to confirm the chromosomal sequence after the *rrnG* operon. Plasmids containing the termination region subclones (see below) were sequenced using *rrnG-P2* and *cat*-specific synthetic DNA oligonucleotides primers. The sequence presented in this paper was obtained for both strands of the DNA, except for nucleotides 241−291 in the 5S gene (as numbered in Figure 3, presented in the Results section of this paper).

Subcloning of the *rrnG* terminator region into test plasmids was done by first purifying a 377 bp AluI fragment from LM49 encompassing the terminator region. This fragment was partially digested with FokI, and the resulting fragments were treated with the Klenow fragment of DNA polymerase I. This produced three convenient fragments; one carried the complex terminator region on a 173 bp fragment (*rrnGtt'*-173), another carried the upstream Rho-independent stem-loop structure (*rrnGt*-40), and the third carried the downstream Rho-dependent element (*rrnGt'*-137). These fragments were cloned into pHBA1 DNA cleaved with HpaI giving rise to pHBA181, pHBA182 and pHBA183. They were also cloned into HpaI-cleaved pHBA2 DNA, giving rise to pHBA281, pHBA282, and pHBA283 (Table I). All constructs were sequenced to verify the integrity of the insert and the cloning site junctions.

In vitro transcription reactions

In vitro reaction mixtures (5 μl) contained 20 mM Tris acetate (pH 7.9), 0.1 mM EDTA, 0.1 mM dithiothreitol, 4 mM magnesium acetate, 50 mM potassium chloride, 1.8 mM ATP, CTP, and UTP, 18 mM GTP, 3.5 μCi α-P^{32}-GTP (410 Ci/mmol), 0.07 units of RNA polymerase (0.12 μg), 0.16 μg of Rho (where indicated), 9.5 units of RNasin, and 0.01−0.014 μg of DNA template. After 15 minutes at 37°C, reactions were stopped by addition of 90 μl of 0.1 M sodium acetate (pH 5.2), 0.4% sodium dodecylsulfate, and 1.3 mg/ml carrier yeast RNA. Following phenol and chloroform extraction, the samples were ethanol precipitated, washed, dried, and re-suspended in 8 μl of formamide plus dyes. 4 μl of the samples were heated at 95°C for 3 minutes and then electrophoresed on 6% acrylamide, 7 M urea gels.

Enzyme assays

The plasmids were expressed in strains grown at 30°C (or 37°C where indicated) in rich medium supplemented with 200 μg/ml ampicillin and 2 μg/ml chloramphenicol. Cell extraction procedures, assay conditions and methods are described elsewhere (2). In this paper, the units for chloramphenicol acetyltransferase (Cat) and β-lactamase (Bla) enzyme activities were expressed as 10 nanomoles substrate converted /min/ml.

RESULTS AND DISCUSSION

To study *rrnG* operon termination, we have subcloned fragments from the end of this operon into the terminator site of our plasmid test system (Figure 1). We examined these fragments for their termination efficiency, their dependency on Rho factor, and their response to the *rrnG-AT* antiterminator. We also sequenced an extended region at the end of the *rrnG* operon to understand the genetic environment of the *rrnG* termination process.

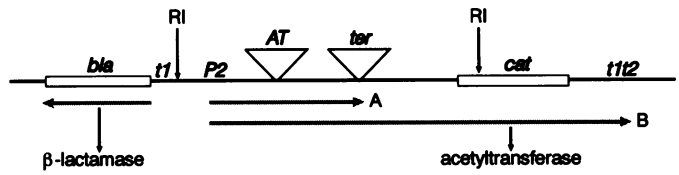

Figure 1. Schematic representation of the plasmid termination test system. *P2*, *AT* and *ter* indicate positions of the *rrnG-P2* promoter, the *rrnG-AT* antiterminator and the terminator sequences. Three different terminator fragments were placed in the *ter* site and their properties have been examined with and without the *rrnG-AT* in the *AT* site. The genes *cat* and *bla* specify chloramphenicol transacetylase and β-lactamase. Shaded right-arrows indicate the extent of terminated **(A)** and read through **(B)** transcripts from the *rrnG-P2* promoter. The read through transcript is translated into chloramphenicol acetyltransferase protein (**acetyltransferase**). The shaded left-arrow is the **β-lactamase** transcript, which originates from the *bla* promoter (not shown). Transcription from external promoters into the *P2-AT-ter-cat* cassette is blocked by *rrnBt1* and *rrnBt1t2* (*t1* and *t1t2*) terminators. The *in vitro* experiments described in this paper use restriction fragment templates derived by cleaving test plasmid DNA at the indicated EcoRI (RI) sites. The sequence of the *rrnGtt' P2-ter-cat* template is shown in Figure 2A. The *AT* insert sequence is shown in Figure 2B.

Table I. Plasmids used in this work. All plasmids have the *rrnG-P2* promoter inserted into the vector promoter site (see Figure 1). The plasmid designations, terminator name and pertinent restriction enzyme sites of the terminator region fragments used in this study are listed in columns 1, 3 and 4. The fragment sizes given in column 5 indicate the number of base pairs from the native *rrnG* termination region sequence that are present in the vector, exclusive of linker or multicloning site sequences. Column 6 lists the derivation of each plasmid used in this study, by reference if the construction has been reported elsewhere, or by the immediate antecedents if the construct is reported here for the first time (see Methods). Plasmids pSL114, pHBA2, pHBA281, pHBA282 & pHBA283 contain a 66 bp *rrnG-AT* sequence (Figure 2B) located between the transcription initiation site and the terminator cloning site (Figure 2A).

plasmid	*AT*	terminator	fragment	size	derived from
pSL102	−	−	−	−	Li *et al.* (1)
pHBA1	−	−	−	−	Albrechtsen *et al.* (2)
pHBA181	−	*rrnGtt'*	FokI-FokI-AluI	173	LM49 & pHBA1
pHBA182	−	*rrnGt*	FokI-FokI	40	LM49 & pHBA1
pHBA183	−	*rrnGt'*	FokI-AluI	137	LM49 & pHBA1
pHBA17	−	*trpt'*	EcoRI-EcoRI	223	Albrechtsen *et al.* (2)
pSL114	+	−	−	−	Li *et al.* (1)
pHBA2	+	−	−	−	pSL114 & HpaI linker
pHBA281	+	*rrnGtt'*	FokI-FokI-AluI	173	pHBA181 & pHBA2
pHBA282	+	*rrnGt*	FokI-FokI	40	pHBA182 & pHBA2
pHBA283	+	*rrnGt'*	FokI-AluI	137	pHBA183 & pHBA2

Characterization of the *rrnG* termination region

Termination strength of rrnG *terminator fragments.* The three *rrnG* terminator fragments used are shown in Table I, and their termination strengths are presented in Table II. The 173 bp FokI-AluI fragment contains the type I (Rho-independent) stem-loop structure plus 137 bp of additional 3′ nucleotides. This fragment stopped transcription with 98% efficiency (Table II, pHBA181). To see if this termination was caused solely by the type I terminator, we examined the type I structure on the 40 bp FokI fragment. We found that it terminated with only 59% efficiency (pHBA182), leaving 39% of the pHBA181 termination unaccounted for. We, therefore, concluded that the type I structure was not the sole terminator present on the 173 bp fragment, and proceeded to examine the termination properties of the downstream 137 bp FokI-AluI fragment. Surprisingly, we found that this fragment was a 98% efficient terminator

Table II. Termination efficiencies of *rrnG* terminator-region fragments. The pertinent control sequences are indicated schematically in column 2. *P2, AT, t, t'* and *cat* represent the *rrnG-P2* promoter, *rrnG-AT* antiterminator, *rrnGt* and *rrnGt'* terminators, and the chloramphenicol acetyltransferase 'reporter' gene. The approximate size of each terminator fragment is indicated by dashes. The plasmid *bla* gene has its own promoter and is used as an internal control. The Cat/Bla ratios (C/B, column 5) correct Cat enzyme activities for possible variations in plasmid copy number (45, 46), protein extraction, etc. The results are presented as percent termination (ie. 100 minus the percent read through) (last column). These values were calculated relative to **pHBA1** (for plasmids pHBA181, pHBA182 & pHBA183) and relative to **pHBA2** (for plasmids pHBA281, pHBA282 & pHBA283). The Cat and Bla values for the controls, pHBA1 and pHBA2, were averages from 10 and 12 independent experiments. Values for terminator containing plasmids were determined in 2 to 6 experiments. Standard errors were less than 14% except for pHBA183 (SE < 18.7%). Cat enzyme activity could not be detected for a pHBA182 derivative from which P2 had been deleted, demonstrating that the low termination value for *rrnGt* is not an artifact caused by the presence of a promoter (data not shown).

plasmid	control sequences			Cat	Bla	C/B	% Termination
pHBA1	*P2*		*cat*	9.70	2.38	**4.08**	**0.0**
pHBA181	*P2*	*-t-t'--*	*cat*	0.10	1.66	0.06	98.5
pHBA182	*P2*	*-t-*	*cat*	2.83	1.69	1.67	59.1
pHBA183	*P2*	*-t'--*	*cat*	0.16	1.98	0.08	98.0
pHBA2	*P2*	*AT*	*cat*	9.45	2.09	**4.52**	**0.0**
pHBA281	*P2*	*AT -t-t'--*	*cat*	0.25	3.77	0.07	98.5
pHBA282	*P2*	*AT -t-*	*cat*	1.76	2.60	0.68	85.0
pHBA283	*P2*	*AT -t'--*	*cat*	0.90	2.54	0.35	92.3

Table III. Rho-dependency of *rrnG* terminator-region fragments. Percent read-through (%RT, columns 6 and 10) of terminator fragments in *rho*$^+$ and *rho115* strains was calculated relative to the values obtained for **pHBA1**. Rho-dependency of a terminator fragment is indicated by a substantial increase of read-through in the *rho115* strain over read-through in the *rho*$^+$ strain (last column). Since *rho* is an essential gene for *E. coli* (47) its function is debilitated but not lacking in the *rho115* strain. Plasmids were grown in DW319 (*rho*$^+$ *recA56*) and M41 (*rho115 recA56*) at 37°C. We find that different strains give different values for absolute Cat and Bla levels, however, the relative values and the ratios are similar (compare% termination and% read-through in Tables I and II). The strain differences therefore do not alter the basic conclusions presented here. The values for pHBA1 are the averages from five experiments. The values for pHBA181, pHBA182 & pHBA183 are averages of two independent experiments. Standard errors were less than 19%.

plasmid	ter	*rho*$^+$				*rho115*				fold
		Cat	Bla	C/B	%RT	Cat	Bla	C/B	%RT	
pHBA1	–	11.47	1.01	11.36	**100.0**	7.71	0.68	11.34	**100.0**	**1.0**
pHBA181	*rrnGtt'*	0.11	0.97	0.11	1.0	0.28	0.64	0.44	3.9	3.9
pHBA183	*rrnGt'*	0.16	0.89	0.18	1.6	0.54	0.65	0.83	7.3	4.6
pHBA182	*rrnGt*	6.12	1.06	5.77	50.8	2.96	0.64	4.63	40.8	0.8

(pHBA183). We have designated the 40 bp fragment, containing the type I terminator, as *rrnGt*; the 137 bp fragment, containing the new terminator, as *rrnGt'*, and the 173 bp fragment, containing both of these sequences, as *rrnGtt'*. Several explanations are possible for the relatively poor termination efficiency displayed by the 40 bp fragment containing *rrnGt*. First, there may be a sequence context effect on the terminator in plasmid pHBA181. Such effects have been described for other terminators by Goliger *et al.* (29) and by Telesnitsky and Chamberlin (30). Alternatively, the *rrnGt* type I terminator may be a relatively poor terminator and require a second terminator to efficiently stop *rrn* transcription.

Measurement of *rrnG* termination thus revealed it to be a more complex process than we had first anticipated. The absence of

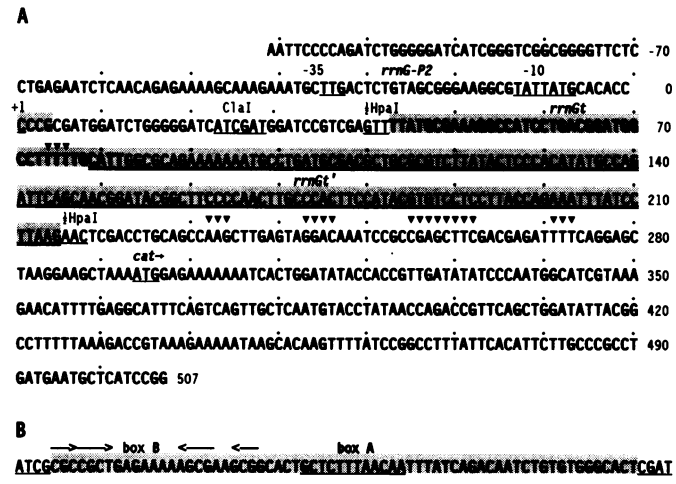

Figure 2A. The pHBA181 (P2- *rrnGtt'* EcoRI fragment sequence. This 618 bp EcoRI fragment contains the *rrnG-P2* promoter and the *rrnGtt'* fragment cloned in the HpaI terminator cloning site. The fragment begins after the *t1* sequence of the vector and ends within the proximal *cat* gene sequence (see Figure 1). The fragment shown here was used as a DNA template for *in vitro* transcription reactions. The sequence is numbered from the +1 transcription start. Gradients (▼) indicate the approximate locations of major *in vitro* transcript ends estimated from Figure 3. The shaded region is the entire *rrnGtt'* sequence; the *rrnGt* and *rrnGt'* sub-fragments are shown by **bold type** and by double-underlining. **Figure 2B.** Sequence of the *rrnG-AT* insertion. Some of the plasmids described in this study also contain the *rrnG-AT* sequence inserted into the ClaI site (Figure 2A). The 66 bp *rrnG-AT* sequence is shaded. Underlined flanking sequences indicate the half-ClaI sites when the fragment is inserted into the plasmid test system (ie. in plasmids pSL114, pHBA2, pHBA281, pHBA282 and pHBA283). BoxB secondary structure (arrows over sequence) and BoxA sequence (underlined) are indicated.

type I features in the 137 bp *rrnGt'* sequence, and the resemblance of *rrnGtt'* terminator organization to that of several amino acid biosynthetic operons, where a type I terminator is followed by a Rho-dependent terminator (31–33), led us to inquire whether this region also contains a Rho-dependent terminator.

Rho-dependency of rrnG *terminator fragments.* Plasmids containing the three terminator fragments and the control plasmid, pHBA1, were put into isogenic *rho*$^+$ and *rho*$^-$ strains and assayed (Table III). Increased read through of *rrnGtt'* and *rrnGt'* occurred in the *rho115* mutant strain. The increase was 4- to 5-fold, and similar to the elevated level of read through of the Rho-dependent *trpt'* terminator under similar conditions (5.7-fold, data not shown). This result suggested that the *rrnGt'* fragment contains a Rho-dependent terminator. We also noted that termination by the type I *rrnGt* terminator fragment (in pHBA182) was not influenced by the *rho* mutation. This was expected because of the fragment's small size and potential for secondary structure (34, 35).

In vitro termination by rrnG *terminator fragments.* The involvement of Rho in *rrnG* termination was examined further using an *in vitro* transcription system. The sequence of the *rrnGtt'* EcoRI fragment template is shown in Figure 2A, and the *rrnGt* and *rrnGt'* EcoRI fragment templates are described in the legend. Addition of Rho to the *in vitro* reactions had a dramatic effect on the fragments containing *rrnGt'*, giving rise to a series of terminated transcripts (Figure 3, lanes 1 & 2 and lanes 5 & 6). The Rho-dependent *trpt'* terminator was included as a control (lanes 7 & 8); type II transcriptional termination was also clearly visible with this fragment when Rho protein was included in the

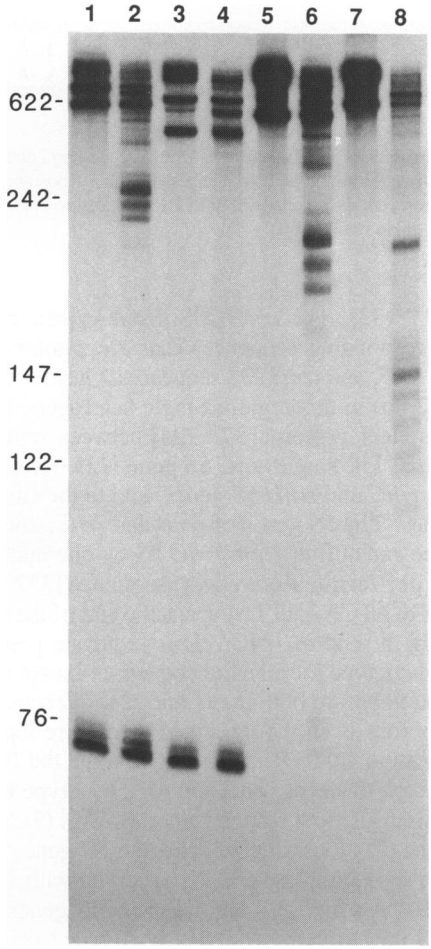

Figure 3. *In vitro* transcription of *rrnG* termination region fragments. Urea-acrylamide gel electrophoresis of *in vitro* transcripts. Templates carrying the *P2 – ter* region (Figure 2A) were transcribed with and without Rho factor (present in lanes 2, 4, 6 and 8). Results for the *rrnGtt'* template are in lanes 1 & 2; *rrn*Gt = lanes 3 & 4; *rrnGt'* = lanes 5 & 6; and *trpt'* = lanes 7 & 8. Band positions of a pBR322 DNA HpaII digest size standard are indicated in the left margin. The type II (Rho dependent) terminator stops are clustered between 220–265 nucleotides for *rrnGtt'*, between 180–225 nucleotides for *rrnGt'*, and between 115–210 nucleotides for *trpt'* Type I terminator stops are at nucleotides 74–76 for both *rrnGtt'* and *rrnGt* templates. The complex pattern in the read through area at the top of the gel is caused, in part, by a background of end-transcription and was diminished or eliminated by addition of cell extracts containing sigma-saturated RNA polymerase and DNA end binding proteins (data not shown).

in vitro reaction mixture. This experiment further showed that *rrnGt* produces a major termination band of approximately 76 nucleotides (lanes 1–4). Two additional lighter bands indicated that RNA polymerase stops at three successive residues, most probably corresponding to uridines specified by three of the five thymidines following the type I terminator stem–loop structure. The *rrnGt* fragment showed no response to the addition of Rho (lanes 3 & 4).

The *in vivo* results with the *rho115* mutant strain, and the clear Rho-dependency demonstrated in this experiment have confirmed that *rrnGt'* contains a type II Rho-dependent terminator. This result indicates a strong parallel between *rrnG* operon termination and the complex *tt'* terminator arrangements found for the tryptophan, isoleucine-valine, and leucine amino acid biosynthetic operons (29–31).

Figure 4. Distal *rrnG* operon sequence from the plasmid pLC23–30. The shaded region indicates the *rrnGtt'* fragment sequence. **Bold** sequence indicates the extent of the *rrnGt* sub-fragment, double-underlining indicates the extent of the *rrnGt'* sub-fragment. Features discussed in the Results are labeled above the first nucleotide of that feature's sequence. Arrows above the *rrnGt* sequence indicate the secondary structure of the type I terminator. The more extensive arrows indicate reading frames discussed in the Results (URF, *witA*, and the TN*1000* transposase). The extent of the short PTR and REP sequences are given in the Results.

Where and how is antiterminated transcription stopped?

The antitermination mechanism present in *rrn* operons creates an antiterminated transcription system that can read through many termination signals. Previous experiments have revealed that this mechanism can antiterminate certain Rho-dependent terminators with high efficiency (2, 3), and in this and other (15) studies we have found that Rho-dependent terminators comprise a prominent part of *rrn* termination regions. These observations pose a dilemma: If antiterminated transcription from the very strong *rrn* promoters persisted through these Rho-dependent terminators, the resulting high level of read through would be inefficient, and disruptive of downstream gene regulation. The *rrn* operons must, therefore, possess some mechanism for stopping antiterminated transcription. Two models address how transcription is stopped in these antiterminated operons. The first, and simplest, scheme holds that antiterminated transcription is stopped by an exceptionally strong terminator (model I). Alternatively, because antitermination mechanisms involve assembly of an antitermination transcription system (36), the second scheme suggests that this assemblage might require disassembly, or cancellation (37) (model II). Once antitermination is canceled, downstream transcription would again become sensitive to terminators. Furthermore, this hypothetical cancellation might depend on sequences located some distance from the termination region, possibly even changing the region where termination occurs.

We have thus tested the three termination region fragments for their *in vivo* response to the *rrn-AT*. We found that the *rrnGtt'* fragment, in the presence of *rrnG-AT*, terminates with 98% efficiency, indicating that the ability to stop antiterminated transcription does reside in this region (Table II, pHBA281). Both of the *rrnGt* and *rrnGt'* sub-fragments also showed considerable resistance to antitermination. The termination efficiency of *rrnGt'* was only reduced from 98% to 92% in the presence of the *rrnG-AT* (pHBA183 vs. pHBA283). This terminator fragment thus

possesses antitermination resistance properties that are nearly as strong as those of *rrnGtt'*. The *rrnGt* fragment also appears to be resistant to antitermination. The *rrnG-AT* actually increased the termination efficiency of the *rrnGt* fragment from 59% to 85% (pHBA182 vs. pHBA282). Although there are no apparent complementary sequences to compete with the terminator structure, it is possible that this 40 bp fragment becomes a better terminator when the additional nucleotides of the antiterminator separate it from the early transcribed region. Alternatively, the fragment may have greater termination properties when confronted with the antiterminated transcription system. This last explanation should be regarded with caution, however, because most other type I terminators that we have examined are somewhat sensitive to antitermination (2, 15).

We have concluded that the *rrnGtt'* fragment alone is sufficient to stop antiterminated transcription, and that distant upstream or downstream antitermination canceling sequences do not appear to be necessary to explain this phenomenon. These results add further support to the notion that *rrnGtt'* is the sole termination region of the operon. These conclusions are supported by recent antitermination studies with other *rrn* and *rpo* operons (15) which provide additional examples of complex terminators that are resistant to *rrnG-AT* antitermination. In most of these new examples, the type II terminator fragments, like *rrnGt'*, are exceptionally resistant to antitermination. Since we have previously reported that several strong Rho-dependent terminators are sensitive to antitermination (eg. *rhoa, trpt'* and the 16S← terminators; ref. 2), we conclude that some quality distinct from terminator strength must be involved in the antitermination resistance of *rrnGt'* and the several *rrn* and *rpo* type II terminators (2, 15). The type I structures probably also play some role in this process, because the complex terminators are consistently better than the type II terminators alone at stopping antiterminated transcription. This suggests that type I terminators should be examined more closely for possible antitermination canceling properties.

Sequence of the *rrnG* terminator region

We have determined distal sequences from the *rrnG* operon carried on a 1828 bp AvaI fragment derived from the Clarke-Carbon plasmid, pLC23−30 (Figure 4) (20). Ellwood and Nomura (38) have shown by electron microscopy that this plasmid contains a large insertion downstream of the *rrnG* operon. Our initial concern was that this insertion might influence *rrnG* termination. However, as described below, we found a Tn*1000* transposon sequence 975 bp downstream of the *rrnG* 5S gene which accounts for the divergence noted by Ellwood & Nomura. The termination region examined in this study is on a 173 bp FokI-AluI fragment commencing 4 bp after the 5S gene. Our sequence contains a recently reported 474 bp *rrnG* termination region sequence (11). That sequence extends from the 5S gene to the beginning of the *wit4* gene (see below). In the following paragraphs we identify some significant features of the pLC23−30 *rrnG*-distal sequence. Position numbers, [*italicized*] or [**emboldened**] within square brackets, refer to the locations of features in Figure 4. In comparing distal *rrn* sequences, we refer to the following published sequences for *rrnB* (7), *rrnC* (=*rrnX* in ref. 8), *rrnD* (9), *rrnE* (10), *rrnG* (this paper and ref. 11), and *rrnH* (incorrectly identified as *rrnG* or *rrnF* in ref. 12). No distal sequence has been published for *rrnA*.

Distal 23S through 5S sequences [1−351]: This region is highly conserved for the *rrn* operon sequences that have been published.

Figure 5. Comparison of URF sequences following the *rrnG* and *rrnE* operons. DNA and single letter amino acid sequences. Possible ribosome binding sites, and start and stop codons are shaded. Both URFs have the opposite orientation with respect to their *rrn* operons.

23S gene [1−132]. The *rrnG* 23S distal sequence is identical with all corresponding sequences that are available for *rrnB, rrnD, rrnE, rrnF*, and *rrnH* 23S sequences. Only *rrnC* is different (6 out of 132 bp) in this region. Single heterogeneities occur in the **23S-5S spacer region** [133−224] between *rrnG* and *rrnC, rrnD* and *rrnH*. The single *rrnG* **5S gene** [225−344] is identical to the single *rrnC* and *rrnH* 5S genes, and to the first of the two *rrnD* 5S genes. The 5S genes of *rrnB* and *rrnE*, and the second *rrnD* 5S gene each differ from *rrnG* 5S by one nucleotide. The 20 bp **PTR (pre-terminator repeat) sequence** [332−351], AACTGCCAGGCATCAAATTA, spans the end of the 5S gene and extends to the 5′ end of *rrnGt*. This sequence precedes all of the published *rrn* type I terminator sequences except that of *rrnC*. A PTR is found before both the *t1* and *t2* sequences of the *rrnB* and *rrnD* operons even though no 5S genes are located before the *t2* terminators. A PTR is also found after the first 5S gene of *rrnD* although this gene is not followed by a type I terminator. The PTR consensus, first recognized by Brosius (7), corresponds roughly to the 3′ processing stalk of the 5S gene, is perfectly conserved in six examples, and is conserved with one and two differences for the first and second *rrnD* 5S genes. The close association of this sequence with nearly all type I *rrn* terminators suggests it might play some role in termination. However, it should be noted that only the final three nucleotides of the PTR are present in the *rrnGt* and *rrnGtt'* fragments examined in this study. Therefore, in our experiments, the PTR is not essential for either termination or for stopping antiterminated transcription.

Termination region [349−521]: The PTR sequence is followed by the *rrnGt* **type I terminator** [349−388]. This sequence possesses 29 contiguous base pairs that are also found in the *rrnBt2* and *rrnDt2* terminators (11, M. Holmes, personal communication). All distal *rrn* sequences diverge abruptly following their type I terminators. The *rrnGt'* **type II terminator** [385−521] commences at this divergence. We have demonstrated the Rho-dependency of *rrnGt'* in this paper. We found no consensus sequences in common between *rrnGt'* and other known Rho-dependent terminators, and we do not know which nucleotides are responsible for the efficient Rho-dependent termination properties of this fragment. The *rut* sequences identified as Rho utilization sites in the λ tR1 terminator (5) are not present in *rrnGt'*.

The rrnGt' sequence does, however, possess two notable similarities to other *E. coli* sequences (in GenBank, release 64), and one of these is closely associated with the 3′ end of another *rrn* operon. First, there is the 26 bp **REP (repetitive extragenic palindromic) sequence** [403−428], TGCCTGATGCGACGCTGCGCGTCTTA, similar to other REP sequences associated with many *E. coli* genes (39−41). REP sequences were not found in any other published *rrn* termination region sequences that extend significantly beyond the type I terminators. These include *rrnB, rrnC, rrnE*, and *rrnF* sequences, and to this degree, REP sequences are not a general feature of *rrn* termination regions.

The contribution of the *rrnG* REP sequence to the termination process has not yet been determined. Second, there is a 63 bp **URF (unidentified reading frame) sequence** [426–521] that is similar to a short sequence following the *rrnE* operon (*rrnE* URF2) (10). The translated *rrn*G and *rrnE* URF sequences share 15 identical, or 17 functionally similar, amino acids (Figure 5). Both URFs are encoded on the DNA strand opposite the *rrn* operons. Additional up-stream translational start signals exist for both sequences; therefore, these URFs may either code for short peptides of 20 and 18 amino acids, or they may represent homologous C-terminal sequences of larger peptides possessing totally divergent N-termini. The URFs are not the same distance from their respective *rrn* operons, lying 40 nucleotides after *rrnGt* and 481 nucleotides after *rrnEt*. Because of its greater distance from the type I terminator, it would be interesting to see if the *rrnE* URF2 plays some role in *rrnE* operon termination, or if sequences closer to the type I terminator influence this process.

Distal sequences [522–1828]: The remaining sequences on the AvaI fragment contain a gene of unknown function, designated as *witA* [667–1215] (11), and part of the transposon **Tn1000** [1216–1828] (42). The beginning of the *witA* nucleic acid sequence is 56% identical with the citrate transport gene *(citB)* (43), and it has been suggested that *witA* specifies a transport protein (44). After 184 codons the *witA* sequence on pLC23–30 is interrupted by the Tn*1000* insertion. Sequences obtained from the Kohara *E. coli* gene bank clone λ436 (22) established the location of the divergence from the *witA* sequence at nucleotide 1216. The most probable explanation for the Tn*1000* insertion in pLC23–30 is that higher copy numbers of *witA* are toxic. Over production of a transport gene could adversely influence cell viability.

CONCLUDING REMARKS

In this paper we have shown that the *rrnG* operon contains a complex termination region, possessing tandem type I (Rho-independent) and type II (Rho dependent) elements. Other experiments from this laboratory have revealed additional examples of complex terminators in *rrn* and RNA polymerase operons (15). These new examples in macromolecular synthetic operons and the already noted examples in amino acid biosynthetic operons (29–31) raise questions concerning how general this complex terminator arrangement might be. For example, do operons that contain housekeeping and carbohydrate utilization genes also possess complex terminators, and are transcripts that originate from weak promoters terminated in the same way as those originating from stronger promoters?

We observed strong termination by *rrnGtt'* in the presence of the *rrnG-AT* antitermination sequence. This suggests that the complex terminator is sufficient to stop antiterminated transcription, and reduces the likelihood that the antitermination mechanism is canceled elsewhere. Of the two sub-fragments present in the complex terminator, the Rho–dependent *rrnGt'* fragment was the better, but not quite as proficient as *rrnGtt'* at stopping antiterminated transcription. Thus *rrnGt'* might be regarded as the strong terminator described in model I above. However, several other strong Rho-dependent terminators are very sensitive to antitermination (2), suggesting that *rrnGt'* is qualitatively different and that termination strength alone is not the factor that determines resistance to antitermination.

ACKNOWLEDGEMENTS

We are grateful to Barbara Stitt and Terry Platt for their generous gifts of purified Rho protein, to Louise Clarke and Masayasu Nomura for the plasmid pLC23–30, and to Akira Ishihama for kindly making the Kohara *E. coli* gene bank available to us. We are indebted to Ciaran Condon for his helpful comments concerning this manuscript and to Birjis Akhund, Leota Mehra and Yasmin Houssein for their initial work identifying and cloning the *rrnG* terminator region. This work was supported by grants from the Carlsberg Foundation and the Danish Natural Science Research Foundation to B.A. and by the National Institutes of Health (GM24751) to C.L.S.

REFERENCES

1. Li,S., Squires,C.L. and Squires,C. (1984) *Cell* **38**, 851–860.
2. Albrechtsen,B., Squires,C.L. and Squires,C. (1990) *J. Mol. Biol.* **213**, 123–134.
3. Aksoy,S., Squires,C.L. and Squires,C. (1984) *J. Bacteriol.* **159**, 260–264.
4. Friedman,D.I., Imperiale,M.J. and Adhya,S.L. (1987) *Annu. Rev. Genet.* **21**, 453–488.
5. Chen,CY. and Richardson,J.P. (1987) *J. Biol. Chem.* **262**, 11292–11299.
6. Galloway,J.L. and Platt,T. (1986) In *Regulation of Gene Expression—25 Years On.* (Higgins,C.F. and Booth,I.R., eds). Cambridge University Press, New York. pp. 155–178.
7. Brosius,J., Dull,T.J., Sleeter,D.D. and Noller,H.F. (1981). *J. Mol. Biol.* **148**, 107–127.
8. Young,R.A. (1979) *J. Biol. Chem.* **254**, 12725–12731.
9. Duester,G.L. and Holmes,W.M. (1980) *Nucl. Acids Res.* **8**, 3793–3807.
10. Liebke,H. and Hatfull,G. (1985) *Nucl. Acids Res.* **15**, 5515–5525.
11. Seol,W. and Shatkin,A.J. (1990) *Nucl. Acids Res.* **18**, 3056.
12. Sekiya,T., Mori,M., Takahashi,N. and Nishimura,S. (1980) *Nucl. Acids Res.* **8**, 3809–3827.
13. Li,S. (1985) Ph.D. thesis, Columbia University, New York.
14. Szeberenyi,J. and Apirion,D. (1983) *J. Mol. Biol.* **168**, 525–561.
15. Albrechtsen,B. (1990) Ph.D. thesis, Columbia University, New York.
16. Berg,K.L., Squires,C.L. and Squires,C. (1989) *J. Mol. Biol.* **209**, 345–358.
17. Friedman,D.I. (1988) In *The Bacteriophages*, vol. II. (Calendar,R. ed.) Plenum Publishing, New York. pp. 263–319.
18. Berg,K.L., Squires,C.L. and Squires,C. (1987) *J. Bacteriol.* **169**, 1691–1701.
19. Guterman,S. and Howitt,L. (1979) *Mol. Gen. Genet.* **169**, 27–34.
20. Clarke,L. and Carbon,J. (1976) *Cell* **9**, 91–99.
21. Messing,J., Gronenborn,B., Muller-Hill,B. and Hofschneider,P. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3642–3646.
22. Kohara,Y., Akiyama,K. and Isono,K. (1987) *Cell* **50**, 495–508.
23. Birnboim,H. and Doly,J. (1979) *Nucl. Acids Res.* **7**, 1513–1523.
24. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982). *Molecular Cloning, A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. p. 391.
25. Davis,R.W., Botstein,D. and Roth, J.R. (1980) *A Manual for Genetic Engineering. Advanced Bacterial Genetics.* Cold Spring Harbor, New York. p. 184.
26. Squires,C., Krainer,A., Barry,G., Shen,W.-F. and Squires,C.L. (1981) *Nucl. Acids Res.* **9**, 6827–6839.
27. Shen,W.-F., Squires,C. and Squires,C.L. (1982) *Nucl. Acids Res.* **10**, 3303–3313.
28. Zagursky,R.J., Baumeister,N. and Berman,M.L. (1985) *Gene Anal. Techn.* **2**, 89–94.
29. Goliger,J.A., Yang,X., Guo,H.-C. and Roberts, J. (1989) *J. Mol. Biol.* **205**, 331–341.
30. Telesnitsky,A.P.W. and Cbamberlin,M. (1989) *J. Mol. Biol.* **205**, 315–330.
31. Mott,J., Galloway,J. and Platt,T. (1985) *EMBO J.* **4**, 1887–1891.
32. Sameshima,J.H., Wek,R.C. and Hatfield,G.W. (1989). *J. Biol. Chem.* **264**, 1224–1231.
33. Rosenthal,E.R. and Calvo,J.M. (1987) *J. Mol. Biol.* **194**, 443–452.
34. Morgan,W.D., Bear,D.G., Litchman,B.L. and von Hippel,P.H. (1985) *Nucl. Acids Res.* **13**, 3739–3754.
35. McSwiggen,J.A., Bear,D.G. and von Hippel,P.H. (1988) *J. Mol. Biol.* **199**, 609–622.
36. Horwitz,J., Li,J. and Greenblatt,J. (1987) *Cell* **51**, 631–641.

37. Barik,S., Ghosh,B., Whalen,W., Lazinski,D. and Das,A. (1987) *Cell* **50**, 885–899.
38. Ellwood,M. and Nomura,M. (1982) *J. Bacteriol.* **149**, 458–468.
39. Newbury,S.F., Smith,N.H., Robinson,E.C., Hiles,I.D. and Higgins,C.F. (1987) *Cell* **48**, 297–310.
40. Higgins,C.F., McLaren,R.S. and Newbury (1988) *Gene* **72**, 3–14.
41. Shyamala,V., Schneider,E. and Feffo-Luzzi Ames,G. (1990) *EMBO J.* **9**, 939–946.
42. Saadi,S., Maas,W.K., Hill,D.F. and Bergquist,P.L. (1987) *J. Bacteriol.* **169**, 1836–1846.
43. Sasatsu,M., Misra,T.K., Chu,L., Laddaga,R. and Silver,S. (1985) *J. Bacteriol.* **164**, 983–993.
44. Seol,W. and Shatkin,A.J. (1990) *J. Bacteriol.* **172**, 4745.
45. Uhlin,B.E. and Nordstroem,K. (1977) *Plasmid* **1**, 1–7.
46. Klotsky,R.-A. and Schwartz,I. (1987) *Gene* **55**, 141–146.
47. Das,A., Court,D. and Adhya,A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1959–1963.