Nucleic Acids Research, 1998, Vol. 26, No. 21 **4895–4900**

A simple polypyrimidine repeat acts as an artificial Rho-dependent terminator *in vivo* and *in vitro*

Martine Guérin, Nathalie Robichon, Johannes Geiselmann[1] and A. Rachid Rahmouni*

Centre de Biophysique Moléculaire, CNRS, rue Charles Sadron, 45071 Orléans cédex 2, France and [1]Laboratoire du Contrôle de l'Expression Génique, Université Joseph Fourier, 460, rue de la Piscine, CERMO, BP 53 X, 38041 Grenoble cédex 9, France

Received July 13, 1998; Revised and Accepted September 21, 1998

ABSTRACT

In this paper, we present evidence that an efficient Rho-dependent terminator can be created by introducing a simple $(AG/TC)_n$ DNA repeat into a transcription unit. The Rho termination activity *in vivo* and *in vitro* is dependent on the length and the orientation of the insert. The transcription of at least 30 bp of the $(AG/TC)_n$ repeat in the orientation encoding the $(rUrC)_n$ sequence on the transcript leads to Rho-dependent termination at a downstream non-terminator site. Our results indicate that the high efficiency of this artificial Rho-dependent terminator is due to optimal interactions between the $(rUrC)_n$ RNA sequence and Rho protein. Thus, our findings strongly suggest that an adequate loading site is the primary determinant for Rho termination activity and provide a more defined system for future investigations.

INTRODUCTION

Transcription termination in *Escherichia coli* occurs at two types of terminators that differ by their mechanisms and DNA sequences (1). Whereas transcript release at intrinsic terminators is a spontaneous process and involves only RNA polymerase, termination at Rho-dependent termination sites requires the participation of Rho protein. A variety of Rho-dependent terminators have been identified in prokaryotic genomes. However, only a few examples have been analyzed in detail, mostly due to their complexity. For instance, the well-studied Rho-dependent terminator λtR1 is composed of two distinct but overlapping parts extending over a region of 200 bp. The upstream part, called *rut*, encodes a segment of RNA to which Rho binds (2–4). Deletion experiments, as well as oligonucleotide competitions (5), have shown that the *rut* element, and especially two regions designated *rutA* and *rutB*, are required for efficient Rho-dependent termination. The downstream part contains several clusters of sites where termination occurs (6,7). Some of these sites have been correlated with the positions where RNA polymerase pauses extensively during transcript elongation (8,9).

Most of our current understanding of Rho function stems from *in vitro* investigations of the structural and enzymatic properties of the protein. Rho acts as a hexamer of identical subunits, each subunit having specific domains for binding RNA and ATP (10,11 and references therein). According to a widely accepted view, the Rho-dependent termination process involves three steps (for reviews see 12 and 13). The protein first binds to the RNA transcript at a loading site located well upstream from the termination endpoints (14). This binding activates the RNA-dependent ATPase activity, leading to a preferential 5′ to 3′ translocation of the protein along the RNA towards the transcription elongation complex stalled at a downstream pause site (15–18). Using its ATP-dependent helicase activity, Rho then dissociates the RNA transcript from the ternary complex (19). The RNA transcript, therefore, plays a key role in Rho function by facilitating a productive interaction between the protein and the transcription elongation complex.

Several features of the RNA have been shown to be important for Rho recognition and action (20 and references cited therein). During the initial step of the termination process, Rho acts as an RNA binding protein. The binding site for the Rho hexamer has been defined as a sequence of ~70–80 nt in length that is rich in cytosines and devoid of secondary structure (20–22). The general preference of Rho for cytosine-rich RNAs led to the postulate that the interactions between cytosine residues and the protein are essential for high affinity binding at the loading site (20,23). However, extensive site directed mutagenesis of the *rut* elements of λtR1 and trp t′ terminators did not reveal any direct correlation between the efficiency of termination and the content or spacing of the cytosine residues within the Rho binding site (24–26). These results suggested that high affinity binding at the loading site is not the only determinant for efficient termination.

The subsequent translocation of Rho depends upon cycles of RNA binding and release that are coupled to rounds of ATP hydrolysis (11,17,18). Using short RNA oligonucleotides as cofactors, Wang and von Hippel (27) have examined in detail the ATPase activity of Rho as a function of the RNA sequence. These studies have shown that mixed (rUrC) oligomers stimulate the ATPase activity of Rho more than do homo-oligomers of rC. The maximal ATPase activity depends on the position of rU residues within the mixed (rUrC) oligomers. Thus, it was suggested that functional interactions between Rho and the transcript that lead to efficient Rho-dependent termination may rely on a combination of tight binding of RNA segments to individual cofactor binding sites within the Rho hexamer and the rate at which these segments are released following ATP hydrolysis. Apparently, in mixed (rUrC) sequences, the rC residues provide strong binding whereas rU residues facilitate the release of RNA during the translocation process (27,28).

In this study, we have tested *in vivo* the validity of the current models of Rho-dependent transcription termination. We asked the

*To whom correspondence should be addressed. Tel: +33 2 38 25 76 08; Fax: +33 2 38 63 15 17; Email: rahmouni@cnrs-orleans.fr

question whether the presence of an $(rUrC)_n$ repeat within a transcript can be sufficient to bring about Rho-dependent RNA release at non-terminator DNA sequences. To this end, we have inserted different lengths of the $(AG/TC)_n$ DNA repeat between a promoter and the coding region of the chloramphenicol acetyltransferase (CAT) gene. Our results show that transcription of the insert in the orientation encoding the $(rUrC)_n$ repeat on the transcript leads to efficient Rho-dependent termination at a downstream site within the CAT gene. The presence of a stretch of 30 nt of the $(rUrC)_n$ sequence on the transcript is sufficient to induce the termination process.

MATERIALS AND METHODS

Enzymes and chemicals

Escherichia coli RNA polymerase was obtained from Sigma (France). Restriction enzymes, the Klenow fragment of DNA polymerase I, T4 DNA ligase, T4 polynucleotide kinase and nuclease S1 were from New England Biolabs (Ozyme, France). The enzymes were used according to the manufacturers' specifications. All the chemicals, including the antibiotics, were from Sigma. The unlabelled NTPs, dNTPs and ddNTPs were bought from Boehringer Mannheim and the ^{32}P-labeled NTPs and dNTPs were from Amersham (France). All oligonucleotides were obtained from Eurogentec Belgium.

Plasmids and bacteria

The plasmid vector pKKTac is a derivative of pKK232-8 (28) in which the Tac promoter was inserted at the *Sma*I site to drive the transcription of the CAT gene. The Tac promoter was subcloned as a 90 bp *Hin*dIII–*Bam*HI fragment from pDR540 (Pharmacia). The pAGn and pTCn plasmid derivatives were constructed by inserting the $(AG/TC)_n$ repeats in both orientations at the filled-in *Sal*I site of pKKTac. For pAG44 and pTC44, the $(AG/TC)_{22}$ repeat was subcloned from an existing plasmid (30). For the other derivatives, the inserts were made by ligating two 10mer complementary oligonucleotides $(AG)_5$ and $(TC)_5$. The ligation products were separated on a 10% polyacrylamide gel and the relevant multimers isolated. The lengths and the orientations of the inserts were determined by dideoxy sequencing. The resulting plasmids were designated according to the sequence of the RNA within the Tac-promoted transcript. The pλtR1 plasmid was constructed by cloning a 233 bp λtR1 fragment containing *rut*A, *rut*B, tI, tII and tIII sequences into the filled-in *Sal*I site of pKKTac. The terminator fragment was prepared from λ DNA by PCR amplification using the primers 5′-AGCCCTTCCCGAGTAAC-3′ and 5′-TAGTGCCTC-GTTGCGTT-3′. All of the above constructions and the preparation of plasmid DNAs for *in vitro* transcription were performed in *E.coli* HB101 according to standard procedures (31). For the *in vivo* studies, the plasmids were introduced into HB101 F′Iq from our laboratory collection. The F′ episome was maintained by growing the cells in the presence of 5 µg/ml of kanamycine. The Tac promoter was induced by addition of IPTG at the final concentration of 1 mM. For the experiments with the Rho mutant, two *Salmonella typhimurium* isogenic strains were used: the wild type rho+ strain and the *rho-111* mutant which has a mutation in Rho coding region that abolishes Rho-dependent termination at non-permissive temperature (37°C). The two strains were from the collection of J. Roth (32) and they were provided to us by L. Bossi (Gif sur Yvette, France).

Isolation and analyses of *in vivo* RNA transcripts

Total bacterial RNAs were extracted from the cells harboring the plasmids as previously described (33,34). The induction with 1 mM IPTG was done 20 min prior to the cell harvest. Northern blot analyses were performed by fractionating 10 µg of total RNA on a 2% agarose gel containing 6% formaldehyde. The RNAs were transferred onto a nylon membrane filter and hybridized with ^{32}P-labeled DNA probes as follows. The membrane was prehybridized for 4 h at 40°C in a solution containing 50% formamide, 5× SSPE, 2× Denhardt solution, 1% SDS and 100 µg/ml denatured sonicated calf thymus DNA. After addition of the denatured DNA probe, the hybridization was allowed to proceed for 18 h at 40°C. The membrane was first washed at room temperature in 2× SSC plus 0.5% SDS for 10 min followed by two washes at 30°C in 0.5× SSC plus 0.1% SDS for 15 min. The different probes used, the *Eco*RI–*Bam*HI Tac fragment, the *Hin*dIII–*Pvu*II CAT proximal fragment and the *Pvu*II–*Sty*I CAT distal fragment were labeled by random priming.

S1 mapping was performed as previously described (34) with some modifications. Briefly, a single-stranded *Hin*dIII–*Pvu*II DNA probe labeled at its *Hin*dIII 3′ end was mixed with 15 µg of total RNA in 40 µl hybridization buffer (40 mM PIPES pH 6.5, 0.4 M NaCl, 1 mM EDTA and 80% formamide) at 37°C for 15 h. After addition of 225 µl of H_2O, 5 µl of 3 M NaCl and 30 µl of 10× S1 buffer (300 mM sodium acetate pH 4.6, 10 mM zinc acetate, 50% glycerol and 200 µg/ml denatured herring sperm DNA), the sample was digested with 160 U of S1 nuclease for 10 min at 37°C. The protected hybrids were precipitated with ethanol and analyzed on a 5% denaturing polyacrylamide gel.

In vitro transcription assays

The *in vitro* transcription mixture contained, in a total volume of 25 µl of buffer (40 mM Tris–HCl pH 8; 5 mM $MgCl_2$, 1 mM DTT and 150 mM KCl), 0.15 pmol of supercoiled DNA plasmid, 0.4 pmol of RNA polymerase and, when present, 1 pmol of purified *E.coli* Rho hexamer (the Rho protein was a generous gift from P. von Hippel; its preparation was previously reported; 27). The sample was incubated for 10 min at 37°C to form open complexes. When desired for competition experiments, 8 pmol of the 18mer oligonucleotides AG18, CT18 or AG18/CT18 duplex were added at this time. The elongation reaction was then carried out for 15 min at 37°C in the presence of three unlabeled NTPs (200 µM each) and one labeled NTP (20 µM non-radiolabeled and 60 nM at 3000 Ci/mmol). The reactions were stopped by the addition of a solution containing 40 µg/ml tRNA, 1 mM EDTA and 0.3 M sodium acetate. After phenol extraction, the samples were ethanol precipitated and resuspended in 80% formamide-dyes solution and analyzed on a 6% denaturing polyacrylamide gel. To study the kinetics of RNA elongation we used the same transcription protocol, except that the KCl concentration was 50 mM, and instead of a supercoiled plasmid we used the *Eco*RI–*Pvu*II restriction fragment from pTC44 as template. The fragment was cut from the plasmid and purified on a 6% native polyacrylamide gel. The transcription reactions were arrested at the indicated times by addition of the stop solution as above.

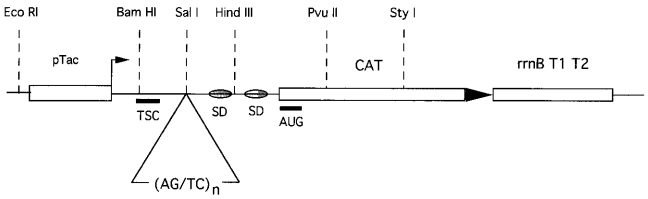

Figure 1. Schematic illustration of the relevant part of the plasmid constructs used in this study. The pAGn and pTCn plasmid series differ by the orientation of the $(AG/TC)_n$ insert relative to the direction of transcription initiated at the Tac promoter. In addition to the relevant restriction sites, the translation stop codons (TSC), the ribosome binding sites (SD) and the CAT start codon are indicated.

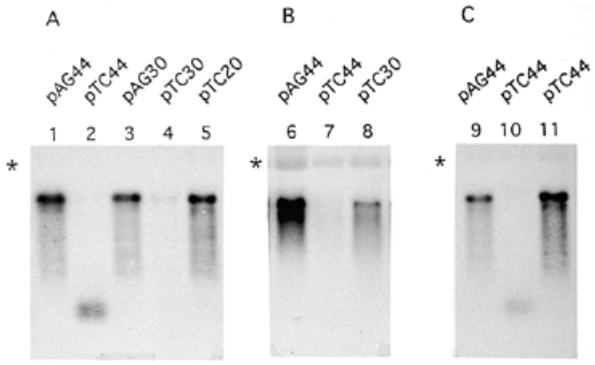

Figure 2. Northern blot analyses of *in vivo* transcripts initiated at the Tac promoter. (**A** and **B**) RNAs extracted from *E.coli* cells harboring the indicated constructs. (**C**) RNAs extracted from wild type *S.typhimurium* (lanes 9 and 10) and from a *rho-111* mutant (lane 11) transformed with the indicated plasmids. All the cells were grown at 37°C (non-permissive temperature for *rho-111* mutant). The blots in (A and C) were hybridized with the *Hin*dIII–*Pvu*II DNA probe whereas the blot in (B) was hybridized with the *Pvu*II–*Sty*I DNA probe. Asterisks indicate slight cross hybridization between the probe and 23S rRNA.

RESULTS

Effect of $(AG/TC)_n$ repeats on transcription *in vivo*

Figure 1 illustrates the relevant features of the plasmid vector pKKTac that was used to study the effect of $(AG/TC)_n$ repeats on transcription *in vivo*. This vector carries the CAT gene driven by the IPTG-inducible Tac promoter. Different lengths (44, 30 and 20 bp) of the $(AG/TC)_n$ repeat were inserted in both orientations at the *Sal*I site which is located between the transcription start site (+1) and the first codon of the CAT gene (AUG). The resulting plasmid series pAGn and pTCn were designated according to the RNA sequence $(rArG)_n$ or $(rUrC)_n$ within the Tac-promoted transcript, respectively.

The effect of the $(AG/TC)_n$ inserts on the expression of the CAT gene *in vivo* was first monitored by testing, under inducing conditions, the chloramphenicol resistance of *E.coli* cells harboring the plasmids (results not shown). As for cells carrying the parent plasmid vector pKKTac, chloramphenicol at 100 µg/ml did not alter the growth of cells harboring pTC20 or any one of the plasmids from the pAGn series (pAG44, pAG30 and pAG20). However, under the same conditions, cells transformed with pTC44 and pTC30 were sensitive to the antibiotic. At lower concentrations of chloramphenicol (10–30 µg/ml), the pTC30-transformed cells grow slowly whereas no growth was observed for cells carrying pTC44. These results suggested a possible interference of the insert at the transcriptional or post-transcriptional levels.

To determine if the inhibition of CAT expression in pTC44 and pTC30-transformed cells was exerted at the level of transcription, we performed northern blot analyses of the RNAs produced *in vivo*. Figure 2A shows the results obtained with a probe that hybridizes to the 5′ part of the CAT RNA transcript (*Hin*dIII–*Pvu*II fragment; Fig. 1). In agreement with the chloramphenicol resistance tests, the probe reveals that a full-length CAT-specific RNA (~1250 nt in length) was made from the plasmids pAG44, pAG30 and pTC20 (lanes 1, 3 and 5). An RNA message of similar size was observed for the parent plasmid vector pKKTac (results not shown). In contrast, a short RNA transcript (~230 nt in length) was detected for pTC44 instead of the full-length CAT message (lane 2). In the case of pTC30, the amount of full-length transcript was drastically reduced (between 50 and 70% loss of signal as compared to pAG30) (lane 4) and a faint band corresponding to the short transcript could be seen upon longer exposure of the blot (results not shown). The blots were further analyzed by hybridization with two other DNA probes. The *Eco*RI–*Bam*HI probe that is specific to the promoter proximal region gave the same patterns as in Figure 2A. The *Pvu*II–*Sty*I probe, which covers the middle of the CAT transcript, revealed the full-length message for pAG44 and, to a lesser extent, for pTC30, but did not detect the short RNA (Fig. 2B). We therefore concluded that the short RNA produced from pTC44, and to some extent from pTC30, corresponds to a Tac-promoted transcript that has been either terminated within, or truncated up to, the region between *Hin*dIII and *Pvu*II sites.

Rho protein is involved in the termination process

The effect of the $(AG/TC)_n$ sequence on Tac-promoted transcription depends on the orientation and the length of the insert and therefore on the nature of the transcript. Furthermore, the putative termination endpoints are located downstream from the repeat. These facts prompted us to consider the involvement of Rho protein in the termination process. To test this possibility, the transcriptional behavior of the plasmid pTC44 was analyzed in the *S.typhimurium rho-111* mutant. The temperature-sensitive *rho-111* allele has been well-characterized *in vivo* and *in vitro* (32,35,36). At the non-permissive temperature (37°C), the mutant strain is defective in Rho-dependent termination, presumably due to an altered interaction of the mutant protein with the transcript (36). Figure 2C shows the northern blot analyses of the RNAs extracted from the wild type and the mutant cells transformed with pTC44. As for *E.coli* cells, the *Hin*dIII–*Pvu*II DNA probe reveals the truncated transcript that was made in the wild type strain harboring the plasmid (lane 10). Conversely, the transcription of the plasmid in the *rho-111* mutant cells leads to the production of a full-length CAT message (lane 11) possessing the same size as the one produced from the control plasmid pAG44 (lane 9). Thus, these data strongly suggest that the truncated RNA results from a Rho-dependent termination event involving the $(rUrC)_n$ repeat on the transcript. This conclusion was supported by additional experiments (not shown) in which the 44 bp repeat was inserted downstream from the start codon (AUG) of the CAT gene (*Pvu*II site). In this case, the Rho-dependent transcription termination associated with the $(rUrC)_n$ repeat on the transcript was observed only in the absence of protein synthesis (10 min treatment with chloramphenicol or tetracycline). These results are consistent

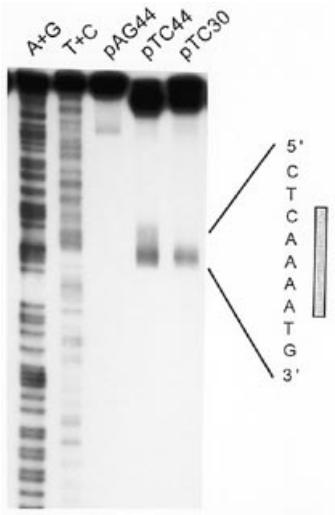

Figure 3. S1 mapping of the Rho-dependent termination end-points. RNAs extracted from *E.coli* cells containing the indicated construct were hybridized with the template strand of the *Hin*dIII–*Pvu*II fragment, digested with S1 nuclease and analyzed on a 6% denaturing polyacrylamide gel. A+G and T+C correspond to the Maxam and Gilbert sequencing ladders obtained from the same DNA fragment.

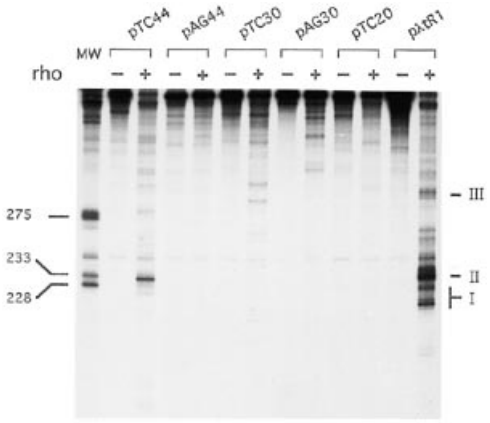

Figure 4. *In vitro* transcription of pAGn and pTCn plasmids in the absence (–) or the presence (+) of Rho protein. The Rho-dependent termination products generated from the control plasmid, pλtR1, are designated by I, II and III and correspond to the preferred subsites within the λtR1 terminator. MW refers to RNA molecular weight markers that have been produced by run-off transcription using plasmids cut at appropriate restriction sites.

with the well-known interference between translation and Rho-dependent transcription termination.

S1 mapping of the termination endpoints

As a further step to characterize the Rho-dependent RNA release within pTC44 and pTC30 *in vivo*, we analyzed the termination endpoints by S1 mapping. Cellular RNAs isolated from *E.coli* cells harboring the plasmids were hybridized to a 3′ end-labeled single-stranded DNA probe (the template strand of the *Hin*dIII–*Pvu*II DNA fragment) and digested with S1 nuclease as described in Materials and Methods. The data in Figure 3 clearly show that the RNAs extracted from cells transformed with pTC44 and pTC30 contain CAT-specific truncated transcripts, their 3′ ends clustering over ~5 nt within the sequence 5′-UUUUG. The amount of truncated transcripts is lower for pTC30 than for pTC44 which is in line with the northern blot analyses. However, the fact that we detect as much truncated RNA for pTC30 may indicate either a higher sensitivity of the S1 mapping experiment or technical limitations of the northern blot analyses when detecting low amounts of short RNAs. The cluster of RNA 3′ ends detected by S1 mapping is obviously related to the sites on the DNA template where Rho-dependent RNA release occurs. These release sites are located 130 bp downstream from the $(AG/TC)_n$ repeat (between positions 230 and 235 relative to the transcription start site in the pTC44 construct). Conceivably, the distribution of RNA 3′ ends may also reflect a heterogeneity of the S1 digest due to thermal fluctuations. Therefore, these results do not exclude the possibility that Rho protein brings about RNA release at a single template position *in vivo*.

In vitro Rho-dependent termination

We next investigated the effect of the $(AG/TC)_n$ DNA repeats on transcription *in vitro* on supercoiled plasmids. The pAGn and pTCn plasmid series, at native superhelical density, were transcribed by purified *E.coli* RNA polymerase in the absence or the presence of Rho protein and the resulting transcripts were analyzed by polyacrylamide gel electrophoresis. As shown in Figure 4, transcription of all plasmids in the absence of Rho gives rise to high molecular weight RNAs, most of which presumably originate at the strong Tac promoter. The ability of Rho to bring about transcription termination under our *in vitro* conditions is clearly seen with the control plasmid pλtR1 which contains the Rho-dependent terminator λtR1 inserted in the functional orientation at the *Sal*I site of the pKKTac vector. The presence of Rho protein caused the RNA polymerase to terminate transcription of pλtR1 at the well-characterized preferred subsites tI, tII and tIII indicated on the right side of the gel (Fig. 4). The presence of Rho had no significant effect on transcription of the plasmids pAG44, pAG30 and pTC20. In contrast, several Rho-dependent termination products were readily detected for pTC44 and, to some extent, for pTC30 (Fig. 4). Hence, in support of the *in vivo* experiments, these results show that the presence of at least 30 nt of the $(rUrC)_n$ repeat on the transcript mediates Rho-dependent termination at downstream non-terminator DNA sequences. Efficient *in vitro* Rho-dependent activity, however, is more clearly revealed with the pTC44 construct for which a predominant termination product is observed. This termination product corresponds to an RNA of ~230 nt in length, indicating that the release site is 130 bp downstream from the repeat. Therefore, the major site at which Rho brings about transcription termination *in vitro* on pTC44 is the same as the one detected by S1 mapping of RNAs produced *in vivo*.

To provide additional proof that the $(rUrC)_n$ sequence on the transcript is involved in the Rho termination activity, we performed transcription experiments in the presence of oligonucleotide competitors. As shown in Figure 5, the Rho termination activity of pTC44 is completely inhibited by the presence of the antisense AG18 oligonucleotide. Under the same conditions, no inhibition was observed with the sense CT18 oligomer (Fig. 5) nor when the two complementary oligonucleotides AG18 and CT18 were added as a duplex (data not shown). Also, the effect is sequence specific since the AG18 oligomer did not interfere with the Rho termination activity of the plasmid pλtR1 (Fig. 5). Thus, these results show that the formation of a heteroduplex between the

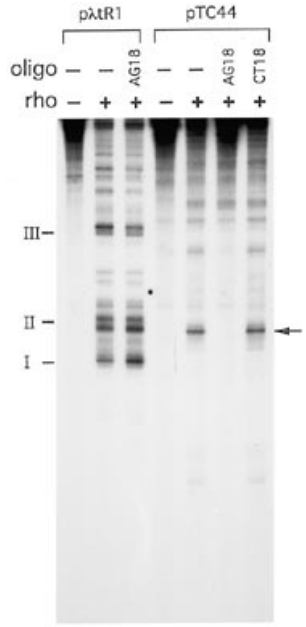

Figure 5. Effect of AG18 and CT18 oligonucleotides on *in vitro* Rho termination of pTC44. pλtR1 and pTC44 were transcribed *in vitro* in the presence (+) or in the absence (–) of Rho protein. AG18, CT18 and AG18/CT18 duplex oligonucleotides were added just prior to transcription elongation. The arrow indicates the major Rho-dependent termination product of pTC44 and I, II, III the Rho-terminated RNAs of pλtR1.

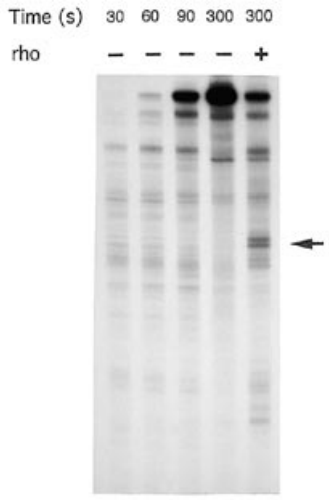

Figure 6. Determination of *in vitro* transcription pause sites in pTC44 construct. A linear DNA fragment from pTC44 was transcribed in the absence (–) or presence (+) of Rho protein for the indicated time. The arrow shows the major Rho-dependent termination product generated from the supercoiled pTC44 plasmid.

$(rUrC)_n$ sequence on the transcript and the AG18 DNA oligomer affects the Rho-dependent termination process.

The Rho termination endpoints coincide with RNA polymerase pause sites

To determine whether the sites of Rho-dependent RNA release correlate with positions on the DNA template where RNA polymerase experiences extensive pausing, we analyzed the kinetics of RNA elongation in the absence of Rho. In these experiments, the *Eco*RI–*Pvu*II restriction fragment from pTC44 was used as template in order to avoid the interference of truncated RNAs originating from other promoters carried on the plasmid. Figure 6 shows the comparison of the Tac-promoted RNAs made after 30, 60, 90 and 300 s of synthesis with the transcripts that have been terminated by Rho protein. The data clearly show that the rate of RNA synthesis across the fragment is not uniform. RNA polymerase pauses at several template positions and most of the Rho-dependent termination sites coincide with pause sites of RNA polymerase. The pattern of Rho-dependent termination obtained with the restriction fragment differs slightly from the one observed with the supercoiled plasmid. The 230 nt RNA (shown by an arrow in Fig. 6) is still the major Rho-dependent termination product but the termination endpoint splits into a doublet of bands. This result suggests that the topological state of the DNA template modulates the Rho-dependent termination process.

DISCUSSION

In this work, we present *in vivo* and *in vitro* data demonstrating that an artificial Rho-dependent terminator is readily generated by the insertion of $(AG/TC)_n$ DNA repeats within a transcription unit. Transcription of the inserts in the orientation encoding the $(rUrC)_n$ repetition on the transcript leads to Rho-dependent termination within the downstream non-terminator DNA sequences. The Rho termination activity is observed with the plasmids harboring $(AG/TC)_n$ inserts of 30 and 44 bp in length and a nearly 100% termination efficiency is obtained *in vivo* with the 44 bp repeat.

The molecular mechanism by which Rho brings about transcript release at Rho-dependent termination sites is still not completely understood. It is clear, however, that the interactions between Rho protein and the nascent transcript are at the basis of the termination process (12,13). The protein first binds to a segment of the transcript (loading site) encoded by the upstream part of the terminator (*rut* element). Following cycles of ATP hydrolysis, Rho acts at a distance by causing transcript release downstream from its initial binding site. Within the framework of this model, our results suggest that the $(rUrC)_n$ repetition on the transcript serves as an efficient loading site for Rho protein.

It is generally accepted that the initial interaction of Rho and the RNA transcript is predominantly determined by the proportion of cytosine residues within this stretch of unstructured RNA (20,24). Thus, the Rho-dependent termination activity of our constructs is presumably reflecting a stable association of Rho with the $(rUrC)_n$ RNA repeat due to the large content of cytosines (50%). The binding site for hexameric Rho has been shown to involve 70–80 contiguous nucleotides of RNA (20–22). Since the size of the repeats in our constructs is smaller, the contribution of RNA flanking sequences to the initial binding cannot be excluded. However, we observed efficient Rho-dependent termination activity when the repeat was inserted within a different sequence context (in the middle of the CAT gene, in the *Pvu*II site). Also, the RNA sequences (up to 40 nt upstream and downstream) surrounding both insertion sites (*Sal*I and *Pvu*II) do not harbor any clusters of cytosines. Therefore, the $(rUrC)_n$ sequence on the transcript seems to play the major role at the initial binding. Additional support to this conclusion is provided by the oligonucleotide competition experiments. In effect, to interfere with Rho termination activity, the heteroduplex must block the initial

interaction of the protein at the loading site, otherwise it would be disrupted by the Rho helicase activity (5,37).

The high efficiency of termination obtained *in vivo* with the 44 bp insert (pTC44) cannot be explained exclusively by a tight binding of Rho at the loading site. Previous studies have shown that there is no direct correlation between the efficiency of termination and the content of cytosine residues within the upstream part of the transcript (24–26). In some instances, an increase of the proportion of cytosines even leads to a reduction of termination efficiency. It is possible that the high efficiency of termination is due to the sequence composition of the repeat. As pointed out by Wang and von Hippel (27,28), the (rUrC) RNA motif may allow faster release of RNA from individual binding sites within the Rho hexamer which could facilitate the translocation of the protein towards the transcription elongation complex. Two models of the translocation process have been proposed. In the simple tracking model, the mechanism involves only reactions of RNA binding and release to individual subunits of the Rho hexamer (17). The tethered tracking model, on the other hand, assumes that certain subunits of the Rho hexamer remain bound to the initial loading site while the enzyme moves towards the 3′ end of the RNA (13,18). The experiments reported here cannot distinguish between these two models but they set a lower limit on the size of the putative primary RNA binding site of the tethered tracking model. An $(rUrC)_n$ repeat of 30 nt is not sufficient to maximally stimulate the termination activity of Rho, but a 44 nt stretch is fully functional. Since the $(rUrC)_n$ repeat clearly dictates the efficiency of Rho-dependent termination of transcription in our system, we can exploit this property in future experiments to assess the influence of the sequences at the entry or tethering site.

The Rho-dependent transcript release in our pTC44 construct occurs predominantly at a site located 130 bp downstream from the repeat, both *in vivo* and *in vitro*. The sequence of this site is rich in uridine residues (5′-UUUUG). A similar sequence preference was observed at certain release sites within natural Rho-dependent terminators, such as tI and tII of λtR1. We have considered the possibility that this release site is related to a neighboring structural motif within the transcript. However, sequence analysis with the Mulfold computer program (38) of the region upstream from the site did not reveal any obvious stem–loop structure. We have also analyzed the kinetics of RNA elongation to determine whether the release site corresponds to a natural pause site of RNA polymerase. Our experiments on linear DNA fragments confirm the well-known relation between RNA polymerase pausing and Rho-dependent RNA release (39–41). However, the results indicate that the efficiency of release at a given site is not directly correlated with the extent of pausing. In particular, we found that the high efficiency of release at the site located 130 bp downstream from the repeat is not the result of a strong transcriptional pausing.

Finally, we note that Hart and Roberts (24,25) have previously reported that efficient Rho-dependent terminators can be obtained by replacing a part of the upstream natural region of λtR1 terminator with relatively long (70–140 bp) DNA repeats encoding (rArC) rich transcripts. Indeed, our results extend these previous studies by showing that the presence of 30–44 nt of the $(rUrC)_n$ repetition on the transcript is sufficient to induce Rho-dependent transcription termination within non-terminator DNA context. Therefore, our findings provide for the first time clear evidence that an adequate loading site is the primary determinant for Rho termination activity. We believe that these artificial terminators will be very useful for future investigations aimed at elucidating the mechanism of Rho-dependent transcription termination.

ACKNOWLEDGEMENTS

We are highly grateful to Dr Marc Leng for continuous encouragement, support and advice. We thank Professor P. von Hippel for the gift of purified Rho protein and Dr L. Bossi for providing the *S.typhimurium* strains. This work was supported in part by la Ligue contre le Cancer, l'ANRS and E.E.C. (Project no. BMH4-CT97-2485).

REFERENCES

1 Richardson,J.P. and Greenblatt,J. (1996) In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2nd ed. ASM Press, Washington, DC, pp. 822–848.
2 Chen,C.-Y.A. and Richardson,J.P. (1987) *J. Biol. Chem.*, **262**, 11292–11299.
3 Faus,I. and Richardson,J.P. (1989) *Biochemistry*, **28**, 3510–3517.
4 Faus,I. and Richardson,J.P. (1990) *J. Mol. Biol.*, **212**, 53–66.
5 Chen,C.-Y., Gallupi,G.R. and Richardson,J.P. (1986) *Cell*, **46**, 1023–1028.
6 Lau,L.F., Roberts,J.W. and Wu,R. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 6171–6175.
7 Morgan,W.D., Bear,D.G. and von Hippel,P.H. (1983) *J. Biol. Chem.*, **258**, 9553–9564.
8 Lau,L.F., Roberts,J.W. and Wu,R. (1983) *J. Biol. Chem.*, **258**, 9391–9397.
9 Morgan,W.D., Bear,D.G. and von Hippel,P.H. (1983) *J. Biol. Chem.*, **258**, 9565–9574.
10 Geiselmann,J., Seifried,S.E., Yager,T.D., Liang,C. and von Hippel,P.H. (1992) *Biochemistry*, **31**, 121–132.
11 Richardson,J.P. (1996) *J. Biol. Chem.*, **271**, 1251–1254.
12 Platt,T. and Richardson,J.P. (1992) In McKnight,S.L. and Yamamoto,K.R. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 365–388.
13 Platt,T. (1994) *Mol. Microbiol.*, **11**, 983–990.
14 Sharp,J.A. and Platt,T. (1984) *J. Biol. Chem.*, **259**, 2268–2273.
15 Lowery,C. and Richardson,J.P. (1977) *J. Biol. Chem.*, **252**, 1381–1385.
16 Galluppi,G.R. and Richardson,J.P. (1980) *J. Mol. Biol.*, **138**, 513–539.
17 Geiselmann,J., Wang,Y., Seifried,S.E. and von Hippel,P.H. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 7754–7758.
18 Steinmetz,E.J. and Platt,T. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 1401–1405.
19 Brennan,C.A., Dombroski,A.J. and Platt,T. (1987) *Cell*, **48**, 945–952.
20 Alifano,P., Rivellini,F., Limauro,D., Bruni,C.B. and Carlomagno,M.S. (1991) *Cell*, **64**, 553–563.
21 Morgan,W.D., Bear,D.G., Litchman,B.L. and von Hippel,P.H. (1985) *Nucleic Acids Res.*, **13**, 3739–3754.
22 Bear,D.G., Hicks,P.S., Escudero,K.W., Andrews,C.L., McSwiggen,J.A. and von Hippel,P.H. (1988) *J. Mol. Biol.*, **199**, 623–635.
23 Richardson,L.V. and Richardson,J.P. (1992) *Nucleic Acids Res.*, **20**, 5383–5387.
24 Hart,C.M. and Roberts,J.W. (1991) *J. Biol. Chem.*, **266**, 24140–24148.
25 Hart,C.M. and Roberts,J.W. (1994) *J. Mol. Biol.*, **237**, 255–265.
26 Zalatan,F. and Platt,T. (1992) *J. Biol. Chem.*, **267**, 19082–19088.
27 Wang,Y. and von Hippel,P.H. (1993) *J. Biol. Chem.*, **268**, 13940–13946.
28 Wang,Y. and von Hippel,P.H. (1993) *J. Biol. Chem.*, **268**, 13947–13955.
29 Brosius,J. (1984) *Gene*, **27**, 151–160.
30 Pulleyblank,D.E., Haniford,D.B. and Morgan,A.R. (1985) *Cell*, **42**, 271–280.
31 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
32 Ciampi,M.S., Schmid,M.B. and Roth,J.R. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 5016–5020.
33 Rahmouni,A.R. and Wells,R.D. (1992) *J. Mol. Biol.*, **223**, 131–144.
34 Guérin,M., Leng,M. and Rahmouni,A.R. (1996) *EMBO J.*, **15**, 5397–5407.
35 Housley,P.R., Leavitt,A.D. and Whitfield,H.J. (1981) *J. Bacteriol.*, **147**, 13–24.
36 Housley,P.R. and Whitfield,H.J. (1982) *J. Biol. Chem.*, **257**, 2569–2577.
37 Steinmetz,E.J., Brennan,C.A. and Platt,T. (1990) *J. Biol. Chem.*, **265**, 18408–18413.
38 Jaeger,J.A., Turner,D.H. and Zucker,M. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 7706–7710.
39 von Hippel,P.H. and Yager,T.D. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2307–2311.
40 Jin,D.J., Burgess,R.R., Richardson,J.P. and Gross,C.A. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 1453–1457.
41 Richardson,L.V. and Richardson,J.P. (1996) *J. Biol. Chem.*, **271**, 21597–21603.