4690–4697 Nucleic Acids Research, 1995, Vol. 23, No. 22

Transcription termination at the *Escherichia coli thra* terminator by spinach chloroplast RNA polymerase *in vitro* is influenced by downstream DNA sequences

Liang-Jwu Chen*, Yuh-Jin Liang, Shih-Tong Jeng[1], Emil M. Orozco[2], Richard I. Gumport[3], Chi-Hui Lin and Ming-Te Yang

Institute of Molecular Biology, National Chung Hsing University, Taichung 40227, Taiwan, Republic of China, [1]Department of Botany, National Taiwan University, Taipei, Taiwan, Republic of China, [2]DEKALB Plant Genetics, 62 Maritime Drive, Mystic, CT 06355, USA and [3]Department of Biochemistry, College of Medicine and School of Chemical Sciences, University of Illinois, Urbana, IL 61801, USA

Received July 24, 1995; Revised and Accepted October 20, 1995

ABSTRACT

We have investigated the mechanism of transcription termination *in vitro* by spinach chloroplast RNA polymerase using templates encoding variants of the transcription-termination structure (attenuator) of the regulatory region of the threonine (*thr*) operon of *Escherichia coli*. Fourteen sequence variants located within its d(G+C) stem–loop and d(A+T)-rich regions were studied. We found that the helix integrity in the stem–loop structure is necessary for termination but that its stability is not directly correlated with termination efficiency. The sequence of the G+C stem–loop itself also influences termination. Moreover, the dA template stretch at the 3′ end of the terminator plays a major role in termination efficiency, but base pairing between the A and U tract of the transcript does not. From the studies using deletion variants and a series of mutants that alter the sequences immediately downstream from the transcription termination site, we found that termination of transcription by spinach chloroplast RNA polymerase was also modulated by downstream DNA sequences in a sequence-specific manner. The second base immediately following the poly(T) tract is crucial for determining the termination efficiency by chloroplast RNA polymerase, but not of the T7 or *E.coli* enzymes.

INTRODUCTION

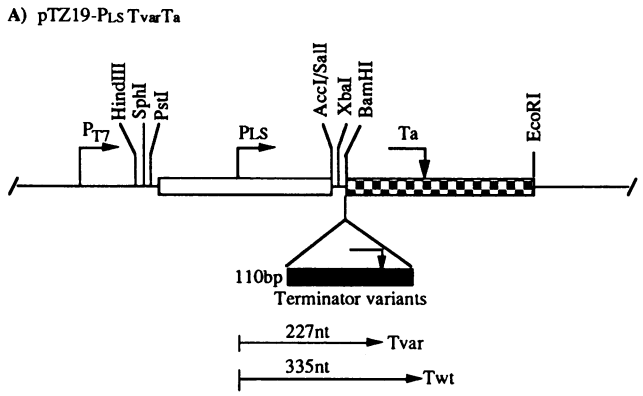
Most plastid genes contain inverted repeat sequences at their 3′ ends that are thought to confer upon their transcripts the ability to form stem–loop structures. Such structures were initially presumed to be transcription terminators, analogous to bacterial factor-independent terminators. However, when they were analyzed in a homologous *in vitro* transcription system, no significant termination activity was observed (1,2). It was subsequently concluded that the 3′ ends of the *rbcL, atpB, psbA, petD* and *rpoA* transcripts from spinach chloroplast are primarily the result of RNA processing and not transcription termination (2,3). In contrast, a recent report indicates that the stem–loop structure of the 3′ ends of the *Chlamydomonas* chloroplast *rbcL* and *psaB* genes appear to have significant termination activity *in vivo* (4). In addition, the prokaryotic factor-independent terminators that contain a typical d(G+C)-rich region of dyad symmetry and are followed by a poly(T) tract in the sense strand, such as the terminator for the threonine attenuator, *thra* from *Escherichia coli*, efficiently terminate transcription *in vitro* by spinach chloroplast RNA polymerase (1,5). This indicates that the transcription termination process in chloroplasts may have at least some features in common with the mechanism used in prokaryotes and that the stem–loop-dependent termination is probably sometimes functional in chloroplasts.

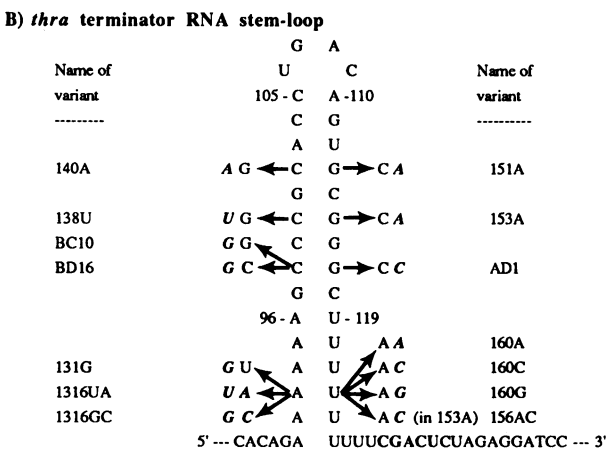
The mechanism of transcription termination for *E.coli* factor-independent terminators has been well characterized. A classical thermodynamic model (6,7) for factor-independent terminators proposes that termination occurs after the G+C-rich region forms a stable base-paired stem–loop structure in the RNA. This causes the polymerase to pause and release the elongating transcript. Release may also be enhanced by the instability of the dA-rU base pairs that often comprise the remainder of the transcript/template hybrid (8,9). Many studies involving base analogues (10,11) heteroduplexes (12) and mutant terminators for *E.coli thra* terminator (13–16) support this model. Transcription termination by the *E.coli thra* terminator with T7 RNA polymerase indicates that the dyad symmetry, the poly(T) tract and many other sequence elements also affect termination (13,14). However, studies indicating that the sequences downstream from the termination site affect termination efficiency (17,18) can not be explained by a simple thermodynamic model. Recent findings about the cleavage of the nascent RNA in the transcriptional complex (19–21), the existence of RNA binding sites on RNA polymerase (19) and the observation of discontinuous movements of DNA and RNA with respect to RNA polymerase during elongation (22–25) and pausing (26) led to an alternative model of transcription termination (27). This inchworming model describes transcription termination as coupling pulsate RNA polymerase movement, and the poly(T) tract that follows the stem–loop structure acts as an inchworming signal (27).

Since the *E.coli thra* terminator is functionally recognized by spinach chloroplast RNA polymerase, the study of the *thra* terminator variants in a transcription system using chloroplast

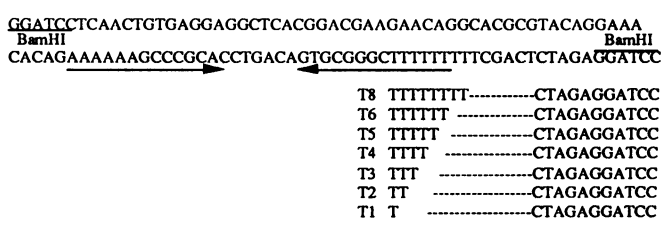
* To whom correspondence should be addressed

RNA polymerase should help us further understand transcription termination in chloroplasts. In this work, we studied the effect of a series of mutations in the *thra* terminator on transcription termination with spinach chloroplast RNA polymerase. Included in this study were 14 mutants with single base substitutions in the d(G+C)-rich and d(A+T)-rich regions and seven nested deletion variants in the poly(T) tract region. Many other variants in the downstream region of the *thra* terminator were also used. The results indicate that the specific sequence as well as the helix stability within the d(G+C)-rich region, and the length of the poly(T) tract affect termination efficiency. In addition, we made the unexpected observation that untranscribed sequence(s) downstream of the *thra* terminator play a significant role in termination efficiency with spinach chloroplast RNA polymerase.

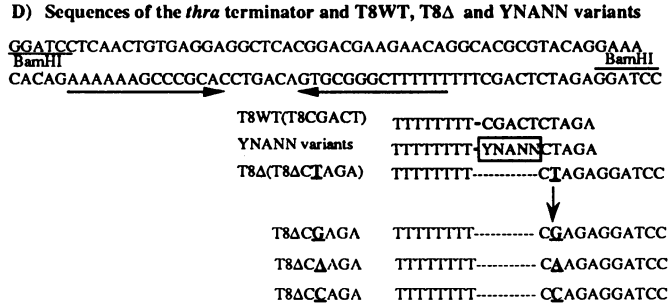




C) Sequences of the 110bp *thra* terminator and its variants



D) Sequences of the *thra* terminator and T8WT, T8Δ and YNANN variants



MATERIALS AND METHODS

Plasmid DNA constructions

The plasmid pTZ19-PLSTa contains the 5′-end of the spinach *rbcL* gene promoter (PLS) and the terminator from the threonine attenuator region (Ta) as previously described (1). The terminator variants were originally from the pTZ-19*thr* variants and pTZ-19T series as described by Jeng *et al.* (13). The 110 bp *Bam*HI DNA fragments containing the various terminator mutants were purified and ligated to the *Bam*HI site of plasmid pTZ19-PLSTa. The orientation of the *Bam*HI DNA fragments in each clone was confirmed. For all of the nested deletion variants, the sequence immediately following the template deoxyadenosine tract had the 5 deoxynucleotides CGACT deleted. In these nested deletion variants, the variant contains an 8 T tract was designated T8Δ. Other similar variants that contain wild-type downstream sequences were designated T8WT. A schematic diagram for the construction of terminator is presented in Figure 1 and their characteristics are summarized in Table 1.

Besides the terminator variants that have sequences changed in d(G+C)-rich hairpin and d(A+T)-rich regions, a series of mutants that have sequences changed downstream of the termination site were constructed. These downstream sequence mutants were created by PCR with a degenerate primer YNANN (Fig. 1D, Y = C or T, N = G, A, C or T) and a reverse primer, the variant T8WT was used as template for DNA fragment amplification. The YNANN primer will theoretically create 128 mutants. The 110 bp DNA fragments that may contain all the 128 different downstream sequence combinations were cut out from the PCR amplified products with *Bam*HI and then cloned into the *Bam*HI site of pTZ-PLSTa. The different downstream sequence combinations in each isolated clone was identified by sequencing. The other set of mutants T8ΔCGAGA, T8ΔCCAGA and T8ΔCAAGA, were also constructed by PCR with a primer that has the nucleotides G, A or C mixed at the position indicated; the variant T8Δ (also designated as T8ΔCTAGA) was used as template.

Figure 1. Schematic diagrams of plasmid DNAs pTZ19-PLSTvarTa and structure of *thra* terminator RNA stem–loop and its variants. (**A**) The terminator variants were constructed by inserting the various 110 bp *Bam*HI DNA fragments from pTZ-19T series, pTZ-19thr variants (13) and YNANN mutants into the *Bam*HI site of plasmid pTZ19-PLSTa (1). The sizes for transcripts initiating at the 'PLS' promoter and terminated at the terminator variants (227 nt) or at the wild-type terminator (335 nt) were indicated. (**B**) The secondary structure of the wild-type *thra* terminator is presented as the conformation that maximizes potential base pairing. The RNA numbering system is same as in Table 2 and the nucleotide positions are relative to the transcription initiation site of T7 RNA polymerase as described previously (13). The changed bases in each variant are indicated as italics and shown in Table 2. (**C**) The nested deletion variants with shortened d(T) tract are shown. Five nucleotides located downstream from termination release sites were deleted as indicated by dash line. (**D**) The relative location of sequences changed for YNANN mutants is shown. A degenerate primer contain mixed nucleotides are indicated by the letter Y = C or T and N = G, A, T or C and the variant T8WT was used as DNA template. The single base change for terminator variants T8ΔCGAGA, T8ΔCAAGA and T8ΔCCAGA are shown.

Table 1. Vectors and the *thra* terminator variants used in this study

Vectors or variants	Characteristics	Reference or source
TZ19T series	Contains the nested deletion variants	(13)
pTZ19 *thr* variants	Contains the 14 variants in G+C rich and A+T rich regions	(13)
pTZ19PLSTa	Contains a chloroplast promoter and the wild-type *thra* terminator	(1)
pTZ19PLSTvarTa	All *thra* terminator variants in this study are cloned between the chloroplast promoter and the wild-type *thra* terminator	This work
YNANN variants	Contains the *thra* terminator variants with various downstream sequences combinations	This work
T8ΔCNAGA variants	Contains four variants that each has only one base changed at the position indicated	This work

Analysis of *in vitro* transcripts

Transcriptionally active extracts of spinach chloroplasts were prepared as previously described (1). Each standard *in vitro* transcription reaction (40 µl) contained 12 mM Hepes–KOH (pH 8), 40 mM KCl, 10 mM $MgCl_2$, 10 mM DTT, 1 U/µl RNase inhibitor, 500 µM ATP, UTP and CTP and 10 µM GTP containing 10 µCi [α-^{32}P]GTP (800 Ci/mmol, 10 µCi/µl; NEN, Dupont) and the following components: one-tenth of the volume (4 µl) was supercoiled or linearized DNA (0.4 µg) in 10 mM Tris–HCl (pH 8), 0.1 mM EDTA and two-fifths of the volume (16 µl) was chloroplast RNA polymerase high-salt extract in DEAE buffer [50 mM Tricine–KOH (pH 8), 50 mM KCl, 0.5 mM EDTA, 0.5 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM benzamidine, 5 mM ε-amino-*n*-caproic acid, 5% (v/v) glycerol] or 12.5 U of T7 RNA polymerase (in DEAE buffer). The DEAE buffer contributed an additional 20 mM of KCl to the final reaction volume. The *in vitro* transcription reaction was incubated at 30°C (for chloroplast RNA polymerase) or at 37°C (for T7 RNA polymerases) for 60 min. The reaction was then stopped by the addition of 40 µl of RNA extraction buffer [6 M urea, 0.36 M NaCl, 20 mM EDTA, 10 mM Tris–HCl (pH 8) and 1% SDS]. *E.coli* tRNA (15 µg/reaction) was added as carrier and the *in vitro* RNA extracted twice with phenol:chloroform and ethanol-precipitated. The nucleic acid pellet was rinsed once with cold 80% ethanol, dried and resuspended in 80% formamide, 0.1% bromophenol blue and 0.1% xylene cyanol. The *in vitro* transcription products were analyzed by electrophoresis through 8% polyacrylamide–8.3 M urea DNA sequencing gels and subsequent autoradiography.

Quantification of termination efficiency

The autoradiograms were scanned with an LKB Ultroscan XL Laser Densitometer or the ^{32}P in RNA-containing gel slices was measured by liquid scintillation counting. Densitometry of multiple autoradiographic film exposures was performed to ensure that the autoradiographic signal was within the linear response range of the film. Each dataset was corrected for background and normalized according to the length and guanosine composition of each transcript. The data from multiple experiments were expressed as the average value ± one standard error. The number of repetitions for each experiment is indicated in the tables. The termination efficiency for each variant was calculated essentially as described by Chen (28). After the absolute efficiency of termination for each variant was obtained, the termination efficiency for each variant relative to the wild-type was calculated (Table 2).

RESULTS

Transcription *in vitro* with *thra* terminator variants by spinach chloroplast RNA polymerase

We used 21 *thra* terminator variants, including point substitutions in the d(G+C) rich region of dyad symmetry, point substitutions in the d(A+T) rich region (Fig. 1B) and a set of nested deletions in the poly(T) tract (Fig. 1C) of the *thra* terminator to study transcription termination by spinach chloroplast RNA polymerase. Each *thra* terminator variant, consisting of a 110 bp *Bam*HI DNA fragment, was cloned into the *Bam*HI site of plasmid pTZ19-PLSTa to get the plasmids designated pTZ19-PLSTvarTa (Fig. 1A). Thus, each construct contained two terminator sequences: (i) an altered *thra* terminator located at the *Bam*HI site, and (ii) immediately downstream, a wild-type *thra* terminator (from the plasmid vector pTZ19-PLSTa) to serve as an internal control (Fig. 1A). The CsCl-prepared supercoiled plasmid DNAs were used as DNA templates. Transcription reactions and their products were analyzed as described in Materials and Methods and the results are shown in Figure 2. All *thra* terminator variants, no matter where the mutation was located, decreased the amount of termination when compared with the wild-type terminator (Fig. 2, lanes 2 and 3–16). Their relative efficiencies of transcription termination by spinach chloroplast RNA polymerase and T7 RNA polymerase are presented in Table 2.

The wild-type *thra*-terminated RNAs (Twt) synthesized from templates encoding the terminator variants in the d(A+T)-rich region (131G, 1316UA, 160A, 160C and 160G) and variant BD16, which does not create a mismatch in the G+C region of the transcript, migrated at the same rate as the transcript formed by the wild-type template (Fig. 2, lane 2 versus lanes 3–8). The RNAs synthesized from variant 1316GC, which created one potential GC base pair in the A+T-rich region and accordingly increased the calculated helix stability, migrated more slowly than the wild-type transcripts (Fig. 2, lane 9 versus lanes 2–8). Conversely, the RNAs synthesized from the terminator variants that caused mismatches in the G+C-rich hairpin region migrated more rapidly than the transcript formed by wild-type template (Fig. 2, lanes 10–16). These mobility anomalies have been attributed to differences in stability of the RNA helices (13,16).

Table 2. Relative termination efficiencies of *thra* terminator variants

Template	Positions altered[a]	Base pair change	ΔG kcal/mol[b]	Spinach %[c]	T7 %[d]
wild-type			-23	100[e]	100[f]
131G	93	AU → GU	-22	77 ± 5	100
153A	115	CG → CA	-16	48 ± 5	30
151A	113	CG → CA	-17	43 ± 2	16
BD16	98;117	CG → GC	-23	29 ± 3	42
138U	100	CG → UG	-21	28 ± 2	33
140A	102	CG→ AG	-17	26 ± 3	14
1316UA	93;122	AU → UA	-23	20 ± 4	44
160A	122	AU → AA	-21	18 ± 2	47
1316GC	93;122	AU → GC	-25	15 ± 3	30
160C	122	AU → AC	-21	13 ± 3	31
160G	122	AU → AG	-21	12 ± 3	30
AD1	117	CG → CC	-16	10 ± 4	9
BC10	98	CG → GG	-16	<10	12
156AC	115	CG → CA			
	122	AU → AC	-14	<10	9

[a]The positions correspond to those shown in Figure 1.
[b]ΔG values were calculated according to Freier *et al.* (29).
[c]Transcription termination was calculated as described in the Materials and methods. Each data represents the average of three individual experiments.
[d]These transcription termination data were derived from Jeng *et al.* (13).
[e]The absolute termination efficiency is 63% (28).
[f]The absolute termination efficiency is 43% according to Jeng *et al.* (13).

Our results confirm that calculated stability was well correlated with relative mobility of the terminator variant transcripts (Fig. 2 and Table 2). The anomalous migrations we observed were obtained for the transcripts produced by terminating at the Twt. The transcripts that terminated at the variant terminators (Tvar) migrated normally. In contrast, others have reported that the anomalous migrations were observed in the smaller transcripts due to the terminator variants, while the longer run-off transcripts migrated normally (13,16). The presence of tandem hairpins, due to the two terminators, may contribute to the different conformations observed in this study.

Termination with variants at the d(G+C)-rich hairpin region

Seven variants (138U, 153A, 140A, 151A, BD16, BC10 and AD1) in the d(G+C)-rich region were studied. The efficiency of termination from these variants ranged from <10 to 48% of the wild-type terminator. Among these variants, BD16, BC10 and AD1 existed in a slightly different sequence context than the remainder of the constructions. In these three variants the template poly(T) tract was shortened from 9 nucleotides (nt), as it in the wild-type *thra* terminator, to 8 nt. The termination efficiencies of two of these variants, BC10 (<10%) and AD1 (10%), were substantially less than those of other mismatch variants 140A (26%), 153A (48%) and 151A (43%). It is plausible that the decreased termination efficiencies in these variants were due to both the mismatches in d(G+C)-rich region and a one-T shortening of poly(T) tract length. However, efficiencies of termination by both *E.coli* and T7 RNA polymerase (13,16) and spinach chloroplast RNA polymerase (Fig. 3, lanes T9WT versus T8WT, and Table 3) were unaffected by changes in poly(T) template from 9-T to 8-T. Hence, the low termination efficiencies for BC10 and AD1 variants are more likely due to the changes in the d(G+C)-rich hairpin and are affected only slightly by the length of the poly(T) tract. The mismatch variants (140A, 153A, 151A, BC10 and AD1) destablize the helix and decrease termination, suggesting that helix integrity in the stem of the stem–loop structure is necessary for efficient termination.

Table 3. The termination efficiency of the *thra* terminator variants transcribed with chloroplast or T7 RNA polymerase

Variants	Composition of downstream sequences	Termination efficiency %[a] Chloroplast	T7
T9WT	T9CGACT	62.6 ± 4.0	64.4 ± 2.4
T8WT	T8CGACT	61.8 ± 4.5	63.3 ± 1.6
T8ΔCTAGA	T8CTAGA	9.2 ± 1.1	63.2 ± 1.9
T8ΔCGAGA	T8CGAGA	54.4 ± 3.7	65.1 ± 1.6
T8ΔCAAGA	T8CAAGA	8.4 ± 1.3	62.3 ± 0.8
T8ΔCCAGA	T8CCAGA	11.7 ± 2.9	58.9 ± 2.0

[a]Transcription termination was calculated as described in the Materials and Methods. Each data represents the average of three individual experiments.

To study the relationship between the termination efficiency and helix integrity, stability values were calculated by the method of Freier *et al.* (29). These calculations indicated that although the variants 138U (stability = –21 kcal/mol) and 140A (–17 kcal/mol)

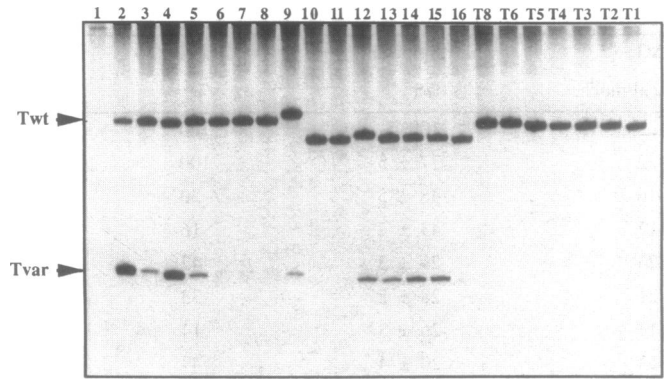


Figure 2. *In vitro* transcription of the *thra* stem–loop terminator variants and the poly (T) tract shortened variants with chloroplast RNA polymerase. The transcripts were prepared and analyzed as described in Materials and Methods. The transcripts terminated at the terminator variants (Tvar) and at the wild-type terminator (Twt) are indicated. Lane 1, no DNA; lane 2, wild-type (T9WT) terminator located at the Tvar site; lane 3, 1316UA; lane 4, 131G; lane 5, BD16; lane 6, 160A; lane 7, 160C; lane 8, 160G; lane 9, 1316GC; lane 10, 156AC; lane 11, BC10; lane 12, 138U; lane 13, 140A; lane 14, 151A; lane 15, 153A; lane 16, AD1; lanes T8, T6, T5, T4, T3, T2, T1 are 8, 6, 5, 4, 3, 2, 1 deoxyadenosine variants, respectively.

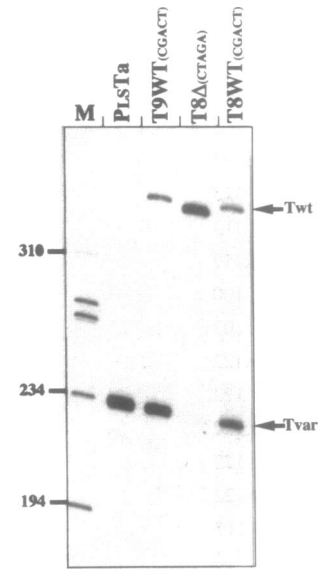


Figure 3. *In vitro* transcription of the *thra* terminator variants T9WT, T8Δ and T8WT with chloroplast RNA polymerase. Lane M: ϕX174–*Hae*III molecular size standards. The lengths of denatured DNA size standards are shown as well as the positions of the transcripts that terminated at the terminator variants (Tvar) and at the wild-type terminator (Twt). The transcript in lane PLSTa, which contain only a wild-type terminator as described in Chen and Orozco (1), is the transcript that terminates at the wild-type terminator.

are thermodynamically stable, they have low termination efficiencies, whereas the variants 151A (–17 kcal/mol) and 153A (–16 kcal/mol) with similar or lower stability, have much higher termination efficiency (Table 2). These finding indicated that there is no direct correlation between the decrease in helix stability and the decrease in termination efficiency. In addition, the variants 138U and 140A, which have their sequences altered at positions #100 and #102, respectively, had lower termination efficiencies than the variants 151A and 153A with the same mismatch base pairing but with their sequences altered at positions #113 and #115 on the opposite side of the stem structure. This finding indicated that not only the mismatch but the sequence itself influenced termination. Consistent with this interpretation, the variant BD16, that forms an inverted GC base pair and has the same calculated stability value (–23 kcal/mol) as the wild-type, had a termination efficiency of only 29% of the wild-type value.

Termination with variants at the d(A+T)-rich region

Seven variants that have sequences altered in both the runs of (dA) and (T) residues were used to determine the effects of changes in the sequence of the d(A+T) rich region of the *thra* terminator (Fig. 1B). Three variants (160A, 160C or 160G), which disrupted the T tract (U tract in transcript) by substituting A, C or G for U at position 122, decreased termination to ≤ 18% of the wild-type value (Table 2). Two variants, 1316UA and 1316GC, which change both the A and U stretches at positions 93 and 122, respectively, allow base pairing at this site; they nevertheless decrease termination efficiencies. The termination efficiency of 1316UA (20%) is nearly the same as that of 160A (18%), which contains a substitution of A with U at position 122 that creates a mismatch. Similarly, variant 160C, which disrupts the poly-U tract and creates a mismatch, terminates with the same efficiency as does the variant 1316GC. In addition, the variant 131G that changes A to G at position 93 in the poly-A tract has a minor effect on termination. These results indicate that the poly-U stretch plays a major role in termination efficiency and that the base pairing between the A and U tract of the transcript does not play a crucial role in termination.

Another variant (156AC), which contains two AC mismatches, one in the G+C-rich region (as in 153A) and one in the A+T-rich region (as in 160C), decreases termination to barely detectable levels suggesting that the two base changes may independently contribute to disrupting termination.

Effects of the length of T tract and of the variants T8Δ and T8WT

Lynn *et al.* (16) reported that the length of the deoxyadenosine tract of the template, encoding the stretch of uridines at the 3′ end of the transcript, affects the termination efficiency with *E.coli* RNA polymerase. A series of nested deletion variants that successively shorten the deoxyadenosine tract in the terminator template (13 and Fig. 1C) were used to test their function with chloroplast RNA polymerase. The nested deletion variants exist in a slightly different sequence context than does the wild-type construction. In all of the nested deletion variants, the sequence of 5 deoxynucleotides that immediately following the template deoxyadenosine tract in the wild-type, d(CGACT), were deleted as a result of the cloning procedures (13). In this study, we have cloned these nested deletion variants of the *thra* terminator in a series of plasmids that also contains the wild-type *thra* terminator immediately downstream as the internal control (Fig. 1A). The results of successively shortening the run of deoxyadenosines in the terminator template are shown in Figure 2 (lanes T8–T1). There was only a background level of termination at Tvar for variant T8, and no detectable signal for any of the other shortened

T variants, while the wild-type *thra* terminator at Twt terminated transcription normally (Fig. 2, bands Twt). Although these data might be explained by extreme sensitivity of the chloroplast enzyme to a shortened T tract, the variant BD16, which also contained a shortened 8-deoxyadenosine tract, had a termination efficiency of ~29% of the wild-type value. BD16 differed from the deletion mutants, however, in that its sequence downstream of the poly-T was intact. Taken together, the loss of termination function in these nested deletion variants with chloroplast RNA polymerase could be attributed to the combination of changes in the 5 deoxynucleotides immediately downstream from the transcription termination site as well as the shortened T tract.

To test how the downstream sequences affect the termination efficiency, a variant T8WT, which contain the wild-type downstream sequences and 8 deoxyadenosine tract, was used along with a T8 variant, one of the nested deletion variants which contained 8 deoxyadenosine tract and has 5 deoxynucleotides CGACT deleted (we have renamed this variant T8Δ for easy comparison; Fig. 1D). These two variants contain an identical structure in the d(G+C)-rich region and the T tract. The only difference between them is that in T8Δ the sequence CGACT has been deleted resulting in the T8 tract being adjacent to the sequence CTAGA (Fig. 1D). T8WT had about the same efficiency of termination as T9WT (Fig. 3 and Table 3), whereas no detectable termination was observed for T8Δ. The wild-type *thra* terminator, that acts as an internal control, functioned normally (Fig. 3). Therefore, changes of deoxynucleotides downstream from the transcription termination site were probably responsible for the decrease in the termination efficiency between T8Δ and T8WT. Such changes can also explain why the chloroplast enzyme failed to terminate at Tvar in the nested deletion variants which lacked the 5-nt downstream sequence (Fig. 2, lanes T8–T1).

In Figure 3, the transcripts that terminate at the wild-type *thra* terminator (Twt) with T8Δ template migrated more rapidly and were more abundant than the transcript terminated at templates T9WT and T8WT. The smaller product could be due to the 5 nt deleted in the Tvar region, and its greater abundance was likely due to the lack of any transcript termination at the Tvar terminator, causing more transcripts to terminate at Twt.

Downstream sequence affects termination in a sequence-specific manner

The 3′ ends of the transcripts produced by chloroplast RNA polymerase at the *thra* terminator corresponded to the 7th or 8th deoxyadenosines downstream with respect to the d(G+C)-rich region (1,31). This site of termination was essentially the same as that produced by *E.coli* RNA polymerase. When the sequences downstream from termination site in the wild-type (T8WT) and the deletion variant (T8Δ) are compared, only 3 out of 5 nt are different (CGACT versus CTAGA). Thus, it becomes of interest to investigate further how these downstream sequences affect the termination efficiency. To address this question, we introduced several mutations within these 5 nt by PCR with a degenerate primer YNANN (Fig. 1D, Y = C or T, N = G, A, C or T). DNA fragments amplified with PCR were isolated and cloned into the *Bam*HI site of pTZ-PLSTa as described in Materials and Methods. More than 80 transformants that contained various sequences in this 5-nt region were identified and tested *in vitro* for transcription termination by chloroplast RNA polymerase.

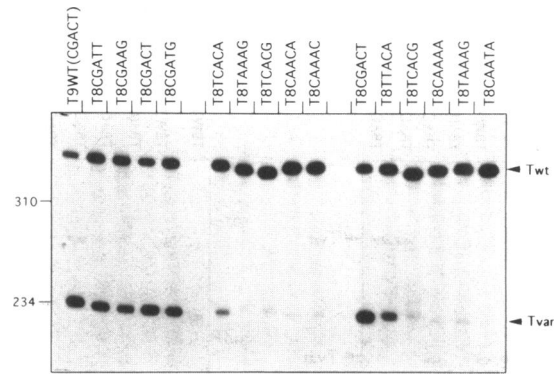


Figure 4. *In vitro* transcription of the *thra* terminator YNANN variants with chloroplast RNA polymerase. Transcription results of some of the YNANN variants are shown. The sequences composition for each YNANN variants is shown on top of each lane. Other designations, see the legend to Figure 3.

Examples of the transcription results are shown in Figure 4. These results indicate that when G appeared in the second nucleotide of this region, transcription terminated efficiently, whereas any other of the 3 nt (A, C, T) in the second position failed to terminate (Fig. 4). A mutant that contained a 10-T tract (2 additional Ts) terminated to some extent, but less efficiently than CGNNN-containing templates. In contrast, mutants having those variations in the 4th or 5th nucleotide contributed no significant effect on termination under the standard reaction conditions.

Changing T to G in T8Δ restores the termination efficiency to near wild-type value

To demonstrate further that the nucleotide located in the second position after the poly(T) tract plays an important role in transcription termination, we constructed three variants designated T8ΔCGAGA, T8ΔCAAGA and T8ΔCCAGA in addition to the original variant, T8ΔCTAGA (Fig. 1D). Their sequences were confirmed to be identical except at the nucleotide underlined (Fig. 1D). These three variants along with T8ΔCTAGA (the same as T8Δ) were subjected to *in vitro* transcription by chloroplast RNA polymerase (Fig. 5A) or T7 RNA polymerase (Fig. 5B). The results (Fig. 5A) indicated that with chloroplast enzyme the variant T8ΔCGAGA had a termination efficiency nearly the same as the wild-type construct, whereas the original template clone T8ΔCTAGA and the other two variants T8ΔCAAGA and T8ΔCCAGA had very inefficient or undetectable termination efficiency. These results clearly demonstrate that the G nucleotide in the second position of the downstream sequence plays an important role in termination efficiency. Whereas T7 RNA polymerase is not influenced by the specific nucleotide present at this position (Fig. 5B and Table 3).

DISCUSSION

The transcription termination model of the factor-independent terminator for *E.coli* RNA polymerase proposed by Yager and von Hippel (6,7) required the d(G+C)-rich hairpin and the downstream poly(T) tract for efficient termination (6,8,30). Requirement of a stable G+C hairrpin has been demonstrated for *E.coli* (16,34) and T7 (13,14) RNA polymerases with various

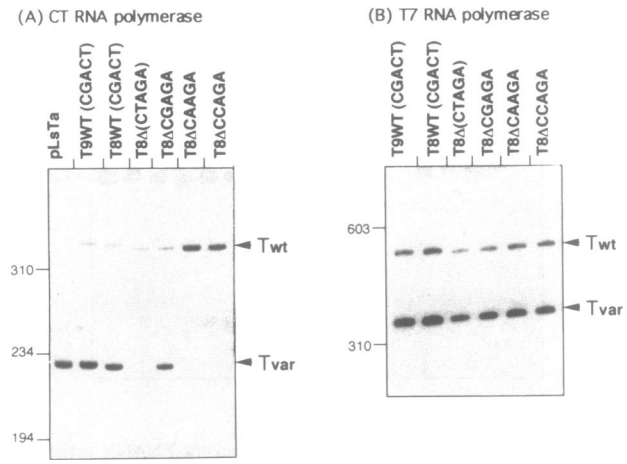


Figure 5. *In vitro* transcription of the *thra* terminator variants derived from T8ΔCNAGA with chloroplast or T7 RNA polymerase. Transcription results of T8ΔCTAGA, T8ΔCGAGA, T8ΔCAAGA, T8ΔCCAGA and T8WT, T9WT with (**A**) chloroplast or (**B**) T7 RNA polymerase are shown. The specific nucleotide changed for each variants are shown on top of each lane. Other designations, see the legend to Figure 3. For the experiment shown in (A), the template T8ΔCTAGA gave less Twt transcript than usually observed (compare with template T8ΔCTAGA in Fig. 3); however, the qualitative result, i.e., the absence of Tvar transcript, is as before.

terminator mutants that destroy the helix integrity and destabilize the hairpin structure. However, the role of the instability of the dA-rU base pairs in the transcription complex has been brought into question by the findings that functioning dA-rU runs may be as short as two base pairs (32) and that there are very strong terminators that lack the poly(T) tract (33).

In this study, using chloroplast RNA polymerase, we show that even though mutations that alter helix integrity and stability in the G+C-rich hairpin region can decrease termination efficiency, there is no consistent correlation between the termination efficiency and helix stability (Table 2). This finding is consistent with previous observations for termination by the chloroplast enzyme using different *E.coli* terminators having varying helix stabilities (1). Transcription by T7 RNA polymerase with the same *thra* terminator variants (13) used in this study and the *in vivo* study of the tR2 terminator of coliphage λ (35) demonstrated that factors in addition to helix stability are necessary for efficient termination.

The lack of direct correlation between helix stability and termination obtained in this study implies that factors other than helix stability might be involved in termination efficiency. For example, variant 138U (–21 kcal/mol) converts the C at position 100 to U to give a UG base pair and results in a termination efficiency value of 28%, whereas variant 153A (–16 kcal/mol) changes this CG base pair to a CA mismatch and surprisingly has a 48% termination efficiency. Another example with a quite different termination efficiency occurred with variants 140A (26%) and 151A (43%). Both have their sequences changed at the same base pair positions (position 102 and 113, respectively, Fig. 1B) and have the same calculated helix stability (–17 kcal/mol). These observations vividly demonstrate that the sequence itself in the G+C-rich region can also influence the termination with chloroplast enzyme. However, the results obtained with chloroplast RNA polymerase are slightly different from those reported with T7 RNA polymerase (13 and Table 2). In that study, using T7 RNA polymerase, the termination efficiency of 138U is similar to that of 153A; likewise, the termination efficiency of 140A is similar to that of 151A (Table 2). Therefore, the differences in termination in these mutants can be attributed to differences between chloroplast RNA polymerase and T7 RNA polymerase.

In addition to the G+C-rich hairpin, the roles of sequence elements of the poly(T) stretch and poly(A) tract preceding the hairpin that might affect the termination efficiency of *thra* terminator have also been reported with *E.coli* (16,34) or T7 (13,14) RNA polymerases. In transcription by T7 RNA polymerase, mismatches in the G+C-rich region (variants 140A, 151A, AD1, BC10, 138U and 153A) resulted in lower termination efficiencies than observed in the variants (160A, 160C, 160G, 1316UA and 1316GC), which contained a mismatch in the A+T-rich region (Table 2). In contrast, the results with the chloroplast enzyme reveal that three mismatch variants 140A, 151A, 153A and variant 138U in the G+C-rich region had termination efficiencies of 26–48%, whereas the variants (160A, 160C, 160G, 1316UA and 1316GC) that each has an altered nucleotide in the middle of T tract terminated transcription with efficiencies between 12 and 20% (Table 2). These observations indicate that, in general, mutations located at the T tract region affect termination efficiency by the chloroplast enzyme more than mutations located in the G+C-rich hairpin.

Studies with the nested deletion variants have been reported with *E.coli* (16) and T7 (13) RNA polymerases. The results indicated that with the T7 enzyme, templates with a (dA) tract of four or fewer fail to terminate transcription, whereas templates with a deoxyadenosine tract of ≥ 5 increase the termination efficiency from 30% (T5) to 90% (T8) of the wild-type value (13). Termination with the T8 construct is similar to that with the wild-type T9, suggesting that the sequences immediately downstream from the transcription termination site do not greatly affect termination with T7 RNA polymerase (13). This view is supported further by transcription with the T7 enzyme of variants T8ΔCGAGA, T8ΔCTAGA, T8ΔCAAGA, T8ΔCCAGA (Fig. 5B and Table 3); the YNANN mutants (data not shown); and the wild-type of T8WT and T9WT (Fig. 5B and Table 3). In contrast with *E.coli* or T7 RNA polymerase, the spinach chloroplast RNA polymerase had little or no termination with any of the shortened T tract variants (Fig. 2, lanes T8–T1). Sequences downstream from the termination release site, however, were crucial for termination efficiency. An observation with *E.coli* RNA polymerase also showed that the sequences between 3 and 7 nt downstream of the T7Te terminator release site affected the strength of that terminator (17). However, the sequence requirement for transcription termination with chloroplast RNA polymerase is much more specific than that with *E.coli* RNA polymerase. Termination efficiency with the chloroplast enzyme was largely influenced by a single nucleotide at a defined position.

Several plausible mechanisms explain how sequences downstream from the release sites could affect termination efficiency. Transcriptional pausing during transcription may play a role in inducing termination, and several reports demonstrate that certain pause sites are encoded by sequences distal to the site of the pause (36,37). Therefore, the specific downstream sequences that are responsible for the termination efficiencies in this study could be those that induce the RNA polymerase to pause and thus to enhance termination. Since transcriptional elongation requires

opening of the DNA double helix, the downstream sequences might also function by affecting the ability of the two DNA strands to be separated during elongation (17).

Levin and Chamberlin (36) have reported that transcriptional pausing can be determined by efficiency with which polymerase utilizes its nucleotide substrates. The utilization of nucleotide substrates is sequence-dependent and the sequence-dependent *Ks* (dissociation constant) differences have been estimated to vary up to 500-fold (36). The mechanism that determines the sequence-dependent *Ks* differences for pausing is unknown; however, changing nucleotide concentrations in the transcription reaction affects the termination efficiencies of various terminators with *E.coli* RNA polymerase (33). Telesnitsky and Chamberlin (17) also reported that sequences downstream of T3Te affect termination efficiency dramatically as a function of the nucleotide concentrations. Therefore, it is possible that specific downstream sequences of the *thra* terminator affect the ability of chloroplast RNA polymerase to utilize substrates efficiently when it is paused at the hairpin, and consequently alter its termination efficiency.

Based on the model for transcription termination of Nudler *et al.* (27), the downstream sequences that are covered in the transcription complex may also function as inchworming signals to induce the termination possibly in conjunction with the poly(T) tract. The specific downstream sequences that affect the termination of chloroplast RNA polymerase may also act as a site for binding to the RNA polymerase, providing the front-end signal that halts the 3′ edge of the transcription complex to facilitate inchworming and termination.

The chloroplast RNA polymerase used in this study is only partially purified and contains proteins other than those strictly required for transcription. Thus, it is possible that protein factors recognizing the downstream sequences might be involved in termination. Further investigation will be required to explore this possibility and the effects of nucleotide concentration variations.

In conclusion, this study demonstrates how changes in sequence and secondary structure can affect the way chloroplast RNA polymerase recognizes a classic rho-independent terminator. In *Chlamydomonas* chloroplast DNA, stem–loop structures can function as transcription terminators *in vivo* (4). This report demonstrates, by site-directed mutagensis, the importance of a stem–loop structure for transcription termination by a higher plant chloroplast RNA polymerase. In many cases the spinach enzyme responded like the *E.coli* and T7 RNA polymerases to variations in terminator sequence and secondary structure. However, in addition to helix stability, the specific sequence of the stem–loop was important for termination efficiency by the chloroplast RNA polymerase. Unexpectedly, we also observed that efficient termination by the spinach enzyme was critically dependent on sequence downstream from the *thra* termination site, in particular, on the presence of guanosine as the second nucleotide 3′ to the oligo-dT tract.

ACKNOWLEDGEMENTS

We thank Tim Setter, David Stern, Rita Monde and Chih-Ning Sun for discussions and critical reading of the manuscript. We thank Li-Wen Fang for DNA sequencing. This work was supported by grants from National Science Council of the Republic of China, to L.-J.C.

REFERENCES

1 Chen, L.-J. and Orozco, E.M., Jr (1988) *Nucleic Acids Res.* **16**, 8411–8431.
2 Stern, D.B. and Gruissem, W. (1987) *Cell* **51**, 1145–1157.
3 Stern, D.B., Jones, H. and Gruissem, W. (1989) *J. Biol. Chem.* **264**, 18742–18750.
4 Blowers, A.D., Klein, U., Ellmore, G.S. and Bogorad, L. (1993) *Mol. Gen. Genet.* **238**, 339–349.
5 Chen, L.-J., Rogers, S.A., Bennett, D.C., Hu, M.-C. and Orozco, E.M., Jr (1990) *Curr. Genet.* **17**, 55–64.
6 Yager, T.D. and von Hippel, P.H. (1987) Transcript elongation and termination in *Escherichia coli* in F.C. Neidhardt, S.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter and H.-E. Umbarger (eds) *Escherichia coli and Salmonella typhimurium*, Vol. 2, pp. 1241–1275, American Society for Microbiology, Washington, D.C.
7 Yager, T.D. and von Hippel, P.H. (1991) *Biochemistry* **30**, 1097–1118.
8 Adhya, S. and Gottesman, M. (1978) *Ann. Rev. Biochem.* **47**, 967–996.
9 Martin, F.H. and Tinoco, I.J. (1980) *Nucleic Acids Res.* **8**, 2295–2299.
10 Farnham, P.J. and Platt, T. (1980) *Cell* **20**, 739–748.
11 Farnham, P.J. and Platt, T. (1981) *Nucleic Acids Res.* **9**, 563–578.
12 Ryan, T. and Chamberlin, M.J. (1983) *J. Biol. Chem.* **258**, 4690–4693.
13 Jeng, S.-T., Gardner, J.F. and Gumport, R.I. (1990) *J. Biol. Chem.* **265**, 3823–3830.
14 Jeng, S.-T., Gardner, J.F. and Gumport, R.I. (1992) *J. Biol. Chem.* **267**, 19306–19312.
15 Lynn, S.P., Gardner, J.F. and Reznikoff, W.S. (1987) *J. Bacteriol.* **152**, 363–371.
16 Lynn, S.P., Kasper, L.M. and Gardner, J.F. (1988) *J. Biol. Chem.* **263**, 472–479.
17 Telesnitsky, A. and Chamberlin, M.J. (1989) *Biochemistry* **28**, 5210–5218.
18 Reynolds, B. and Chamberlin, M.J. (1992) *J. Mol. Biol.* **224**, 53–63.
19 Altmann, C.R., Solow-Cordero, D.E. and Chamberlin, M.J. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3784–3788.
20 Surratt, C.K., Milan, S.C. and Chamberlin, M.J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7983–7987.
21 Orlova, M., Newlands, J., Das, A., Goldfarb, A. and Borukhov, S. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4596–4600.
22 Krummel, B. and Chamberlin, M.J. (1992) *J. Mol. Biol.* **225**, 221–237.
23 Krummel, B. and Chamberlin, M.J. (1992) *J. Mol. Biol.* **225**, 239–250.
24 Nudler, E., Goldfarb, A. and Kashlev, M. (1994) *Science* **265**, 793–796.
25 Chamberlin, M.J. (1994) *Harvey Lect.* **88**, 1–21.
26 Wang, D., Meier, T., Chan, C.L., Feng, G., Lee, D.N. and Landick, R. (1995) *Cell* **81**, 341–350.
27 Nudler, E., Kashlev, M., Nikiforov, V. and Goldfarb, A. (1995) *Cell* **81**, 351–357.
28 Chen, L.-J. (1995) *Bot. Bull. Acad. Sin.* **36**, 95–100.
29 Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. and Turner, D.H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
30 Platt, T. (1986) *Ann. Rev. Biochem.* **55**, 339–372.
31 Gardner, J.F. (1982) *J. Biol. Chem.* **257**, 3896–3904.
32 Bean, B., Koren, R. and Mildvan, A. (1977) *Biochemistry* **16**, 3322–3333.
33 Reynolds, R., Bermudz, R.M. and Chamberlin, M.J. (1992) *J. Mol. Biol.* **224**, 31–51.
34 Yang, M.-T. and Gardner, J.F. (1989) *J. Biol. Chem.* **264**, 2634–2639.
35 Cheng, S.C., Lynch, E.C., Leason, K.R., Court, D.L., Shapiro, B.A. and Friedman, D.I. (1991) *Science* **254**, 1205–1207.
36 Levin, J.R. and Chamberlin, M.J. (1987) *J. Mol. Biol.* **196**, 61–84.
37 Lee, D.N., Phung, L., Stewart, J. and Landick, R. (1990) *J. Biol. Chem.* **265**, 15145–15153.