 Nucleic Acids Research, 1999, Vol. 27, No. 14 **2883–2888**

Terminator element mutations affect both the efficiency and position of RNA polymerase I termination in *Schizosaccharomyces pombe*

Philip S. Shwed and Ross N. Nazar*

Department of Molecular Biology and Genetics, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Received March 22, 1999; Revised and Accepted May 26, 1999

ABSTRACT

RNA polymerase I transcripts, purified from *Schizosaccharomyces pombe* cells, terminate at three sites that precede 'Sal box'-like termination element (TE) sequences. Essential features in these elements were investigated by the *in vivo* expression of targeted mutations. RNA analyses confirmed a functional significance for two of the elements (Boxes 1 and 3), but indicated that the third, less related, sequence (Box 2) does not function as a termination signal. The results further indicated that the most conserved residues in the two active TEs, as well as adjacent regions, are also most critical to function. Furthermore, some mutations in these elements or in immediately flanking sequences affect not only the efficiency of termination, but also alter the position of termination by as much as 35 nt. Since the element is able to influence the site of termination over a surprisingly long stretch of DNA sequence, these observations suggest that the TE does not act simply as a pause element by fixing the termination factor.

INTRODUCTION

In eukaryotes, the termination of RNA polymerase I (pol I) pre-rRNA transcripts is mediated by the binding of a protein transcription termination factor (TTF) to a termination element (TE) located 12–20 bp downstream of the actual termination site (reviewed in 1–5). Based primarily on *in vitro* studies, a two-step termination mechanism has been proposed in which the terminator protein acts to pause the elongating complex allowing an upstream release element (URE) to mediate the release of the transcript at the termination site (6). While a URE has not been found in every organism (7) and the need for this element has been questioned experimentally (8), the TE and its interaction with a TTF has been documented clearly in several cell models, including the mouse, frog and budding and fission yeasts (reviewed in 1–5). Recently, a mouse release factor has also been cloned. This protein appears to bind specifically to uridylate residues encoded by the URE sequence (9).

The termination element itself was first documented in the mouse as an 18 bp repeating sequence motif termed the 'Sal box' because it contained a *Sal*I restriction enzyme recognition sequence (10). Although the sequence homology between TEs is rather limited in diverse organisms, functionally equivalent repeating elements have been observed to be associated with all termination sites and, where studied, a cognate binding protein has been identified. Unexpectedly, in *Saccharomyces cerevisiae* this protein was found to be Reb1p, a protein also linked to the control of gene expression (reviewed in 11).

In *Schizosaccharomyces pombe* rRNA transcripts, three termination sites have been identified by the mapping of termini with S_1 nuclease digestion (12). Three 'Sal box'-like TE sequences have also been linked with these sites (12) and a protein factor with affinity for two of these elements has been cloned and found to be related to the *S.cerevisiae* Reb1p protein (13). In the same two *S.pombe* TE sequences, however, the sequence homology is not limited to the core protein binding site. Rather, it appears to extend almost continuously from the termination site and even several residues downstream of the functional TE core that has been identified in *S.cerevisiae*. To evaluate the significance of these extended similarities, in this study targeted mutations were introduced throughout the region of sequence homology and the mutated rDNAs were expressed *in vivo* for transcript analyses.

MATERIALS AND METHODS

Strains used

Escherichia coli C490 (from C600, recA$^-$, rk$^-$, mk$^-$, thr$^-$, leu$^-$, met$^-$) (14) was used as a host for a pTZ19R plasmid template containing only the 3'-ETS sequence of rDNA (12). *Schizosaccharomyces pombe* (h$^-$, *leu* 1-32, *ura*4-D18) (15) was used as a host for the various pFL20 yeast shuttle vector derivatives containing 'tagged' rDNA repeats and mutated terminator elements.

Construction and expression of mutant rRNA intergenic sequences

Site-specific base pair substitutions were introduced into the intergenic region of *S.pombe* rDNA by the 'megaprimer' technique (16) with an extended annealing time (17). The template for mutagenesis (Fig. 2A) was a 2.4 kb fragment of *S.pombe* rDNA encompassing the 3'-end of 25S rDNA and 800 bp of downstream 3'-ETS and intergenic sequences that was subcloned into the pTZ19R vector (12). Two markers were contained in the fragment: a unique *Not*I restriction enzyme recognition sequence within the last variable domain of 25S rDNA and a

*To whom correspondence should be addressed. Tel: +1 519 824 4120; Fax: +1 519 837 2075; Email: rnnazar@uoguelph.ca

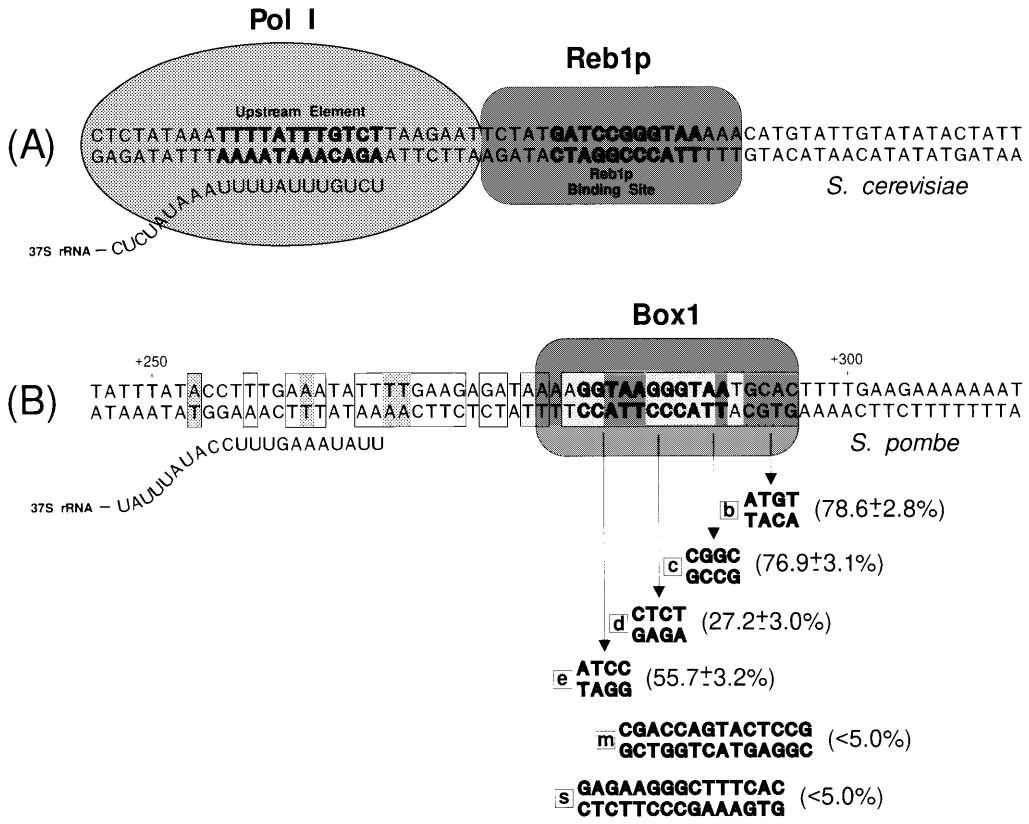

Figure 1. Effect of TE sequence substitutions on the efficiency of pol I termination in *S.pombe*. (**A**) Sequences affecting termination in *S.cerevisiae* as proposed by Reeder and Lang (4). Sequence elements required for Reb1p binding (Reb1p binding site) or the release of pol I (upstream element) are indicated in bold with the 3'-end of the released transcript shown below. (**B**) Equivalent regions at the first termination site (Site 1) in the *S.pombe* rDNA. Shaded residues indicate identical nucleotides in all three *S.pombe* termination sites, while the enclosed residues are identical only in Sites 1 and 3. The primary protein binding site (Box1) was previously defined by genomic footprinting (12). Bold sequences indicate the putative core TE sequence as well as mutations in the TE sequence at Site 1; the efficiency of termination with each mutant sequence is indicated in parentheses. The extent of transcription is indicated by the 37S pre-rRNA.

22 bp deletion in an extended hairpin structure in the 3'-ETS. Mutations were introduced into the fragment by a two-step PCR-based procedure (17). Mutations were identified by differential hybridization (18) and confirmed by dideoxy sequencing (19). Fragments containing appropriate changes were cleaved from the intermediate and used to replace the normal sequences in the shuttle vector containing the tagged *S.pombe* rDNA transcriptional unit (Fig. 2B) previously described by Melekhovets and co-workers (12).

S_1 nuclease mapping of transcript termini in rRNA precursors

The termini of plasmid-derived mature and precursor rRNA transcripts were detected and quantified by S_1 nuclease mapping as described by Aldea *et al*. (20) using PCR-generated plasmid-specific probes for normal or mutant rDNA that spanned the 3'-ETS and intergenic sequences. The purified probes were digested with *Not*I restriction endonuclease (Fig. 2) and radio-labeled by filling in the recessed 3'-end with [α-^{32}P]dCTP and Klenow enzyme (20). Appropriate probes (~100 000 c.p.m.) were incubated with total cellular RNA extracts (20 µg) in 20 µl hybridization buffer (3 M NaTCA, 5 mM Na_2EDTA, 50 mM PIPES pH 7.0) at 30°C for 12 h. The hybrids were then digested for 30 min at 37°C with S_1 nuclease (50–200 U) in 15 mM $ZnCl_2$, 250 mM NaCl, 40 mM NaAcO, pH 5.5 (200 µl total volume). Digests were terminated with SDS/EDTA, extracted with phenol/chloroform/isoamyl alcohol (25:24:1), ethanol precipitated and analyzed on 6% polyacrylamide sequencing gels alongside standard chain termination sequence reaction products applied to the analytical gels as fragment length markers.

Scanning densitometry

Quantitative analysis of the fractionated bands corresponding to rRNA termini was performed using a Umax Astra 600P scanning densitometer (Umax Technologies, CA) and Molecular Analyst® PC software v.1.0 (BioRad Laboratories, CA). The nuclease protection signals from duplicate electrophoretic fractionations of two independent transformants of a given construct were analyzed and expressed as a percentage of the total radioactivity in all transcripts.

RESULTS

Although past studies indicate that the termination and release of pre-rRNA in *S.cerevisiae* is dependent on two short sequence elements, the Reb1p binding site and an upstream release element (Fig. 1), sequence comparisons in *S.pombe* indicate two striking differences. As also indicated in Figure 1,

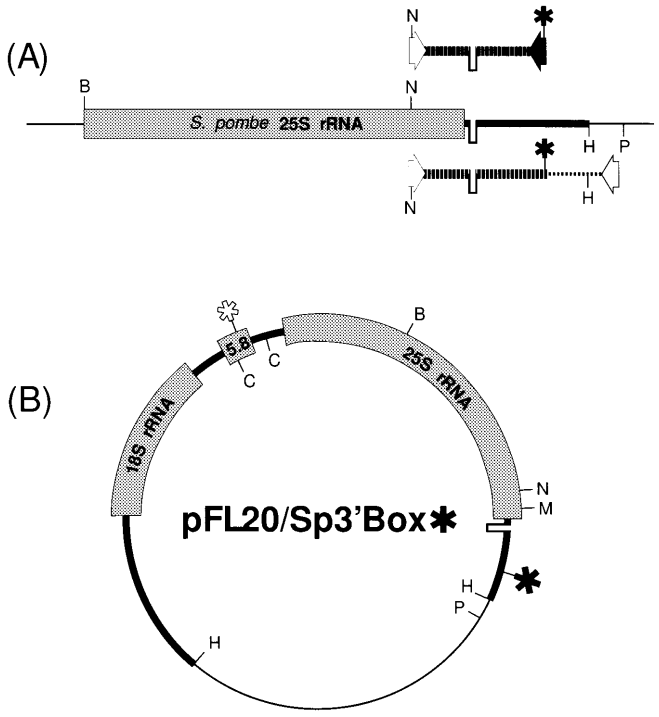

Figure 2. PCR-mediated mutations in the termination sequences of rDNA for *S.pombe*. (**A**) A *Bgl*II (B)–*Hin*dIII (H) endonuclease digestion fragment of *S.pombe* rDNA containing the 3'-end of the mature 25S rRNA sequence (shaded box) and the 3' spacer region (thick continuous line) with three 'Sal box'-like terminator elements and a 22 nt deletion (12) in the conserved extended hairpin of the 3'-ETS (blunt hairpin) was subcloned in pTZ19R (thin continuous line). Appropriate mutagenic primers (filled asterisk, closed arrow) overlapping targeted sequence elements were used to PCR amplify an intermediate mutagenic primer (thick broken line) which was again used with a second plasmid-specific primer (open arrow) to PCR amplify the full mutant sequence. (**B**) The mutant DNA was digested with *Not*I (N) and *Pvu*II (P) restriction endonucleases and subcloned into a pFL20 yeast shuttle vector (thin line) containing a *S.pombe* rDNA transcription unit (pFL20/Sp25Not), which also contained a *Not*I restriction site (12) and a tag (open asterisk) in the 5.8S rRNA sequence. Recombinants, which contained the mutated termination sequences, were selected by differential hybridization (18) using the mutagenic oligonucleotide primers.

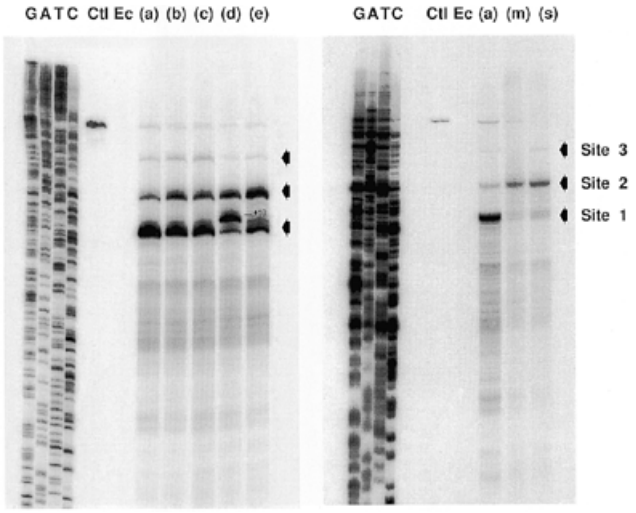

Figure 3. Pol I transcript termini in nascent RNAs from *S.pombe* rDNA containing TE sequence substitutions. Four base substitutions were introduced into the first terminator element, as described in Figure 2, and mutant rDNAs were used to transform *S.pombe* cells. RNA was prepared by SDS/phenol extraction from cells transformed with normal TE sequences (a) and the ATGT/TACA (b), CGGC/GCCG (c), CTCT/GAGA (d), ATCC/TAGG (e), GACCAGTACTCCG/GCTGGTCATGAGGC (m) or GAGAAGGGCTTCAC/CTCTTCCCGAAAGTG (s) substitutions as described in Figure 1. One kilobase *Not*I–*Hin*dIII digestion fragments of DNA overlapping the terminal regions (Fig. 2) in each of the plasmids were labeled at the 3'-end of the coding strand with [α-^{32}P]dCTP and Klenow fragment for plasmid-specific probes. The labeled fragments were incubated at 90°C for 5 min to separate the strands and then hybridized with RNA aliquots (20 µg) at 30°C for 12 h. The resulting hybrids were digested with 100 U of S_1 nuclease at 37°C for 30 min and, after extraction with phenol/chloroform, the fragments were fractionated on a 6% polyacrylamide sequencing gel. An undigested probe (Ctl) and a digest with RNA from *E.coli* cells (Ec) were also fractionated together with standard dideoxy sequencing reaction products (G, A, T and C) as residue markers. The positions of the major termini with normal *S.pombe* RNA are indicated at the right.

a comparison of the three known termination sites reveals a more extended sequence homology between Sites 1 and 3 (enclosed residues) and a substantially reduced level of homology in the third, Site 2 (shaded residues). To further evaluate the significance of these similarities and differences, site-specific mutations, using a PCR-based mutation strategy (17), were introduced into the regions of highest homology and the termination of mutant transcripts was examined *in vivo* using an efficiently expressed, tagged shuttle vector system (22). As illustrated in Figure 2 and fully documented in past studies (8,12), in this case the *S.pombe* rDNA transcriptional unit contained tags in the 5.8S and 25S rDNA sequences as well as a deletion in the highly conserved 3'-ETS stem, which has been shown to inhibit pre-rRNA processing (12). As demonstrated in Figure 3 (lane a), such a deletion permits the termination sites to be easily visualized and quantified. With normal gene termination sequences ~90% of the transcripts end at Site 1 and most of the remaining molecules end at Site 2.

The first series of changes were directed at the 'Sal box'-like element, previously shown to be essential for termination by complete deletion (12) and presumably functionally equivalent to the Reb1p binding site sequence in *S.cerevisiae* (Fig. 1). As illustrated by the example analyses in Figure 3 and summarized in Figure 1, large sequence changes again resulted in an essentially complete inhibition of termination. This was true even when the sequences were the known or putative termination sequence elements in mouse and *S.cerevisiae* (lanes m and s).

In contrast, the effects of smaller 4 bp block substitutions were relatively modest; in most cases, the reduction in the efficiency of termination was usually ~25%. The most striking changes occurred when the more conserved residues were changed. For example, the AGG/TCC sequence appears in all three TEs and corresponds to the left border of the Reb1p binding site in *S.cerevisiae*. When this cluster was altered, the efficiency of termination dropped by almost 50%. A cluster of three G/C box pairs also appears in all three TEs and even in the Reb1p binding site of *S.cerevisiae*. When this cluster was disrupted the overall efficiency for the first element only dropped by ~25% but, as shown in Figure 3 (lane d), the actual site of termination was altered for most transcripts with ~75% terminating

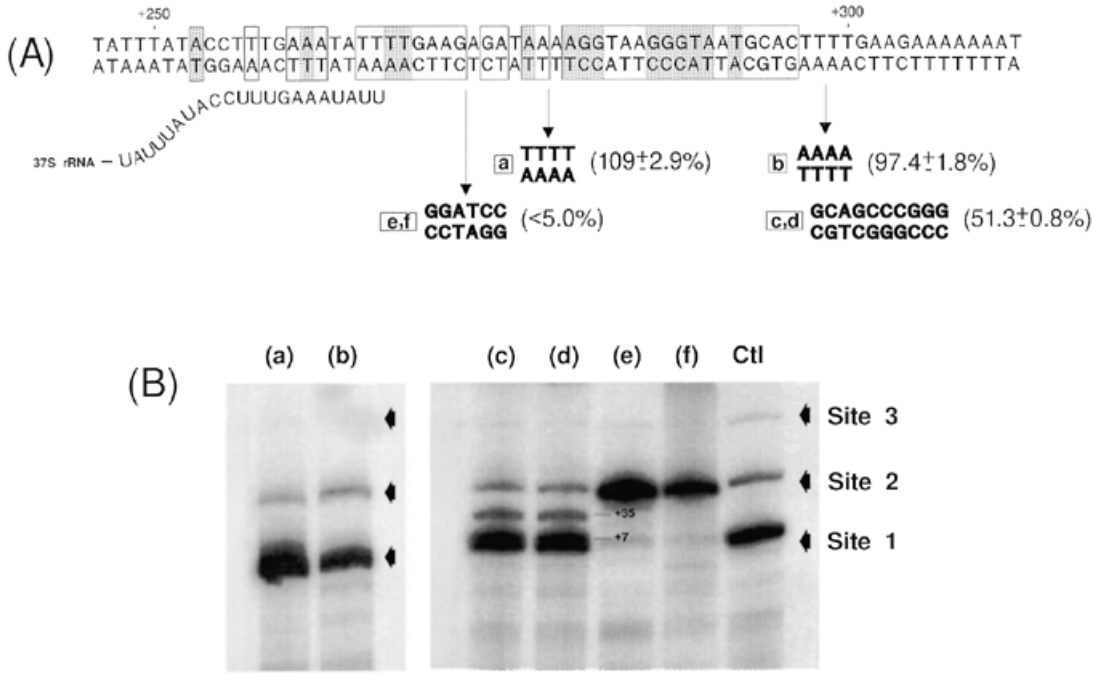

Figure 4. Effect of sequence substitutions in nucleotides surrounding the termination factor binding site in *S.pombe*. (**A**) First termination site in the *S.pombe* rDNA. Shaded residues indicate identical nucleotides in all three termination sites, while the enclosed residues are identical in Sites 1 and 3. Bold sequences indicate the changes which were made together with the efficiency of termination at Site 1 relative to the normal sequence (in parentheses). (**B**) Pol I transcript termini in nascent RNAs from *S.pombe* cells expressing the mutant rDNAs. RNA was prepared from cells transformed with normal TE sequences (Ctl) and the TTTT/AAAA (a), AAAA/TTTT (b), GGATCC/CCTAGG (c and d) and GCAGCCCGGG/CGTCGGGCCC (e and f) substitutions. Termini were detected after S_1 nuclease digestion and the relative concentrations were determined after fractionation as described in Figure 3. The positions of the three known sites of termination are indicated at the right.

~18 nt beyond the normal Site 1 and only ~25% of the termini ending at Site 1.

Because the sequence homology for at least two of the TEs (Boxes 1 and 3) was observed to extend further upstream and downstream, changes were also introduced in these flanking regions. As illustrated in Figure 4, simple sequence inversions (lanes a and b) had no effect, but more extensive changes, both upstream and downstream of the TE, had surprisingly dramatic effects. When the region between the termination site and protein binding site was altered, with GGATCC being substituted for the highly conserved TTGAAGAGATAA region, a very severe reduction in the termination efficiency was evident (Fig. 4, lanes e and f). When the downstream region was altered, the overall efficiency of termination was not significantly affected, but only ~50% of the molecules terminated at the normal site. The remaining 40% of the transcripts terminated 7 nt downstream and 10% terminated 35 nt after the normal site.

The sequence differences which are associated with the second mapped termination site also raise questions about the critical sequence elements in *S.pombe*. When the Box 1 sequence is fully deleted, ~90% of terminations occur at Site 2 (8), an observation which suggests that the Box 2 sequence is equally effective in bringing about termination. However, an alternative explanation, which has been raised in the past (8), is the possibility that the mapping of termini by S_1 nuclease digestion is misleading, due to non-specific cleavage in the extended thymidylic acid residue cluster that follows Site 2. To test this possibility further, the conserved central core of the second TE was extensively changed as indicated in Figure 5. As also shown in Figure 5, when this mutant gene was expressed *in vivo*, the pattern of termini was entirely normal and there was no elevation in the termini ending at Site 3, as might be anticipated if the second putative TE actually functions to terminate transcripts at Site 2.

DISCUSSION

The present studies, while further supporting the presence of TEs at two of the known termination sites in *S.pombe* rRNA, raise a number of new questions regarding functionally important features in this yeast, and perhaps other organisms. The most highly conserved feature in all three termination sites in *S.pombe* and even *S.cerevisiae* is a cluster of three G/C base pairs at or near the center of the TE (Fig. 1). As also indicated in Figures 1 and 3, a four base substitution in this region dramatically altered the position of termination in ~75% of the transcripts. Changes in a second conserved region (AGG/TCC) had no effect on the positioning of termination, but reduced the efficiency by ~50%. Other equally large changes in less conserved regions had modest effects (~25% reduction) and only changes at many positions in the TE reduced termination to very low or undetectable levels. Taken together, these changes suggest that stability in protein binding is dependant on interactions across the entire sequence element. However, residues which are conserved in these yeasts appear to have greater

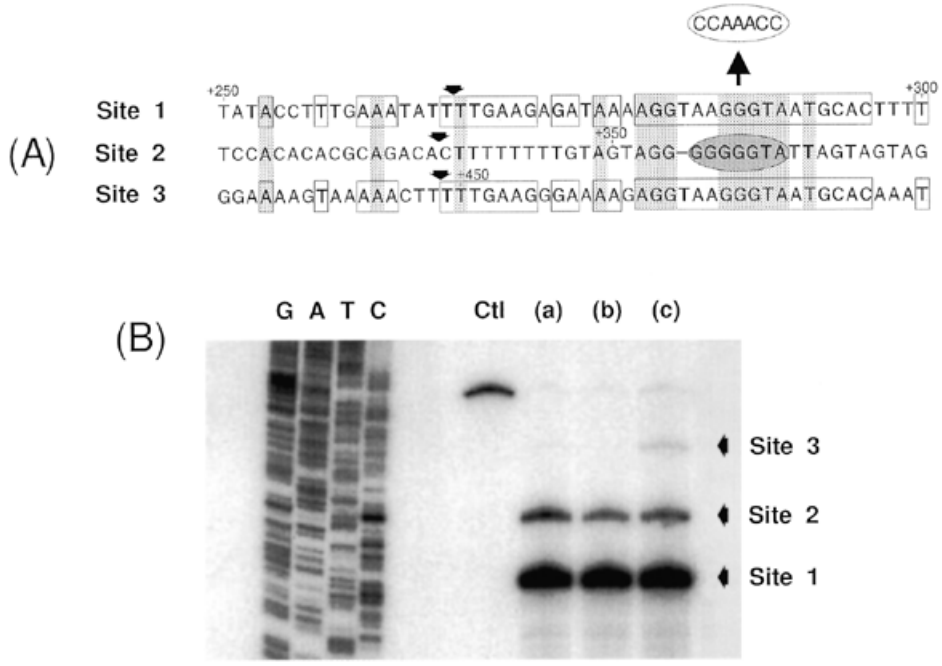

Figure 5. Effect of TE sequence substitutions at the second termination site in *S.pombe* rDNA. (**A**) A sequence comparison between the known pol I termination sites in *S.pombe* rRNA, as described by Melekhovets and co-workers (12). Enclosed stippled areas indicate identical nucleotides at Sites 1 and 3; the lightly stippled areas indicate nucleotides which are also identical at Site 2. Small arrows indicate the actual sites of termination and the large arrow indicates the changes which were targeted into Site 2. (**B**) Pol I transcript termini in nascent RNAs from *S.pombe* cells expressing rDNAs with a normal or mutated TE sequence at Site 2. RNA was prepared from cells expressing normal rDNA (lane c) or two different transformants expressing the mutant rDNA (lanes a and b) and the termini were detected after S_1 nuclease digestion and fractionation on a 6% polyacrylamide gel. An undigested probe (Ctl) was also fractionated together with standard dideoxy sequencing reaction products (G, A, T and C) as residue markers. The three known termini with normal *S.pombe* RNA are indicated at the right.

functional significance and can affect both the efficiency and even the position of termination.

Changes downstream of the termination factor binding site and in the binding site itself which result in alternative termination sites are especially interesting. As shown in Figures 3 (lane d) and 4 (lanes c and d), such changes result in transcripts which terminate up to 35 nt beyond the factor binding site. The reason for these extended termini is not obvious, but they raise the possibility that the termination factor may not act simply to pause the polymerase complex, but actually may interact actively with the transcription complex to induce termination either rapidly upstream from the site of binding or downstream of it when it acts less efficiently. In mouse, Grummt and co-workers (23) have found that the binding of TTFI induces DNA binding. Such a conformational alteration may underlie the changes which we observe in the termination position. While other possibilities can be proposed, these present observations are clearly difficult to reconcile with the prevailing model for termination involving a simple pause mechanism (6).

The present study has also demonstrated that a functional role for the conserved 5'-flanking sequences is indicated by the abolition of termination with sequence changes between termination Site 1 and the TE (Fig. 4, lanes e and f). While more distant upstream elements have been the target of past experimentation (9,24,25), in general, the significance of immediately flanking sequences to the pol I termination process is even less clear. The present study indicates that the TE sequence, at least in *S.pombe*, may be more extended than previously proposed in *S.cerevisiae* (Fig. 1). *In vitro* studies, carried out in a mouse cell model, have also shown that downstream sequences contribute to the efficiency of termination and that the position of the 3'-end of pre rRNA was affected by deletions in the 5'-flanking sequences (26). Therefore, taken together, both *in vivo* and *in vitro* analyses are consistent with a more extended TE, which also appears to do more than simply pause the transcription complex.

As noted earlier, previous sequence comparisons raised the possibility that in *S.pombe* the second largest transcript did not actually represent a true termination event, but is simply an artifact of S_1 nuclease digestion due to preferential cleavage of the extended A/T cluster which follows this putative termination site. As illustrated in Figure 3 (lanes m and s), when the first TE sequence was significantly altered, the second transcript became dominant. In contrast, as shown in Figure 5, when a large change was introduced into the conserved residues which constitute the second putative factor binding site, there was no effect on the transcripts terminated at Site 2 and there was no elevation in transcripts terminated at Site 3. These observations strongly support the possibility that termini ending at Site 2 are preferential cleavages in transcripts terminating at Site 3 or beyond. We conclude that termination in *S.pombe* rDNA actually occurs at only the two highly conserved termination sites with 90% of transcripts released at Site 1 and the remainder at Site 3.

ACKNOWLEDGEMENTS

The authors wish to thank Dr Yuri F. Melekhovets for advice and helpful discussions. This study was supported by the National Sciences and Engineering Council of Canada.

REFERENCES

1. Paule,M.R. (1993) *Gene Expr.*, **3**, 1–9.
2. Richardson,J.P. (1993) *Crit. Rev. Biochem. Mol. Biol.*, **28**, 1–30.
3. Reeder,R. and Lang,W. (1994) *Mol. Microbiol.*, **12**, 11–15.
4. Reeder,R. and Lang,W. (1997) *Trends Biochem. Sci.*, **22**, 473–477.
5. Mason,S.W., Sander,E.E., Evers,E. and Grummt,I. (1998) In Paule,M.R. (ed.), *Transcription of Ribosomal RNA Genes by Eukaryotic RNA Polymerase I*. Springer-Verlag and R.G.Landes, Austin, TX, pp. 179–194.
6. Lang,W.H., Morrow,B.E., Ju,Q., Warner,J.R. and Reeder,R.H. (1994) *Cell*, **79**, 527–534.
7. Labhart,P. (1995) *Nucleic Acids Res.*, **23**, 2252–2258.
8. Melekhovets,Y.F., Shwed,P.S. and Nazar,R.N. (1997) *Nucleic Acids Res.*, **25**, 5103–5109.
9. Jansa,P., Mason,S.W., Hoffmann-Rohrer,U. and Grummt,I. (1998) *EMBO J.*, **17**, 2855–2864.
10. Grummt,I., Maier,U., Öhrlein,A., Hassouna,N. and Bachellerie,J.-P. (1985) *Cell*, **43**, 801–810.
11. Planta,R.J., Gonçalves P.M. and Mager W.H. (1995) *Biochem. Cell Biol.*, **73**, 825–834.
12. Melekhovets,Y.F., Good,L., Abou Elela,S. and Nazar,R.N. (1994) *J. Mol. Biol.*, **239**, 170–180
13. Zhao,A., Guo,A., Liu,Z. and Pape,L. (1997) *Nucleic Acids Res.*, **25**, 904–910.
14. Hanahan,D. (1983) *J. Mol. Biol.*, **166**, 557–580.
15. Heyer,W., Sipiczki,M. and Kohli,J. (1986) *Mol. Cell. Biol.*, **6**, 80–89.
16. Sarkar,G. and Sommer,S. (1990) *BioTechniques*, **8**, 404–407.
17. Good,L. and Nazar,R.N. (1992) *Nucleic Acids Res.*, **20**, 4394–4395.
18. Zoller,M.J. and Smith,M. (1984) *DNA*, **3**, 479–488.
19. Sanger,F., Nicken,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
20. Aldea,M., Claverie-Martin,F., Diaz-Torres,M. and Kushner,S. (1988) *Gene*, **65**, 101–110.
21. Wartell,R.M. and Reznikoff,W.S. (1980) *Gene*, **9**, 307–319.
22. Abou Elela,S., Good,L., Melekhovets,Y.F. and Nazar,R.N. (1994) *Nucleic Acids Res.*, **22**, 686–693.
23. Smid,A., Finsterer,M. and Grummt,I. (1992) *J. Mol. Biol.*, **227**, 635–647.
24. Lang,W. and Reeder,R. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 9781–9785.
25. Mason,S.W., Sander,E. and Grummt,I. (1997) *EMBO J.*, **16**, 163–172.
26. Kuhn,A., Normann,A., Bartsch,I. and Grummt,I. (1988) *EMBO J.*, **7**, 1497–1502.