Nucleic Acids Research, 1996, Vol. 24, No. 22 **4577–4583**

In vitro transcription of a poly(dA)·poly(dT)-containing sequence is inhibited by interaction between the template and its transcripts

Ryoiti Kiyama* and Michio Oishi

Institute of Molecular and Cellular Biosciences, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113, Japan

Received June 27, 1996; Revised and Accepted September 20, 1996

ABSTRACT

Transcription of poly(dA)·poly(dT)-containing sequences was investigated *in vitro* using plasmids carrying a $(dA)_{34}$·$(dT)_{34}$ tract in the coding region of the *lacZ* gene. The efficiency of transcription of the $(dT)_{34}$ sequence on the transcribing strand by *Escherichia coli* RNA polymerase was substantially lower (~60%) than that of the $(dA)_{34}$ sequence or of the control *lacZ* gene. Analysis of the transcription process of the $(dT)_{34}$ sequence by T3 RNA polymerase showed that the transcription was frequently arrested or terminated at the middle as well as immediately proximal of the $(dA)_{34}$·$(dT)_{34}$ tract, and it occurred more prominently following accumulation of transcription products. This inhibition was strongly enhanced by the addition of the oligonucleotide $(dT)_{34}$ or poly(U) to the reaction mixture, while $(dA)_{34}$ and the duplex $(dA)_{34}$·$(dT)_{34}$ suppressed the inhibition. A similar transcriptional inhibition was also observed in transcription mediated by T7 RNA polymerase and eukaryotic RNA polymerase II. We also demonstrated RNA·DNA complex formation of the $(dA)_{34}$·$(dT)_{34}$ tract with poly(U), but not with poly(A). These findings strongly suggest that poly(dT)-containing template sequences interact and form a complex with its transcription products, possibly an RNA·DNA triplex, which blocks further transcription. This would explain the instability of the plasmids transcribing mRNAs with poly(U) but not poly(A) tracts and the underrepresentation of poly(U) but not poly(A) tracts in mRNAs.

INTRODUCTION

Transcription is generally controlled through interactions between template DNA and RNA polymerases with the help of various types of protein factors collectively referred to as transcription factors (1–6). These factors interact with specific DNA sequences in the transcription-regulatory regions. The promoter region located immediately upstream of the genes accommodates binding sites for these factors, while enhancer and silencer elements modulate the transcriptional activity from upstream or downstream of the genes. On the other hand, simple sequences such as $(dA)_n$, $(dG)_n$, $(dA–dC)_n$, which constitute microsatellite DNA are also known to affect the efficiency of transcription. The presence of A+T-rich sequences in the promoter region for example, enhances or suppresses transcription in *Escherichia coli* and yeast genes (7–10), and short poly(dA–dT) sequences can substitute for the function of promoter elements (9,11,12). Sigma factor, an essential component of the transcription machinery in *E.coli*, is released at the poly(dA–dT) tract once the transcription is initiated (13). These microsatellite sequences, which are distributed ubiquitously along genomic DNA in most eukaryotes, are likely to be associated with other biological reactions including recombination and replication (14–17). Although the exact roles played by such sequences remain to be determined, microsatellite DNA sequences could affect these reactions through formation of unique topological structures or by interactions with specific binding proteins (12,17–19).

Other than specific nucleotide sequences, specific or 'unusual' DNA structures affect the transcription. For example, $(dA–dC)_n$ sequences affect transcription through Z-DNA formation (20,21). Bent DNA, formed by short poly(dA) tracts at an interval of 10–11 bp, can alter the transcriptional efficiency (22,23). Because of low melting temperatures, A+T-rich sequences tend to melt during transcriptional elongation causing transcriptional slippage (10,24).

Controlling gene expression could be achieved by several strategies. Antisense RNA or DNA against mRNA is one of the methods to control gene expression at translation. In contrast, gene expression could be controlled by the formation of triplex DNA using specific DNA or RNA oligonucleotides which bind to the regulatory region, to inhibit the initiation of transcription (25–27). Transcription elongation can also be inhibited by oligonucleotides that form triplex DNA (28,29) or by peptide nucleic acids that form a D-loop (30).

Previously, we reported that *E.coli* plasmid clones containing microsatellite poly(dA)·poly(dT) exhibited marked instability in their maintenance in host cells (31). This observation was further investigated by employing a series of plasmids containing a $(dA)_{34}$·$(dT)_{34}$ tract placed downstream of the *lacZ* promoter, and the results suggested that the instability was likely to be caused by transcription of the poly(dT) strand (32). In this report, we show that the *in vitro* transcription of the $(dA)_{34}$·$(dT)_{34}$ tract was inhibited as a result of the interaction between the poly(U)-containing transcripts and their templates. The possible involvement of an RNA·DNA triplex structure in the inhibition is discussed.

*To whom correspondence should be addressed. Tel: +81 3 3812 2111; Fax: +81 3 3818 9437; Email: kiyamar@hgc.ims.u-tokyo.ac.jp

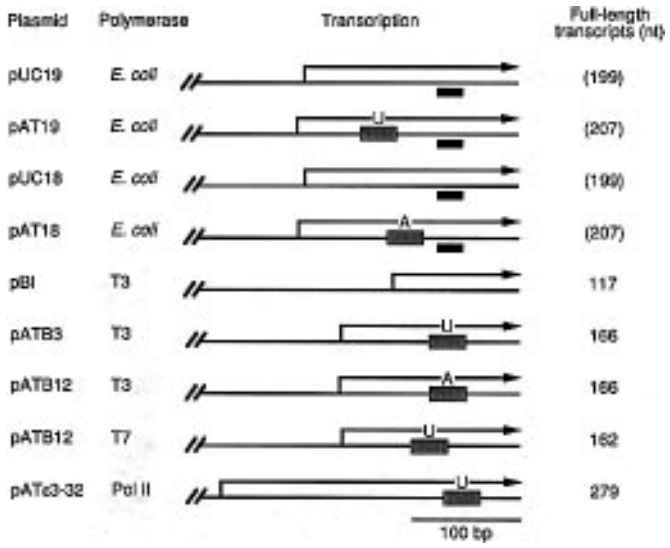


Figure 1. Summary of transcription assays. Combinations of plasmids and polymerases, and the lengths of full-length transcripts appeared in this report are listed. Poly(dA)·poly(dT) tracts are shadowed and the positions of *lacZ* probe were shown by solid boxes. pBl: pBluescript SK(–), U: $(U)_{34}$, A: $(A)_{34}$.

MATERIALS AND METHODS

Materials

Oligonucleotides were synthesized by a Millipore Cyclone DNA synthesizer and purified with Milligen Oligo-Pak columns. Poly(U) and poly(A) were purchased from Pharmacia.

Plasmid construction

Plasmids pAT19 and pAT18 were constructed from vector pUC19 or pUC18 (33) by replacing the *Bam*HI–*Hin*dIII sequence with GGATCC(*Bam*HI)–$(A)_{32}$–AAGCTT(*Hin*dIII) (32). pATB3 and pATB12 were constructed by inserting the GGATCC–$(A)_{32}$–AAGCTT sequence into the *Eco*RV site of pBluescript SK(–) (Stratagene, USA) in the direction of the $(dT)_{34}$ (pATB3) or $(dA)_{34}$ (pATB12) sequence on the transcribing strand for T3 RNA polymerase (therefore, on the template strand for T7 RNA polymerase). The plasmid pATε3-32 was constructed by inserting the *Eco*RV–*Ssp*I fragment of the human ε-globin gene (positions –274 to +185 relative to the cap site, thus containing the whole promoter region) into the *Sma*I site of pATB3 in the direction that the transcription by eukaryotic RNA polymerase II proceeds to the $(dA)_{34}$·$(dT)_{34}$ tract. Plasmids used in the transcription assays are summarized in Figure 1.

In vitro transcription with *E.coli* RNA polymerase

To avoid read-through transcription, purified plasmid DNA (pUC19, pAT19, pUC18 or pAT18) was digested with *Pvu*II and *Sca*I, to separate the *bla* and *lacZ* genes. Approximately 1 µg of plasmid DNA was incubated with 1 U of *E.coli* RNA polymerase (Boehringer Mannheim, Germany) in 20 µl of a mixture of 40 mM Tris–HCl (pH 7.5), 6 mM $MgCl_2$, 5 mM NaCl, 2 mM spermidine, 10 mM DTT, 0.5 mM each of rNTPs and 1 U RNase inhibitor (Takara, Kyoto) for 30 min at 37°C, followed by incubation with 70 U DNase I (Takara) for 15 min at 37°C. The reaction was terminated by addition of SDS (0.1%). The reaction mixture was then spotted on Hybond-N+ (Amersham) nylon membranes. Two identical membranes were made. Hybridization with ^{32}P-labeled 24mer sequencing primers #1224 (CGCCAGGGTTTTCCCAG-TCACGAC, New England Biolabs, USA) as the *lacZ* probe, and with ^{32}P-labeled 20mer oligonucleotides (TATGCGGCGACCG-AGTTGCT, positions 2208–2227) as the *bla* probe were performed according to Geliebter *et al.* (34). Oligonucleotide probes were labeled with [γ-^{32}P]ATP (4500 Ci/mmol, ICN, USA) using T4 polynucleotide kinase (New England Biolabs) to a specific activity of 1.5×10^8 c.p.m./µg. After hybridization, the membranes were washed twice with 3× SSC, 10 mM sodium phosphate buffer (pH 7.2), 10% (v/v) Denhardt solution, 5% (w/v) SDS for 1 min, once with the same buffer for 1 h and then with 1× SSC, 1% (w/v) SDS for 1 h. Hybridization and washing membranes was performed at 60°C. The membranes were autoradiographed with a BAS2000 Image Analyzer (Fujix, Japan), and the radioactivities of the RNA–oligonucleotide complexes were quantitated.

In vitro transcription (run-off assay) with T3 or T7 RNA polymerase, or with nuclear extracts from HeLa cells

One hundred ng of pATB3 or pATB12 was mixed in 20 µl of the transcription mixture described above except that rCTP was replaced by 5 µCi of [α-^{32}P]rCTP (650 Ci/mmol), and transcription was initiated by adding 1 µl of T3 (50 U/µl; BRL, USA) or T7 (10 U/µl; Wako, Japan) RNA polymerases. After incubation at 37°C for the indicated lengths of time, the reaction was terminated by addition of 1 µl of 10% (w/v) SDS. The transcripts were resolved by electrophoresis through a 6% polyacrylamide–7 M urea gel under denaturing conditions. Transcription with the nuclear extract (3 mg/ml proteins) from HeLa cells was carried out with 1 µg of template (*Xho*I digests of pATε3-32 DNA), in 20 µl of buffer containing 12 mM HEPES (pH 7.9), 12% (v/v) glycerol, 0.3 mM DTT, 0.12 mM EDTA, 60 mM KCl, 12 mM $MgCl_2$, 0.5 mM rNTPs except for rGTP and 5 µCi of [α-^{32}P]rGTP (650 Ci/mmol) at 30°C for the indicated lengths of time (35). The nuclear extract was prepared according to the method described by Dignam *et al.* (36) and detailed previously (37).

Detection of RNA·DNA complex

RNA·DNA complex formation was assayed first by incubating 400 µM (nucleotide-equivalent) of ^{32}P-labeled third strand, poly(A) or poly(U) and 20 ng/µl of *Xmn*I-digested plasmid DNA in 25 µl of 10 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl (TMN buffer) for 30 min at 37°C, followed by electrophoresis on an agarose gel (0.7%) in 50 mM Tris–borate (pH 8.3) and 10 mM $MgCl_2$ overnight at 4°C (37). The gel was then soaked in 10% (w/v) trichloroacetic acid, dried under paper towels and autoradiographed. Poly(A) and poly(U) were labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP (4500 Ci/mmol).

RESULTS

Inhibition of transcription by *E.coli* RNA polymerase

The efficiency of transcription of the $(dA)_{34}$·$(dT)_{34}$ tract was first examined *in vitro* with *E.coli* RNA polymerase using several template plasmids (pUC19, pAT19, pUC18 and pAT18) producing either poly(U)-containing (pAT19) or poly(A)-containing (pAT18)

transcripts, or *lacZ* (control) transcripts (pUC19 and pUC18). These plasmids were used for the analysis of plasmid instability (29), and among them, only pAT19 exhibited a marked instability. This instability was shown to be closely related with the transcription of *lacZ* gene (29). Full-length run-off transcripts were detected by hybridization with *lacZ* probe (#1224 oligonucleotide) located just downstream of the $(dA)_{34}\cdot(dT)_{34}$ tract (Table 3). By this assay, the transcripts containing the poly(U) or poly(A) tract were quantitated. We found that the level of *lacZ* transcripts with pAT19 was significantly lower than the others; ~60% of those with the other plasmids. Meanwhile, the level of transcription of the control (*bla*) gene was the same among all the plasmids (data not shown) and was used to normalize the *lacZ* transcripts. This result suggested that transcription was inhibited when a $(dT)_{34}$ sequence was present on the transcribing strand. The inhibition by the presence of $(dT)_{34}$ sequence apparently occurred only when it was located in a *cis* position because transcription of $(dT)_{34}$ did not affect *bla* gene transcription in the same reaction mixture. Since a number of strong and weak transcription initiation sites exist in the pUC plasmids for *E.coli* RNA polymerase (data not shown) which may complicate analysis of the results, we examined transcription by T3 RNA polymerase.

Table 3. *In vitro* transcription of poly(dA)·poly(dT)-containing sequences by *E.coli* RNA polymerase

	Relative rate of transcription[a]			
Plasmid	pUC19	pAT19	pUC18	pAT18
Experiment				
1	1.00	0.65	1.08	0.90
2	1.00	0.53	1.09	1.04
Average	1.00	0.59	1.09	0.97

[a]Transcripts of *lacZ* gene were quantitated by dot hybridization with ^{32}P-labeled #1224 oligonucleotides as described in Materials and Methods. The radioactivity was counted and the ratios of the counts of the *lacZ* probe to those of the *bla* probe were calculated and normalized by the ratio for pUC19 as 1.00.

Inhibition of transcription by T3 RNA polymerase

In vitro transcription of the $(dA)_{34}\cdot(dT)_{34}$ tract was further examined with plasmid pBluescript and its derivatives (pATB3 and pATB12) where the tract was placed between the T3 and T7 promoters (Fig. 2). When the plasmid DNA linearized with *Kpn*I was used as a template for run-off assay, T3 RNA polymerase produced 166 nt $(U)_{34}$-containing (for pATB3) or $(A)_{34}$-containing (for pATB12) transcripts, while control pBluescript vector produced 117 nt transcripts (Fig. 1). The results with pATB3 indicated a strong termination site in the middle as well as at the immediate proximal region of the $(dT)_{34}$ sequence on the transcribing strand (lower bracketed regions in Fig. 2A). On the other hand, transcription of pATB12 with T3 RNA polymerase showed little inhibition at these regions. The kinetics of the appearance of radioactivity in the major bands (corresponding to 166 and 117 nt full-length transcripts) (Fig. 2B) indicated that incorporation for pATB3 reached a plateau earlier than for pATB12 or the control vector. As shown in Figure 2C, while the ratios of the full-length to the premature transcripts for pATB12 increased up to 60 min of incubation, the ratio for pATB3 started to decline after 10 min of incubation, reaching 50% within the next 90 min. These results may be best explained by inhibition of transcription of the full-length RNA midway when the $(dT)_{34}$ was present in the transcribing strand, and the effect became more prominent when the transcripts were accumulated at the later stage of incubation. Although there were several other minor bands corresponding to the premature transcripts including the ~130 nt transcripts with pATB12 (Fig. 2A), none showed this type of kinetics, indicating that they are not inhibited by the full-length transcripts but represent the products of premature termination.

Involvement of complex formation between transcripts and their templates in the inhibition

To investigate further the mechanism of the observed transcriptional inhibition, we analyzed the transcription process in the presence of oligodeoxyribonucloetides, $(dT)_{34}$ or $(dA)_{34}$, or an oligoribonucleotide, poly(U) (~40–50 nt). As shown in Figure 3A, when at least 10 nM of $(dT)_{34}$ was added in advance to the transcription mixture, the level of the premature transcripts increased. On the other hand, addition of at least 10 nM $(dA)_{34}$ decreased the level of inhibition. Poly(U) caused an effect similar to that seen with $(dT)_{34}$ (Fig. 3A). Note that the concentration of poly(U) (2 μM), which was nucleotide-equivalent, was roughly equivalent to 40–50 nM in the number of poly(U) molecules. When the transcription was compared as a function of time among the reactions with 1 μM $(dT)_{34}$ or $(dA)_{34}$ (Fig. 3B) or without oligonucleotides (Fig. 2A), the ratio of full-length to premature transcripts for the transcription with $(dT)_{34}$ decreased without a lag (summarized in Fig. 3C). In contrast, the ratio for the transcription with $(dA)_{34}$ first increased to 2.4-fold between 1 and 10 min, but decreased later to the original level. These results indicate that transcription was also inhibited in the presence of $(dT)_{34}$, although the inhibition occurred without a delay. We found, however, that when $(dA)_{34}$ was present in the reaction mixture, the degree of inhibition was apparently reduced. This could be explained by the oligonucleotide forming a duplex with the $(U)_{34}$ sequence in the transcripts and, as a result, effectively removing the free transcripts. If this was the case, $(U)_{34}$-containing transcripts interacted with the $(dA)_{34}\cdot(dT)_{34}$ tract, which would result in inhibition of transcription.

We also examined the transcription in the presence of the 44 bp duplex DNA containing a $(dA)_{34}\cdot(dT)_{34}$ tract in the middle (Fig. 3D). While there was no effect of this DNA at 0.01 and 0.1 μM (data not shown), the suppression of inhibition was observed at 1 and 4 μM (Fig. 3D, left or right panel, respectively). Furthermore, an apparent reduction of some of the premature transcripts (shown by arrowheads) was observed between them.

Inhibition of transcription by T7 RNA polymerase and RNA polymerase II

Inhibition of transcription by the $(dA)_{34}\cdot(dT)_{34}$ tract was also examined with other RNA polymerases. Figure 4 shows transcription with T7 RNA polymerase and eukaryotic RNA polymerase II using pATB12 and pATε3-32, containing the promoter from the human ε-globin gene, respectively. In both cases, transcription was inhibited in the middle and/or in front of the $(dT)_{34}$ tract. However, the degrees and the positions of the inhibition varied among the polymerases.

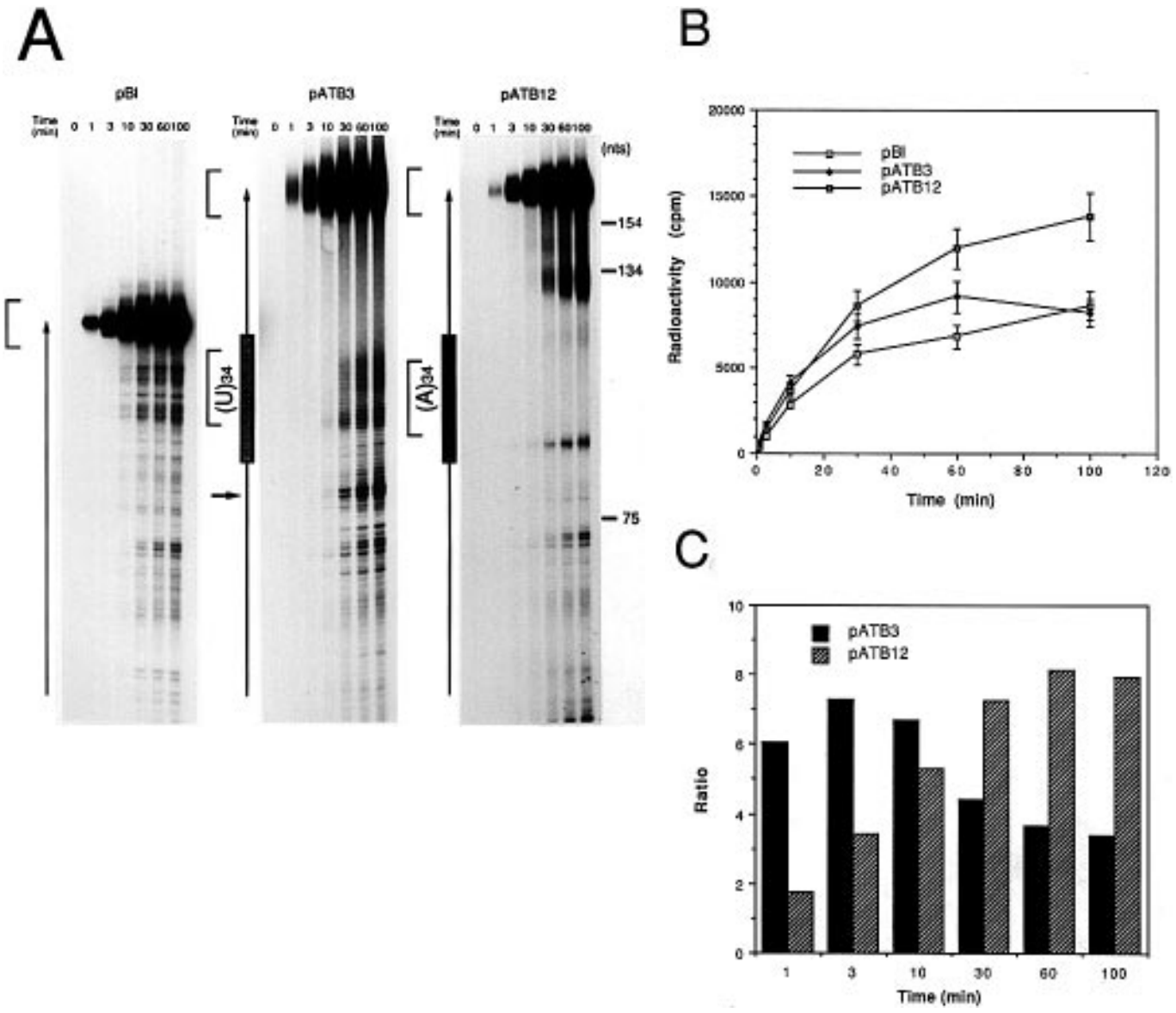


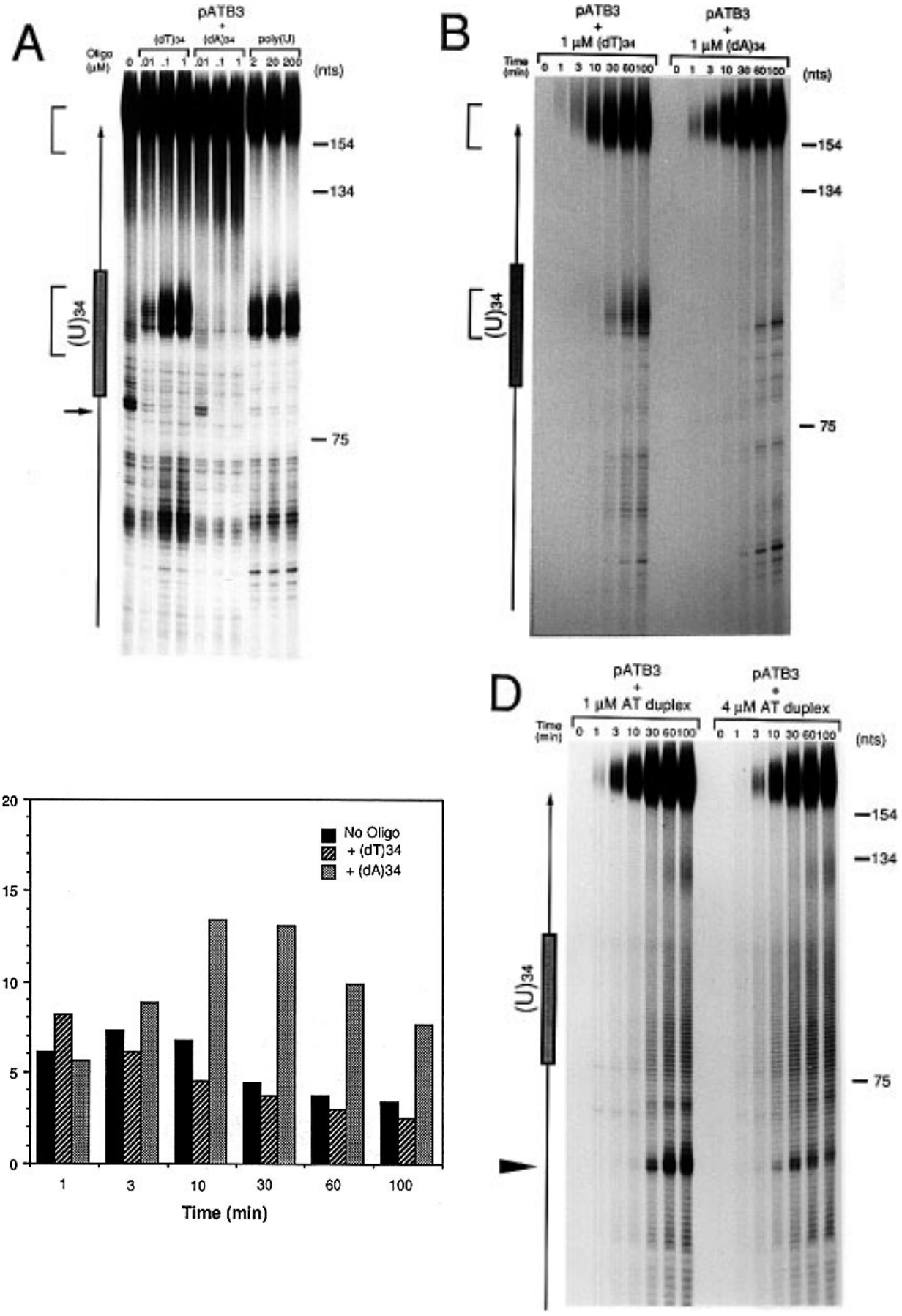
Figure 2. Run-off assay with T3 RNA polymerase. (**A**) *In vitro* transcription (run-off assay) of $(dA)_{34}\cdot(dT)_{34}$-containing plasmids (pATB3 and pATB12) or the control vector (pBl). One hundred ng of plasmid DNA was used for transcription, and the transcripts were resolved on a 6% polyacrylamide–7 M urea gel. The full-length and the premature transcripts within the tract for pATB3 are bracketed. The premature transcripts are indicated by an arrow. (**B**) The radioactivities of the full-length transcripts [upper bracketed regions in (A)] were plotted against the reaction time. The vertical bars indicate the deviation (± 5%) between two identical experiments. (**C**) The molecular ratios of the full-length to premature transcripts for pATB3 and pATB12 [lower bracketed region in (A)] are plotted. The whole region of the $(dA)_{34}\cdot(dT)_{34}$ tract was used as the premature transcripts for pATB12.

RNA·DNA complex formation

To examine the strand specificity of the interaction between the plasmid DNA and transcripts containing a poly(U) sequence, the plasmid (pUC19, pAT19, pUC18 or pAT18) digested with *Xmn*I was mixed with ^{32}P-labeled poly(A) or poly(U) and subjected to electrophoresis in a buffer containing Mg^{2+} (Fig. 5). Complex formation was observed only in the combination of pAT19 or pAT18 and poly(U) (Fig. 5, lanes 8 and 10), indicating that RNA transcripts containing a poly(U) sequence, but not poly(A), can interact or form an RNA·DNA complex with the template plasmid DNA at the $(dA)_{34}\cdot(dT)_{34}$ tract. pAT18 has the ability to form a complex if poly(U) is supplied. Mg^{2+} was required for this complex formation (data not shown).

DISCUSSION

In this paper, we have provided evidence suggesting that accumulated transcription products containing poly(U) sequences inhibit further transcription. This inhibitory effect could be explained as a result of either of two mechanisms: (i) that a complex was formed between templates and the transcripts and blocked further transcription; or (ii) that RNA polymerases detached from the templates at the $(dA)_{34}\cdot(dT)_{34}$ run. We are, however, inclined to believe that the effect was caused by the former mechanism for


A
pATB3
+
(dT)34
(dA)34
poly(U)
Oligo (μM)
0 .01 .1 1 .01 .1 1 2 20 200
(nts)
154
134
75
(U)34
B
pATB3
+
1 μM (dT)34
pATB3
+
1 μM (dA)34
Time (min)
0 1 3 10 30 60 100
0 1 3 10 30 60 100
(nts)
154
134
75
(U)34
C
Ratio
20
15
10
5
0
No Oligo
+ (dT)34
+ (dA)34
1 3 10 30 60 100
Time (min)
D
pATB3
+
1 μM AT duplex
pATB3
+
4 μM AT duplex
Time (min)
0 1 3 10 30 60 100
0 1 3 10 30 60 100
(nts)
154
134
75
(U)34

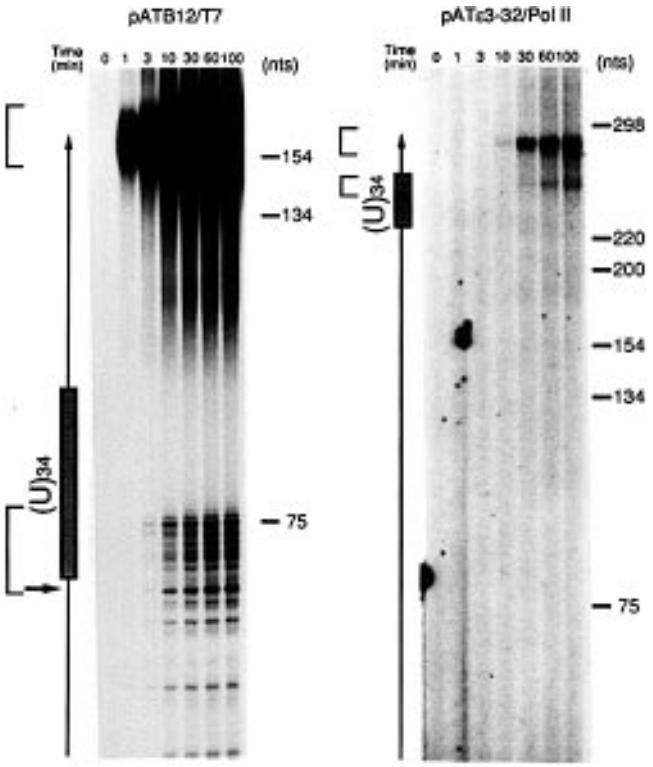


Figure 4. Run-off assay with T7 RNA polymerase or eukaryotic RNA polymerase II. One hundred ng of pATB12 linearized with *Kpn*I or 1 µg of pATε3-32 linearized with *Xho*I was used for *in vitro* transcription for the indicated lengths of time. The positions of the full-length (162 nt for pATB12 and 279 nt for pATε3-32) and the premature transcripts (indicated by an arrow for those terminated before the $(dA)_{34}\cdot(dT)_{34}$ tract and by brackets for those within the tract and the full-length transcripts) are indicated.

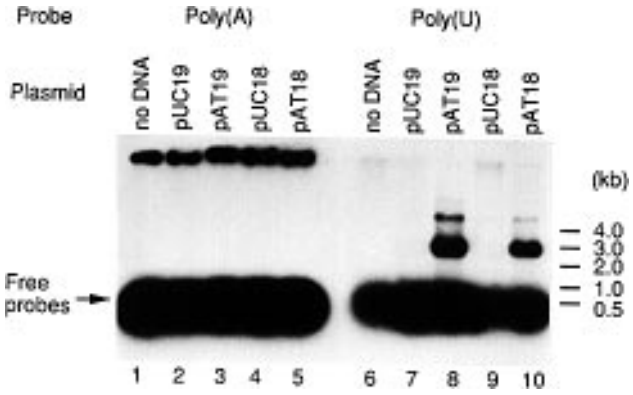


Figure 5. Detection of RNA·DNA complex. The linearized plasmid DNA, pUC19, pAT19, pUC18 or pAT18 (2.95 kb), or no DNA as a control, was mixed with ^{32}P-labeled poly(A) or poly(U) (40–50 nt), incubated for 30 min at 37°C and electrophoresed at a neutral pH in the presence of Mg^{2+}. The gel was then fixed and autoradiographed.

the following reasons. First, the relative ratio of the premature to full-length transcripts increased as the transcription products accumulated (Fig. 2B and C). Secondly, when $(dA)_{34}$ or duplex $(dA)_{34}\cdot(dT)_{34}$ was present in the reaction mixture, the degree of inhibition was reduced (Fig. 3). Thirdly, addition of $(dT)_{34}$ caused transcriptional inhibition, but without a time lag. If the former is the case, a complex between the poly(U)-containing transcripts and the poly(dA)·poly(dT) tract in the templates is likely to be responsible for the observed inhibition.

The molecular structure of the complex between poly(U) and poly(dA)·poly(dT) is probably an RNA·DNA triplex. As described above, the inhibition was observed between the template poly(dA)·poly(dT) and the transcript poly(U), but not poly(A) (Table 3 and Fig. 2). It has been reported that such an RNA·DNA triplex between poly(dA)·poly(dT) and poly(U) can be stably formed in the presence of magnesium ions (38,39). We also observed strand-specific complex formation between the plasmid DNA and RNA by a gel assay (Fig. 5). Since poly(dA)·poly(dT) tracts adopt a rigid structure (40), it is not likely to form an intramolecular triplex or H-DNA at the tracts which would leave a part of the tract unpaired and create the binding site for poly(U). However, since a complex between the template plasmid DNA and nascent synthesized transcripts was also detected (data not shown), it is possible that 5′-part of the transcript can interact with the tract and form a stable complex while the reaction is still proceeding. Such a complex would be stabilized by additional hydrogen bonds and thus have a higher melting temperature than that for the complex formed by the completely intermolecular triplex formation. Meanwhile, the presence of the premature transcripts immediately proximal to the poly(dA)·poly(dT) tract (Fig. 2A) could be due to the complex formation throughout the tract, which blocks the transcription machinery to enter there. Adding $(dT)_{34}$, however, decreased the premature transcripts at the proximal position, indicating that there seems to be a difference between the complex formed with the transcripts and the one with $(dT)_{34}$ or poly(U). This difference could be due to the presence of the franking non-poly(U) region in the transcripts.

The inhibition became quite prominent when the transcripts were accumulated to a certain level. The concentration of transcripts must have reached ~4 nM after 10 min incubation when the inhibition started to occur. This was also the case in the reaction in the presence of at least 10 nM $(dT)_{34}$ (Fig. 3A). We estimated previously that the apparent dissociation constant of $(dT)_{34}\cdot(dA)_{34}\cdot(dT)_{34}$ to duplex and single-stranded DNA is ~2 × 10^{-8} M (37). Provided that half of the template plasmid DNA (100 ng/20 µl, 2.6 nM) in the reaction mixture formed a triplex and that the dissociation constant is roughly equivalent to that between RNA and DNA, this implies that roughly a concentration of the third strand of 20 nM is required for triplex formation, which is well in the range of the concentration of $(dT)_{34}$ observed to cause transcriptional inhibition (between 10 and 100 nM; Fig. 3A).

Figure 3. Run-off assay with T3 RNA polymerase in the presence of oligodeoxyribo- or oligoribonucleotides. (**A**) *In vitro* transcription (60 min at 37°C) of pATB3 in the presence of the indicated concentrations of $(dT)_{34}$, $(dA)_{34}$ or poly(U). Molecule-equivalent values for $(dT)_{34}$ and $(dA)_{34}$, and nucleotide-equivalent values for poly(U) (40–50 nt long) are indicated. The full-length and the premature transcripts are bracketed. The premature transcripts terminated before the tract are indicated by an arrow. (**B**) Time course of transcription in the presence of 1 µM $(dT)_{34}$ (on the left) or $(dA)_{34}$ (on the right). (**C**) The time courses of the molecular ratio (full-length to premature transcripts) were compared between transcription in the presence of $(dT)_{34}$ or $(dA)_{34}$, and in the absence of the oligonucleotides (reproduced from Fig. 2C). (**D**) Time course of transcription in the presence of 1 µM (on the left) or 4 µM (on the right) of the 44 bp duplex DNA with the nucleotide sequence AGCTT$(A)_{34}$GGATC (AT duplex). Arrowheads indicate the premature transcripts that showed a decrease of intensity in the presence of 4µM AT duplex compared with that in the presence of 1 µM AT duplex.

Biological effects of the inhibition

One of the interesting suggestions from the results presented here is that among many mRNA species, those containing poly(U) may be underrepresented. A survey of the frequencies of poly(U) or poly(A) sequences in the coding region of mRNAs from various species (total 61248 peptide sequences) indicated that polyphenylalanines (derived from UUU codon) were underrepresented while polylysines (derived from AAA codon) appeared ~20 times more frequently. Although this could simply be due to the underrepresentation of phenylalanine residues in proteins compared with lysines, there is a possibility that the synthesis of poly(U)-containing mRNAs is unfavorable, as presented here. Furthermore, long poly(U) sequences in the coding region tend to be located in the middle or in the 3′ region of the coding region (data not shown). As reported previously, the effect of such tracts is low when they are placed far from the promoter region (32). In any case, inefficient mRNA synthesis appears to have been avoided in many species by underrepresenting these sequences.

The major source of the poly(dA)·poly(dT) tracts in the eukaryotic genomes is poly(dA) tails of peudogenes and retroposons. In the human genome, *Alu* sequences appear approximately every 3 kb and 80% of them contain perfect A stretches of >10 bp (41). Although the biological and evolutional significance of these elements are not known, one of the potential functions would be the regulation of the transcriptional direction. In the human β-globin locus, where all five active genes including the ε-globin gene face the same direction, six of the seven *Alu* sequences having a poly(dA) tail of at least 10 bp face the same direction as that of the globin genes (42). Several alternative transcription initiation sites are located at up to –4.5 kb region of the ε-globin gene (43). Although two *Alu* sequences located at –2.5 and –1.5 kb from the ε-globin gene face in opposite directions to each other, only the one facing the same direction as that of the globin genes contains a long stretch of polypurine·polypyrimidine sequences induding $(A)_{14}$. Therefore, the transcription in the reverse direction could be disfavored by the transcription that creates poly(U)-containing transcripts.

Cloning strategy using *lacZ* expression vectors

The relationship of the transcriptional inhibition presented here with the *in vivo* events merits further examination. As we reported previously, however, the replication of *E.coli* plasmids containing a poly(dA)·poly(dT) tract was inhibited when the poly(dT) strand was transcribed by a strong expression system, *lacZ* for example (32). This instability was further enhanced by the presence of IPTG (data now shown) which is widely used for *lacZ* expression. Therefore, the general cloning strategy using the vectors with multiple cloning sites located within the *lacZ* coding region could cause a selection among the cloned sequences. The stability was restored by the usage of pBR322 as a cloning vector (32), which is a low copy-number plasmid and lacks a strong and inducible expression system.

ACKNOWLEDGEMENTS

This work was supported by Grants-in-Aid from the Ministry of Education of Japan. We thank Yuko Wada-Kiyama for information regarding the eukaryotic *in vitro* transcription system.

REFERENCES

1 Greenblatt, J. (1991) *Trends Biochem. Sci.* **16,** 408–411.
2 Johnson, P. F. and McKnight, S. L. (1989) *Annu. Rev. Biochem.* **58,** 799–839.
3 Kornberg, R. D. and Lorch, Y. (1992) *Annu. Rev. Cell Biol.* **8,** 563–587.
4 McClure, W. R. (1985) *Annu. Rev. Biochem.* **54,** 171–204.
5 Roeder, R. G. (1991) *Trends Biochem. Sci.* **16,** 402–408.
6 von Hippel, P. H., Bear, D. G., Morgan, W. D. and McSwiggen, J. A. (1984) *Annu. Rev. Biochem.* **53,** 389–446.
7 Chan, P. T. and Lebowitz, J. (1990) *J. Biol. Chem.* **265,** 4091–4097.
8 Chen, W., Tabor, S. and Struhl, K. (1987) *Cell* **50,** 1047–1055.
9 Struhl, K. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 8419–8423.
10 Wagner, L. A., Weiss, R. B., Driscoll, R., Dunn, D. S. and Gesteland, R. F. (1990) *Nucleic Acids Res.* **18,** 3529–3535.
11 Auble, D. T., Allen, T. L. and deHaseth, P. L. (1986) *J. Biol. Chem.* **261,** 11202–11206.
12 Collis, C. M., Molloy, P. L., Both, G. W. and Drew, H. R. (1989) *Nucleic Acids Res.* **17,** 9447–9468.
13 Hansen, U. M. and McClure, W. R. (1980) *J. Biol. Chem.* **255,** 9564–9570.
14 Asai, T., Takanami, M. and Imai, M. (1990) *EMBO J.* **9,** 4065–4072.
15 Lapidot, A., Baran, N. and Manor, H. (1989) *Nucleic Acids Res.* **17,** 883–900.
16 Schultes, N. P. and Szostak, J. W. (1991) *Mol. Cell. Biol.* **11,** 322–328.
17 Weinreb, A., Katzenberg, D. R., Gilmore, G. L. and Birshtein, B. K. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 529–533.
18 Gut, S. H., Bischoff, M., Hobi, R. and Kuenzle, C. C. (1987) *Nucleic Acids Res.* **15,** 9691–9705.
19 Yee, H. A., Wong, A. K. C., van de Sande, J. H. and Rattner, J. B. (1991) *Nucleic Acids Res.* **19,** 949–953.
20 Naylor, L. H. and Clark, E. M. (1990) *Nucleic Acids Res.* **18,** 1595–1601.
21 Nordheim, A. and Rich, A. (1983) *Nature* **303,** 674–679.
22 Crothers, D. M., Haran, T. E. and Nadeau, J. G. (1990) *J. Biol. Chem.* **265,** 7093–7096.
23 Hagerman, P. J. (1990) *Annu. Rev. Biochem.* **59,** 755–781.
24 Milligan, J. F. and Uhlenbeck, O. C. (1989) *Methods Enzymol.* **180,** 51–62.
25 Maher, L. J. III, Dervan, P. B. and Wold, B. (1992) *Biochemistry* **31,** 70–81.
26 Skoog, J. U. and Maher, L. J. III (1993) *Nucleic Acids Res.* **21,** 2131–2138.
27 Duval-Valentin, G., Thuong, N. T. and Helene, C. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 504–508.
28 Young, S. L., Krawczyk, S. H., Matteucci, M. D. and Toole, J. J. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10023–10026.
29 Skoog, J. U. and Maher, L. J. III (1993) *Nucleic Acids Res.* **21,** 4055–4058.
30 Hanvey, J. C., Peffer, N. J., Bisi, J. E., Thompson, S. A., Cadilla, R., Josey, J. A., Ricca, D. J., Hassman, C. F., Bonham, M. A., Au, K. G., Carter, S. G., Bruckenstein, D. A., Boyd, A. L., Noble, S. A. and Babiss, L. E. (1992) *Science* **258,** 1481–1485.
31 Kiyama, R., Nishikawa, N. and Oishi, M. (1994) *J. Mol. Biol.* **237,** 193–200.
32 Kiyama, R. and Oishi, M. (1994) *Gene* **150,** 57–61.
33 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33,** 103–119.
34 Geliebter, J., Zeff, R. A., Schultze, D. H., Pease, L. R., Weiss, E. H., Mellor, A. L., Flavell, R. A. and Nathenson, S. G. (1986) *Mol. Cell. Biol.* **6,** 645–652.
35 Wada-Kiyama, Y., Peters, B. and Noguchi, C. T. (1992) *J. Biol. Chem.* **267**, 11532–11538.
36 Dignam, J. D., Lebowitz, R. M. and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
37 Kiyama, R. and Camerini-Otero, R. D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10450–10454.
38 Han, H. and Dervan, P. B. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 3806–3810.
39 Riley, M., Maling, B. and Chamberlin, M. J. (1966) *J. Mol. Biol.* **20,** 359–389.
40 Fox, K. R. (1990) *Nucleic Acids Res.* **18,** 5387–5391.
41 Economou, E. P., Bergen, A. W., Warren, A. C. and Antonarakis, S. E. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2951–2954.
42 Behe, M. J. (1987) *Biochemistry* **26,** 7870–7875.
43 Allan, M., Lanyon, W. G. and Paul, J. (1983) *Cell* **35,** 187–197.