

Specific Inhibition of *in Vitro* Transcription Elongation by Triplex-Forming Oligonucleotide–Intercalator Conjugates Targeted to HIV Proviral DNA[†]

Carine Giovannangeli,^{*,‡} Loïc Perrouault,[‡] Christophe Escudé,[‡] Nguyen Thuong,[§] and Claude Hélène[‡]

Laboratoire de Biophysique, Muséum National d'Histoire Naturelle, INSERM U 201, CNRS URA 481, 43 rue Cuvier, 75231 Paris Cedex 05, France, and Centre de Biophysique Moléculaire, CNRS UPR 4301, Rue Charles Sadron, 45071 Orléans Cédex 02, France

Received December 19, 1995; Revised Manuscript Received April 8, 1996[®]

ABSTRACT: A 16-base pair oligo(purine)•oligo(pyrimidine) sequence present in the coding region of two HIV1 proviral genes (*pol* and *nef*) was chosen as a target for triplex-forming oligonucleotides in *in vitro* transcription assays. Inhibition of transcription elongation was observed with triplex-forming oligonucleotide–acridine conjugates (Acr-15-TCG:5'-Acr-T₄CT₄G₆-3' and Acr-9-TC:5'-Acr-T₄CT₄-3' where C is 5-methylcytosine) under conditions where the unsubstituted oligomers did not exhibit any inhibitory effect. Both SP6 bacteriophage RNA polymerase and eukaryotic RNA polymerase II were physically blocked by such a triplex barrier. The polymerase arrest is caused by the triple-helical complex involving the hydrogen-bonded oligonucleotide stabilized by the intercalated moiety and not solely by the acridine molecule specifically intercalated at the duplex–triplex junction. The stability of the triple-helical complex formed by the 15-mer containing thymines, cytosine, and guanines (15-TCG) and involving the formation of six contiguous C•GxG base triplets was strongly enhanced in the presence of a benzopyrindole derivative (BePI), which intercalates in triplex structures. This improvement of the binding affinity led to an increased inhibition of transcription elongation. The present results demonstrate the necessity to use triplex-forming oligonucleotides with high binding affinity and a long residence time on their double-stranded target to efficiently inhibit transcription elongation. These data provide a rational basis for the optimization and the development of triplex-forming oligonucleotides as transcriptional blockers, even when they are targeted to the transcribed portion of a gene, downstream of the transcription initiation site.

Oligonucleotide-directed triplex formation could be expected to inhibit all stages of the transcriptional cycle, i.e., formation of the active promoter complex, initiation, and elongation. A large number of studies have focused on the ability of triple-helical complexes to inhibit transcription initiation by preventing the binding of transcription factors [Cooney et al., 1988; Maher et al., 1992; Grigoriev et al., 1992, 1993; for a review see Thuong and Hélène (1993) and Kim and Miller (1995)]. A 13-nt triplex-forming oligonucleotide targeted to a sequence located just downstream of the *bla* promoter has been shown to prevent the initiation of transcription by *Escherichia coli* RNA polymerase *in vitro* (Duval-Valentin et al., 1992). Only a few studies have focused on intermolecular triple helices positioned far downstream of the transcription initiation site, so that they are presumed to interfere with transcription elongation by acting as a roadblock to RNA polymerase moving along the DNA. With bacteriophage RNA polymerases, these studies describe no arrest of SP6 RNA polymerase using 2 μ M 22-nt (T,G)-third strand at 37 °C (Skoog & Maher, 1993), moderate inhibitory effects on T3 RNA polymerase using 4 μ M 45-nt (T,G)-third strand at 37 °C, with both specific and nonspecific inhibitions (Rando et al., 1994), and inhibition of T7 RNA polymerase using 15 μ M 20-nt (T,C)-third strand

containing several phosphorothioate groups at 25 °C, but without evidence of any truncated transcript (Xodo et al., 1994). With eukaryotic RNA polymerase II, a low efficiency of arrest was described using 20 μ M 15-nt third strand forming a pH-independent triplex within a G-free transcription cassette at 30 °C (Young et al., 1991). In contrast, several studies have reported that intra- and intermolecular triplexes can inhibit DNA polymerase elongation (Dayn et al., 1992; Giovannangeli et al., 1993; Samadashwily et al., 1993; Samadashwily & Mirkin, 1994; Hacia et al., 1994).

Thus, triple helices do not appear, until now, to be very effective at blocking elongation by either prokaryotic or eukaryotic RNA polymerases, whereas triplex effects on transcription initiation are described to be much more robust. This is the reason why it is commonly thought that inhibitors based on triple-helix formation may be limited to the gene control region. This idea is supported in part by the fact that RNA polymerases are multiprotein complexes replete with DNA melting and helicase activities. Moreover, the phage RNA polymerases are known either to displace a number of protein complexes or to bypass tight-binding proteins with substantial levels of read-through transcription existing beyond the protein block. These include the lac repressor–operator complex (Dubendorf & Studier, 1991; Giordano et al., 1989), the *Xenopus* 5S gene transcription complex (Wolffe et al., 1986), and a tight-binding *EcoRI* mutant protein (Pavco & Steege, 1991).

The transcribed region is usually much longer than the associated regulatory sequences of a given gene and provides more target sites for triplex formation. Therefore, the

[†] Supported by the French Agency for AIDS (ANRS).

^{*} Author to whom correspondence should be addressed: e-mail, giovanna@mnhn.fr.

[‡] Muséum National d'Histoire Naturelle.

[§] Centre de Biophysique Moléculaire.

[®] Abstract published in *Advance ACS Abstracts*, August 1, 1996.

development of biological applications of triplex-forming oligonucleotides would be strongly reinforced if it was possible to design oligonucleotides acting as efficient roadblocks for elongating RNA polymerases.

The sequence that we have chosen as a target in our studies is a 16-base pair (bp) oligo(purine)•oligo(pyrimidine) sequence (called the poly(purine) tract sequence and abbreviated as PPT)¹ that is conserved in all HIV1 strains and is present twice in the HIV proviral DNA: (i) in the 3'-part of the *pol* gene coding for the integrase (IN) viral protein and (ii) in the *nef* gene upstream of the 3'-LTR. Our goal was to use triplex-forming oligonucleotides to specifically inhibit proviral gene expression at the transcriptional level.

We present here two *in vitro* transcription model systems that we have developed using either a phage RNA polymerase (SP6) or eukaryotic transcriptional machinery (RNA polymerase II). We have examined whether and how intermolecular triplex structures formed on the PPT sequence affect RNA synthesis of the viral HIV1 genes under the control of either the bacteriophage SP6 promoter or the eukaryotic CMV promoter. The use of these model systems has allowed us to determine conditions under which a 15-mer oligonucleotide, containing T, C, and G nucleotides, acts as a specific transcriptional inhibitor. We have achieved this goal by the simultaneous use of an oligonucleotide–intercalator conjugate and a benzo[e]pyridoindole derivative (BePI) acting as a triplex stabilizer (Mergny et al., 1992). Both the covalent modification and the noncovalent interaction induced an enhancement in triplex stability. Under these conditions, the triplex was able to physically block RNA polymerases tracking along the DNA with great efficiency.

MATERIALS AND METHODS

Reagents. Restriction endonucleases were purchased from New England Biolabs. SP6 RNA polymerase and ribonucleoside triphosphates were from Epicentre Technologies. [α -³²P]GTP was purchased from ICN and RNase inhibitor from Promega.

The benzopyridoindole derivative (see Figure 5 for description) was synthesized as previously described (Nguyen et al., 1992).

Oligonucleotides and Transcription Templates. The unmodified oligonucleotides used in this study were obtained from Eurogentec. The oligonucleotide–acridine conjugates were synthesized as previously described (Asseline et al., 1984; Thuong & Chassignol, 1988).

The plasmids called pRP159 and pRP156 (gifts from Dr. D. van Gent, Netherlands Cancer Institute, Amsterdam) were constructed by insertion of HIV1 and HIV2 integrase genes (*IN*) (~0.9 kbp), respectively, into the pSP65 host vector (Promega) behind the SP6 promoter. Plasmid pSG-F47 (a gift of O. Schwartz, Institut Pasteur, Paris) containing the *nef* gene under the control of the cytomegalovirus (CMV) promoter was obtained from pTG-1147 plasmid (Schwartz et al., 1993).

In Vitro Transcription Assays. For the SP6 promoter, transcription reactions were performed in a 10 μ L solution containing 40 mM Tris-HCl, 6 mM MgCl₂, 2 mM spermidine, 4 mM dithiothreitol, and 1 unit/ μ L RNase inhibitor with

linearized template DNA (0.5 μ g ~ 10 nM). The final nucleotide concentrations were 500 μ M for ATP, CTP, and UTP, 100 μ M GTP, and 0.3 μ M [α -³²P]GTP (2 μ Ci, 650 Ci/mmol). After incubations with different oligonucleotides, transcription was initiated by the addition of SP6 RNA polymerase (25 units ~ 25 nM). Transcription reactions were terminated by ethanol precipitation. Transcription templates pRP159 and pRP156 were cut with the restriction endonucleases *Nde*I and *Hind*III, respectively, to create a terminus for run-off transcripts. To create RNA markers, templates were partially digested by *Dra*I and transcribed under the same conditions (see Figure 1 for details).

For the CMV promoter, *in vitro* transcription reactions were performed using a HeLa cell nuclear extract (Promega). The DNA template was obtained by digestion of pSG-F47 by *Eco*RI and *Nhe*I; *Nhe*I was used to prevent transcription from another promoter (murine LTR) present in the retroviral vector. This transcription template (1 μ g) was incubated in the absence or presence of different oligonucleotides, and transcription was initiated by the addition of HeLa nuclear extracts (eight standardized transcription units) to a final volume of 25 μ L containing 20 mM Hepes (pH 7.9), 100 mM KCl, 3 mM MgCl₂, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol, 400 μ M ATP, CTP, and UTP, 20 μ M GTP, and 0.4 μ M [α -³²P]GTP. Transcription was allowed to proceed for 60 min at 25 °C and then stopped by a solution containing 0.3 M Tris-HCl (pH 7.4), 0.3 M NaOAc, 0.5% SDS, 2 mM EDTA, and 3 μ L/mL tRNA. Transcripts were extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and ethanol precipitated.

Radiolabeled RNA transcripts were analyzed by electrophoresis on 6% denaturing polyacrylamide gels. Quantification (\pm 10%) of the specific inhibition of transcription elongation was achieved by phosphorimager analysis. Inhibition was expressed in terms of percentage of truncated transcript; every quantitative analysis presented in this work assumed that radioactivity was proportional to the length of the RNA fragments.

Spectroscopic Methods. Absorption spectra were recorded with a Uvikon 940 spectrophotometer. Melting curves were obtained by decreasing and then increasing the temperature of the 500 μ L samples at a rate of 0.1 °C/min. Triplex mixtures were obtained by mixing equimolar amounts of the three strands (1.0 μ M) in the absence or presence of 10 μ M BePI.

RESULTS

Specific Inhibition of Transcription Elongation by Triplex-Forming Oligonucleotide–Intercalator Conjugates. The effects of triple-helical DNA complexes on transcription elongation were tested in two model systems: (i) the HIV1-*IN* gene containing the 16-base pair oligo(purine)•oligo(pyrimidine) PPT sequence placed under the control of the bacteriophage SP6 promoter in the pRP159 plasmid; (ii) the HIV1-*nef* gene also containing the PPT sequence placed under the control of the eukaryotic CMV promoter in the pSG-F47 plasmid.

In previous studies (Giovannangeli et al., 1992a,b), we have demonstrated that this 16-bp tract (PPT) of HIV proviral DNA allowed triplex formation with an oligonucleotide containing the three bases thymine (T), 5-methylcytosine (C), and guanine (G) (15-TCG:5'-T₄CT₄G₆-3'). The 15-TCG

¹ Abbreviations: BePI, 1-[(amino)propylamino]-4-methyl-5H-9-methoxybenzo[e]pyrido[4,3-b]indole; PPT, polypurine tract; HIV, human immunodeficiency virus; Pol II, RNA polymerase II.

BePI	Acr-(5')	9-TC	9-TC	16-TC	15-TCG
—	—	8	12	22	34
—	+	—	36	—	52
+	—	26	—	30	56

oligomer was bound in a parallel orientation with respect to the oligo(purine) strand and formed a triplex in an almost pH-independent manner, in contrast to the standard T₃C- or T₃ C-containing third strands (16-TC:5'-T₄CT₄C₆T-3') whose binding was strongly pH dependent (Giovannangeli et al., 1992a). At neutral pH the six contiguous cytosines did not

In the bacteriophage system, the two partners (DNA duplex or triplex and SP6 RNA polymerase) were used in a

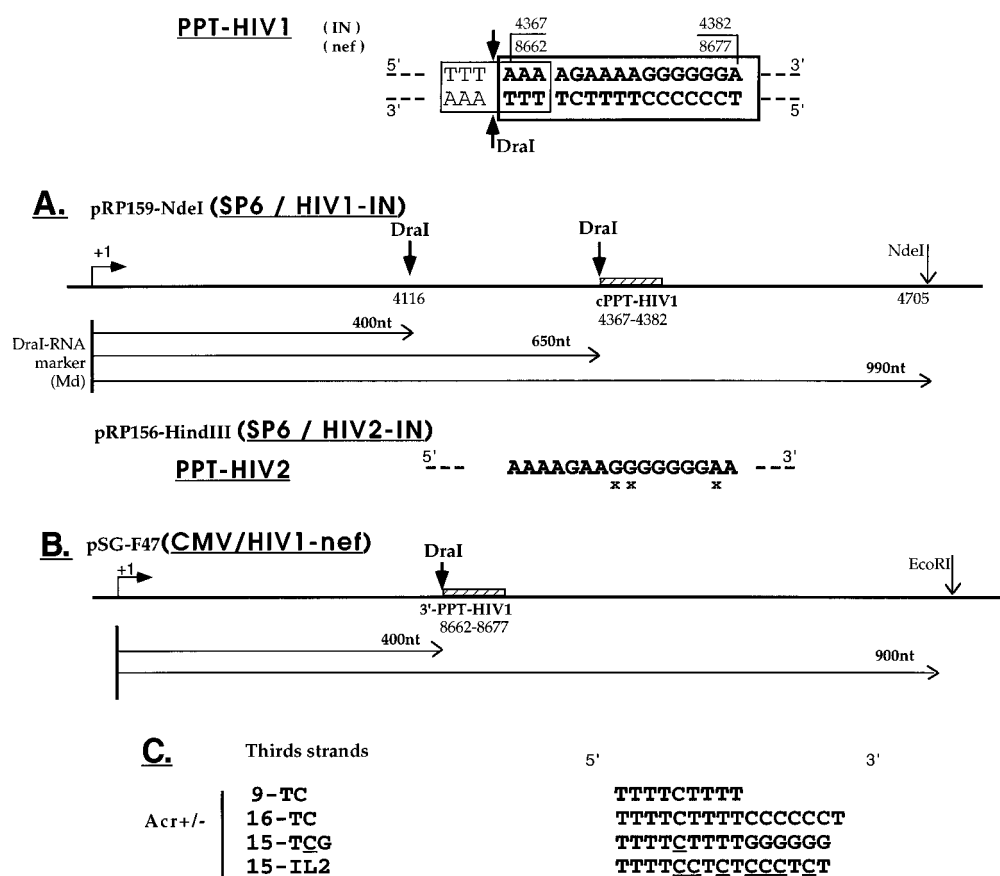


FIGURE 1: Experimental design. The 16-bp oligo(purine)·oligo(pyrimidine) sequence of HIV proviral DNA that we have chosen as a target in the present work is indicated. It is named the poly(purine) tract (PPT) sequence. This sequence is present in two HIV genes: the *pol* gene within the region coding for the integrase (*IN*) (cPPT: 4367–4382) and the *nef* gene (3'-PPT: 8662–8677). The *Dra*I recognition sequence (TTTAAA) is boxed with a thin line. Indicated numbers refer to positions on the HIV1 genome (HIV1-BRU strain; Wain-Hobson et al., 1985). (A) Top: Scheme of the DNA template (pRP159, linearized at the *Nde*I restriction site, position 4705) used for *in vitro* transcription assays showing the site of transcription initiation (+1) and the 16-bp oligo(purine)·oligo(pyrimidine) target (PPT) present in the *IN* gene (cPPT: 4367–4382). The two *Dra*I cleavage sites at positions 4116 and 4366 are indicated by arrows. One of these sites is located exactly at the 5'-end of the oligo(purine) target for triple-helix formation. This allowed us to generate a radiolabeled RNA size marker that localized exactly one extremity of the triplex-binding site. When the *Nde*I-linearized plasmid was partially digested by *Dra*I and used for a run-off transcription reaction in standard conditions, three transcripts ("Dra-I-RNA marker", abbreviated as Md) were obtained whose lengths are 400, 650, and 990 nucleotides (nt), as shown on the scheme. Bottom: Sequence of the mutated sequence in the pRP156 plasmid that contains the HIV2-*IN* gene downstream of the SP6 promoter. HIV2-*IN* contains a 16-bp oligo(purine)·oligo(pyrimidine) sequence (called PPT-HIV2) at the same position as in the HIV1-*IN* gene; this HIV2 sequence differs from that of HIV1 at three positions (indicated by x). (B) Eukaryotic promoter (CMV). The pSG-F47 plasmid contains the HIV1-*nef* gene. The 16-bp oligo(purine)·oligo(pyrimidine) sequence (3'-PPT: 8662–8677) present in the *nef* gene is positioned 400 nt downstream of the transcription initiation site. (C) Sequences of the third strands used in this work. Cytosine, that are methylated at carbon 5 are underlined (C). Acr indicates 2-methoxy-6-chloro-9-aminoacridine; a pentamethylene linker was used to link the 9-amino group to the 5'-phosphate of the oligonucleotide (Asseline et al., 1984; Thuong & Chassignol, 1988).

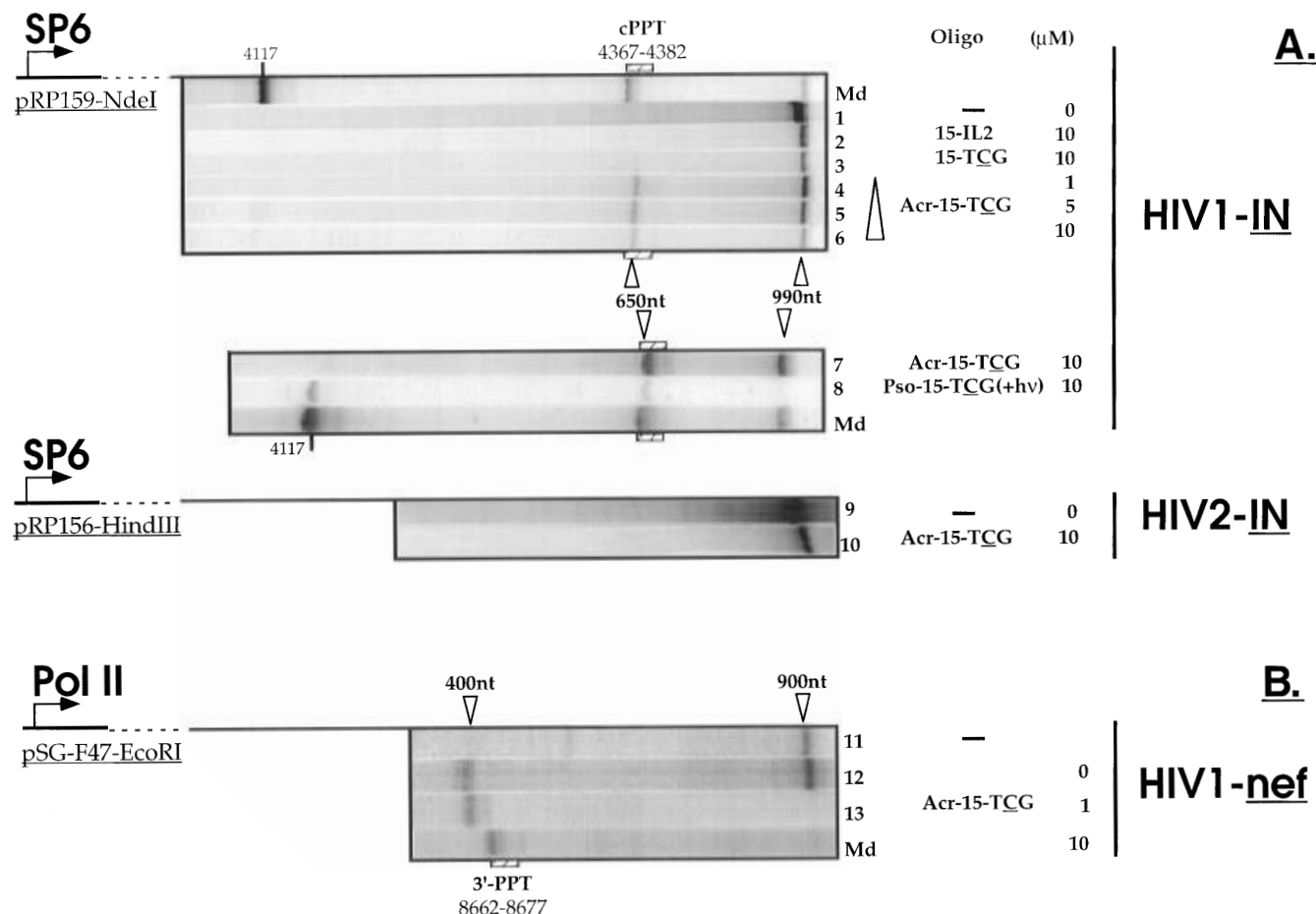


FIGURE 2: Specific inhibition of transcription elongation by triplex-forming oligonucleotide–intercalator conjugates. (A) Bacteriophage system. Linear plasmids (pRP159-NdeI/HIV1 and pRP156-HindIII/HIV2, see Figure 1) were incubated in the absence (lanes 1 and 9) or presence of unmodified oligonucleotide 15-TCG (10 μ M, lane 3), the acridine conjugate Acr-15-TCG (1, 5, 10 μ M, as indicated near the gels, lanes 4–7 and 10), a control oligonucleotide–acridine conjugate, Acr-15-IL2 (10 μ M, lane 2), and Pso-15-TCG (10 μ M, cross-linked (+hv) on the plasmid before the transcription reaction, lane 8). The plasmids were used as templates in an *in vitro* transcription reaction using SP6 RNA polymerase for 60 min at pH 7.0 and 25 $^{\circ}$ C (see Materials and Methods for other conditions). (B) Eukaryotic system. The linear plasmid (pSG-F47) was incubated in the absence (lane 11) or presence of increasing concentrations of Acr-15-TCG (as indicated near the gels, lanes 12 and 13) and used as a template in *in vitro* transcription reactions under standard conditions (25 $^{\circ}$ C, see Materials and Methods). Open arrows denote the migration of full-length and truncated transcripts. Location of the PPT triplex site is indicated (hatched box). Lane Md: *Dra*I-RNA marker (see Figure 1).

controlled medium. The only constraints were to choose conditions that allowed us to obtain (i) a detectable 990-nt full-length transcript (in particular, a pH value higher than 7.0 and a temperature of at least 25 $^{\circ}$ C were required) and (ii) efficient binding of the oligonucleotide by formation of a stable triplex. We thus determined the effect of a triple-helical structure on transcription elongation (at pH 7.0 and 25 $^{\circ}$ C) and compared the ability of triplexes of different stabilities to inhibit the SP6 RNA polymerase moving along the DNA. To address this point, the pRP159 plasmid was linearized with *Nde*I (pRP159-*Nde*I) and annealed with the triplex-forming oligonucleotide prior to its use as a template for *in vitro* transcription in standard conditions (as described in Materials and Methods). In this system the SP6 promoter directed the transcription of a 990-nt RNA (Figure 1). As shown in the gel of Figure 2A, templates harboring some of the triple-helical complexes directed the synthesis of a truncated transcript of \sim 650 nt. Thus, Acr-15-TCG was able to inhibit RNA synthesis (Figure 2A, lanes 4–6). The unsubstituted oligonucleotide (15-TCG) did not exhibit any detectable inhibitory effect (Figure 2A, lane 3) under these conditions. Under the conditions of Figure 2A (25 $^{\circ}$ C, pH 7.0) it was impossible to increase the oligonucleotide

concentration far above 10 μ M, because this led to an unspecific inhibitory effect on global RNA synthesis, most likely due to the binding of a high excess of oligonucleotide to the enzyme (compare lanes 1 and 2 and 9 and 10 in Figure 2A).

The specificity of triple-helix-mediated inhibition of transcription was demonstrated according to several criteria. First, RNA synthesis was arrested at the site corresponding to the expected triplex-forming site; the latter was localized by a run-off transcription reaction using the pRP159-*Nde*I plasmid partially digested by *Dra*I, a restriction enzyme that cleaves exactly at the 5'-boundary of the PPT oligo(purine) sequence [Figure 1A (top) and Figure 2A (top), lane Md]. Second, no effect was observed with an unspecific acridine-substituted oligonucleotide (Acr-15-IL2: Acr-5'-T₄C₂T-CTC₃TCT-3' with the same T content as 15-TCG), which has been previously described to inhibit the initiation of transcription from the IL2-R α promoter (Grigoriev et al., 1993) but did not bind to the HIV1-PPT sequence. Third, we used the linearized pRP156 plasmid as a transcription template in the presence of the Acr-15-TCG oligonucleotide; this plasmid contains the 16-bp PPT sequence of HIV2, which differs from the HIV1-PPT sequence at three positions

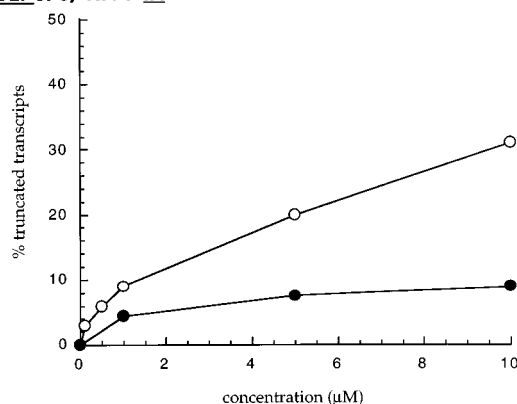
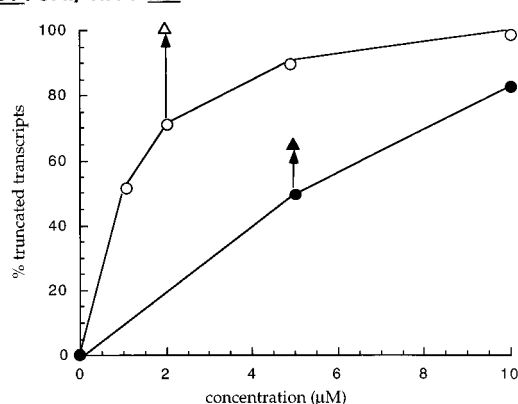
A. SP6/ HIV1-IN**B. Pol II/ HIV1-nef**

FIGURE 3: Quantitation of the truncated transcription products as a function of oligonucleotide concentration: Acr-15-TCG (open symbols) and Acr-9-TC (filled symbols). (A) Bacteriophage system (25 °C, pH 7.0, 60 min). (B) Eukaryotic system (standard transcription conditions: 25 °C, 60 min; for other conditions, see Materials and Methods); 0 μM BePI (circles) and 25 μM BePI (triangles).

(see Figure 1A, bottom). Under conditions (pH 7.0, 25 °C, 10 μM) where a truncated 650-nt transcript was obtained with the pRP159-*Nde*I plasmid (Figure 2A, lane 6), only the full-length transcript was obtained with the pRP156-*Hind*III plasmid (Figure 2A, lane 10).

We then investigated the effect of the triplex formed on the PPT sequence with the oligonucleotide–intercalator conjugates on the eukaryotic transcriptional machinery. The transcription template (pSG-F47-*Eco*RI) containing the HIV1-*nef* gene was incubated in the absence or presence of the third strand. Transcription was initiated by the addition of HeLa nuclear extract as a source of RNA polymerase II in standard conditions (as described in Material and Methods). Transcription from the CMV promoter led to the synthesis of a 900-nt RNA (Figure 2B, lane 11). As in the bacteriophage system, when transcription took place in the presence of triplex-forming oligonucleotide, shorter products appeared (Figure 2B, lanes 12 and 13). They corresponded to an arrest around the PPT site (8662–8677), which was localized by the run-off transcript obtained on the pSG-F47 transcription template digested by *Dra*I (Figure 2B, lane Md). IC₅₀ values of 1 and 5 μM were determined from the concentration dependence of transcription inhibition at 25 °C by Acr-15-TCG and Acr-9-TC, respectively (Figure 3B). It should be noted that eukaryotic RNA polymerase II was more efficiently blocked by a triplex barrier than the bacteriophage SP6 enzyme. In the presence of 1 μM Acr-

15-TCG, 50% of the Pol II transcription was stopped compared to 10% in the SP6 assay. Similarly, for 5 μM Acr-9-TC the values were 50% and 8%, respectively (Figure 3).

Blockage of Transcription Elongation by a Triplex Barrier. The fact that the unsubstituted 15-TCG failed to block both RNA polymerases (SP6 and Pol II) moving along the DNA under conditions (10 μM, 25 °C) where triplex was formed (considering *T_m* values) suggested that the arrest of transcription by acridine–oligonucleotide conjugates might occur when the polymerase encountered the acridine molecule specifically intercalated at the duplex–triplex junction. To test this hypothesis, we used the shorter triplex-forming oligonucleotide Acr-9-TC as a tool in the bacteriophage system. When the transcription reaction took place in the presence of Acr-9-TC, a new phenomenon was observed as compared to Acr-15-TCG (Figure 4A). The expected 650-nt truncated transcript was obtained (as with Acr-15-TCG), but two other shorter products (~400 and 570 nt) appeared. As depicted in Figure 4B, this result could be explained if Acr-9-TC was bound not only to the expected PPT triplex site where nine base triplets could form (called site I on Figure 4) but also to two other sites (called sites II and III) where eight contiguous triplets could be formed. The different patterns obtained with Acr-9-TC and Acr-15-TCG might be explained by the decreased binding of Acr-15-TCG at sites III and II, as compared to Acr-9-TC, due to an electrostatic repulsion between the unbound 3'-most six nucleotides of 15-TCG and the template. We have observed (unpublished results) and it has been reported (Cheng & van Dyke, 1994) that unbound "tails" decrease the stability of triple-helical complexes. Only at site I could Acr-15-TCG form 15 base triplets instead of 9 for Acr-9-TC, and enhanced binding was observed (see Table 1, Δ*T_m* = 16 °C).

Several characteristics of sites I and III should be emphasized. (i) They are almost equivalent for triplex formation and stability with Acr-9-TC, since they allow the formation of nine and eight triplets, respectively, and intercalation of the acridine molecule occurs at the 5'-TpA-3' sequence present in both cases at the duplex–triplex junction on the 5'-side of the purine-rich strand. (ii) However, the two sites exhibit different acridine positions with respect to RNA polymerase tracking along the DNA: on site III the enzyme reaches the triplex *via* the oligonucleotide portion, whereas on the PPT site I it reaches the intercalator portion first (see the schematic representation in Figure 4B). (iii) Both triplex sites have an overlap of 3 bp with a recognition site of *Dra*I, which exactly cuts the sequence at the intercalation site of the acridine (5'-T↓A-3', as shown in Figure 4). The use of the run-off transcription products called *Dra*I markers (Figure 1) allowed us to exactly localize this 5'-TpA-3' sequence [at positions 4366–4367 (site I) and 4116–4117 (site III)]. (iv) The template strand was the pyrimidine-rich strand for the PPT site I and the purine-rich one for site III.

The positions of the 3'-ends of the truncated RNAs (open arrows near the gel in Figure 4A) around the triplex regions I and III were compared to the *Dra*I marker RNAs (lane Md, filled arrows). It is clear that the triplex-induced truncated transcript was shorter than the ~400-nt fragment of the *Dra*I marker for site III. This arrest pattern was consistent with a physical blockage of the RNA polymerase by the hydrogen-bonded structure of the triplex rather than

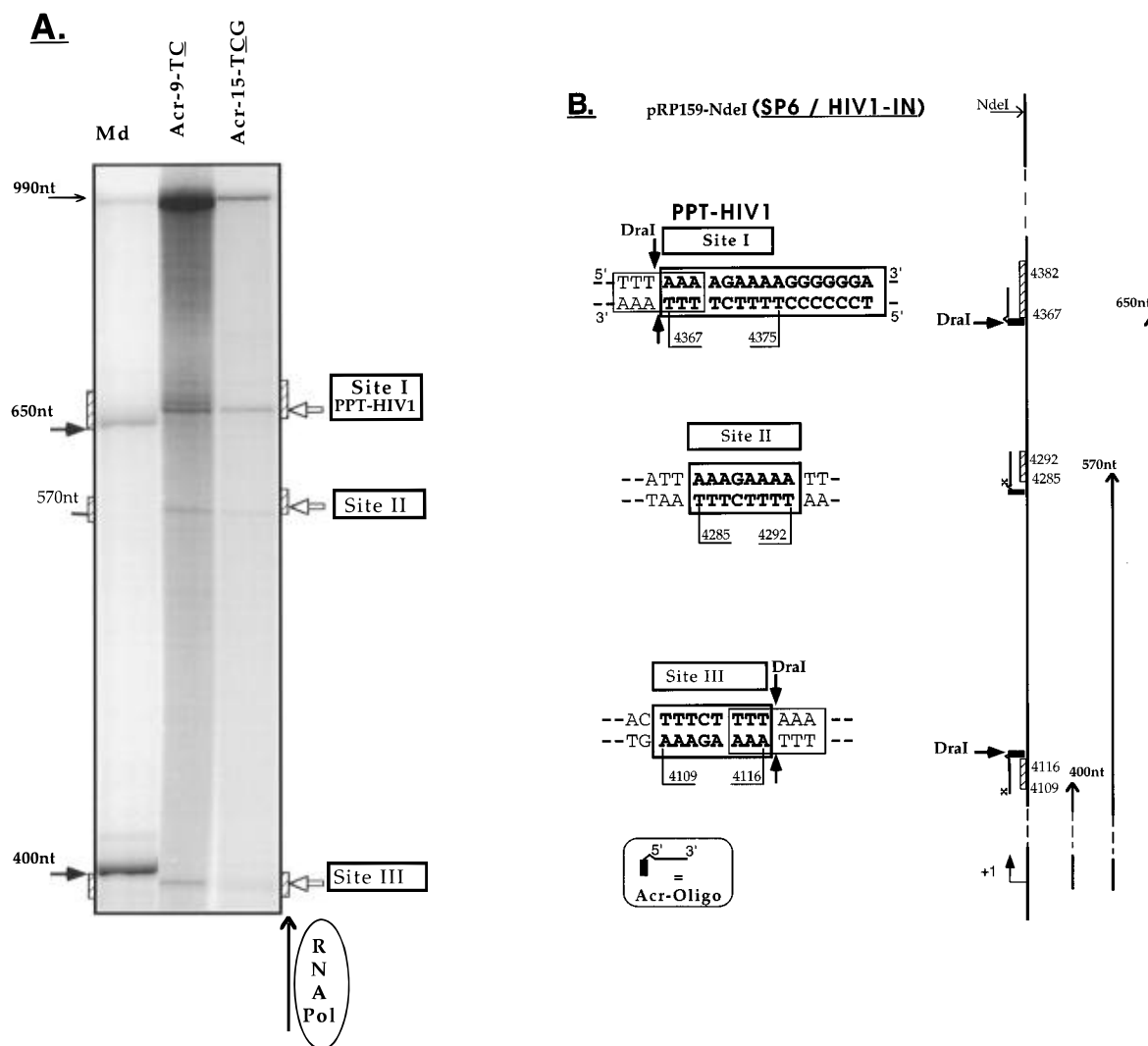


FIGURE 4: Physical arrest of RNA polymerase by a noncovalent triplex barrier. (A) Arrest of transcription elongation by Acr-9-TC. The positions of the acridine intercalation sites at sites III and I determined by the *DraI*-RNA marker (lane Md) are indicated by filled arrows near the gel. The 3'-ends of the triplex-induced truncated RNAs, obtained in the presence of Acr-9-TC or Acr-15-TCG, are reported near the gel using open arrows. The Acr-9-TC lane was overloaded to detect the truncated products obtained with low abundance. (B) Left: Scheme of the three potential binding sites of the 9-nt triplex-forming oligonucleotide Acr-9-TC in the HIV1-IN gene (I, 4367–4375 (PPT); II, 4285–4292; III, 4109–4116); *DraI* cleavage sites are indicated (bold arrows). Right: The positions of the three triplex sites (hatched boxes) with the bound Acr-9-TC are depicted schematically. The Acr-9-TC bound to its target is represented with acridine (filled rectangle) covalently attached to the 5'-end of the oligonucleotide. Crosses are for mismatched bases. Acridine orientation: Site III allows for the formation of eight consecutive base triplets with the acridine derivative intercalated at the duplex–triplex 5'-TpA-3' junction (position 4116–4117) as observed on the PPT site I (intercalation at position 4366–4367 and nine base triplets); the mismatched 3'-terminal base in Acr-9-TC is indicated by a cross (x). Site II also allows for the formation of eight base triplets, but there is a mismatched base next to the acridine (indicated by a x). The RNA polymerase encounters the triplex by the oligonucleotide part on site III and by the intercalator part on PPT site I. Coding strand: For sites I and II the oligo(pyrimidine)-containing strand is the coding strand giving rise to an oligo(purine) RNA sequence (identical to the top strand), contrary to site III where the oligo(purine)-containing strand is the template.

by the intercalated acridine. At site I, the truncated transcript was longer than expected if the polymerase had stopped at the intercalation site (corresponding to the *DraI* cleavage site). This phenomenon can be ascribed to a property of the phage enzyme, which is known to incorporate a few nucleotides onto the ends of truncated transcripts obtained, for example, at psoralen-cross-linked sites (Shi et al., 1988; see discussion). To validate this hypothesis, the covalent triplex formed with the cross-linked oligonucleotide–psoralen conjugate was used as a tool to generate a perfectly well-positioned block just at the entry of the triplex: the cross-link took place at the 5'-TpA-3' site [position 4366–4367, as described in Giovannangeli et al. (1992b)], preventing any strand separation at this position. The truncated transcripts around the PPT site obtained with the noncovalent

and the covalent triplex have exactly the same size (Figure 2, compare lanes 7 and 8); this observation is consistent with a physical arrest of RNA polymerase by the fully hybridized triple-helical complex, taking into account the fact that phage RNA polymerases have been reported to incorporate extra nucleotides onto the 3'-end of run-off RNA (Milligan et al., 1987; Sampson & Uhlenbeck, 1988; Shi et al., 1988; White & Phillips, 1988, 1989; Pavco & Steege, 1991). It should be noted that Pso-15-TCG could also bind to site III and give rise to photoinduced cross-links at the 5'-TpA-3' site at position 4116–4117 (see Figure 2, lane 8).

The pattern obtained with the eukaryotic transcription machinery confirms this interpretation. The truncated transcript was shorter than the RNA *DraI* marker: RNA synthesis was arrested by the triple-helix-forming oligo-

nucleotide a few bases before the triplex site (Figure 2B, compare lanes 12 and Md). This termination pattern can be easily explained by a physical arrest of the RNA polymerase II when its forefront contacts the triplex, which maintains the active site a few nucleotides upstream, until the RNA is released.

The comparative results presented here illustrate that the termination patterns induced by a triplex barrier are dependent on the nature of the RNA polymerase. The truncated products arising from the transcription arrest at the triplex site were shorter for the eukaryotic than for the phage polymerase. Similar polymerase-dependent patterns of termination have already been reported (Pavco & Steege, 1991): a cleavage-defective *EcoRI* protein blocked RNA synthesis ~10 nt upstream for the *E. coli* bacterial enzyme compared to the phage polymerases (T7 and SP6).

It should be noted that elongation by the SP6 RNA polymerase was blocked independently of whether the template strand contained the oligo(purine) (site III) or the oligo(pyrimidine) (site I) sequence. The oligonucleotide third strand recognizes and binds the oligo(purine) strand of an oligo(purine)·oligo(pyrimidine) sequence through Hoogsteen hydrogen bond formation. These data eliminate the possibility that the enzyme was blocked by Hoogsteen duplex formation between the purine strand and the oligonucleotide, which might occur when the two strands of the DNA template are separated during the transcription process.

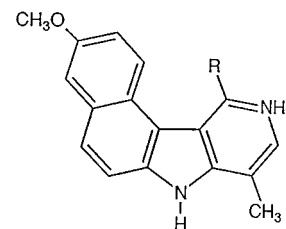
The identification of the weak truncated product corresponding to triplex site II was not studied in detail, but it was clear that the arrest efficiency was lower (~4-fold) than the one obtained at sites III and I. This result can be explained by the unfavorable situation for intercalation of the acridine moiety at site II since the thymine at the 5'-end of the 9-nt third strand did not interact by Hoogsteen hydrogen bonding with the 8-bp target (as depicted in Figure 4B), and this probably decreased the intercalation efficiency at the 5'-TpA-3' sequence by lengthening the distance between the last triplet and the intercalating site.

The preceding observations allowed us to conclude that a triplex-forming oligonucleotide–intercalator conjugate was a strong enough hydrogen-bonded ligand to act as a roadblock for both the SP6 and the Pol II RNA polymerases moving along the DNA, whatever the orientation of the acridine moiety and whatever the template strand.

Enhancement of the Efficiency of Transcription Inhibition by a Triplex-Specific Ligand. Benzopyridoindole (BePI) derivatives (Figure 5A) have been previously shown to stabilize triple-helical complexes as a result of intercalation between T·AxT base triplets (Mergny et al., 1992). The binding of 16-TC was only slightly increased in the presence of BePI ($\Delta T_m = 8^\circ\text{C}$) because positive charges in C·GxC⁺ triplets prevent the binding of a cationic molecule such as BePI. In contrast, BePI induced a strong enhancement of the binding of the 15-TCG ($\Delta T_m = 22^\circ\text{C}$).

At 25 °C transcription reactions were carried out after the addition of Acr-15-TCG or Acr-9-TC and 25 μM BePI. The percentages of truncated transcripts as a function of oligonucleotide concentration are reported in Figure 5B. For example, ~30% inhibition of RNA synthesis by SP6 RNA polymerase was obtained with 10 μM Acr-15-TCG in the absence of BePI, and enhanced inhibition (~60%) was observed in the presence of 25 μM BePI. A weaker

A. BePI



B. SP6/ HIV1-IN

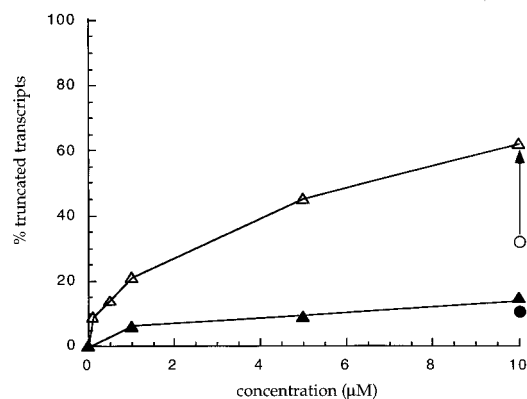


FIGURE 5: Triplex-specific ligand (BePI) and enhancement of triplex-induced inhibition. (A) Structure of 1-[(amino)propylamino]-4-methyl-5H-9-methoxybenzo[e]pyrido[4,3-b]indole (BePI), a tetracyclic compound able to stabilize triplex structures [R = (CH₂)₃NH₃⁺] (Nguyen et al., 1992). (B) Quantitation of triplex-arrested transcripts in the presence of 25 μM BePI in the bacteriophage system (25°C, pH 7.0, 60 min): Acr-15-TCG (open triangles); Acr-9-TC (filled triangles). For comparison, the percentages obtained in the absence of BePI are reported for one concentration (circles).

improvement in inhibition efficiency was obtained for the triplex formed by Acr-9-TC. These results were in agreement with triplex stabilities measured by the thermal dissociation experiments (see Table 1). The same enhancement of inhibition was obtained with the eukaryotic polymerase in the presence of 25 μM BePI (as indicated in Figure 3B): complete inhibition (i.e., 100% of truncated transcript) was obtained at a concentration of 2 μM Acr-15-TCG as compared to 10 μM in the absence of BePI; with 5 μM Acr-9-TC, the inhibition increased from 50% (no BePI) to 65% (25 μM BePI).

Termination Efficiency and Steady-State Occupancy of the Target. Even in the presence of BePI (25 μM) the unsubstituted 15-TCG (10 μM) did not arrest transcription elongation by the SP6 RNA polymerase (25 °C, pH 7.0) by more than 1–3%. However, under the conditions used for the transcription reaction (10 μM 15-TCG, 25 μM BePI, 25 °C), 100% of the 15-TCG should be bound to its double-stranded target since the T_m value of the corresponding triplex (using a 1.0 μM mixture of each strand of the triplex) was around 56 °C. It was not possible to increase the oligonucleotide and BePI concentrations above 20 and 25 μM , respectively, because SP6 transcription was globally inhibited and the full-length transcript was no longer observed under our standard elongation conditions (25 °C, pH 7.0). This inhibition is probably due to nonspecific sequestering of the enzyme by the high excess of single-stranded DNA or BePI molecules.

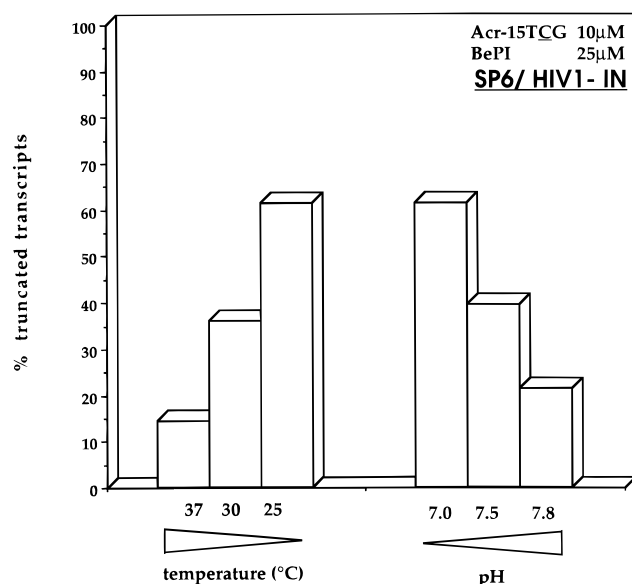


FIGURE 6: Influence of transcription conditions on the efficiency of transcription inhibition. The percentage of triplex-arrested transcripts was determined for different temperatures and pH values used during the transcription reaction with SP6 RNA polymerase (10 μ M Acr-15-TCG and 25 μ M BePI).

The bacteriophage system allowed us to adjust the reaction conditions. The efficiency of transcription inhibition thus was measured as a function of pH and temperature (Figure 6). The arrest of transcription elongation was more efficient when pH was lowered or temperature was decreased, as expected from the corresponding increase in triplex stability. Whatever the pH or temperature values in the range investigated, complete steady-state occupancy of the double-helical target site by Acr-15-TCG was determined from spectroscopic experiments (results not shown).

These observations demonstrated that, even with levels of third strand that gave quantitative binding to the triplex region, substantial amounts of read-through transcription beyond the triplex barrier continued to occur with the bacteriophage enzyme. This phenomenon was particularly significant when the conditions were optimized to enhance transcription by SP6 RNA polymerase (pH 7.8 and 37 °C).

DISCUSSION

Physical Arrest of Elongating RNA Polymerases. In the present work, we have demonstrated that an oligonucleotide–intercalator conjugate targeted to a sequence located in a coding region far downstream of the start site is able to arrest transcription elongation by both SP6 RNA polymerase and RNA polymerase II and leads to the production of a discrete truncated RNA reflecting the position of the bound third strand. As it is well established that some intercalators by themselves are able to induce the termination of transcription by phage RNA polymerases and prokaryotic *E. coli* RNA polymerase (White & Phillips, 1988, 1989), we investigated whether the progression of RNA polymerase was blocked (i) by the third strand itself, tightly hydrogen-bonded to a specific triplex site on the DNA template, or (ii) by the acridine molecule present at the duplex–triplex junction. We used the bacteriophage model system as a tool to elucidate this molecular mechanism. The polymerase arrest was observed independently of whether the intercalator or the triple helix was met first by the elongating enzyme. The

analysis of the 3'-end position of the triplex-arrested RNAs exhibited a termination site positioned a few bases downstream of the "entry" of the triplex site (depicted in Figures 2 and 4A): this observation is consistent with an arrest induced by the fully hybridized triplex, considering that extra nucleotides are incorporated by the phage polymerase.

The study of the three nonequivalent binding sites of Acr-9-TC that exist in the *IN* gene (Figure 4B) allowed us to qualitatively highlight other characteristics of the triplex blockage. The enzyme arrest was observed whatever the template strand [i.e., oligo(purine)- or oligo(pyrimidine)-containing strand] and whatever the distance between the transcription start and the triplex site. These data confirmed that the blockage was due to the triple-helical complex itself.

Stabilization by the BePI Molecule of the Triplex Formed with a (T,G)-Containing Third Strand; Termination Efficiency. There is a correlation between the relative stability of the triple-helical complexes and the triplex-induced termination efficiency. Stabilization of the oligonucleotide on its DNA target (by acridine substitution, by lengthening of the nucleotidic portion, or by addition of BePI molecule) increased its blocking effect on both RNA polymerases.

It is important to emphasize the remarkably strong enhancement of stability induced by the BePI molecule on the pH-independent triplex formed by 15-TCG since a 22 °C increase in T_m values was obtained in the presence of this molecule. It was previously shown that a long stretch (≥ 4 base triplets) of T·AxT triplets was necessary for maximal triple-helix stabilization by BePI (Mergny et al., 1992; Pilch et al., 1993). This observation was consistent with the density of negative charges around T·AxT triplets and the geometry of the triplets, which were important parameters to determine the binding of this triplex-specific cationic ligand. The high improvement in binding affinity induced by BePI on the triplex formed with 15-TCG extends the range of triplexes that can be stabilized. Interestingly, this result holds for pH-independent triplexes formed with G-containing third strands that interact with the oligo(purine) target by formation of either (i) Hoogsteen hydrogen bonds, leading to a parallel binding of the oligonucleotide with respect to the oligo(purine) strand of the target (as previously described for 15-TCG), or (ii) reverse Hoogsteen hydrogen bonds (Escudé et al., 1996), leading to antiparallel binding, which has been previously described for (T,G)-containing third strands (Durland et al., 1991). These data are of importance because pH-independent triplexes are the best candidates for specifically preventing *en vivo* transcriptional activity at any step (initiation or elongation).

Rational Design of a Triplex-Forming Oligonucleotide as an Efficient Blocker for Inhibition of Transcription Elongation. Tight, specific binding is necessary, but it is not sufficient since, even under conditions where the template is totally covered by the oligonucleotide, substantial amounts of read-through transcription continue to occur. This result demonstrates the crucial role played by kinetic parameters of the complex. The use of a psoralen-modified oligonucleotide allowed us to form an irreversible triplex after irradiation (> 310 nm) (Giovannangeli et al., 1992b; Grigoriev et al., 1993a; Duval-Valentin et al., 1992). This covalent modification, which is nearly quantitative, prevents oligonucleotide dissociation. When such a modified DNA was used as a transcription template, a dramatic decrease in the full-length product was observed, whereas the amount

of truncated transcripts increased (Figure 2, lane 8). These data suggest that the partial inhibition obtained in the absence of covalent adduct reflects the dissociation of the triple-helical complex. Preliminary kinetic data demonstrate that the increase in thermodynamic stability of the triplex obtained with the acridine-oligonucleotide conjugate compared to the unmodified one is caused by a large decrease (~ 10 -fold, at 25 °C) in the dissociation rate constant (C. Giovannangeli, unpublished results). This means that the oligonucleotide has a 10 times longer residence time on its double-stranded target.

The triplex-induced termination efficiency is determined by the properties of both the blocking triplex and the elongation complex and may reflect a constant competition between the rate of elongation and the rate of RNA release. It is thus interesting to compare triplex-induced termination efficiencies using the bacteriophage SP6 RNA polymerase and the eukaryotic RNA polymerase II. The data that we have reported here using the same well-defined triplex have allowed us to directly compare the phage and the eukaryotic systems. It clearly appears that the triplex barrier is much more efficient at arresting the eukaryotic transcriptional machinery than the bacteriophage one. These results are consistent with the different stabilities of their ternary complexes (Donahue et al., 1994; Liu et al., 1994) and their completely different elongation rates (Butler & Chamberlin, 1982; Erie et al., 1992). Moreover the totally different termination patterns obtained with the two enzymes likely reflect the different structures of the paused ternary elongation complexes for which some models have been proposed (Altmann et al., 1994; McDowell et al., 1994; Chamberlin, 1995), but no high-resolution structures have been reported yet.

Our results illustrate that the design of thermodynamically and kinetically stable triplex-forming oligonucleotides will be necessary to succeed in efficiently inhibiting transcription elongation under physiological conditions. They provide a framework for the design and optimization of triplex-forming oligonucleotides as efficient elongation terminators. This would greatly expand the repertoire of genes amenable to repression with triplex-forming oligonucleotides, in addition to target sites located within regulatory sequences of the genes. Moreover, a DNA target in a transcribed region may be more accessible to an oligonucleotide when compared to regulatory regions, which are known to be involved and regulated by highly complex chromatin structures (Wolffe, 1994; Roberts & Green, 1995; Sauer et al., 1995; Gallego et al., 1995).

ACKNOWLEDGMENT

We thank Drs. Plasterk and Van Gent (Amsterdam) and Drs. O. Schwartz and P. Charneau (Institut Pasteur, Paris) for the gift of the plasmidic constructions carrying various HIV inserts, G. Burban for their preparations, and Prof. E. Bisagni for the gift of the BePI derivative.

REFERENCES

Altmann, C. R., Solow-Cordero, D. E., & Chamberlin, M. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3784–3788.
Asseline, U., Delarue, M., Lancelot, G., Toulmé, F., Thuong, N. T., Montenay-Garestier, T., & Hélène, C. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3297–3301.

Butler, E. T., & Chamberlin, M. J. (1982) *J. Biol. Chem.* **257**, 5772–5778.
Chamberlin, M. J. (1994) *Harvey Lect.* **88**, 1–21.
Chen, A. J., & Van Dyke, M. W. (1994) *Nucleic Acids Res.* **22**, 4742–4747.
Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science* **241**, 456–459.
Dayn, A., Samadashwily, G. M., & Mirkin, S. M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 504–508.
Donahue, B. A., Yin, S., Taylor, J. S., Reines, D., & Hanawalt, P. C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8502–8506.
Dubendorf, J. W., & Studier, F. W. (1991) *J. Mol. Biol.* **219**, 45–59.
Durland, R. H., Kessler, D. J., Gunnel, S., Duvic, M., Pettit, B. M., & Hogan, M. E. (1991) *Biochemistry* **30**, 9246–9255.
Duval-Valentin, G., Thuong, N. T., & Hélène, C. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 504–508.
Erie, D. A., Yager, T. D., & Hippel, P. H. V. (1992) *Annu. Rev. Biophys. Biomol. Struct.* **21**, 379–415.
Escudé, C., Sun, J. S., Nguyen, C. H., Bisagni, E., Garestier, T., & Hélène, C. (1996) *Biochemistry*, in press.
Gallego, F., Fernandez-Busquets, X., & Daban, J.-R. (1995) *Biochemistry* **34**, 6711–6719.
Giordano, T. J., Deusehle, U., Bujard, H., & Allister, W. T. (1989) *Gene* **84**, 209–219.
Giovannangeli, C., Rougée, M., Montenay-Garestier, T., Thuong, N. T., & Hélène, C. (1992a) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8631–8635.
Giovannangeli, C., Thuong, N. T., & Hélène, C. (1992b) *Nucleic Acids Res.* **20**, 4275–4281.
Giovannangeli, C., Thuong, N. T., & Hélène, C. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10013–10017.
Grigoriev, M., Praseuth, D., Robin, P., Hemar, A., Saison-Behmoaras, T., Dautry-Varsat, A., Thuong, N. T., Hélène, C., & Harel-Bellan, A. (1992) *J. Biol. Chem.* **267**, 3389–3395.
Grigoriev, M., Praseuth, D., Guieysse, A. L., Robin, P., Thuong, N. T., Hélène, C., & Harel-Bellan, A. (1993a) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3501–3505.
Grigoriev, M., Praseuth, D., Guieysse, A. L., Robin, P., Thuong, N. T., Hélène, C., & Harel-Bellan, A. (1993b) *C.R. Acad. Sci. Paris* **316**, 492–495.
Hacia, J. G., Dervan, P. B., & Wold, B. J. (1994) *Biochemistry* **33**, 6192–6200.
Kim, H. G., & Miller, D. M. (1995) *Biochemistry* **34**, 8165–8171.
Liu, B., Wong, M. L., & Alberts, B. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10660–10664.
Maher, L. J. I., Dervan, P. B., & Wold, B. (1992) *J. Am. Chem. Soc.* **114**, 70–81.
McDowell, J. C., Roberts, J. W., Jin, D. J., & Gross, C. (1994) *Science* **266**, 822–825.
Mergny, J. L., Duval-Valentin, G., Nguyen, C. H., Perrouault, L., Faucon, B., Rougée, M., Montenay-Garestier, T., Bisagni, E., & Hélène, C. (1992) *Science* **256**, 1681–1684.
Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
Nguyen, C. H., Bisagni, E., Lavelle, F., Bissery, M. C., & Huel, C. (1992) *Anti-Cancer Drug Des.* **7**, 219–233.
Pavco, P. A., & Steege, D. A. (1991) *Nucleic Acids Res.* **19**, 4639–4646.
Pilch, D. S., Martin, M. T., Nguyen, C. H., Sun, J. S., Bisagni, E., Garestier, T., & Hélène, C. (1993) *J. Am. Chem. Soc.* **115**, 9942–9951.
Rando, R. F., DePaolis, L., Durland, R. H., Jayaraman, K., Myrick, M., & Hogan, M. E. (1994) *Nucleic Acids Res.* **22**, 678–685.
Roberts, S. G. E., & Green, M. R. (1995) *Nature* **375**, 105–106.
Samadashwily, G. M., & Mirkin, S. M. (1994) *Gene* **149**, 127–136.
Samadashwily, G. M., Dayn, A., & Mirkin, S. M. (1993) *EMBO J.* **12**, 4975–4983.
Sampson, J. R., & Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1033–1037.
Sauer, F., Fondell, J. D., Ohkuma, Y., Roeder, R. G., & Jäckle, H. (1995) *Nature* **375**, 162–164.
Schwartz, O., Rivière, Y., Heard, J. M., & Danos, O. (1993) *J. Virol.* **67**, 3274–3280.

- Shi, Y. B., Gamper, H., & Hearst, J. E. (1988) *J. Biol. Chem.* 263, 527–534.
- Skoog, J. U., & Maher, L. J. (1993) *Nucleic Acids Res.* 21, 4055–4058.
- Sun, J. S., Giovannangeli, C., François, J. C., Kurfurst, R., Montenay-Garestier, T., Asseline, U., Saison-Behmoaras, T., Thuong, N. T., & Hélène, C. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 6023–6027.
- Sun, J. S., François, J. C., Montenay-Garestier, T., Saison-Behmoaras, T., Roig, V., Thuong, N. T., & Hélène, C. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 9198–9202.
- Thuong, N. T., & Chassignol, M. (1988) *Tetrahedron Lett.* 29, 5905–5908.
- Thuong, N. T., & Hélène, C. (1993) *Angew. Chem., Int. Ed. Engl.* 32, 666–690.
- Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., & Alizon, M. (1985) *Cell* 40, 9–17.
- White, R. J., & Phillips, D. R. (1988) *Biochemistry* 27, 9122–9132.
- White, R. J., & Phillips, D. R. (1989) *Biochemistry* 28, 6259–6269.
- Wolffe, A. P. (1994) *Cell* 77, 13–16.
- Wolffe, A. P., Jordan, E., & Brown, D. D. (1986) *Cell* 44, 381–389.
- Xodo, L., Alunni-Fabbroni, M., Manzini, G., & Quadrifoglio, F. (1994) *Nucleic Acids Res.* 22, 3322–3330.
- Young, S. L., Krawczyk, S. H., Matteucci, M. D., & Toole, J. J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 10023–10026.

BI952993X