Inhibition of RNA polymerase II transcription by oligonucleotide–RecA protein filaments targeted to promoter sequences

(human immunodeficiency virus long terminal repeat/parvovirus/minute virus of mice)

EFIM I. GOLUB*, CHARLES M. RADDING, AND DAVID C. WARD

Departments of Genetics, and Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510

Communicated by Donald M. Crothers, May 20, 1993

ABSTRACT In the presence of RecA protein, which plays a major role in genetic recombination in *Escherichia coli*, an oligodeoxyribonucleotide can find its homologous counterpart in double-stranded DNA and form triple-stranded structures. A triple-stranded structure formed by an oligonucleotide with a sequence overlapping essential regulatory elements of a viral promoter, such as TATA or GC boxes, inhibited *in vitro* transcription driven by RNA polymerase II. An oligonucleotide with eight nucleotides homologous to its target suppressed RNA polymerase II activity in HeLa cell extracts. This procedure offers a potential alternative to the usual mutational analysis of transcriptional promoters.

Eukaryotic transcription is regulated by complex interactions of various sequence-specific DNA-binding proteins with promoters (1). Polypurine or polypyrimidine oligodeoxyribonucleotides (ODNs) can nonenzymatically form a sequence-specific triple helix by hydrogen binding to polypyrimidine or polypurine tracts within duplex DNA (2–4). Formation of a triple helix in a region of DNA that is involved in interactions with proteins of the RNA polymerase II transcription complex can prevent the binding of protein factors and thereby suppress transcription (5–9). However, the inhibition of transcription by oligopurine or oligopyrimidine sequences can be achieved only if the appropriate tracts exist in the regulatory sequences. By contrast, ODNs of arbitrary sequence, when coated by the RecA protein of *Escherichia coli* (10) in the presence of ATP or its nonhydrolyzable analog adenosine 5'-[γ -thio]triphosphate can pair with homologous duplex DNA–RecA protein and form a triple-stranded structure which is stable in the presence of RecA protein (11). These nucleoprotein complexes, which include RecA protein and three strands of DNA, have been shown to prevent the original duplex DNA sequence from being methylated or digested by restriction enzymes (12).

Recently, we found that a triple-stranded structure formed between ODN–RecA complexes and duplex DNA at specific sequences upstream or downstream of a transcriptional start site could inhibit either initiation or elongation of transcription by T3 or T7 phage RNA polymerases (13). In this communication, we describe experiments which show that specific ODN–RecA protein filaments can markedly inhibit initiation of transcription by eukaryotic RNA polymerase II.

MATERIALS AND METHODS

Plasmids and ODNs. Three different DNAs were used as a template for transcription: (i) a 490-bp *Ava* I–*Eco*RI fragment (Fig. 1A) of plasmid pEG602 (14) which carries sequences from the long terminal repeat (LTR) of human immunodeficiency virus type 1 (HIV-1); the LTR contains an HIV-1

promoter and several regulatory elements which control proviral DNA expression (15); (ii) a 600-bp *Eco*RI–*Hind*III fragment of plasmid pYT104-1 (16, 17), which includes the P6 promoter of the human parvovirus B19; part of the B19 sequence relevant to our study is shown in Fig. 1B; and (iii) a plasmid, pMB415 digested by *Pst* I (18), which carries the P4 promoter of minute virus of mice (MVM). Run-off transcripts made from the HIV-1 LTR, and the P6 and P4 parvovirus promoters were 335, 250, and 210 nucleotides long, respectively.

The ODN sequences used in this study are listed in Table 1. ODN PBS-T3₄₂, which does not share sequence homology with any of the templates, served as the heterologous control. ODNs were synthesized on a model 394 DNA synthesizer (Applied Biosystems) and purified by separation in a polyacrylamide gel or by using NENSorbPrep cartridges (New England Nuclear).

Formation of Oligonucleotide Joints by RecA Protein. RecA protein was purified as described (19). ODN at 10 μ M was mixed with 5 μ M RecA protein in a solution containing 25 mM Tris acetate (pH 7.2), 2 mM magnesium acetate, and 1 mM adenosine 5'-[γ -thio]triphosphate. The 10- μ l mixture was incubated for 5 min at 37°C to allow RecA protein to coat the ODN. Then, 2 μ l of template DNA (490-bp *Ava* I–*Eco*RI fragment of pEG602, 600-bp *Eco*RI–*Hind*III fragment of pYT104-1, or *Pst* I-digested pMB415 was added ($\approx$ 30 μ M) in 30 mM magnesium acetate. After 15–20 min at 37°C, the mixture was used as a source of template for *in vitro* transcription.

Run-Off Transcription. The transcription mixture contained 5 μ l ($\approx$ 8 transcriptional units) of HeLa nuclear extract, 6 μ l of transcription buffer [20 mM Hepes, pH 7.9/100 mM KCl/0.2 mM EDTA/0.5 mM dithiothreitol/20% (vol/vol) glycerol], 1 μ l of ribonucleotide mixture (10 mM ATP/10 mM CTP/10 mM UTP/0.4 mM GTP), 1 μ l of [α -³²P]GTP (3000 Ci/mmol, 10 mCi/ml; New England Nuclear; 1 Ci = 37 GBq), and 0.2 μ l (50 ng) of plasmid DNA used as internal standard. The mixture was added to 12 μ l of the template mixture prepared as described above. After 30 min of incubation at 30°C, transcription was stopped by adding 175 μ l of stop buffer (0.3 M Tris-HCl, pH 7.4/0.3 M sodium acetate/0.5% SDS/2 mM EDTA with tRNA at 3 μ g/ml). Reaction products were extracted with 200 μ l of phenol/chloroform/isoamyl alcohol (25:24:1, vol/vol) and precipitated with 3 volumes of 95% ethanol. Half of the RNA sample was mixed with an equal volume of formamide dye mix (98% formamide/10 mM EDTA/0.1% xylene cyanol/0.1% bromophenol blue), heated 5 min at 90°C, and subjected to electrophoresis in a denaturing 8% polyacrylamide gel. The gel was autoradiographed for 12–72 hr at –70°C with an intensifying screen. To quantitate *in vitro* transcripts, densitometric scanning was carried out

Abbreviations: HIV-1, human immunodeficiency virus type 1; LTR, long terminal repeat; MVM, minute virus of mice; ODN, oligodeoxyribonucleotide.

*To whom reprint requests should be addressed.

FIG. 1. Maps of fragments of DNA that were used as a template for transcription, and position of ODNs. (A) *Ava*I-*Eco*RI fragment of HIV-1 LTR (from plasmid pEG602). (B) Part of *Eco*RI-*Hind*III fragment of plasmid pYT104-1 that carries promoter P6 of human parvovirus B19. Arrow shows the transcriptional start site. Numbers depict nucleotide position relative to the transcriptional start site. Lowercase letters indicate ODN sequences that are not homologous to duplex DNA.

by using a gel scanner (Molecular Dynamics). To estimate the amounts of specific transcripts, their band intensities were normalized by dividing by the intensity of the band produced by transcription of the internal standard. HeLa cell nuclear extract and unlabeled ribonucleotides were purchased from Promega.

RESULTS

Inhibition of Transcription from an HIV-1 LTR Promoter.

Fig. 2 shows results of an experiment with *in vitro* run-off transcription in which a 600-bp fragment of the HIV-1 LTR

was used as a template. In the presence of RecA, ODN EG100₅₀, whose sequence overlaps a TATA box (see Fig. 1A and Table 1), strongly inhibited transcription (Fig. 2, compare lanes 2 and 3). Similarly, ODN EG106₃₃, whose sequence overlaps three GC boxes implicated in transcription from the HIV-1 promoter, also inhibited transcription in a RecA-dependent fashion, albeit to a lesser degree (compare lanes 4 and 5). In contrast, two other ODNs, PBS-T3₄₂ and EG107₃₀, had no effect on transcription. The former of these ODNs does not share any homology with the template, whereas the latter overlaps HIV-1 enhancers which are not functional in uninduced or uninfected HeLa cells (see below). ODN

Table 1. ODNs used in this study

Name	No. of nucleotides	Sequence	Comment
PBS-T3 ₄₂	42	ATTACGCCAAGCTCGAAATTAACCTCACTAAAGGGAACAAA	Has no homology to any template used in this study ("heterologous control")
EG100 ₅₀	50	GATGCTGCATATAAGCAGCTGCTTTTGCCTGTACTGGGTCTCTCTGGTT	Homologous to -37 to +13 of HIV-1 LTR; overlaps TATA box
EG106 ₃₃	33	<u>GAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCG</u>	Homologous to -78 to -46 of HIV-1 LTR; overlaps GC boxes
EG107 ₃₀	30	CTACAAGGGACTTTCCGTTGGGGACTTTCC	Homologous to -111 to -82 of HIV-1 LTR; overlaps enhancers
EG108 ₄₂	42	CCTTTTAAAGACCGTAAAGAAAAATAAGCAC AAGTTTATC	Homologous to +244 to +285 of HIV-1 LTR downstream from HIV-1 promoter
EG110 ₃₃	33	GTTAGTTTTGTAACGGTTAAAAATGGGCGGAGCG	Homologous to -81 to -49 of P6 promoter; overlaps distal GC box
EG111 ₃₃	33	CTACAGTATATATAGCACGGTACTGCCGCAGCT	Homologous to -38 to -6 of P6 promoter; overlaps TATA box
EG113 ₃₀	30	CTTTCTTTCTGGGCTGCTTTTTCTGGACT	Homologous to -5 to +25 of P6 promoter
EG114 ₂₅	25	CCCCCTCCCCCCCCCGGGCGGAGCG	Sequence of last 10 nucleotides is homologous to -58 to -49 of P6 promoter; overlaps part of distal GC box
EG118 ₂₅	25	CCCCCCCCCCCCCTTTGTTAAAAATGG	Sequence of last 10 nucleotides is homologous to -66 to -57 of P6 promoter; overlaps 3 nucleotides of distal GC box
EG115 ₂₀	20	CGGTTAAAAATGGGCGGAGCG	Homologous to -68 to -49 of P6 promoter; overlaps distal GC box

Sequences homologous to TATA boxes, GC boxes, or HIV-1 enhancers are underlined.

FIG. 2. Inhibition of transcription from HIV-1 promoter by homologous RecA oligonucleoprotein joints. Arrows show the position of bands whose mobility is consistent with that expected for a full-length transcript from HIV-1 promoter (335 bases) or P4 promoter of MVM (210 bases). Transcription from the P4 promoter was used as an internal standard. In lanes 2 and 4, RecA protein was omitted and replaced by an equal volume of R buffer (50 mM Tris-HCl, pH 7.5/5 mM dithiothreitol/0.3 mM EDTA/10% glycerol), which is used for dilution of RecA protein. Lane 1, 32 P-end-labeled pBR322 digested by *Msp* I to provide size markers; numbers at left identify the lengths of the fragments in nucleotides.

EG108₄₂, containing a sequence homologous to a downstream region (nucleotides 244–285), also did not influence transcription.

Inhibition of Transcription from Promoter P6 of the Human Parvovirus B19. Transcription from the P6 promoter was strongly inhibited by ODNs EG110₃₃ and EG111₃₃ (Fig. 3A, compare lanes 1–3), which overlap a distal GC box and the TATA box, respectively (see Fig. 1B). ODN EG115₂₀, which overlaps the distal GC box but consists of only 20 nucleotides (see Fig. 1B), caused a considerable but weaker effect on transcription than the 33-mer EG110₃₃ (Fig. 3B). ODNs EG114₂₅ and EG118₂₅ consist of 25 nucleotides of which only a sequence of 10 nucleotides is homologous to the template (Fig. 1B). ODN EG114₂₅, which overlaps 8 out of 9 nucleotides of the distal GC box, caused a profound effect on transcription, while ODN EG118₂₅, which overlaps only 3 nucleotides of the GC box, did not effect transcription (Fig. 3C). ODN EG113₃₀, which forms a triple-stranded structure in a region overlapping the transcription initiation site, showed only a very small effect on transcription (Fig. 3B).

Inhibition of Transcription from Promoter P6 of MVM. The region upstream of the P4 promoter of MVM has a TATA box

FIG. 4. Inhibition of transcription from the P4 promoter of MVM. Arrows show position of bands that correspond to transcripts from P4 (210 bases) or HIV-1 (335 bases) promoter. Transcription from HIV-1 promoter was used as an internal standard. Lane 1, 32 P-labeled pBR322 digested by *Msp* I.

containing the sequence AAAGTATATAAGCAA, of which a sequence of 9 nucleotides is homologous to ODN EG100₅₀, and 8 nucleotides are homologous to ODN EG111₃₃. As shown in Fig. 4, both ODNs were effective in inhibition of transcription directed by the P4 promoter. The upstream region of the P4 promoter containing a GC box sequence of TGGGCGTGGTT shares homology with a sequence of 7 nucleotides of ODN EG106₃₃. This ODN, however, did not disturb transcription from the P4 promoter (Fig. 4, compare lanes 2 and 5).

DISCUSSION

Transcription directed by eukaryotic RNA polymerase II depends on various protein factors which bind to the promoter region in a sequence-specific way. Different agents and conditions which prevent those factors from binding to duplex DNA could critically change transcription patterns or even completely suppress them. As was shown previously, an essential role in basal transcriptional activity of the HIV-1 promoter is played by a TATA box which is a target of cellular factor TFIID (20) and by three GC boxes which interact with cellular factor Sp1 (21). In full accordance with those data, we found that ODNs EG100₅₀ and EG106₃₃ which overlap the TATA box and GC boxes, respectively, and evidently formed nucleoprotein joints in those regions, effectively inhibited transcription from the HIV-1 promoter (Fig. 2). Cellular factor NF- κ B in its active form binds to two HIV-1 enhancers and drastically increases promoter activity (22). Such activation occurs only in cells exposed to mitogens, cytokines, phorbol esters, and some heterologous viruses. Under nonstimulated conditions, however, NF- κ B is inactive and regions occupied

FIG. 3. Inhibition of transcription from the P6 promoter of human parvovirus B19. Arrows show position of bands that correspond to transcripts from P6 (250 bases) or P4 (210 bases) promoters. Transcription from P4 promoter was used as an internal standard. Lane 4 in A and lane 1 in B, 32 P-labeled pBR322 digested by *Msp* I.

by the enhancers do not participate in transcription. Correspondingly, ODN EG107₃₀, whose sequence overlaps the two HIV-1 enhancer elements, did not influence transcription from the HIV-1 promoter directed by nuclear extract from uninduced HeLa cells (Fig. 2).

Unlike the HIV-1 LTR promoter, the P6 promoter of human parvovirus B19 has not been studied in detail. It was shown by footprinting analysis that some cellular component specifically binds to a distal GC box (16). In agreement with that observation, ODN EG110₃₃, which overlaps the distal GC box, was found to inhibit transcription from this promoter by >90%. ODN EG111₃₃, which overlaps a TATA box, also blocked transcriptional activity of the P6 promoter (Fig. 3A).

A mutational study had shown previously that a TATA box is essential for *in vitro* transcription from the P4 promoter of MVM (18). Correspondingly, ODNs EG100₅₀ and EG111₃₃, which both overlap 6 out of 7 nucleotides of the TATA box, were effective in inhibiting transcription.

In previous experiments with phage T3 and T7 RNA polymerases, we found that triple-stranded structures formed by an ODN in the presence of RecA protein that were located downstream of the RNA polymerase promoter inhibited elongation of transcription and caused accumulation of a truncated transcript (13). In sharp contrast, ODN EG108₄₂, which forms a triple-stranded structure located downstream of the HIV-1 promoter, did not stop elongation of transcription (Fig. 2). However, this ODN prevented digestion of template DNA by the restriction enzyme *Dra* I, whose recognition site was overlapped by the ODN (data not shown), thus indicating that the triplex was indeed formed. Obviously, RecA nucleoprotein joints located downstream of the promoter cannot stop the procession of the eukaryotic transcriptional machinery through such joints.

In contrast to ODNs that overlap TATA or GC boxes, ODN EG113₃₀, which spans the initiation transcription site, did not have an appreciable effect on transcription (Fig. 3B). This result was unexpected, since complexes formed on a promoter by general transcription factors and RNA polymerase II protect such regions from DNase digestion (23). However, formation of these initiation complexes starts with the binding of transcription factor TFIID to a TATA element, and this binding leads to DNA bending (24). It is possible that this bending destroys the triple-stranded structure formed in the region.

The experiments reported here demonstrate that ODNs that bind to elements which are targets for binding cellular proteins responsible for activation of transcription inhibit transcription. On the basis of these data, obtained with three unrelated viral promoters, we assume that ODNs could be used as a general tool for finding regulatory elements within promoter regions of the polymerase II-dependent genes and evaluating their relative importance in transcription. Compared with the widely used genetic analysis of promoter activity, whereby specific mutations are introduced in a suspected regulatory element, the present method of analysis has the advantage of being rapid and simple. It is also interesting to speculate that the triplex analytical approach could be applied to systems other than transcription where regulation of biological activity is achieved by the interaction of protein factors with sequence-specific binding sites (for example, for probing DNA replication origins or matrix attachment regions).

How long should the ODN be, and how much homology should there be for the ODN to be effective in inhibition of transcription? We observed virtually complete inhibition of transcription with ODN EG110₃₃, which consists of 33 homologous nucleotides and overlaps the distal GC box of the P6 promoter (Fig. 2A). ODN EG115₂₀ also overlaps the GC box but contains only 20 nucleotides. This ODN was not as effective in transcription inhibition as ODN110₃₃ (Fig. 2).

Previously we described experiments with an ODN consisting of 20 nucleotides which caused strong inhibition of transcription by phage RNA polymerase (13).

ODN EG114₂₅, a 25-residue oligomer which shares sequence homology with only 10 nucleotides of the target duplex DNA, nevertheless decreased transcription by 80% (Fig. 2B). ODNs EG100₅₀ and EG111₃₃, which share homology with 9- and 8-nucleotide sequences of the targeted template TATA box of P4 promoter, respectively, also were capable of suppressing transcription (Fig. 4). On the other hand, we found no inhibition of transcription by ODN EG106₃₃, whose sequence is homologous to only 7 nucleotides in the GC box of the P4 promoter (Fig. 3). This GC box is important for *in vitro* transcription from the P4 promoter (18). Apparently, homology in 8 nucleotides is necessary and sufficient for an ODN to inhibit transcription provided that the total length of the ODN is adequate. Thus, while an oligonucleotide of ≥ 20 nucleotide residues is required for the stable binding of RecA protein (25), an 8-bp region of homology is sufficient for formation of a nucleoprotein triple-stranded structure. These data are consistent with the work of Hsieh *et al.* (11), who, using a series of 20-base ODNs, found that RecA protein can pair (albeit at low efficiency) as few as 8 bases of homology.

ODNs EG114₂₅ and EG118₂₅ have a similar structure; both consist of a sequence of 15 nucleotides which does not share homology with the template, followed by a sequence of 10 nucleotides homologous to the template. ODNs EG114₂₅ and EG118₂₅ overlap 8 and 3 nucleotides of the distal GC box of the P6 promoter, respectively. As a result, EG114₂₅, but not EG118₂₅, actively suppressed transcription from the promoter. These results suggest that the RecA-coated oligonucleotide must specifically cover the recognition site to interfere and thus may provide a reasonably discriminatory tool in mapping such sites.

We thank Ning Ye for the purification of RecA protein and Dr. K. Krady and Dr. Z. Pitluck for the generous gift of plasmids pYT104-1 and pBM415. This research was supported by grants from the National Institutes of Health, GM40633 and HG00338.

- Mitchell, P. J. & Tjian, R. (1989) *Science* **245**, 371-378.
- Moser, H. E. & Dervan, P. B. (1987) *Science* **238**, 645-650.
- Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241**, 456-459.
- Beal, P. A. & Dervan, P. B. (1991) *Science* **251**, 1360-1363.
- Postel, E. H., Flint, S. J., Kessler, D. J. & Hogan, M. E. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8227-8231.
- Orson, F. M., Thomas, D. W., McShan, W. M., Kessler, D. J. & Hogan, M. E. (1991) *Nucleic Acids Res.* **19**, 3435-3441.
- McShan, W. M., Rossen, R. D., Laughler, A. H., Trial, J. A., Kessler, D. J., Zengdegi, J. C., Hogan, M. E. & Orson, F. M. (1992) *J. Biol. Chem.* **267**, 5712-5721.
- Maher, L. J., III, Dervan, P. B. & Wold, B. (1992) *Biochemistry* **31**, 70-81.
- Mergny, J. L., Duval-Valentin, G., Nguyen, C. H., Perrouault, L., Faucon, B., Rougee, M., Monteny-Garestier, T., Bisagni, E. & Hélène, C. (1992) *Science* **256**, 1681-1684.
- Radding, C. M. (1991) *J. Biol. Chem.* **266**, 5355-5358.
- Hsieh, P., Camerini-Otero, C. S. & Camerini-Otero, R. D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6492-6496.
- Ferrin, L. J. & Camerini-Otero, R. D. (1991) *Science* **253**, 1494-1497.
- Golub, E. I., Ward, D. C. & Radding, C. M. (1992) *Nucleic Acids Res.* **20**, 3121-3125.
- Golub, E. I., Li, G. & Volsky, D. J. (1990) *J. Virol.* **64**, 3654-3660.
- Wong-Staal, F. & Gallo, R. C. (1985) *Nature (London)* **317**, 345-403.
- Blundell, M. C. & Astell, C. R. (1989) *J. Virol.* **63**, 4814-4823.
- Shade, R. O., Blundell, M. C., Cotmore, S. F., Tattersall, P. & Astell, C. R. (1986) *J. Virol.* **58**, 921-936.

18. Ahn, J. K., Gavin, B. J., Kumar, G. & Ward, D. C. (1989) *J. Virol.* **63**, 5425–5439.
19. Shibata, T., Cunningham, R. P. & Radding, C. M. (1981) *J. Biol. Chem.* **256**, 7557–7564.
20. Starcich, B., Ratner, L., Josephs, S. F., Okamoto, T., Gallo, R. C. & Wong-Staal, F. (1985) *Science* **227**, 538–540.
21. Jones, K. A., Kadonaga, J. T., Luciw, P. A. & Tjian, R. (1986) *Science* **232**, 755–759.
22. Nabel, G. & Baltimore, D. (1987) *Nature (London)* **326**, 711–713.
23. Buratowski, S., Hanh, S., Guarente, L. & Sharp, P. A. (1989) *Cell* **56**, 549–561.
24. Hozikoski, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T. & Roeder, R. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1060–1064.
25. Leahy, M. C. & Radding, C. M. (1986) *J. Biol. Chem.* **261**, 6954–6960.