Volume 251, number 1,2, 79-83 FEB 07358 July 1989

Human nuclear protein interacting with a conservative sequence motif of Alu-family DNA repeats

N.V. Tomilin and V.M. Bozhkov

Institute of Cytology, Academy of Sciences of the USSR, Tikchoretskii Av.4, Leningrad 194064, USSR

Received 16 May 1989

Human retrotransposons, Alu-family DNA repeats (AFRs), have variable nucleotide sequence but conservative short elements, which may have important functions, are also present. In our previous reports we have described human nuclear DNA-binding protein interacting with AFRs and evidence was presented that the protein recognizes sequence motif 5'-GGAGGC-3' which is conserved in the spacer of RNA polymerase III promoter of AFRs and in the SV40 T-antigen-dependent replication origin of AFRs. In this study it was found that double-stranded synthetic oligonucleotides containing indicated conservative sequences of AFRs actually have high-affinity binding site for HeLa nuclear protein. The data suggest that non-infected human cells contain nuclear DNA-binding protein which recognizes the conservative sequence motif of AFRs – GGAGGC.

Alu repeat; Conservative motif; DNA-binding protein

1. INTRODUCTION

DNA-binding proteins interacting with conservative sequence elements of eukaryotic promoters play an important role in the up and down regulation of transcription. Alu-family DNA repeats (AFRs) which are human retrotransposons [1-3] contain bipartite internal promoter for RNA polymerase III which is active in vitro [4] but is transcriptionally silent in HeLa cells in vivo [5]. The nucleotide sequence of human AFRs is rather variable [6] but some conserved elements are also present [7]. One of the conservative regions is situated at the left end of AFRs covering the enhancing element of polIII promoter and a part of spacer between enhancing and directing elements (fig.1).

In our previous reports [8-12] we have described human nuclear protein interacting selectively with a typical AFR present in the plasmid BLUR8 [13]. Differential binding of restriction subfragments of BLUR8 ARF [11], and DNase footprinting experiments [12] suggest that two regions within AFR are important for protein binding: (i) a spacer between enhancing and directing elements of the RNA polymerase III (polIII) promoter, (ii) a central region containing the SV40 T-antigen-binding site and the replication origin [14]. Common conservative sequence for these regions is hexamer GGAGGC [12]. To obtain further evidence for the specificity of the Alu-binding protein, we have used in this study two synthetic double-stranded oligonucleotides containing the hexamer motif, and one cloned oligonucleotide with a double hexamer sequence (table 1), as well as synthetic oligonucleotides (AP1 and AP3) having no common sequences with AFRs.

Correspondence address: N.V. Tomilin, Institute of Cytology, Academy of Sciences of the USSR, Tikchoretskii Av. 4, Leningrad 194064, USSR

2. MATERIALS AND METHODS

Nuclear 0.4 M NaCl extract from HeLa cells isolated as described before [9,10] was used in all experiments of this study as a source of Alu-binding protein. The sequences of oligonucleotides used in this study are shown in table 1. All the oligonucleotides except T2 were synthesized using Gene Assembler (Pharmacia). Oligonucleotide T2 containing the SV40 T-antigen-binding site I (table 1) was isolated after cleavage of the plasmid pT2 with *Eco*RI and *Hin*dIII. Plasmid pT2, a gift from Dr S. Mirkin (Institute of Molecular Genetics AS USSR,

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

Moscow), was constructed by subcloning the double-stranded T2 oligonucleotide (table 1) tailored with the *Bam*HI and *Eco*RI cohesive ends into *Bam*HI, *Eco*RI-cleaved plasmid pUC19. A 58 bp *Eco*RI-*Hin*dIII fragment from pT2 containing a hexamer doublet and a part of pUC19 polylinker was 3′-end-labeled with ^{32}P using Klenow DNA polymerase and [α-^{32}P]dATP, and purified in a 5% polyacrylamide gel. Non-cloned synthetic oligonucleotide AluB4 was 5′-end-labeled using [γ-^{32}P]ATP and polynucleotide kinase. The double-stranded substrate was prepared by mixing ^{32}P-AluB4 with 3-fold molar excess of non-labeled AluB3, heating for 1 min at 90°C and then slowly cooling to room temperature in aDNA-binding buffer [10]. Non-labeled competitor oligonucleotides were prepared in the same way after mixing equimolar amounts of complementary oligonucleotides. A mobility shift assay and other methods were described before [9–12].

3. RESULTS AND DISCUSSION

Working with partially purified preparations of Alu-binding protein ($Q_{0.4}$ fraction, see [10]) and oligonucleotide T2 (table 1), we have detected two DNA-protein complexes, X1 and X2 [12], in a mobility shift assay and it was suggested that in the low mobility complex, X2, both GGAGGC hexamers are bound to protein and in the high mobility complex, X1, only one. To check whether human cells contain other proteins interacting with T2, we have used a mobility shift assay with this oligonucleotide, no partially purified Alu-binding protein, but a crude nuclear extract from HeLa and have observed complexes indistinguishable (fig.2) from the X1 and X2 complexes observed with the partially purified Alu-binding protein [12]. The data are consistent with the view that human cells contain only one major GGAGGC-binding protein and that there are no other DNA-binding proteins interacting with high affinity with T2 oligonucleotide. As expected, of the two AT-rich mammalian DNAs used (calf thymus and human placenta) human DNA showed the higher competitor activity (fig.2, lane 3), and the GC-rich DNA from *Micrococcus luteus* (fig.2, lane 2) was a better competitor than the DNA from *Escherichia coli* (fig.2, lane 4).

^{32}P-labeled double-stranded oligonucleotide AluB (table 1) containing a single GGAGGC hexamer was incubated with nuclear extract from HeLa cells in the presence of a fixed amount (100 ng) of calf thymus DNA and different amounts of non-labeled competitor oligonucleotides which were either non-ligated (fig.3A) or treated with T4 phage polynucleotide kinase and DNA ligase (fig.3B). It is seen from fig.3A that AluB oligonucleotide forms a single complex with protein, which is stable in the presence of 3–10-fold molar excess of GGAGGC-free oligonucleotide AP1 (fig.3, lanes 1 and 2). The same amount of complex was observed in the absence of oligonucleotides (not shown). However, the addition of AluB non-labeled oligonucleotide suppresses the

Table 1

Oligonucleotides used in this study

Name of ds oligonucleotide	Name of ss oligonucleotide	Sequence	Origin of sequence
AluA	AluA1	5′-GGAGGCTGAGGCA-3′	Alu consensus T-antigen-binding site
	AluA2	3′-CCTCCGACTCCGT-5′	
AluB	AluB4	5′-TGGGAGGCCAAGGA-3′	Alu left end conserved region
	AluB3	3′-ACCCTCCGGTTCCT-5′	
AP1	AP11	5′-GATCGTGACTCAGCGC -3′	transcription factor AP1 site in hMTII alpha promoter
	AP12	3′- CACTGAGTCGCGCTAG-5′	
AP3	AP31	5′-GATCTGGGACTTTCCACA -3′	SV40 early promoter factor AP3-binding site
	AP32	3′- ACCCTGAAAGGTGTCTAG-5′	
T2[a] (cloned)		5′-...GGAGGCTTTTTTGGAGGC...-3′ 3′-...CCTCCGAAAAAACCTCCG...-5′	SV40 T-antigen-binding site I from early region

[a] This synthetic oligonucleotide tailored with *Bam*HI (left) and *Eco*RI (right) cohesive ends was cloned into *Bam*HI-*Eco*RI cleaved pUC19 to get pT2. 3′-end-labeled *Hin*dIII-*Eco*RI 58 bp fragment from pT2 was used in the binding experiment (fig.2). Plasmid pT2 was kindly supplied to us by Dr S. Mirkin (Moscow)

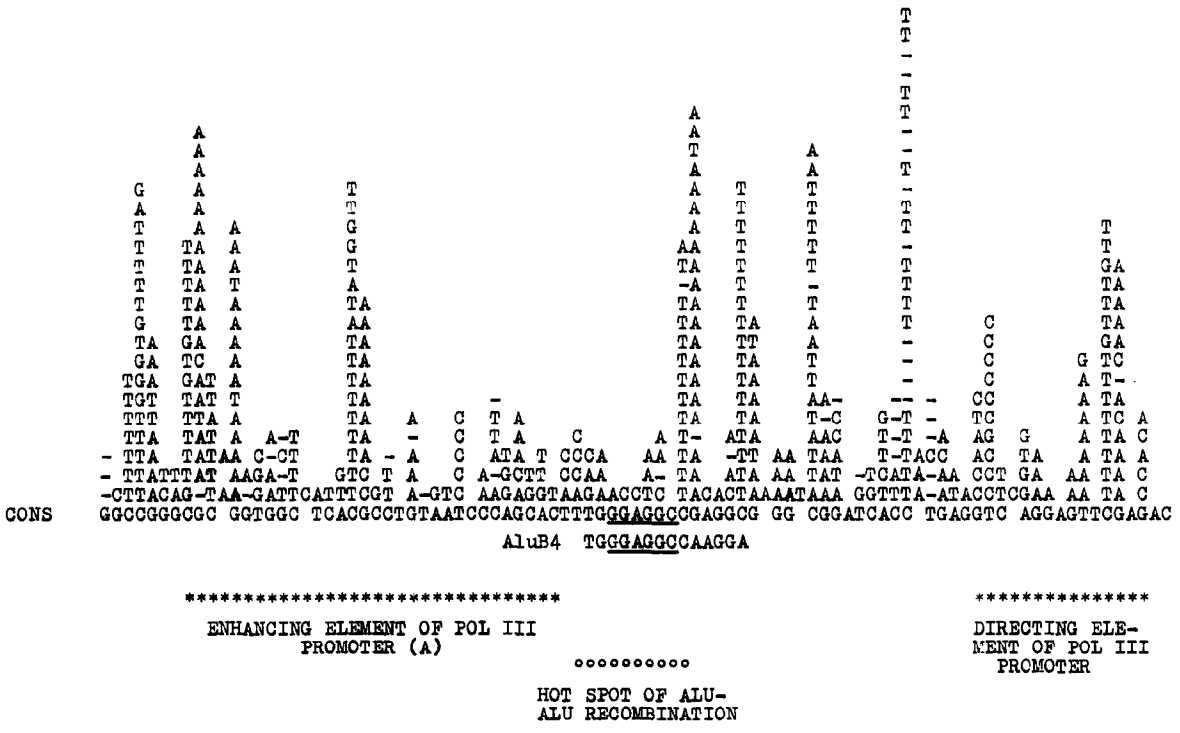

Fig.1. Conservative and variable regions at the left end of Alu repeats. CONS, consensus Alu sequence is taken from [7] and all deviations from the consensus in 40 ARFs from GenBank are plotted; −, nucleotide deletion. The position of a hot spot of Alu-Alu recombination is taken from [15], and of the polIII promoter from [4]. Note a high degree of variability at methylatable CG sequences of the consensus.

formation of the complex under the conditions used (fig.3, lanes 3 and 4). Similar but lower inhibiting effect showed non-labeled oligonucleotide AluA (fig.3, lane 6) which also contains GGAGGC (table 1). The same result was obtained with the ligated competitor oligonucleotides (fig.3B). In this experiment a longer incubation time was used which resulted in a decrease of the apparent competitor activity as compared to the experiment shown in fig.3A. The protein-binding site clearly present in AluA seems to have an affinity 2–3 times lower to protein than the site present in AluB, which was estimated from the relative concentration of competitor oligonucleotides leading to 50% binding inhibition. Again, GGAGGC-free oligonucleotide AP3 showed a much lower affinity to the GGAGGC-binding protein as compared to AluA and AluB (fig.3B).

We examined the binding inhibition of AluB by some non-labeled genomic DNAs (fig.4) and found that human DNA is about 3-times better a competitor than calf thymus or *E. coli* DNA, consistent with the view that the density of GGAGGC-binding sites in human DNA is higher than in other DNAs. In a random DNA sequence any definite hexamer is present on average once in 4 kb of DNA but human DNA contains additional GGAGGC sites which are conserved in Alu-repeats.

The results presented in this paper as well as those published before [11,12] strongly suggest that human cells contain high-affinity DNA-binding protein which recognizes the sequence 5′-GGAGGC-3′ conserved in the spacer region of the RNA polymerase III promoter of the Alu-family DNA repeats as well as in the T-antigen-dependent origin of replication of the repeats (fig.1 and [7]). The existence of such a protein anticipates the importance of the GGAGGC motif in the life cycle of the retrotransposon. Although other explanations are apparently possible, we would like to suggest the

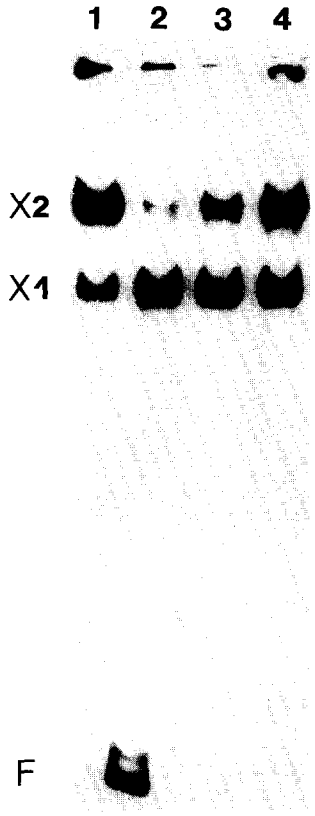

Fig.2. Gel retention pattern after binding of T2 oligonucleotide. ^{32}P-labeled T2 oligonucleotide (0.5 ng) was incubated for 30 min at room temperature with 1 μl of HeLa nuclear extract and 100 ng of calf thymus DNA (lane 1), *M. luteus* DNA (lane 2), human placental DNA (lane 3) and *E. coli* DNA (lane 4) in the binding buffer [10], runned in a 4% polyacrylamide gel, then gel was dried and radioautographed, F, free fragment; X1 and X2, complexes with protein.

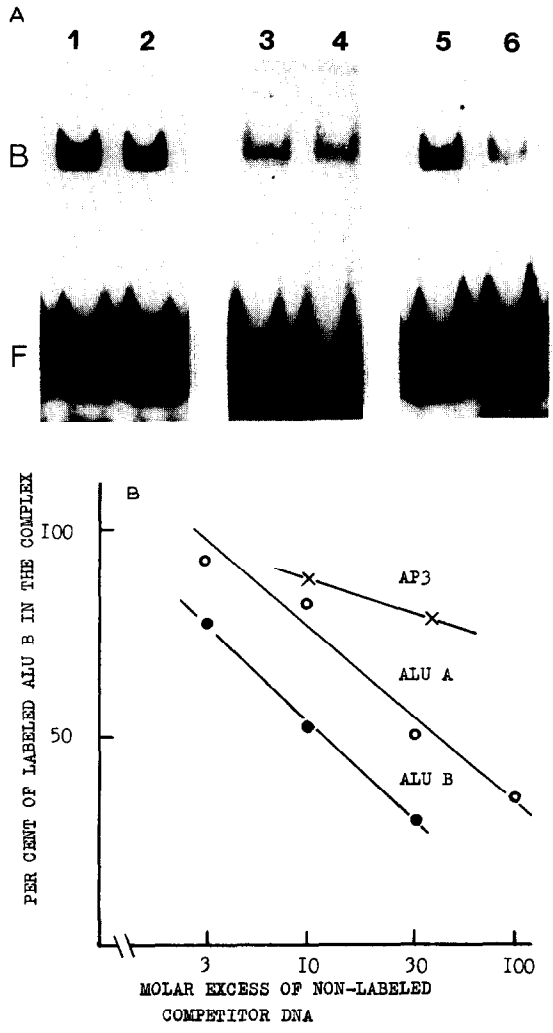

Fig.3. The inhibition of the binding of ^{32}P-labeled AluB oligonucleotide by different non-labeled oligonucleotides. 0.12 ng of ^{32}P-labeled AluB (table 1) was incubated for 10 min (A) or 30 min (B) at room temperature with 1 μl of HeLa nuclear extract in the presence of 100 ng of non-labeled calf thymus DNA and of different amounts of competitor oligonucleotides either non-ligated (A) or treated with T4 polynucleotide kinase and DNA ligase (B). In A, lanes 1 and 2 show the influence of 3- and 10-fold molar excesses of AP1 (no inhibition of binding), lanes 3 and 4, 3- and 10-fold excesses of AluB, lanes 5 and 6, 3- and 10-fold molar excesses of AluA. F, free oligonucleotide; B, complexes. For other methods see legend to fig.2.

hypothesis that the human GGAGGC-binding protein might be a repressor of RNA polymerase III-driven transcription of Alu-repeats, since binding sites for the protein are situated in the polIII promoter, and since AFRs are transcriptionally silent in HeLa cells in vivo [5] where the protein is present. The complexes of the GGAGGC-binding protein and Alu-repeats might be stabilised by cooperative binding at the two indicated hexamer sites [11] which are separated by 12.9 helical turns and may bind a single dimer of protein [11,17].

Interestingly, the left part of the spacer between enhancing and directing elements of polIII promoter (fig.1) is a hot spot of Alu-Alu recombination [15] which might be an indirect consequence of nucleotide sequence conservation in this region, since non-homologous recombination in non-conserved regions of Alu-repeats should be suppressed in somatic cells [16].

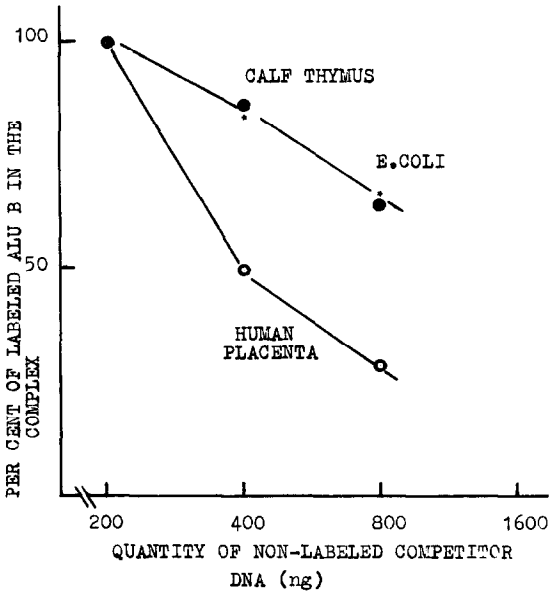

Fig.4. Competition of different non-labeled genomic DNAs for the binding of labeled AluB oligonucleotide. For methods see legend to fig.3. (○) Competitor DNA from human placenta, (●) DNA from calf thymus, (*) DNA from *E. coli*. The amount of complex (measured by densitometry of radioautograph) at 200 ng of a given competitor DNA is taken as 100%.

Acknowledgements: We thank Dr S. Mirkin for the gift of the plasmid pT2, and Dr F. Vikchanskaya who has prepared HeLa cells.

REFERENCES

[1] Jelinek, W.R. and Schmid, K. (1982) Annu. Rev. Biochem. 51, 823-844.
[2] Lin, C.S., Goldthwait, D.A. and Samols, D. (1988) Cell 54, 153-159.
[3] Rogers, J. (1985) Int. Rev. Cytol. 93, 187-279.
[4] Perez-Stable, C., Ayres, T.M. and Shen, C.K.J. (1984) Proc. Natl. Acad. Sci. USA 81, 5291-5295.
[5] Poulson, K.E. and Schmid, C.W. (1986) Nucleic Acids Res. 14, 6145-6158.
[6] Bains, W. (1986) J. Mol. Evol. 23, 189-199.
[7] Willard, C., Nguyen, H.T. and Schmid, C.W. (1987) J. Mol. Evol. 26, 180-186.
[8] Perelygina, L.M., Tomilin, N.V. and Podgornaya, O.I. (1986) Molekulyarnaya Genetika (Russian) 12, 15-19.
[9] Perelygina, L.M., Tomilin, N.V. and Podgornaya, O.I. (1987) Mol. Biol. (Russian) 21, 1610-1619.
[10] Perelygina, L.M, Tomilin, N.V. and Podgornaya, O.I. (1987) Mol. Biol. Rep. 12, 111-116.
[11] Podgornaya, O.I., Perelygina, L.M. and Tomilin, N.V. (1988) FEBS Lett. 232, 99-102.
[12] Tomilin, N.V., Perelygina, L.M. and Podgornaya, O.I. (1989) in: Proceedings of the 6th Sovjet-Italian Symposium 'Macromolecules in the Functioning Cells', Moscow, 1988, in press.
[13] Rubin, C.M., Houck, C.M., Deininger, P.L., Friedmann, T. and Schmid, C.W. (1980) Nature 284, 372-374.
[14] Johnson, E.M. and Jelinek, W.R. (1986) Proc. Natl. Acad. Sci. USA 83, 4660-4664.
[15] Lehrman, M.S., Goldstein, J.L., Russel, D.W. and Brown, M. (1987) Cell 48, 827-835.
[16] Radman, M. (1988) in: Genetic Recombination (Kucherlapati, R. and Smith, G. eds) pp. 169-192, Washington.
[17] Kramer, H., Niemoller, M., Amoyal, M., Rever, B., Von Wicken-Bergmann, B. and Muller-Hill, B. (1987) EMBO J. 6, 1377-1381.