POLYPYRIMIDINE TRACTS ISOLATED FROM INVERTED REPEAT SEQUENCES OF RAT DNA CONTAINING THE REPEAT SEQUENCE d(CTC)

S. SZALA, H. PATERAK and M. CHORAŻY
Department of Tumor Biology, Institute of Oncology, 44-100 Gliwice, Poland

Received 24 March 1980
Revised version received 12 May 1980

1. Introduction

After denaturation of duplex DNA some sequences with inverted repeats can reassociate by intramolecular reactions to form different types of hairpins or 'foldback DNA' [1]. The arrangement and distribution of inverted repeats occurs in DNA from a wide variety of eukaryotic sources [2–11]. However, the biological function of these sequences is still not explained. Inverted repeats may take part in recombination [12–14], in replication of DNA [15–17] or in the translocation of some DNA sequences during differentiation [18]. It has also been suggested that transcribed repeat sequences found in double-stranded regions of a primary transcript comprise processing sites within pre-mRNA [19–21].

In a search for sequence peculiarities of the foldback DNA, we found that a portion of inverted repeats in rat DNA contained long pyrimidine tracts (and complementary purine sequences) with an average length of about 80 nucleotides [22]. This report provides the evidence that the polypyrimidine · polypurine tracts in rat foldback DNA are mainly composed of the short repeat sequence:

$$d\begin{pmatrix} G-A-G \\ C-T-C \end{pmatrix}$$

We have also found that heterogeneous nuclear RNA contains pyrimidine tracts complementary to purine sequences in foldback DNA and that purine cRNA (cRNA(pur)) transcribed in vitro from pyrimidine tracts of foldback DNA can form complexes with the latter.

2. Materials and methods

2.1. *Isolation of foldback DNA*

DNA was isolated from Wistar rat liver and from rat kidney cells as in [22]. DNA dissolved in 0.12 M sodium phosphate buffer (pH 6.8) was denatured for 10 min at 100°C, rapidly cooled in an ice-bath, and the 'self-annealed' DNA isolated on a hydroxylapatite (HAP) column at 70°C [1]. To remove single-stranded DNA, foldback DNA was digested with endonuclease S_1, deproteinized with chloroform and again chromatographed on a hydroxylapatite column as above [22]. Foldback DNA prepared by the hydroxylapatite method, and deprived of single-stranded ends by digestion with S_1 endonuclease, retained the ability of instantaneous reassociation ($C_0t < 10^{-4}$) during at least 3 courses of thermal denaturation. Presumably foldback DNA prepared in this way consists of inverted sequences closely adjacent to each other, thus the 'loop' of snap-back structures consists of only a few nucleotides [1–3,7,22] and therefore is not accesible to S_1 attack.

2.2. *Isolation of polypyrimidines*

Long polypyrimidine tracts (poly[d(C,T)]) were isolated from S_1-treated foldback DNA according to [23].

2.3. *Isolation of nuclear RNA*

The steady-state nuclear RNA from rat liver nuclei was prepared according to [24]. The size of the RNA, determined by gel electrophoresis on 2% agarose under denaturing conditions (6 M urea, pH 3.5) [25], was between 28 and 35 S, i.e., in the range described [24].

2.4. *Transcription of DNA by* E. coli *RNA polymerase*

The purified polypyrimidine tracts and the foldback DNA (treated with endonuclease S_1) were transcribed into complementary, purine RNA (cRNA-(pur)) with RNA polymerase from RNase-free mutant MRE 600 of *Escherichia coli*. The incubation mixture contained in 32 μl total vol.: 0.033 M Tris · HCl (pH 7.9); 0.09 M KCl, 6 mM 2-mercaptoethanol, 33 mM $MgCl_2$, 5 units RNA polymerase (P-L Biochem., Milwaukee, WI), 0.1–1 μg polypyrimidine tracts or foldback DNA, 0.165 mM each of triphosphonucleotides (Sigma Chem., St. Louis, MO) containing 50 μCi [^{3}H]ATP and [^{3}H]GTP (spec. act. 20 Ci/mmol, The Radiochemical Centre, Amersham). Incubation time was 30 min at 37°C. Then 2 μg DNase (RNase free, Worthington Biochem., Freehold, NJ) was added and incubation continued for 15 min. The mixture was diluted to 100 μl with 0.5% sodium dodecyl sulphate (SDS) and extracted 2 times with phenol saturated with 1% SDS and 1 mM EDTA. RNA was precipitated 3 times with ethanol at –20°C. The size of cRNA was evaluated by electrophoresis on 7.5% polyacrylamide gel in a buffer containing 7 M urea, 0.09 M Tris–borate/2.5 mM EDTA (pH 8.3), with yeast tRNA markers (Sigma Chem., St Louis, MO). After electrophoresis, gels were frozen, sliced and each fraction counted with soluene–POP–POPOP–toluene scintillator.

2.5. *Fingerprinting of cRNA*

About 10^5 cpm of each ^{3}H-labeled cRNA (transcribed either from polypyrimidine tracts or from foldback DNA) were digested with RNase T_1 as in [7]. The oligonucleotides were separated in a two dimensional system: electrophoresis on cellulose acetate at pH 3.5, and thin-layer chromatography on cellulose impregnated with polyethyleneimine developed with pyridine–formate buffer (pH 3.5) [7]. The oligonucleotide spots were located by fluorography at –60°C and radioactivity of each spot was quantified in a scintillation counter.

2.6. *Thermal dissociation determinations*

Thermal chromatography of cRNA(pur):DNA hybrids was performed on HAP according to [26]. The column does not retain pyrimidine tracts of foldback DNA or cRNA(pur), which were eluted with 0.1 M sodium phosphate (NaP) buffer at 45–60°C. About 2×10^4 cpm of cRNA(pur) and 10 μg DNA or nRNA were mixed in 100 μl 0.12 M NaP at a cRNA:DNA ratio of ~1:1000. The mixture was denatured at 100°C for 5 min and incubation at 60°C for C_0t (mol.s/l) 0.1. Under these conditions >90% of cRNA(pur) was annealed to 10 μg DNA. Hybrids were then applied at 45°C to 1 cm^3 HAP equilibrated with 0.1 M NaP (pH 6.8). Non-hybridized cRNA was eluted 2 times with 6 ml of 0.1 M NaP buffer. Fractions of 6 ml from each 5°C increment were collected. Then 150 μg bovine serum albumin was added to each fraction and cRNA(pur) was precipitated with 5% trichloroacetic acid. Precipitates were collected on Whatman GF/B filters and radioactivity determined in a liquid scintillation counter. To avoid artefacts in determination of T_m, a 'window diagram' was constructed as in [27].

2.7. *R-loop formation*

To exclude reassociation between DNA strands and to assure that the reaction involved hybridization between RNA and DNA strands, R-loop formation was done according to [28]. Foldback DNA and cRNA(pur) samples in 4 M $NaClO_4$ were incubated first at 90°C for 1 min, then at 67.5°C for 30 min, and finally at 40°C for 30 min. The solution was diluted with 0.1 M NaP and applied on a HAP column as above.

3. Results and discussion

Although inverted repeat sequences have been found in a variety of organisms, their internal organization is not fully elucidated. Looking for internal high repeat sequences of polypyrimidine tracts isolated from foldback DNA of the rat, we have analysed the distribution and the molar ratio of oligonucleotides obtained by RNase T_1 digestion of complementary, polypurine RNA (cRNA(pur)). To avoid nearest-neighbour transfer of radioactivity from triphosphates labeled with ^{32}P [7], we used ATP and GTP labeled with ^{3}H. The size of cRNA transcribed from foldback DNA was 40–80 nucleotides, whereas the size of cRNA transcribed from polypyrimidine tracts was 50–200 nucleotides (fig.1). Both cRNAs were digested with RNase T_1 and oligonucleotides were separated on a two-dimensional system as in section 2 (fig.2a,b). The molar ratios of G, AG, A_2G etc. found in both cRNAs are presented in table 1.

cRNA(pur) is composed mainly of G and AG

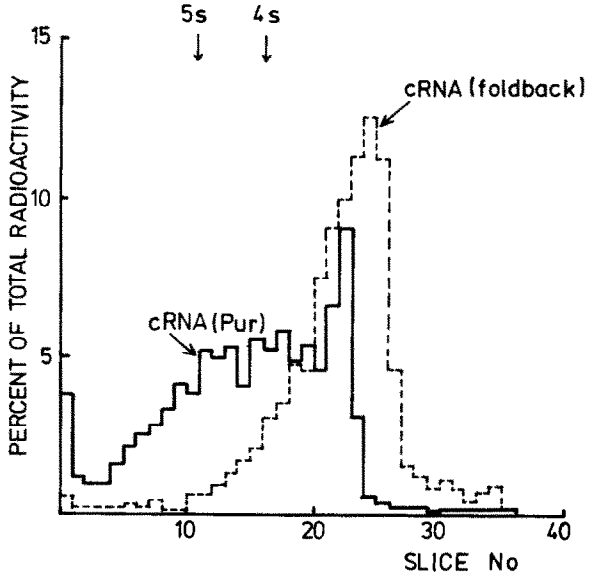

Fig.1. Size distribution of cRNA(pur) transcribed from polypyrimidine tracts isolated from the foldback DNA, and cRNA transcribed from total foldback DNA in 7.5% polyacrylamide gel under denaturing conditions. 5 S and 4 S denote 5 S RNA and tRNA markers, respectively.

Table 1
Molar ratio of oligonucleotides following RNase T_1 digestion of cRNAs

	Molar ratio				
	G	AG	A_2G	A_3G	A_4G
cRNA(pur)	1.00	1.19	0.12	0.015	–
cRNA(foldback)	1.00	0.38	0.05	0.01	0.002

Molar oligonucleotide ratio was calculated from the radioactivity of each spot and the relative specific activity of adenine and guanine nucleotide

(G+AG account for 89% of the total radioactivity) with G:AG ratio close to 1. Thus the major sequence in cRNA(pur) is (GAG)n corresponding to the sequence d(CTC) in foldback DNA. The different molar ratio of A and G in cRNA transcribed from the foldback DNA suggests that completely different, or additional, sets of sequences containing C and T were transcribed from the foldback DNA.

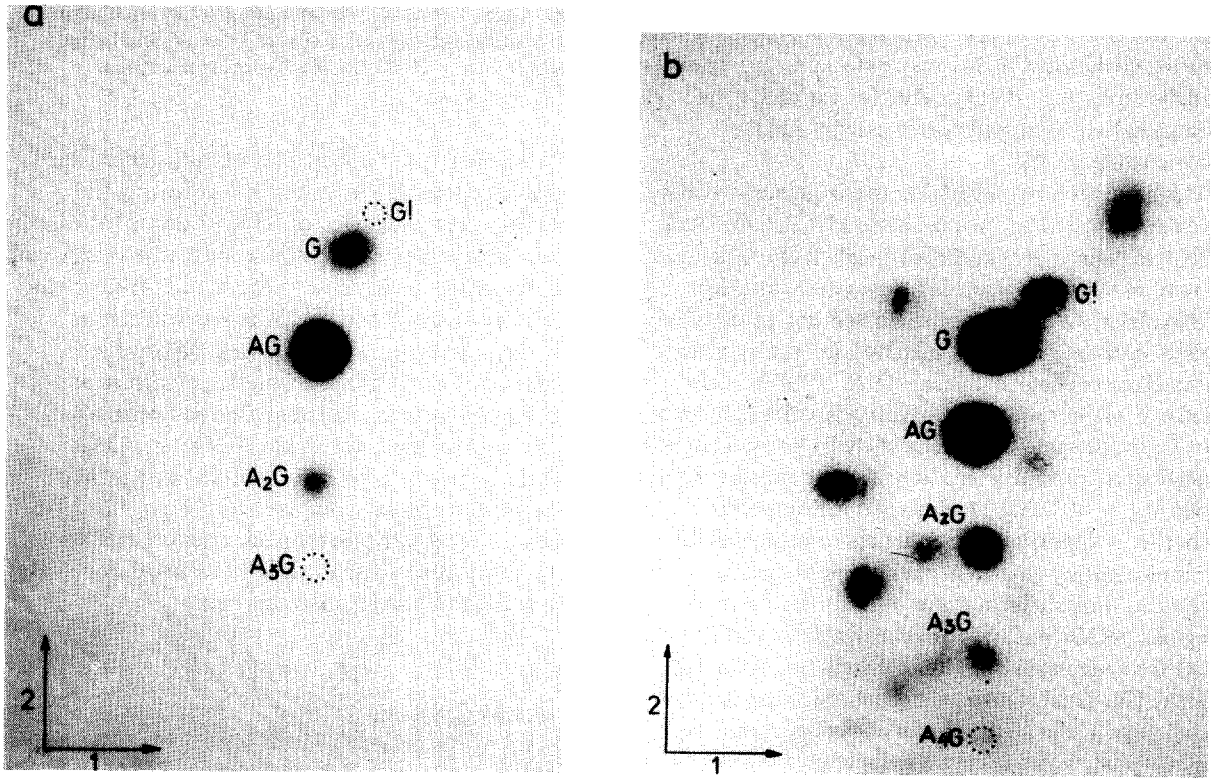

Fig.2. Fingerprints of RNase T_1 digest of cRNA uniformly labeled with [^{3}H]ATP and [^{3}H]GTP synthesized on polypyrimidine tracts isolated from the foldback DNA (a), and cRNA transcribed from foldback DNA (b). Oligonucleotides were separated in two dimensions: (1) high-voltage electrophoresis on acetate cellulose at pH 3.5; (2) on thin-layer plate of cellulose impregnated with polyethyleneimine developed with 1.5 M pyridine–formate buffer (pH 3.5). G! marks the position of cyclic phosphate.

The comparison of cRNA(pur) sequences with cRNA transcribed from foldback DNA implies that the fraction of d(CTC) tracts in foldback DNA is rather small. From this data it is not possible to estimate the internal arrangements of polypurine and polypyrimidine tracts in foldback DNA.

The internal periodicity of polypyrimidine tracts is not unique to the rat genome. Satellite DNA isolated from *Drosophila melanogaster* contains repeated sequences [29–32]:

$$\text{poly}\left[d\binom{A-A-G-A-G}{T-T-C-T-C}\right]$$

$$\text{poly}\left[d\binom{A-A-G-A-G-A-G}{T-T-C-T-C-T-C}\right]$$

$$\text{poly}\left[d\binom{A-A-A-G-G}{T-T-T-C-C}\right]$$

The conservation of polypyrimidine:polypurine tracts during evolution [33,34] may have some biological implications, e.g., the potential for forming triple-stranded helices of polypurine:polypyrimidine sequences [35,36].

A basic problem is whether the sequences in question are transcribed. According to [20,21], foldback DNA sequences are indistinguishable from sequences found in hairpin regions of heterogeneous nuclear RNA. Here we observed that labeled polypyrimidine tracts isolated from foldback DNA do not hybridize with 'heterogeneous' nuclear RNA: >90% of radioactivity was eluted from a hydroxylapatite column at 45°C. In the same conditions of hybridization labeled cRNA(pur) formed stable hybrids with nuclear RNA with T_m = 79°C (fig.3). However, it is not clear whether polypyrimidine sequences of nuclear RNA complementary to cRNA(pur) reside in double-stranded (hairpin) structures or in single-stranded regions. cRNA(pur) was obviously able to form 'hybrid' complexes with double-stranded, foldback DNA. This is substantiated by the following observations (fig.3): the HAP column does not retain cRNA(pur), which is eluted completely at 45–60°C; DNA:RNA hybridization in the presence of a vast excess of foldback DNA favours snap-back of foldback DNA rather than formation of a proper

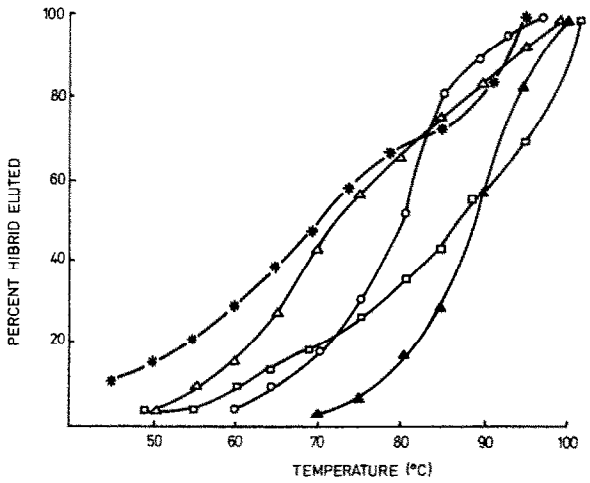

Fig.3. Hydroxylapatite thermal elution curve of hybrids formed between cRNA(pur) and different nucleic acids: (○) steady state nuclear RNA and cRNA(pur); (△) cRNA(pur) and foldback DNA (before reassociation with cRNA(pur) foldback DNA was denatured at 100°C and rapidly cooled; (*) cRNA(pur) and non-denatured, double-stranded foldback DNA (pre-treated with S_1); (▲) thermal elution curve of foldback DNA which was formely hybridized with cRNA(pur) followed by absorbance; (□) thermal elution curve of hybrids between foldback DNA and cRNA(pur) formed in 4 M $NaClO_4$ by the R-looping technique.

DNA:cRNA(pur) duplex. The strong interaction of complementary DNA strands is revealed by the melting curve with a T_m of 88°C observed by absorbance measurements. cRNA(pur) interacts weakly with foldback DNA, whether denatured at 100°C or not, and melts out at 72–73°C. When hybridization between cRNA(pur) and foldback DNA was performed in 4 M $NaClO_4$, i.e., under conditions where RNA:DNA interactions are preferred over DNA:DNA reassociation, the more stable cRNA(pur):DNA hybrid is observed, with a T_m of 86°C. One may assume that a 'weak complex' with T_m = 72–73°C may represent a triple-stranded structure (cRNA double-stranded foldback DNA); however more data are needed to justify this suggestion.

Acknowledgements

We thank Professor David Shugar for reading the manuscript. This work profited from the grant no 1301 of the Polish National Cancer Programm PR-6.

References

[1] Wilson, D. A. and Thomas, C. A. jr (1974) J. Mol. Biol. 84, 115–144.
[2] Schmid, C. W., Manning, J. E. and Davidson, N. (1975) Cell 5, 159–172.
[3] Cech, T. R. and Hearst, J. E. (1975) Cell 5, 429–446.
[4] Perlman, S., Philips, C. and Bishop, J. O. (1976) Cell 8, 33–42.
[5] Deininger, P. L. and Schmid, C. W. (1976) J. Mol. Biol. 106, 773–790.
[6] Bell, A. J. and Hardman, N. (1977) Nucl. Acids Res. 4, 247–268.
[7] Szala, S., Michalska, J., Paterak, H., Bieniek, B. and Chorąży, M. (1977) FEBS Lett. 77, 94–98.
[8] Bazetoux, S., Jouanin, L. and Huguet, T. (1978) Nucl. Acids Res. 5, 751–769.
[9] Deumling, B. (1978) Nucl. Acids Res. 5, 3589–3602.
[10] Hardman, N., Jack, P. L., Brown, A. J. P. and McLachlan, A. (1979) Eur. J. Biochem. 94, 179–187.
[11] Drahovsky, D., Boehm, T. L. J. and Kreis, W. (1979) Biochim. Biophys. Acta 563, 28–35.
[12] Sobel, H. M. (1973) Adv. Genet. 17, 411–490.
[13] Wagner, R. E. jr and Radman, M. (1975) Proc. Natl. Acad. Sci. USA 72, 3619–3622.
[14] Doyle, G. G. (1978) J. Theor. Biol. 70, 171–184.
[15] Cavalier-Smith, T. (1974) Nature 250, 467–470.
[16] Bateman, A. J. (1975) Nature 253, 379–379.
[17] Murakami, H., Taira, S. and Mori, H. (1977) J. Theor. Biol. 68, 183–197.
[18] Dickinson, D. G. and Baker, R. F. (1978) Proc. Natl. Acad. Sci. USA 75, 5627–5630.
[19] Robertson, H. D., Dickson, E. and Jelinek, W. (1977) J. Mol. Biol. 115, 571–589.
[20] Jelinek, W. (1977) J. Mol. Biol. 115, 591–602.
[21] Jelinek, W., Ewans, R., Wilson, M., Salditt-Georgieff, M. and Darnell, J. E. (1978) Biochemistry 17, 2776–2783.
[22] Szala, S., Bieniek, B., Michalska, J. and Chorąży, M. (1976) Biochim. Biophys. Acta 432, 129–144.
[23] Straus, N. A. and Birnboim, H. C. (1974) Proc. Natl. Acad. Sci. USA, 71, 2992–2995.
[24] Sippel, A. E., Hynes, N., Groner, B. and Schütz, G. (1977) Eur. J. Biochem. 77, 141–151.
[25] Woo, S. L. C., Rosen, J. M., Liarakos, C. D., Choi, Y. C., Busch, H., Means, A. R., O'Malley, B. W. and D. L. Peterson (1975) J. Biol. Chem. 250, 7027–7039.
[26] Miyazawa, Y. and Thomas, C. A. jr (1965) J. Mol. Biol. 11, 223–237.
[27] Martinson, H. G. and Wagenaar, E. D. (1977) Biochim. Biophys. Acta 474, 445–455.
[28] Chien, Y. H. and Davidson, N. (1978) Nucl. Acids Res. 5, 1627–1637.
[29] Sedoroff, R., Lowenstein, L. and Birnboim, H. C. (1975) Cell 5, 183–194.
[30] Endow, S. A., Polan, M. L. and Gall, J. G. (1975) J. Mol. Biol. 96, 665–692.
[31] Brutlag, D., Carlson, M., Fry, K. and Hsieh, T. S. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 1137–1146.
[32] Fry, K. and Brutlag, D. (1979) J. Mol. Biol. 135, 581–593.
[33] Straus, N. A. and Birnboim, H. C. (1976) Biochim. Biophys. Acta 454, 419–428.
[34] Birnboim, H. C. (1978) J. Mol. Biol. 121, 541–559.
[35] Johnson, D. and Morgan, A. R. (1978) Proc. Natl. Acad. Sci. USA 75, 1637–1641.
[36] Lee, J. S., Johnson, D. A. and Morgan, A. R. (1979) Nucl. Acids Res. 6, 3073–3091.