Volume 5 Number 7 July 1978 Nucleic Acids Research

Periodic organisation of foldback sequences in Physarum polycephalum nuclear DNA

Norman Hardman and Peter L.Jack

Department of Biochemistry, University of Aberdeen, Marischal College, Aberdeen AB9 1AS, UK

Received 3 April 1978

ABSTRACT

Nuclear DNA from the slime mould Physarum polycephalum is shown to contain interspersed inverted repeat sequences, such that denatured fragments of DNA containing pairs of these sequences form intra-chain duplexes under appropriate conditions. The organisation and distribution of the nucleotide sequences responsible for the formation of foldback structures in Physarum DNA have been investigated using the electron microscope. The majority of foldback duplexes have sizes ranging up to 800 base pairs, and about 60-80% of DNA molecules 2.2 x 10^4 bases in length contain interspersed foldback elements. The size of individual foldback duplexes, and also the length of the intervening sequences which separate them, are non-random. The results can best be explained by a model in which separate foldback foci in Physarum DNA are spaced periodically at regular intervals. The regions containing foldback foci are thought to contain smaller, tandemly-arranged sequences of discrete sizes, in some cases related to other nucleotide sequences of a similar nature in the same locality in Physarum DNA.

INTRODUCTION

A significant proportion of nuclear DNA from eukaryotic cells, when denatured and annealed under suitable conditions, forms duplexes very rapidly and independently of DNA concentration. This special fraction has been termed foldback DNA by Britten and Smith [1] because it is understood to contain DNA molecules with duplex segments which result from intra-chain annealing of inverted, complementary sequences. Consistent with this interpretation, foldback DNA observed using the electron microscope has been shown to contain a preponderance of molecules with a foldback configuration [2-5]. Most of the duplexes in foldback molecules in DNA from animal cells form a distribution of

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

lengths ranging from 100-1000 base pairs [2,3,6], and the sequences forming foldback duplexes are widely distributed amongst single-copy sequences [2].

Lower eukaryotic organisms such as the slime moulds Physarum polycephalum and Dictyostelium discoideum have proved useful as model systems in the study of biochemical changes which occur during the cell 'cycle' or differentiation. Consequently, it is of interest to know whether the kinds of mechanism which operate in these organisms to control the expression of genetic information are directly comparable to those in animal cells. Similar to higher eukaryotic organisms, nuclear DNA both from Dictyostelium and Physarum contain repetitive sequences [7,8]. Accordingly, due to the recognised importance of repetitive sequences in eukaryotic DNA as potential control elements in gene expression in animal cells, experiments have been carried out in our laboratory to characterise the repetitive component in Physarum polycephalum nuclear DNA.

In our previous studies [9,10], it has been shown that the character of Physarum foldback duplexes is very similar to the corresponding fraction from human DNA [11]. In the present report we have extended this earlier work and determined the distribution of foldback duplexes in longer chains of Physarum DNA using the electron microscope. The majority of these longer chains of Physarum DNA contain foldback structures, many containing multiple foldback foci. Successive foldback duplexes are separated by intervening sequences of unexpectedly uniform lengths, suggesting that the interspersed sequences from which they are derived are located at regular intervals throughout a substantial proportion of the Physarum genome.

MATERIALS & METHODS

Growth of Physarum cultures

The macroplasmodial form of Physarum polycephalum was used in this study, and was grown on filter papers over growth medium as described previously [9], except that in the present work the Colonia Leicester (CL) strain of

Physarum was used [12]. The organism was a generous gift from Dr. J. Dee, Department of Genetics, University of Leicester, U.K.

Labelling and isolation of DNA

Labelling of DNA in macroplasmodia was carried out by the addition of 2 μCi of [methyl-^{3}H]thymidine/ml (specific activity 20 Ci/mol) to the growth medium, as described previously [9]. Nuclei were isolated from G2-phase plasmodia following the third post-fusion mitosis in the presence of [^{3}H]thymidine, using the method of Mohberg and Rusch [13]. Plasmodia were homogenised for 30 s at low speed using an Ato-Mix blender fitted with a Polytron BEW-5 head (Bronwell Scientific Inc., Rochester, N.Y.) mounted in an MSE 800 ml stainless steel chamber.

Isolated nuclei from five macroplasmodia were suspended in 5 ml of lysis buffer containing 30 mM NaCl, 10 mM EDTA and 50 mM Tris-HCl, pH 8.0. Sarkosyl NL was added to a final concentration of 1%(v/v) followed by 200 μg Proteinase K/ml. The mixture was then incubated at 45°C overnight. The total volume was increased to 10.5 ml by the addition of lysis buffer, then analytical grade CsCl was added to give a final density of 1.710 g/ml. The solution was divided into three equal portions which were centrifuged to equilibrium in a Type 40 rotor in a Beckman L2 preparative ultracentrifuge at 40,000 rev/min for 60 h at 20°C. Fractions were collected from the bottom of the gradients and samples assayed for radioactivity. Appropriate fractions were pooled and dialysed into 20 mM NaCl, 1 mM EDTA and 10 mM Tris-HCl, pH 8.0.

Alkaline sucrose gradient fractionation of DNA

0.5 ml of a solution containing 10 μg [^{3}H]-labelled Physarum DNA/ml was denatured by incubation at room temperature in the presence of 0.3 M NaOH for 30 min. This solution was layered over a 5-10%(w/v) sucrose gradient containing 0.15 M NaCl, 0.05 M NaOH and 0.005 M EDTA, and the gradient centrifuged in an SW27 rotor in a Beckman L2 ultracentrifuge at 27,000 rev/min. Following centrifugation

twenty fractions were collected and assayed for radioactivity. The labelled material sedimented as a unimodal peak. A 100 µl sample from a fraction taken from the leading edge of the radioactivity profile, containing 14% of the total recovered radioactivity, was neutralised by the addition of an equal volume of 0.4 M sodium phosphate, pH 6.8, at 4°C. A 30 µl sample of this material was then taken and prepared for observation in the electron microscope, as described below.

Electron microscopy

DNA obtained from alkaline sucrose gradients was prepared immediately for observation by electron microscopy without intentional annealing. In the 3-5 min taken to prepare and spread molecules it is calculated that the equivalent $C_o t$ value [$C_o t$ is the product of DNA concentration and time of annealing] assuming optimal conditions for annealing in the hyperphase spreading solution, is 0.06-0.1 mmol s/l. The final DNA concentration in the hyperphase did not exceed 0.1 µg/ml.

Preparations were made using the Kleinschmidt monolayer technique, incorporating the modifications of Davis and Hyman [14] and additional minor modifications. The hyperphase solutions contained DNA, 0.12 M sodium phosphate, 0.03 M NaCl, 7%(w/v) sucrose, 10 mM EDTA, 30 µg cytochrome C/ml, 55%(v/v) formamide and 0.1 M Tris-HCl, pH 8,5. The hypophase contained 20%(v/v) formamide in 10 mM Tris-HCl, pH 8.5. Using the relationship between formamide concentration and equivalent temperature [15] it is calculated that 55% formamide produces an equivalent effect to raising the temperature to 60°C, room temperature being 20°C. Monolayers were mounted on parlodian-coated copper grids and rotary shadowed with Pd/Pt (1:4, w/w). Samples were observed in a Philips EM400 electron microscope.

Length data taken from electron micrographs were converted to nucleotide base equivalents using the factors 1 µm = 3050 base pairs for duplexes, and 1 µm = 3930 bases for single chains. These factors were determined by spreading molecules of Lambda duplex DNA and ØX174 single

chain circles respectively as standards under similar conditions to test molecules. Measurements were taken from micrographs using a Depose H-C map measurer.

RESULTS & DISCUSSION

In a previous report [9] it was not possible from the available data to determine whether foldback foci in Physarum DNA were located randomly or arranged in clusters in the genome, in the latter case in accord with the distribution of foldback sequences determined for some higher organisms [2]. The approach to the problem adopted in this present study has been to investigate the distribution offoldback foci in longer single chain DNA fragments.

In our hands it has proved difficult to isolate native DNA from Physarum cultures completely free of hidden single chain breaks. When DNA from our preparations is denatured, single chains with a wide spectrum of lengths is obtained, such that the proportion of usefully long DNA molecules containing multiple foldback foci is small. In the present work this difficulty has been overcome by the selection of long single chains of Physarum DNA by alkaline sucrose gradient fractionation, as described in experimental methods. Appropriate gradient fractions were pooled, neutralised and spread for observation in the electron microscope after annealing had taken place to an estimated $C_o t$ value of 0.1 mmol s/l, which is expected to allow the formation of foldback duplexes but preclude significant annealing of intermolecular complementary sequences [9].

Table 1 shows the classification of foldback duplex-containing structures observed in electron micrographs of alkaline sucrose gradient-fractionated DNA molecules of average length 2.2×10^4 bases ($\pm 0.47 \times 10^4$ bases standard deviation). The various classifications of hairpin structures have been described previously [9]. Examples of some multiple hairpin structures are shown in Figure 1 A-D.

One of the problems associated with experiments of this type is that it is difficult to extend very long

Table 1 : Classification of foldback structures obtained from Physarum polycephalum nuclear DNA

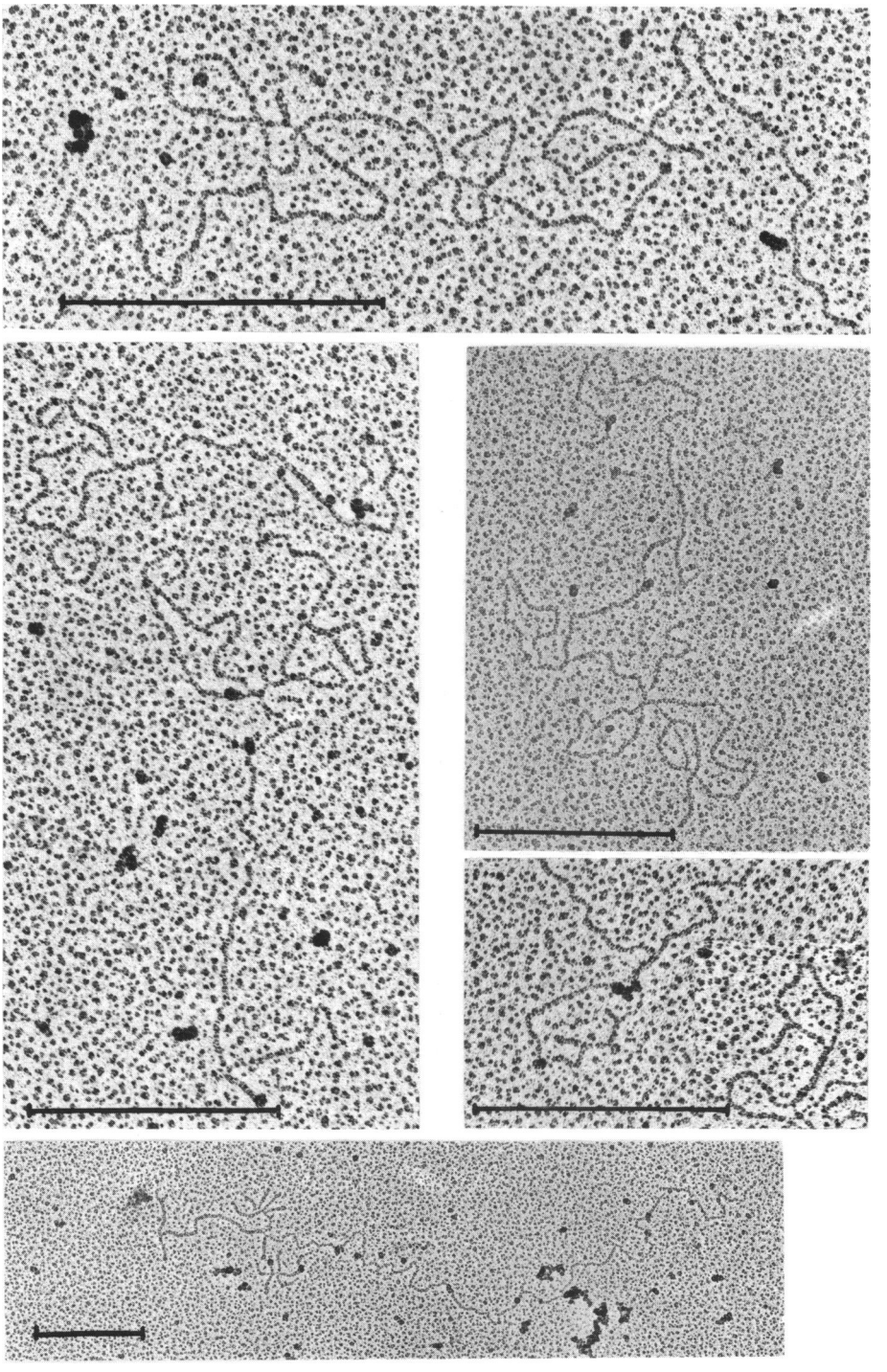
Structural type	Amount in : Traceable Molecules	Tangled Molecules
	% of Total	
Single chains	27.9	-
Solitary hairpins :		
Looped	13.6	-
Unlooped	15.6	-
Multiple hairpins :		
2*	8.2	9.5
3	3.4	3.5
4	2.1	1.3
5	1.3	1.1
Forks, complex or abiguous structures	-	12.4

*Figures refer to the number of hairpins observed per DNA chain

Physarum nuclear DNA was fractionated on alkaline sucrose gradients and prepared for electron microscopy as described in experimental methods. The data shown were obtained from 578 structures from a fraction containing molecules with an average chain length of 22,000 ± 4700 bases, determined from the chain length of 40 molecules traced at random from electron micrographs. Examples of unlooped and looped hairpins are shown in Figure 1A and Figure 1B respectively. Multiple hairpins are classified according to the number of hairpins present in the DNA chain. Tangled molecules are those in which chain overlaps precluded an unambiguous determination of the distance between any two successive foldback duplexes. The category containing forks and complex structures is referred to in the text.

DNA fragments containing many separate duplexes and obtain totally unambiguous structures for length measurements. Although all structures were classified, as shown in Table 1,

structures in which chain crossovers precluded a complete interpretation were not used for length analyses, and are classified separately in Table 1 as tangled molecules. Due to the comparatively large numbers of these partly ambiguous molecules it is necessary to exclude the possibility that a particular class of structures is selected for length measurements. First, it can be seen from Table 1 that the distribution of hairpin types in the various classifications is very similar for both traceable and tangled molecules. Second, measurements of traceable duplex stems on hairpin segments in tangles (data not shown) gives a similar size distribution to the measurements presented below for completely traceable structures. It is concluded that tangled structures probably do not contain a special class of foldback duplexes responsible for their more complex appearance, but rather they result as a consequence of the failure of some DNA molecules to unfold properly during the spreading procedure. A separate classification in Table 1 includes a minor fraction of molecules containing non-foldback structures, and other complex structures. The class includes forks, and long, linear duplexes which may have an intermolecular origin. This class also includes those foldback molecules characteristic of Physarum ribosomal DNA [16], an example of which is shown in Figure 1 E. In spite of the fact that the appearance of Physarum rDNA is highly characteristic, because of random fragmentation and the known variability in the configuration of rDNA structures [16] the possibility cannot be completely excluded that occasional foldback structures in the main classifications may be derived from rDNA fragments. However, since rDNA accounts for only 1-2% of the nuclear DNA complement in Physarum, and because 60-83% of fragments 2.2×10^4 bases long contain foldback elements, from Table 1, it is concluded that the overwhelming proportion of foldback duplexes result from the annealing of sequences derived from the remaining 98% of Physarum nuclear DNA. It is considered very unlikely that alkaline sucrose gradient fractionation would lead to an enrichment of particular



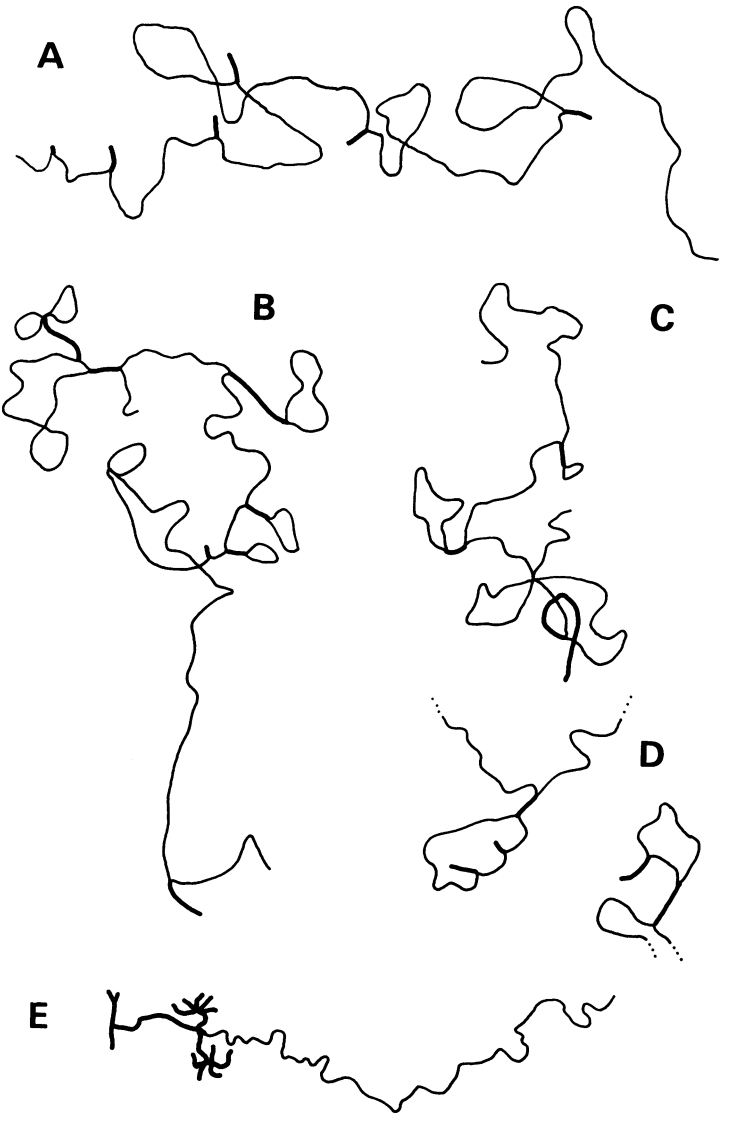


Figure 1 : Electron micrographs of Physarum foldback DNA

The foldback structures shown are examples of those obtained from nuclear DNA molecules 22,000 bases in length, referred to in Table 1. Figure 1A-D are molecules containing multiple looped and unlooped hairpins. Figure 1E is a presumptive fragment of the extrachromosomal nucleolar satellite rDNA, which is an expected minor component of Physarum nuclear DNA. Molecules with this characteristic appearance comprised about 0.8% of the total structures. Length bars are 0.5 μm.

sequences of DNA, but this possibility cannot be excluded.

Figure 2a shows the frequency distribution of hairpin loop lengths measured from looped hairpins on traceable molecules classified in Table 1. Similarly, Figure 2b shows the distribution of interhairpin distances on traceable multiple hairpin structures. The average measured distances between pairs of unlooped hairpins (UU-pairs), looped hairpin pairs (LL-pairs) and pairs of alternate

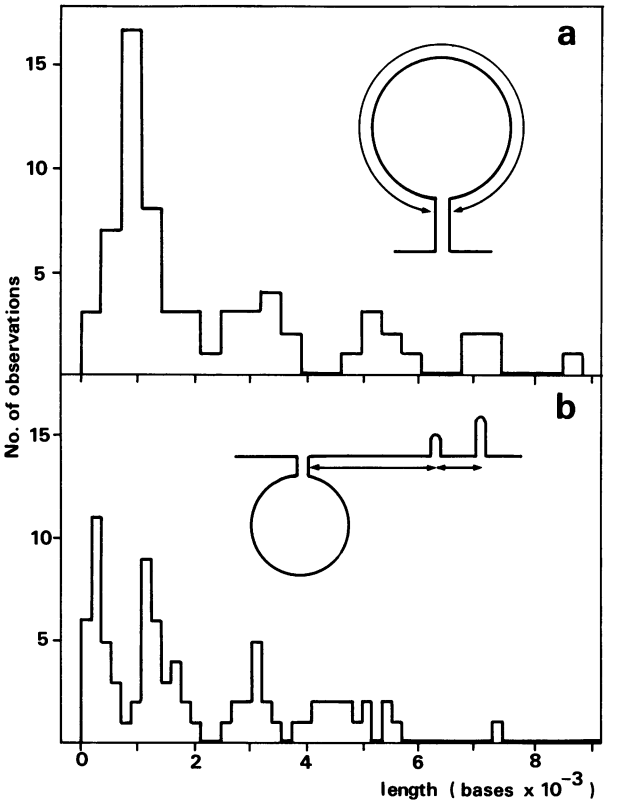


Figure 2 : Hairpin loop lengths and interhairpin distances in foldback DNA

Figure 2a shows the distribution of loop lengths measured from looped hairpin structures selected at random. The distribution of interhairpin distances in multiple hairpin structures is shown in Figure 2b. As described in the text, no significant difference was obtained in the measurements of the average distance between unlooped hairpin pairs, looped hairpin pairs, or looped and unlooped hairpins.

hairpin types (UL-pairs) was 2520, 2670 and 2400 bases respectively, which are considered to be the same within experimental error. 40% of the double hairpins were UU-pairs, 34% UL-pairs and 26% LL-pairs, giving a ratio of UU : UL : LL of 1 : 0.86 : 0.67. These ratios are different from those expected for a completely random distribution of looped and unlooped hairpins based on the observed frequency of solitary hairpins in traceable structures (1 : 1.79 : 0.79), the looped : unlooped hairpin ratio being 0.88 : 1 from Table 1.

Ordered arrangement of foldback sequences in Physarum DNA

The results presented in Figure 2 clearly show that the distribution of measurements for loops and interhairpin distances is non-random. The periodicities evident from the data suggest that the nucleotide sequences forming foldback structures are spaced at regular intervals in segments of Physarum DNA. Figure 3 shows schematically the interpretation of the observed periodicities.

The first period, 0-700 bases in Figure 2b, is represented in Figure 3 by a 700 base-long segment indicating a region of high probability of locating two separate sequences each capable of independently forming adjacent foldback foci with their respective complementary inverted sequences. Using the same approach the second period, 900-2100 bases from Figure 2b, is shown as representing cases where separate segments of DNA of equal size, 600 bases [(2100-900)/2] each have a high probability of containing a foldback sequence, and these are separated from each other by a 900 base-long spacer segment devoid of foldback duplexes. However, to be consistent with the data the 900 base intervening element could be located at any point within the 2100 base boundary. The longer-range, 2500-3700 base and 3800-5600 base periods in Figure 2b are interpreted in a similar manner, shown in Figure 3. It can also be seen that the most probable lengths for loops in looped hairpins (Figure 2a) form a set of similar, overlapping periodicities, also shown in Figure 3. In contrast to the data for interhairpin distances, however,

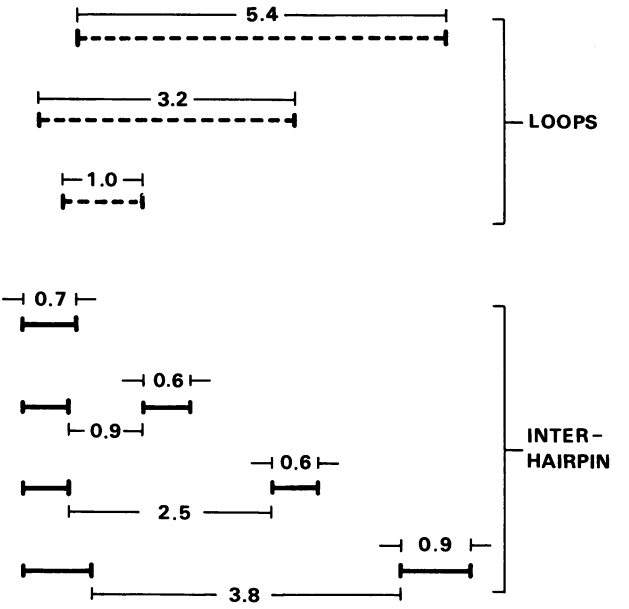


Figure 3 : Interpretation of periodicities in the organisation of foldback sequences in Physarum DNA

Dotted lines show the most probable lengths for loops on looped hairpins, derived from the data in Figure 2a. Solid bars define the regions of high probability of locating nucleotide sequences capable of forming separate foldback duplexes based on the distribution of interhairpin distances shown in Figure 2b. The scheme illustrates a regular arrangement in which separate regions (R-regions) containing foldback sequences are spaced at intervals of about 1000, 3200 and 5400 bases. Most probable values for loop lengths coincide with the periodic arrangement deduced independently from interhairpin distances. Longer periods consist of approximately integral multiples of the shorter periodic units, as described in the text. The figures are bases x 10^{-3}.

this latter data suggests that related inverted nucleotide sequences, capable of forming looped foldback duplexes, are located at separate, discrete sites in Physarum DNA. The data in Figure 2 is thus consistent with the idea that separate foci containing potential foldback sequences, some of which are related, are spaced by more or less discrete distances of about 1000, 3200 and 5400 bases in Physarum DNA. There is additional tentative evidence in Figure 2a of longer overlapping periodicities, but

these are less frequently observed than the shorter-range periods.

Properties of foldback duplexes

Figure 4a and 4b show respectively the distribution of duplex stem sizes determined from looped and unlooped hairpins. Unlooped hairpin duplexes, and most looped hairpin duplexes, form a narrow distribution of lengths between 70-600 base pairs. Occasional looped hairpins contain far longer duplexes, up to 4500 base pairs long. These results confirm our earlier observations in a different strain of Physarum [9]. However, in the present study, larger numbers of structures have been analysed and narrower windows have been taken for the histogram size distributions. There is evidence from the data that hairpin duplexes form a discontinuous distribution of lengths. Common lengths of 190 and 370 base pairs are found for looped hairpin duplexes shown in Figure 4a, and 120 or 370 base pairs for unlooped hairpin duplexes, indicated in Figure 4b. These estimates are nearly equal in magnitude to the terms of an arithmetic series of lengths forming multiples of an approximately 60 base pair sequence element (60, 120, 180, 240, 300, 360 ....).

A model for the sequence organisation of Physarum DNA

Physarum foldback duplexes are shown here to form discrete collections of lengths, longer duplexes seemingly consisting of reassociation registers forming an arithmetic series of lengths differing by multiples of about 60 base pairs. If the analysis is correct, then it follows that the separate sequences from which foldback duplexes are derived have a similar periodic nature. A model sequence arrangement which can simply, and most satisfactorily account for this observation assumes that foldback duplexes might be formed from sequences which are composed of smaller tandemly-arranged units, each unit being of a common size of around 60 bases in length. An illustration of the model is shown in Figure 5, which attempts to account for the observed distribution of foldback sequences and the nature

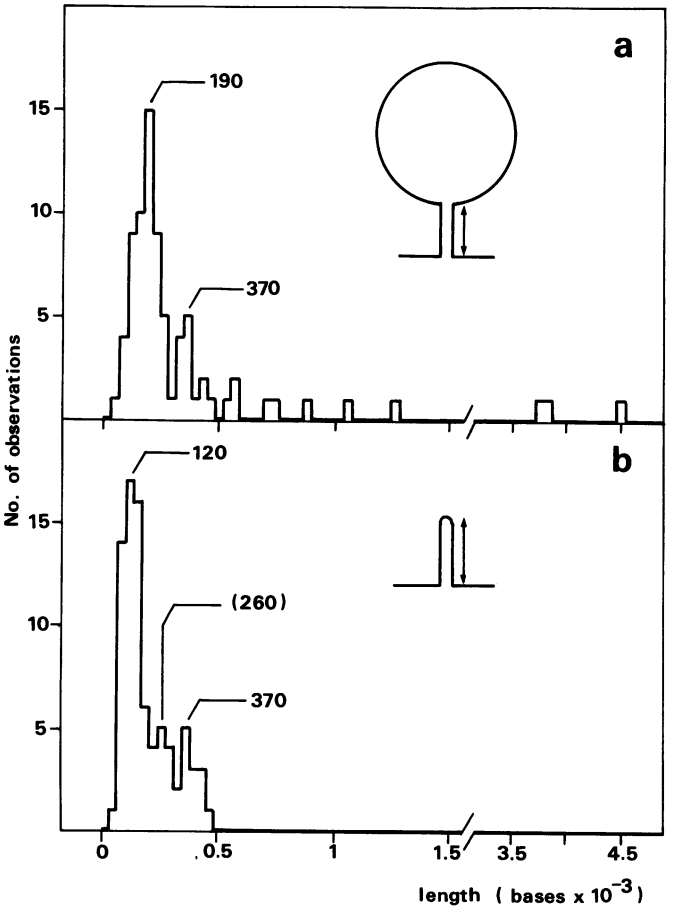


Figure 4 : Distribution of hairpin stem lengths

The distribution of stem lengths for foldback duplexes on looped and unlooped hairpins is shown in Figure 4a and Figure 4b respectively. No distinction has been made between the duplexes on solitary hairpins, or those taken from multiple hairpin structures.

of foldback duplexes. Foldback sequences are envisaged to be located in discrete regions, denoted by the segments R, which consist of tandemly-arranged sequences of unit length s. In order to account for the properties of unlooped hairpins we suppose that some R-regions contain immediately adjacent s-units which are related in sequence and inverted between the two chains of the native duplex,

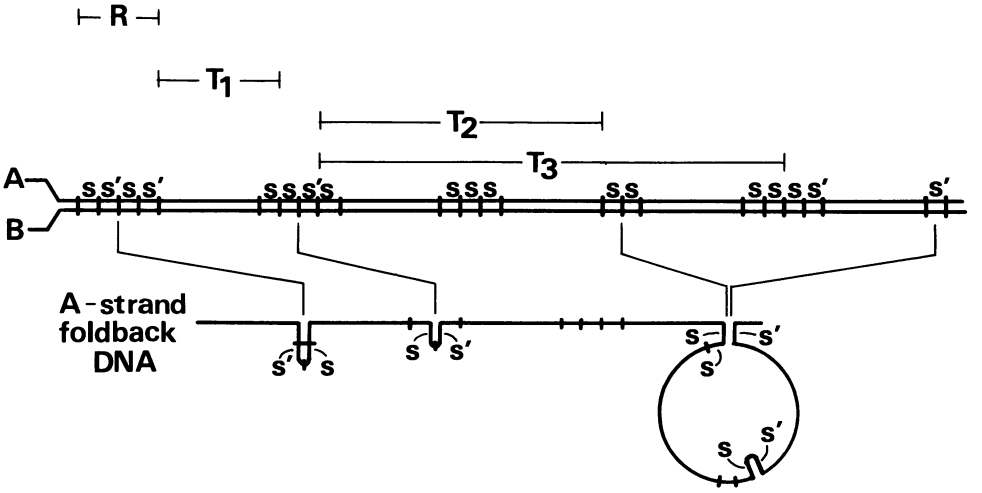


Figure 5 : A model for the organisation of foldback sequences in Physarum nuclear DNA

The figure illustrates a duplex fragment of Physarum DNA containing segments, R , containing potential foldback sequences. Separate R-regions are thought to consist of a number of smaller internal units of a common length, s . Some of the s-units in a given R-region, and in different R-regions in the same locality in the DNA, are related, and some of these are inverted between the two strands A and B of the native duplex, such that units of s and their respective inverted sequence complements s' are sometimes present in the same DNA chain and are thus capable of duplex formation when the DNA is denatured. The lower chain diagram illustrates one possible configuration that might be adopted by intramolecular annealing of the denatured A-strand. Depending upon the sequence relationship between the units s and s' in the same or in different R-regions, alternative foldback configurations might result. Unlooped hairpins result from the annealing of juxtaposed, inverted complementary repeated sequences from within a single R-region. Looped hairpin duplexes bring together inverted complementary repeats from separate R-regions. Potential foldback duplexes are spaced by discrete distances of 1000, 3200 and 5400 bases along the DNA chain, corresponding to the distances T_1, T_2 and T_3, which represent the length of segments between regularly spaced R-regions separated by 1, 2 and 3 interspersed non-foldback sequence spacers respectively (see text), possibly together with intervening R-regions which, for reasons discussed in the text, may be unable to participate in foldback duplex formation.

giving rise to the possibility of intra-chain annealing of copies of s with their inverted, adjacent sequence

complements, denoted by s'. Permitted lengths for the duplexes formed in this way would be s, 2s, 3s, 4s ...etc. but the actual lengths would depend on the sequence relationship between successive segments s in an R-region, and also upon local arrangements of s and s'.

R-regions are shown in Figure 5 to be separated from sequences of a similar nature by linker sequences of discrete lengths, representing segments of distinct sequence devoid of foldback structures. The chain diagram in Figure 5 indicates the different kinds of foldback structure which might result from the sequence arrangement described above, showing how looped hairpins could form by annealing complementary sequences from separate R-regions. This could also account for the overlapping periodicities evident from the independent measurements of interhairpin distances and loop lengths shown in Figure 2, as analysed in Figure 3. Experiments in which foldback duplexes from Physarum DNA are eluted from hydroxyapatite crystals as a function of temperature [10] indicate that the duplexes in foldback DNA are much less stable than native DNA, and may be composed of imperfectly-matched sequences. Hence, foldback duplexes could be derived from related, but not identical, sequences ; mismatched but sufficiently stable to be seen as duplexes in the present study.

In the version of the model shown in Figure 5 it is supposed that the intervening sequences forming the longer periodicities shown in Figure 3 need not be devoid of additional R-regions to be consistent with the data. These additional R-regions would have to be unable to form hairpins with neighbouring R-regions, as illustrated in Figure 5. Conceivably, this might be due to either some R-regions in close proximity possessing unrelated or poorly-matched s-units, or because some adjacent R-regions might contain direct register repeated units $\ldots s_1 s_2 s_3 \cdots\cdots s_1 s_2 s_3 \cdots\cdot$ rather than inverted, complementary repeats $\ldots s_1 s_2 s_3 \cdots\cdots s_3' s_2' s_1' \cdots\cdot$ thus possessing the wrong sequence configuration for duplex formation. In this context, and consistent with this idea, it is noteworthy that the spacings of the

longer-range, 2500-3700 and 3800-5600 base periods illustrated in Figure 3 overlap with the 900-2100 base period. A further consideration which lends credibility to the above idea is that some hairpin structures contain secondary hairpin duplexes spaced along the single chain loops of some looped hairpins, as shown in Figures 1B and 1D.

To be consistent with the data analysis in Figure 3 most R-regions in native DNA might be expected to have lengths in the range 0-800 base pairs. Reference to Figure 4 shows that the size of most of the hairpin duplexes are in accord with this estimate. However, also present are occasional, far longer duplexes in looped hairpins, presumably derived from longer stretches of inverted complementary sequence. These duplexes, although few in number, nevertheless make an important contribution to the foldback duplex yield. It will be of interest to know if these longer inverted repeats also possess an internal structure similar to that proposed for shorter foldback sequences.

Implications of the model

In accord with the results presented here, experiments involving the binding of Physarum foldback duplexes to hydroxyapatite show an increase in duplex-containing yield with DNA chain length, interpreted as evidence for the interspersion of foldback sequences with other sequence components (Jack and Hardman, unpublished results). The present study also shows that sequences forming looped and unlooped hairpins are mutually interspersed in Physarum DNA, as indicated in Figure 1. Provided traceable molecules can be regarded as typical, as discussed above, then the data presented in Table 1 shows that about 70% of DNA molecules 22,000 bases in length contain foldback sequences. Analysis of nuclear DNA by reassociation kinetics indicates that about 40-45% of the Physarum genome contains middle-repeated sequences [8]. Hence it is reasonable to assume that foldback elements are present in regions of the Physarum genome occupied both by single-copy and repetitive sequences. It should be stressed that the sequence complexity of

foldback DNA in Physarum, or of the intervening non-foldback sequences, has not been investigated in the present study. Although each foldback duplex must be formed from sequences which occur at least twice in the Physarum genome, it is not known whether foldback sequences are part of the single-copy or repetitive component in the normal sense.

With regard to other studies, there is an aspect of the model presented here which is appealing in its consistency with previous observations on the properties of foldback sequences in DNA from higher organisms, and may therefore be of relevance to the organisation of similar sequences in animal-cell DNA. Earlier studies on foldback DNA from several sources [2] has shown that sequences forming foldback duplexes are arranged in spaced clusters. Such an observation is in accord with the present model since foldback duplexes derived from within a given R-region would be clustered together, and spaced at regular intervals from adjacent clusters derived from separate R-regions.

It is a surprising observation that foldback sequences so widely distributed in Physarum DNA should be arranged in such a similar and uniform manner. It remains to be elucidated if such regular sequence arrangements are a general feature in the structural organisation of DNA in eukaryotic organisms. In this regard, experiments of a similar type to those described here indicate that hamster-cell DNA contains foldback sequence elements which display a similar pattern of periodic organisation to those in Physarum DNA (Hardman, Bell and McLachlan, unpublished results). If, as we suspect, a regular pattern of organisation of foldback sequences has been conserved during the evolution of diverse eukaryotic organisms such as hamster and Physarum, then it might be reasonable to suppose that it is a cause or consequence of its importance in relation to the structure or function of eukaryotic chromosomes.

Acknowledgements

The authors are grateful to Mr. Henry Eichelberger and Mrs. Doris Duncan for preparation of electron micrographs. P.L.J. is endebted to the Science Research Council for

the award of a studentship. Financial support from the Medical Research Council and the Scottish Hospitals Endowment Research Trust is gratefully acknowledged.

REFERENCES

1 Britten, R.J. and Smith, J. (1970) Carnegie Inst. Washington Yearbook 68, 378-386
2 Wilson, D.A. and Thomas, C.A. Jr. (1974) J. Mol. Biol. 84, 115-138
3 Cech, T.R. and Hearst, J.E. (1975) Cell 5, 429-446
4 Schmid, C.W., Manning, J.E. and Davidson, N. (1975) Cell 5, 159-172
5 Bell, A.J. and Hardman, N. (1977) Nucleic Acids Res. 4, 247-268
6 Dott, P.J., Chuang, C.R. and Saunders, G.F. (1976) Biochemistry 15, 4120-4125
7 Firtel, R.A. and Bonner, J. (1972) J. Mol. Biol. 66, 339-361
8 Fouquet, R.A., Bierweiler, B. and Sauer, H.W. (1974) Eur. J. Biochem. 44, 407-410
9 Hardman, N. and Jack, P.L. (1977) Eur. J. Biochem. 74, 275-283
10 Jack, P.L. and Hardman, N. (1977) Biochem. Soc. Trans. 5, 678-679
11 Deininger, P.L. and Schmid, C.W. (1976) J. Mol. Biol. 106, 773-790
12 Cooke, D.J. and Dee, J. (1975) Genet. Res. Camb. 24, 175-187
13 Mohberg, J. and Rusch, H.P. (1971) Exp. Cell Res. 66, 305-316
14 Davis, R.W. and Hyman, R.W. (1971) J. Mol. Biol. 62, 287-301
15 McConaughy, B.L., Laird, C.D. and McCarthy, B.J. (1969) Biochemistry 8, 3289-3293
16 Vogt, V.M. and Braun, R. (1976) J. Mol. Biol. 106, 567-587