Sequence organisation in nuclear DNA from Physarum polycephalum: methylation of repetitive sequences

Paul A.Whittaker, Alan McLachlan and Norman Hardman

Department of Biochemistry, University of Aberdeen, Marischal College, Aberdeen AB9 ^l AS, UK

Received 22 December 1980

ABSTRACT

Nuclear DNA from the slime mould Physarum polycephalum is digested by the restriction endonuclease HpaII to generate a high molecular weight and a low molecular weight component. These are referred to as the M+ and the M- compartment, respectively. Sequences that are present in the M+ compartment are cleaved by MspI, the restriction enzyme isoschizomer of HpaII, thus showing that the recognition sequences for these enzymes in M+ DNA contain methylated CpG doublets. The distribution of repetitive sequences in the M+ and M- DNA compartments was investigated by comparison of the 'fingerprint' patterns of total Physarum DNA and isolated M+ DNA after digestion using different restriction endonucleases, and by probing for the presence of specific repetitive sequences in Southern blots of M+ and M- DNA by the use of cloned DNA segments. Both types of experiment indicate that many repetitive sequences are shared by both compartments, though some repetitive sequences appear to be considerably enriched, or are present exclusively, either in M+ DNA or in M- DNA.

INTRODUCTION

DNA reassociation kinetic studies have revealed that the eukaryotic genome contains both single-copy and repetitive sequences¹. In the case of the acellular slime mould Physarum polycephalum the repetitive component comprises about one-third of the genome². The majority, if not all, of these repetitive sequences are closely interspersed with the single-copy sequences in a large portion of the genome, with a pattern of organisation resembling that found in <u>Xenopus</u> DNA³. Due to their ubiquity in eukaryotic DNA, and their possible involvement in genetic regulation and other nuclear functions, repetitive sequences have attracted considerable interest. However, with the possible exception of a few cases $4-6$ their functional significance is still obscure.

5-methylcytosine (5meC) is the major modified base found in the DNA of higher eukaryotes⁷. Although the biological function of 5meC in DNA

remains unknown, a number of suggestions have been made to explain its presence in organisms throughout the animal kingdom . Direct experimental evidence in support of a particular biological role for DNA methylation is scarce, however. Moreover, in some cases contradictory reports have appeared on tissue-specific variations in the 5meC content of DNA^{9,10}, and on the possible correlation between DNA methylation and chromatin structure 11,12 . However, in general, studies have shown that repetitive sequences have a higher 5meC content than single-copy DNA sequences 8,13-16.

Recently, we have examined the distribution of 5meC in Physarum polycephalum nuclear DNA by using the restriction endonuclease isoschizomers 17 HpaII and MspI¹⁷. Both enzymes cleave the sequence $5'$ CpCpGpGp3' but only MspI cleaves the methylated analogue $5'$ CpmCpGpGp3'¹⁸. Using HpaII, and another restriction enzyme , HhaI, which is inactive on DNA which contains a methylated recognition sequence 5'GpmCpGpCp3'¹⁹ we have been able to demonstrate that the Physarum genome, in common with those of a number of other eukaryotes such as the sea urchin ${\tt Echnius}^{ZO_Z/1}$, is composed of two compartments with respect to the methylation of HpaII restriction sites. These compartments are referred to as "M+" and "M-" and correspond to segments of DNA which contain either methylated or unmethylated HpaII sites, respectively. In the case of the Physarum genome the M+ compartment consists of about 20% of the total nuclear DNA. The presence of discrete DNA bands in MspI digests of total DNA which were shown to be absent, or present in undetectably small amounts, in corresponding HpaII restriction digests of Physarum DNA, has indicated that certain repetitive sequences might be considerably enriched in the M+ fraction.

In the present study we have extended these previous observations on the distribution of repetitive sequences in the M+ and M- compartments of Physarum DNA using two experimental approaches. The first approach involved restriction analyses of M+ DNA, purified using sucrose density gradient centrifugation, and comparison with similar analyses of total Physarum nuclear DNA. The second involved qualitative hybridisation analysis using cloned segments of Physarum nuclear DNA which are known to contain specific repeated sequences. The results reported show that many repeated sequences are equally distributed in the M+ and M- compartments. However, some repeated elements appear to be present almost exclusively either in M+ DNA or in M- DNA. The results also indicate that the recognition sites for the restriction nucleases TaqI and MspI occur less frequently in M+ DNA than in total Physarum DNA. Possible explanations

for these observations are discussed.

MATERIALS & METHODS

Culture methods and isolation of DNA

Physarum polycephalum strain LU648 x LU688²², a generous gift from Dr J. Dee, Department of Genetics, University of Leicester, U.K., was used throughout. Microplasmodia were grown as described previously $^{23}.$

DNA was labelled to a high specific activity by growth of the organism in the presence of 33 microcuries of 32 P-acid orthophosphate/ml of growth medium. Nuclei were isolated by the method of Mohberg and Rusch, as modified by Hardman and Jack 24 . DNA was isolated from purified nuclei by isopycnic centrifugation as described $previously²⁴$. Density gradient fractions were pooled and dialysed versus 20 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0.

Cloned segments of Physarum DNA

DNA clones containing Physarum repetitive DNA sequences were constructed by cleavage of the DNA using restriction nucleases BamHl and HindIII, followed by ligation into the plasmid vector pBR322, cleaved using the same restriction nucleases. Recombinant DNA clones were designated pPH 1-67. A detailed description of the methods used for the preparation of these cloned DNA segments is in preparation (McLachlan, A. and Hardman, N., unpublished results). Selected DNA clones, which were shown to contain repetitive DNA sequences (pPH14, 29, 32, 53, 54, 55; ref. 34) were used in this study.

Recombinant plasmids were grown in the host E.coli HB101 by chloramphenicol amplification²⁵. Cleared lysates were prepared 26 from cultures of recombinant DNA-containing clones, and supercoiled DNA isolated by one cycle of purification using CsCl-ethidium bromide density gradient centrifugation $^{27}.$

Manipulation of bacteria containing recombinant plasmids was carried out using a Category II containment laboratory according to the guidelines and procedures of the Genetic Manipulation Advisory Group. Restriction nucleases

Enzymes were used under conditions recommended by the supplier. Restriction nucleases HaeIII, MspI and EhaI were obtained from CP Laboratories Ltd., P.O. Box 22, Bishop's Stortford, Herts., England, and AluI, BamEl, HindIII, HpaII and TaqI were obtained from the Boehringer Corporation (London) Ltd., Bell Lane, Lewes, East Sussex, England.

Electrophoresis

0.5-1% agarose gels were electrophoresed in buffer containing 5 m M sodium acetate, 1 mM EDTA, 40 mM Tris-HCl, pH 7.4. 5%(w/v) polyacrylamide gels containing 5% acrylamide/0.125% bisacrylamide were electrophoresed in buffer containing 30 mM NaH₂PO₄, 1 mM EDTA, 36 mM Tris-HCl, pH 7.7. After electrophoresis gels were stained in electrophoresis buffer containing 1 microgramme of ethidium bromide/ml, and then photographed under short-wave ultraviolet light through an orange filter using Kodak Tri-X film.

For autoradiography, gels were dried on to Whatman 3MM paper using a Biorad Model 224 slab gel drier, then exposed to Fuji X-ray film at -70° C using a Mach 2 dose-reducing X-ray screen (Hanimex, U.K., Ltd.). Southern transfer hybridisations

DNA was denatured in agarose gels and blotted on to nitrocellulose filters as described by Southern²⁸. Before hybridisation filters were washed at 65° C overnight using 5 x SSCP (SSCP = 0.15 M NaCl, 0.015 M sodium citrate, 0.02 M sodium phosphate, pH 7.0) containing 5 x Denhardt's solution²⁹, 0.1% SDS, 10%(w/v) dextran sulphate, 2.25 mM EDTA. Sonicated probe DNA (1 microgramme), about 300 nucleotide pairs in length, was 5'-end labelled by the use of polynucleotide kinase (Boehringer Mannheim, London, Ltd.) in the presence of gamma-³²P-ATP (7000 microcuries/mM, New England Nuclear Corp.). The probe DNA was denatured by boiling for 10 min at 100° C and added immediately to the hybridisation mixture. Filters were hybridised overnight at 65° C, after which they were washed using 5 x SSCP, 0.1% SDS for 30 min at 65° C, followed by two washes using 2 x SSCP, 0.1% SDS at 65° C, the first for 60 min and the second for 30 min. Filters were air-dried, baked at 80 $^{\circ}$ C for 10 min, then autoradiographed as described above for dried gels.

Sucrose density gradient centrifugation

For purification of M+ DNA sequences, preparative $5-20$ % (w/v) sucrose gradients (5 ml) in 5mM NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 7.4, were formed in cellulose nitrate tubes. DNA (10 microgrammes) in 0.1 ml of restriction buffer, was loaded on to each gradient. Gradients were centrifuged at 40,000 r.p.m. for 90 min at 20° C in the SW50.1 rotor of the Beckman L5-50 preparative ultracentrifuge. Appropriate gradient fractions were pooled and DNA precipitated by the addition of 2 volumes of 95%(v/v) ethanol.

RESULTS

When Physarum nuclear DNA is digested with a variety of restriction nucleases, for example AluI, HaeIII and TaqI, and the resultant DNA fragments electrophoresed either on agarose or polyacrylamide gels, a smear of unresolved fragments is obtained upon which a series of sharp DNA bands is superimposed. These bands are assumed to arise due to cleavage of repeated DNA sequences containing internal sites for cleavage with these nucleases, by analogy with similar results obtained in several other $30-33$. These bands, which are just detectable in ethidium bromide stained gels, correspond to DNA fragments which vary in length from about 200 base pairs to approximately 10 kilobase pairs (kbp). Many of the low molecular weight bands can be detected only when fairly large amounts of DNA are loaded on gels , and many of the high molecular weight bands are difficult to resolve from the background smear of DNA. Greater sensitivity and resolution is afforded by autoradiography of 32 P-labelled DNA fragments (Fig.1) since as little as 0.1 microgramme of 32 P-labelled DNA can be electrophoresed. In the present study Physarum DNA was used which was labelled uniformly in vivo to a high specific activity (5 x 10^4 c.p.m/ microgramme of DNA) using 32 P-orthophosphate. For all restriction digests completeness of digestion was monitored by co-digestion of bacteriophage lambda DNA and the generation of the characteristic terminal restriction digest pattern for each enzyme on ethidium bromide staining of the gels. Restriction analysis of M+ DNA

It has been shown previously that Physarum nuclear DNA can be resolved into two components after digestion using the restriction nuclease HpaII¹⁷. This resolution can be improved by further digestion using the enzyme $Hhat¹⁷$. Utilising the size difference between M+ and M-DNA sequences after digestion with these enzymes it is possible to separate the two components by centrifugation on neutral sucrose density gradients (Fig.2a). Using this method it is possible to obtain M+ DNA free of contaminating M- sequences, as indicated by gel electrophoresis of samples of pooled M+ and M- DNA fractions taken from gradients (Fig.2b).

The M+ DNA fraction isolated in this manner was subjected to digestion using the restriction nucleases AluI, HaeIII and TaqI, and the DNA fragments electrophoresed on 1% agarose gels (Fig.3a) or 5% polyacrylamide gels (Fig.3b), together with similar digests of total DNA for comparison. After digestion of M+ DNA with these enzymes a series of sharp bands can be seen against a background smear of unresolved DNA fragments. The distribution of

Figure 1: Separation of restriction fragments of Physarum nuclear DNA

DNA was digested with restriction nucleases and subjected to electrophoresis on A.5% polyacrylamide gels, and B.1% agarose gels, as described in the text, with reference to restriction fragments of bacteriophage lambda and bacteriophage PM2 as as size markers.

1, HaeIII restriction fragments; 2, AluI restriction fragments; 3, TaqI restriction fragments; 4, Undigested DNA.

Approximately 0.1 microgramme of $32P-$ labelled DNA was loaded on to each gel slot (specific activity 5 x 10^4 c.p.m./microgramme). Gels were dried and autoradiographed as described in the text.

fragment sizes differs markedly for each enzyme, reflecting the frequency with which the enzymes cleave the DNA chains in the M+ fraction. A large proportion of the Taql-digested DNA is found in fragments of high molecular weight (2-10 kbp), whereas in AluI digests much of the DNA is found in fragments of low molecular weight (0.5-2 kbp). In contrast with both these enzymes, HaeIIIdigests M+ DNA to give a broad spectrum of fragment sizes with a size distribution resembling that of HaeIII-digested total DNA. Thus, TaqI restriction sites appear to occur less frequently in M+ DNA than HaeIII and AluI sites, when compared with total DNA.

Figure 2: Purification of M+ DNA sequences from Physarum DNA by using sucrose density gradient centrifugation

(a) 10 microgrammes of $32P$ -labelled Physarum nuclear DNA was digested to completion using the enzymes HpaII and HhaI. DNA fragments were separated by sucrose density gradient centrifugation. A separate gradient included undigested DNA for comparison. The fractions that were pooled to provide purified M+ fragments and M- fragments are indicated. A HpaII/HhaI digested DNA; \blacksquare - - Undigested DNA.

(b) Electrophoresis of fractionated DNA on a 1% agarose gel. M+ and Msequences were obtained as described in (a). Size standards were provided by HindIII restriction fragments of bacteriophage lambda and PM2 DNA. 1, M- sequences; 2, M+ sequences; 3, total unfractionated DNA.

Depending on which of the three enzymes is used to digest Physarum DNA, between about 20-30 discrete bands can be visualised in the gels. Many of the DNA bands are common to the digests of M+ and total DNA, and they appear to be present in similar amounts. However, certain DNA bands in digests of total DNA are undetectable in parallel digests of M+ fragments. Similarly, M+ digests appear to be enriched in particular bands when compared with total DNA digests. Some of the major differences are indicated in Fig. 3 by arrows.

The general conclusion from such experiments is that many bands are common to both M+ and total DNA, but DNA bands corresponding to some

repetitive sequences appear to be enriched either in the M+ fraction in some cases, or in the M- fraction in others. Trivial explanations for these results have been discounted. Care was taken in each case to ensure that equivalent amounts of radioactively-labelled DNA were loaded into parallel lanes of the gels, and that the characteristic terminal digestion pattern of the bacteriophage marker DNA was obtained. Probing M+ and M- DNA fractions for the presence of specific repeated sequences

The experiments described above were complemented by investigating the distribution in M+ and M- DNA of sequences displaying homology with specific repeated sequences in the Physarum genome. These sequence probes were obtained by molecular cloning of segments of Physarum DNA. Six cloned segments of Physarum DNA were used in this study, all of which are known to contain repetitive sequences. These are clones pPH 14, 29, 32, 53, 54 and 55 (ref. 34 and unpublished results of A. McLachlan & N. Hardman). Paired HpaII and MspI restriction digests of total Physarum nuclear DNA were electrophoresed on 1% agarose gels, transferred to nitrocellulose filters, and hybridised to $32P-$ labelled probe sequences obtained from sonicated fragments of each DNA clone by end-labelling in vitro. The autoradiographs from a series of experiments are shown in Fig.4. Appropriate control experiments were carried out to ensure that hybridisation patterns did not result from homology between vector DNA sequences and Physarum DNA.

Clones pPH29, pPH53 and pPH55 hybridised to a heterogeneous smear of relatively high molecular weight DNA fragments in MspI digests, with a limited indication of a superimposed pattern of bands. In sharp contrast, HpaII failed to digest the genomic sequences which were complementary to these cloned repetitive elements, with the possible exception of some pPH55 homologous segments, where some hybridisation to HpaII-digested fragments of lower molecular weight than M+ DNA segments was observed (Fig.4a). Thus,

Figure 3: Gel electrophoresis of total Physarum DNA and isolated M+ DNA after digestion with restriction nucleases

 32 P-labelled total DNA or purified M+ DNA was digested to completion using restrictton nucleases. DNA fragments were analysed using A. 1% agarose gels, or B. 5% polyacrylamide gels. Equivalent amounts of radioactivity $(2 \times 10^4 \text{ c.p.m.})$ were loaded on to each slot of the gels. Gels were dried and autoradiographed as described in the text. Size markers were HindIII restriction fragments of bacteriophage lambda and PM2 DNAs. 1, Alur restriction fragments; 2, TaqI restriction fragments (two different exposures of the same gel are shown in A); 3, HaeIII restriction fragments.

Figure 4: Hybridisation of selected Physarum DNA clones containing repetitive elements to total Physarum DNA digested using restriction nucleases HpaII (H) or MspI (M)

Restriction fragments from genomic DNA were separated on 1% agarose gels and transferred to nitrocellulose using the Southern blotting procedure. DNA blots were hybridised separately to $32P-$ labelled probe sequences prepared using clones pPH14, 29, 32, 53, 54 and 55, as described in the text. (a) Clones showing predominant hybridisation to M+ DNA fragments; (b) Clones showing predominant hybridisation to M- DNA fragments.

for these DNA sequences, hybridisation is confined almost exclusively to the M+ compartment of the Physarum genome. Interestingly, the pattern of hybridisation of these cloned DNA fragments to MspI-digested DNA shows that many of the complementary genomic DNA fragments that are cut by this enzyme remain at high molecular weight, indicating that these DNA segments possess sparsely distributed MspI-sensitive sites that occur far less frequently than in M- DNA.

In contrast to pPH29, 53 and 55, clones pPH14, 32 and 54 display

detectable hybridisation both to M+ DNA sequences, and to a wide range of DNA fragment sizes in M- DNA (Fig.4b). Some weak bands of hybridisation superimposed on the hybridisation smear are again visible. The size of the genomic DNA fragments in the regions corresponding to M+ compartment sequences which display complementarity with the labelled probes are reduced in size when digested using MspI, indicating that the presence of these sequences in HpaII-resistant material results from DNA methylation.

DISCUSSION

The portion of the nuclear DNA from Physarum polycephalum that is referred to here as the M+ fraction contains DNA segments about 20-50 kb in length which are devoid of cleavage sites for the restriction nucleases HpaII and HhaI 17 . Both of these enzymes are inactive on sequences which contain 5meC in the CpG doublets which form part of their respective recognition sites in DNA. Since no isoschizomer to HhaI has been found which cleaves the methylated recognition sequence analogue, as in the case of HpaII, it is not possible in any simple way to determine whether HhaI recognition sequences are methylated, or merely absent from M+ DNA. The observation that M+ DNA, after digestion using MspI, generates DNA fragments that are still of comparatively high molecular weight is noteworthy, since M+ sequences also appear to contain fewer sites than total DNA for cleavage by the nuclease TaqI (Fig.2b) which recognises the sequence 5'TpCpGpAp3' and which is still active on sequences which contain mCpG doublets at these sites³⁵. On the assumption that there are no other major sites of modification around these sequences which might affect the action of the enzymes MspI and TaqI, this indicates that CpG doublets may be more sparsely distributed in the M+ fraction of Physarum DNA than in the remainder of the genome. It is also interesting to note that this property is not shared by the enzymes HaeIII or AluI, which cut M+ DNA as frequently as sequences derived from the rest of the genome. These results are in sharp contrast to those of Bird & Taggart 21 in their study of Echinus DNA, who demonstrated that methylated MspIsensitive sites in the M+ fraction occurred as frequently as in M- DNA. There may be a number of possible explanations for this phenomenon which remain to be investigated, but it could conceivably be explained if mCpG sequences were concentrated in the M+ fraction, and, as has been suggested $^{36}\,$ they were highly mutable leading to loss of CpG doublets and generation of TpG doublets.

Distribution of repetitive sequences

In our previous study we showed that at least some repetitive sequences in Physarum DNA are methylated and enriched in the HpaII-resistant, M+ compartment, since DNA bands indicative of cleavage products of repetitive sequences are clearly visible in ethidium bromide-stained gels of DNA restriction fragments generated using the enzyme MspI, but not the enzyme $H_{DAII}¹⁷$. In the present study we have extended this work by carrying out -a more detailed examination of the patterns of DNA bands generated on digestion of purified M+ DNA, by the use of restriction nucleases that preliminary work had shown to cleave families of repetitive sequences in Physarum DNA, namely HaeIII, AluI and TaqI. Thus it is shown that whereas most of the repetitive sequences amenable to this analysis are found in similar proportions in M+ and total DNA, some bands appear to be enriched, or present exclusively, either in the M+ fraction or in the M- fraction. Hence, some repetitive sequences exhibit specific compartmentation with respect to the pattern of methylation of HpaII sites in the DNA segments from which they are derived.

This conclusion is supported by experiments involving the hybridisation of specific, radioactively-labelled repeated sequence probes, derived from Physarum nuclear DNA by molecular cloning (Fig.4). Some DNA clones containing repetitive elements hybridise to sequences in both M+ and M-DNA. Others anneal to the M+ compartment specifically. However, none of the DNA clones selected as probes showed homology with M- DNA exclusively. That the observed hybridisation to M+ DNA is to regions containing methylated HpaII restriction sites was demonstrated by the observation that the complementary genomic sequences are cleaved by MspI. As with total M+ DNA, MspI-cleaved sequences showing homology with pPH29, pPH53 and pPH55 are of high molecular weight, in accord with the idea that these restriction sites possibly occur infrequently in the M+ fraction, as indicated above. Compartmentation of the Physarum genome by use of the enzyme HpaII

Since only a small number of potentially methylated CpG doublet sequences are being probed by the type of restriction analysis reported here, it should be borne in mind that the results have no bearing on the distribution of the bulk of the 5meC residues in M+ or M- DNA. However, whatever the distribution of 5meC in sites other than those suceptible to study by the use of HpaII and MspI, it is evident that certain repetitive sequences are compartmented in the M+ and M- fractions of the Physarum genome.

Only one other study has been reported in which specific repetitive sequences have been used to probe methylated and unmethylated sequences in eukaryotic DNA²⁰. In this case the probe was a DNA clone derived from a tandemly-repeated, satellite-like sequence in Echinus DNA. This sequence was shown to reside exclusively in the M+ compartment, and may be similar in some respects to those highly repetitive sequences in other organisms that are also believed to be hypermethylated. Two lines of evidence point to the conclusion that none of the cloned repetitive sequences chosen as probes for this study represent simple, tandemly-repeated DNA sequences. The first stems from the fact that the hybridisation pattern of the cloned sequences to genomic DNA that has been cleaved using a wide variety of restriction nucleases is complex and smeared, rather than taking the form of discrete 'ladders' of hybridisation, which would be consonant with a satellite-type pattern of organisation. The second stems from the observation that the restriction maps of all the DNA clones are complex, and they show little evidence of regularly-spaced restriction sites consistent with a simple tandemly-repeated structure³⁴. The results are, however, consistent with the idea that the DNA clones contain interspersed repetitive DNA sequences. It will be of further interest to investigate whether any relationship exists between the properties and genomic organisation of these cloned repetitive sequences and their pattern of methylation. It is also of interest that Reilly et al. 37 have shown recently that the extrachromosomal rDNA sequences in Physarum reside in the M- compartment, as defined here, indicating that they are not methylated.

ACKNOWLEDGEMENTS

The authors are grateful to the Cancer Research Campaign and the Medical Research Council for the support of this work, and to Professor H M Keir for the provision of facilities in the Department of Biochemistry.

REFERENCES

2 Hardman, N., Jack, P.L., Fergie, R.C. & Gerrie, L.M. (1980) Eur. J. Biochem. 103, 247-257

- 3 Davidson, E.H., Hough, B.R., Amenson, C.S. & Britten, R.J. (1973) J. Mol. Biol. 77 1-23
- 4 Jelinek, W.R., Toomey, T.P., Leinward, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Deininger, P.L. & Schmid, C.W. (1980) Proc. Nat. Acad. Sci. (U.S.) 77, 1398-1402
- 5 Kimmel, A.R. & Firtel, R.A. (1979) Cell 16, 787-796

