Volume 3 no.10 October 1976 Nucleic Acids Research

A possible structure for calf satellite DNA I

Gérard Roizes

Centre de Recherches de Biochimie Macromoléculaire du C.N.R.S., Route de Mende, B.P. 5051, 34033 Montpellier Cedex, France

Received 9 August 1976

ABSTRACT :

Calf satellite DNA I (ρ = 1.715) has been hydrolysed by a number of restriction endonucleases. It consists of a repeating unit of 1460 nucleotide pairs within which the sites of Eco R II, Mbo I, Sac I, Alu I, Ava II and Hha I were localised in comparison with those of Eco R I and Hind II. The distribution of the Hpa II, Sac I, Hha I, Hinf I and Mbo II sites within calf satellite DNA I, as well as that of several restriction endonuclease sites within calf satellite DNA III (ρ = 1.705) allowed me to define subsatellite fractions. Furthermore, some of the sites of the CpG containing restriction enzymes Hpa II and Hha I are lacking. The possible implications of these results are discussed.

INTRODUCTION

Satellite DNAs in eucaryotes are composed of very simple nucleotide sequences and on the basis of the partial sequence analysis of guinea pig satellite DNA Southern (1) proposed a model by which the formation and evolution of the characteristic internal repeats occur as the result of a system of multiplication and divergence. Reassociation kinetics and analysis with restriction endonucleases have shown the different internal repeats predicted by this model exist in all the satellite DNAs studied so far (2,3). The mechanisms which give rise to these internal repeats are unknown although there is some evidence that unequal crossing over occurs frequently in their formation (3).

Very simple original sequences have been shown to occur in Drosophila melanogaster (4), Drosophila virilis (5), HS β Kangoroo rat (6), mouse (7) satellite DNAs and it is likely to be the case for most of satellite DNAs.This is certainly true of Calf satellite DNA I (ρ = 1.715) which also exhibits two other internal repetitions : one of 200-250 nucleotide pairs which was shown to exist on the basis of reassociation kinetics and one of 1400-1500 nucleotide pairs identified by restriction endonuclease analysis (2,8).

In this paper, I present evidence based on analysis with a

number of different restriction endonucleases that calf satellite DNA I has a very complex fine structure. This structure consists of several subsatellite fractions which have different distributions of the Hpa II, Hha I, Sac I and Hinf I restriction endonuclease sites. Furthermore, in each of these subsatellite fractions, some sites are sometimes lacking (Hpa II, Hha I and Sac I). This could be due to the methylation of the CpG dinucleotides which occur in the restriction sites of Hpa II and Hha I. A less probable explanation is that there is modification of the site itself as a result of base changes.

The possible implications of these results are discussed.

MATERIALS and METHODS

1) Calf satellite DNA preparations :

Calf satellite DNAs I (ρ = 1.715) and III (ρ = 1.705) were purified from calf thymus DNA by a procedure derived from the one described by Fillipski et al. (9). The satellite DNA fractions were considered to be pure when their profiles looked unimodal in analytical cesium chloride centrifugations using a MSE analytical ultracentrifuge equipped with a scanning system. Satellite DNA I was obtained pure after one centrifugation in Cs_2SO_4-Ag^+ (Rf = 0.35) and three in CsCl. For satellite DNA III, one centrifugation in Cs_2SO_4-Ag^+ (Rf = 0.35) and one in CsCl were necessary.

2) Restriction endonucleases :

Eco R I and Eco R II were prepared using the procedure developed by Yoshimori (10); Hind by the method of Smith et al. (11). Hind III was a gift of H.Cooke. Hpa II was a gift of P.Nardeux and Hpa I of A.Prunell. Alu I, Hae III, Hha I, Mbo I, Mbo II, Xma I, Hinf I, Sac I were prepared by procedures developed by Rich Roberts et al. (personal communication; CSH.Laboratory, N.Y.) and Ava I and Ava II by procedures developed by S.Hughes and K.Murray (personal communication).

3) DNA Hydrolysis :

The DNA samples were digested in the Eco R I buffer (10). Variable times and quantities of enzymes were necessary to completely hydrolyse the DNA samples. To check if the reaction was complete, parallel hydrolysis of SV40 or λ DNA was carried out with the same enzymes. In the case of Hpa II, Hha I and Sac I, where it was essential to know if the hydrolysis was complete or not, large excess of enzyme and long digestion times were used. However, this was not possible with Alu I because of the presence of a contaminating exonuclease, so the

results obtained with this restriction endonuclease are the product of a partial hydrolysis.

4) Gel electrophoresis :

The DNA fragments of the restriction endonuclease digests were analysed by gel slab electrophoresis. The gels were either 1.4% or 0.3% agarose (Sigma for electrophoresis) in Tris 36 mM, pH 7.5, 30 mM NaH_2PO_4, 1 mM EDTA Na_2 , ethidium bromide at 0.5 µg/ml. The gel slab apparatus was similar to the one described by Studier (12) for the 1.4% gels. For a 0.3% agarose, a horizontal apparatus was necessary in order to prevent the gels from contracting or tearing. After electrophoresis, the gels were illuminated by a UV source and a photograph was taken with either polaroid P/N 57 or ilford FP4 films. Microdensitometer tracings were made from the negatives on a Joyce-Loebl microdensitometer. The markers used to estimate the lengths of the satellite DNA fragments in the gel electrophoresis were the SV 40 DNA strain 776-24IO (this was a gift from P.Nardeux, Institut de Recherches Scientifiques sur le Cancer, Villejuif, France) by digestion with the restriction endonucleases Hind, Hae III, Eco R II and Eco R II + Hae III (13,14,15).

With the 1.4% agarose gel, it was possible to estimate, by interpolation of the satellite DNA bands with the SV40 DNA markers, between 1550 and 250 nucleotide pairs accurately. For shorter fragments, it was quite impossible to determine their sizes with any accuracy, and, therefore they had to be calculated. For example, one band of the Mbo I digest of calf satellite DNA I (which was 405 nucleotide pairs long) gives two bands when hydrolysed by Eco R I. One has 260 nucleotide pairs so the second is calculated to have 145 nucleotide pairs. They were not measured in acrylamide gels, because of the difficulty of preparing the large quantities of satellite DNAs which would have been necessary for these experiments.

RESULTS AND DISCUSSION

1) Digestion of Calf satellite DNA I by restriction endonculeases Eco RII, Mbo I, Hae III, Alu I, Ava II, Ava I, Xma I and Mbo II :

Calf satellite DNA I when digested by Eco R I and Hind gives products which are resolved into bands by agarose gel electrophoresis (2,8). With Eco R I one band of approximately 1400 to 1500 nucleotide pairs in length is obtained. The Hind restriction site occurs twice, the fragments adding up to the original length of 1460 nucleotide pairs (this is the value obtained in the present work and it will be used as such). Calf satellite DNA I therefore consists of internal repeats tandemly arranged

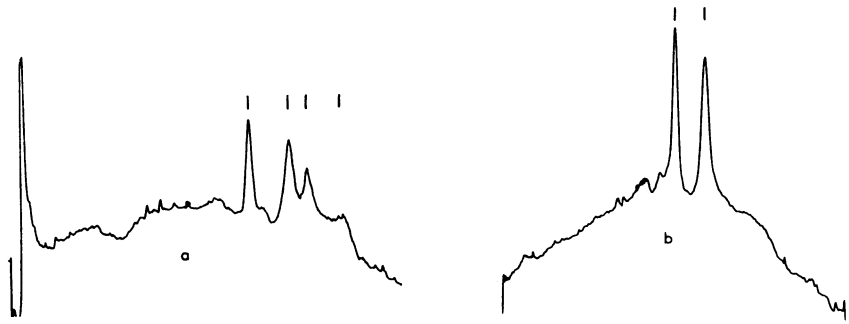

Figure 1 : DNA fragments from Calf satellite DNA I hydrolysed with Eco R II and Mbo I :

The three largest fragments obtained with Eco R II (a) had 540, 360 and 300 nucleotide pairs. In the Mbo I (b) digest, the fragments were 650 and 405 n.p. long.

along the satellite DNA molecules.

It was interesting to look at the distribution within the repeating unit of the sites of other different restriction endonucleases. When digested by Ava I and Xma I, Calf satellite DNA I was not significantly reduced in size. The Eco R II and Mbo I restriction sites are present in a manner equivalent to those of Eco R I and Hind II (Fig. 1a and 1b). At least four fragments were obtained with Eco R II (The 1.4% agarose gel does not resolve DNA fragments smaller than 200 nucleotide pairs) and two with Mbo I. The fragment of 405 nucleotide pairs obtained by the Mbo I digestion looked like a non resolved double band. The satellite DNA I was therefore analysed on agarose gel after double digestion with Eco R I and Mbo I. New bands appeared whilst the fragment corresponding to 405 nucleotide pairs was still present (Fig.2). The Mbo I site was therefore present three times within the 1460 nucleotide pairs repeating unit. It was not determined if very small fragments were present in both cases as this would have required very large quantities of cold DNA before they could be visualised.

The Hae III site was present a great number of times (14 fragments have been found in CSH laboratory, CSH annual report, 1975). The largest fragment obtained was approximately 330 nucleotide pairs in length (as determined both by gel electrophoresis and electron microscopy).

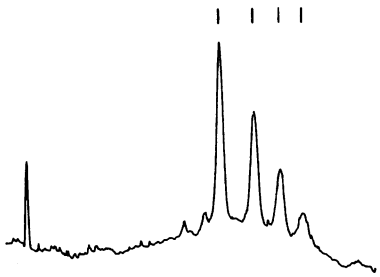

Figure 2 : Double digestion of Calf satellite DNA I with Eco R I and Mbo I

The two largest fragments obtained had 650 and 405 nucleotide pairs. As they were obtained from a fragment of 405 n.p. (Fig. 1b), the sum of the smallest fragments must be 405 n.p. The length of the larger of the two was 260 n.p., therefore, the other had 145 n.p.

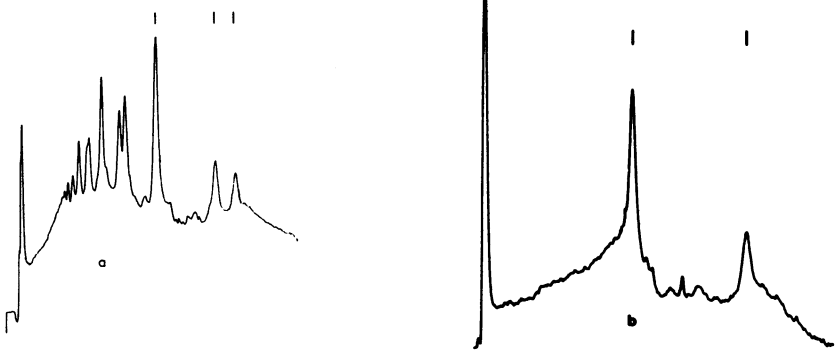

Figure 3 : Digests of Calf satellite DNA I with Alu I and Ava II :

The digest with Alu I (a) was partial and the sizes of the fragments were 645 , 815 and 1460 for the shortest. The two fragments of the Ava II digest (b) had 1160 and 300 n.p.

The Alu I site occured twice in the repeating unit of Calf satellite DNA I, as judged from the products of a partial hydrolysis (Fig.3). The two basic fragments obtained add up to 1460 nucleotide pairs. It was not possible to go further in the digestion because of the presence of a contaminating exonuclease in the preparation. Two fragments were obtained when Calf satellite DNA I was digested by Ava II (Fig.3b), showing that the Ava II site was present twice in the repeating unit.

2) Digestion of Calf satellite DNA I by restriction endonucleases Hpa II, Hha I, Sac I, Hinf I and Mbo II :

Surprisingly the Hpa II restriction site was present, although infrequently, at intervals which corresponded to the dimer, trimer, tetramer... of the repeating unit of 1460 nucleotide pairs of Calf satel-

lite DNA I (Fig. 4). No monomer was detectable with the quantity of DNA used in one experiment (one microgram). These results suggest that this satellite DNA is composed of at least two subfractions which evolved independently after the formation of the repeat of 1460 nucleotide pairs.

One could, at this stage, think that the fraction of DNA which was digested by Hpa II in the manner described in Fig. 3 was in fact, another satellite, contaminating the satellite DNA I in spite of extensive purification. That this is not the case is clear if one examines the

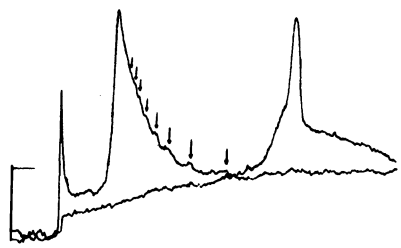

Figure 4 :Digestion of Calf satellite DNA I with Hpa II :

The fragments corresponding to the dimer, trimer, tetramer... of the repeating unit of 1460 nucleotide pairs are shown with arrows. For comparison, the microdensitometer tracing of a complete digest of Calf satellite DNA I with Eco R I is also drawn.

digests of satellite DNA I with Hha I. What is more, the Hha I sites were present in such a way that several series of bands were detectable in agarose gel electrophoresis (Fig.5a). The monomer, dimer, trimer of the 1460 nucleotide pairs repeating unit were present, as well as four bands (560, 600, 860 and 900 nucleotide pairs) and the corresponding intermediate bands between the monomer and dimer, dimer and trimer, trimer and tetramer...It is noticed, at this stage, that the four fragments add up two by two (900 and 560, and 860 and 600) to give 1460 nucleotide pairs. Therefore, it is possible to consider that this distribution of the Hha I sites defines two types of satellite DNA I molecules, or two subfractions of the same satellite DNA.

By an analysis of the digest of satellite DNA I with Hha I in 0.3% agarose gel electrophoresis (Fig.5b), it was possible to detect up to the 7th mer.

The pattern obtained when satellite DNA I is hydrolysed by Sac I (Fig.6a) is also striking. Although most of the fragments contained 1460

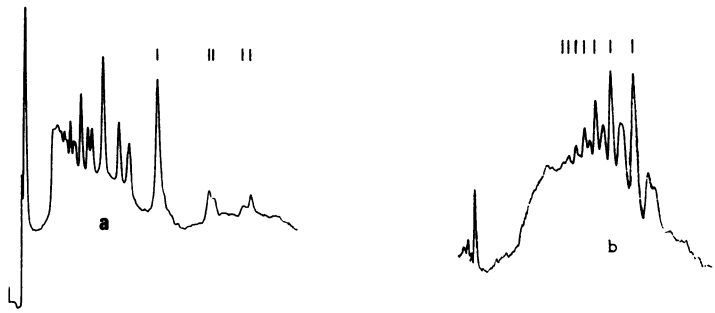

Figure 5 : Digest of Calf satellite DNA I with Hha I :

The sizes of the fragments of the Hha I digest were estimated in 1.4% agarose gel (a) as 560,600,860,900 and 1460 n.p. when taken together, the 560 and 900 n.p. fragments on the one band, and the 600 and 860 n.p. on the other, occured in isomolecular yields. Intermediate fragments are easily visible between one n mer and the (n + 1) mer of the 1460 n.p. repeating unit. 0.3% agarose (b) was used to resolve the longer fragments.

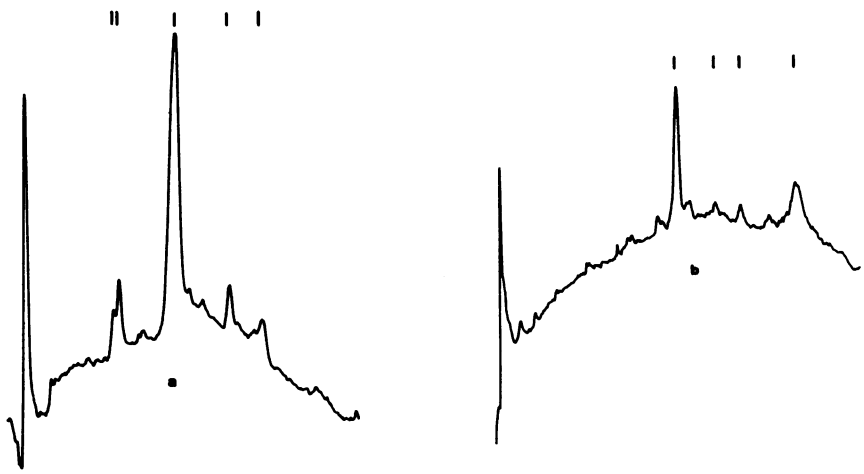

Figure 6 : Digests of Calf satellite DNA I with Sac I and Sac I + Eco R I:

The sizes of the fragments of the Sac I digests (a) were estimated as 610, 850 and 1460 n.p. In the double digest with Sac I and Eco R I (b), the 1460 n.p. fragment gave two fragments of 1100 and 360 n.p. The 850 n.p. was also reduced in size by 40 n.p. whilst the 610 n.p. was left intact.

nucleotide pairs, some fragments of 850 and 610 nucleotide pairs and the corresponding intermediate bands between the monomer and dimer and the dimer and trimer also occured. The digestion of DNA fractions containing high proportions of other Calf satellite DNAs showed that these minor

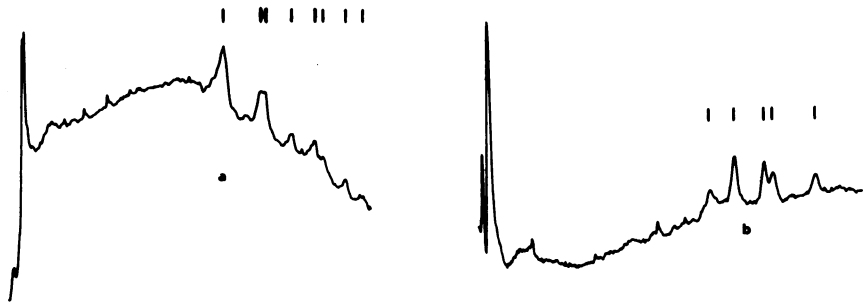

Figure 7 : Digests of Calf satellite DNA I by Hinf I (a) and Mbo II (b):

bands originate from satellite DNA I and not from other contaminating fractions.

Both the 850 and 1460 n.p. fragments obtained with Sac I contain one Eco R I site (Fig. 6b). However the distance between the Eco R I site and the Sac I sites on the 1460 n.p. and 850 n.p. fragments were different (360 and 40 n.p. respectively). This strongly suggests that there are two new subfractions, a major one in which all the molecules of Calf satellite DNA I contain the Sac I site once and a minor one in which it is contained twice.

When hydrolysed by Hinf I (Fig. 7a) satellite DNA I gives a complex banding pattern, which also suggests the existence of several subsatellite fractions which contain the Hinf I sites in different positions. Although less complex, the pattern given by Mbo II (Fig. 7b) could again be explained by considering at least the existence of two subfractions in satellite DNA I.

3) Localisation of the Eco R II, Sac I, Mbo I, Alu I, Ava II, Hha I, Hind II and Eco R I sites on the satellite DNA I molecules :

In establishing a physical map the Eco R I restriction site was taken as the origin and the positions given by Botchan (2) for the two Hind II sites were used.

a) Localisation of the Eco R II sites :

A double digest of satellite DNA I with Eco R I and Eco R II (Fig. 8) localises the Eco R I site within the largest fragment given by Eco R II alone. Satellite DNA I, after digestion with Eco R I, was then

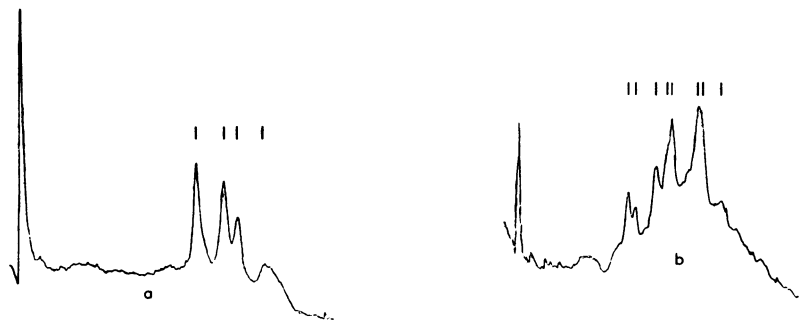

Figure 8 : Double digest of Calf satellite DNA I with Eco R I and Eco R II :
The largest fragments obtained in a total double digest with Eco R I and Eco R II (a) had 510, 360 and 300 nucleotide pairs. The Eco R I site, therefore, occured in the 540 n.p. fragment of the Eco R II digest (Fig.1a). A partial digest with Eco R II of satellite DNA I completely digested with Eco R I (b) gave the bands shown : 1460, 1430 (not resolved), 1320, 1020, 870, 810 ... It was impossible to estimate the lengths of the fragments occuring between 1020 and 870 n.p.

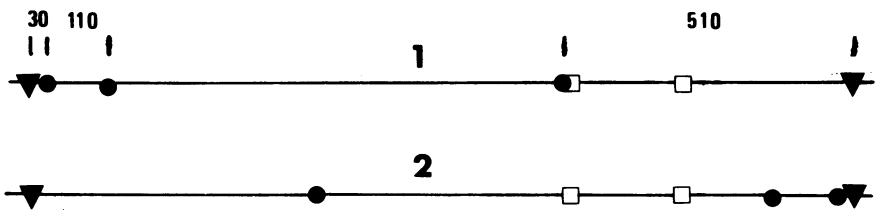

Figure 9 : The two possible localisations of the 540 and 110 n.p. fragments of the Eco R II digest :
(▼) Eco R I, (●) Eco R II, (□) Hind II.

partially hydrolysed by Eco R II (Fig. 8). The apparition of a band of 1320 nucleotide pairs demonstrates the presence of fragments of 110 nucleotide pairs and of 150 nucleotide pairs which were not resolved in Fig. 1a. Two localisations of the 540 and 110 nucleotide pair fragments are compatible with these results (Fig. 9)

The first was shown to be the correct one, since, in a double digest of satellite DNA I by Eco R II and Hind II, only the 540 nucleotide pair fragment disappeared in the gel.

b) Localisation of the Sac I sites :

As shown in Fig. 6b, the Eco R I site is situated 1100 nucleotide pairs from the Sac I site of the major fraction. Two possible localisations of this Sac I site are compatible with this result (Fig. 10a).

In order to decide whether position 1 or position 2 was the correct one, a double digest of satellite DNA I with Sac I and Eco R II was performed (Fig. 11). The 540 nucleotide pair fragment obtained by digest with Eco R II alone was not affected by Sac I. However, the amount of the 360 nucleotide pair fragment present was considerably reduced showing that the Sac I site of the major subfraction was localised within it. This result also confirms that only the major fraction of satellite DNA I contains a Sac I site in this position. In the same experiment, all the fragments

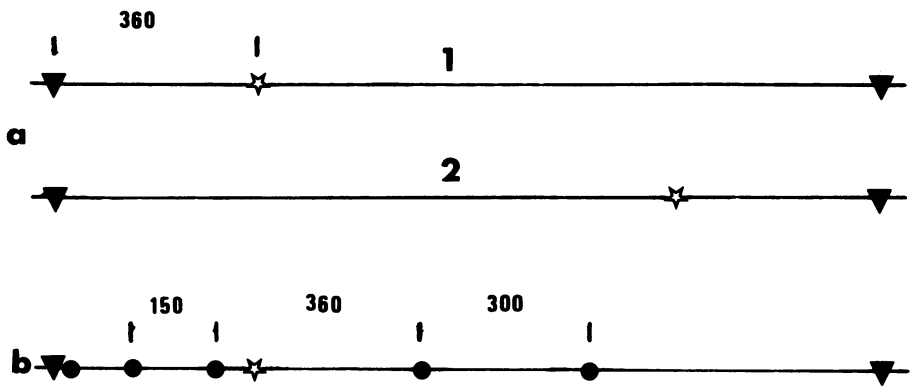

Figure 10 : The localisation of the Sac I and Eco R II sites :
(a) two possible localisations of Sac I (✡). (b) Physical map for (▼) Eco R I, (✡) Sac I and (●) Eco R II.

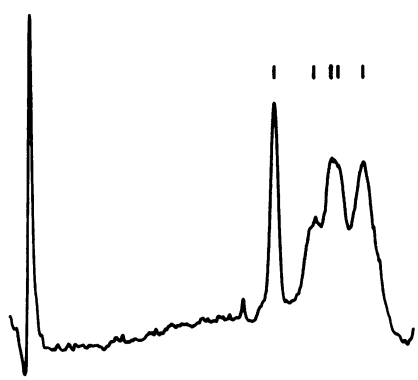

Figure 11 : Double digest of Calf satellite DNA I with Sac I and Eco R II:
The 4 longest fragments obtained were measured and contained 540,360,300 and 290 n.p. respectively. The 290 n.p. fragment was a Sac I cleavage product of the 360 n.p. fragment obtained with Eco R II alone.

of an Eco R II digest could be localised on the physical map (Fig. 10b).

c) Localisation of the Mbo I sites :

As already shown in Fig. 2, the Eco R I site is localised within one of the two 405 nucleotide pair fragments. The two Hind II sites were localised after a double digestion with Hind II and Mbo I (Fig.12a) within the second 405 nucleotide pair fragment. Two possible distributions of the three Mbo I sites were compatible with this result (Fig.13a). Therefore, a double digest of satellite DNA I with Sac I and Mbo I was performed.The 650 nucleotide pair fragment of a digest with Mbo I alone was converted to a 550 nucleotide pair fragment showing that the first distribution was the correct one (Fig.13b).

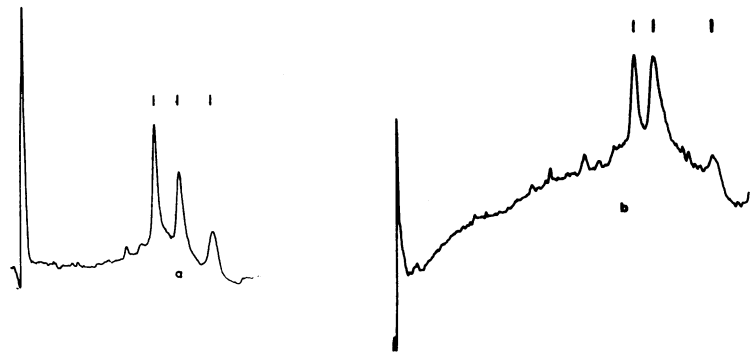

Figure 12 : Double digest of Calf satellite DNA I with Mbo I and Hind II and Mbo I and Sac I :

The two longest fragments obtained in a double digest of satellite DNA I with Hind II and Mbo I (a) were 650 and 405 n.p.A double digest with Mbo I and Sac I (b) generated fragments of 550 and 405 n.p. and several smaller ones which were not estimated.

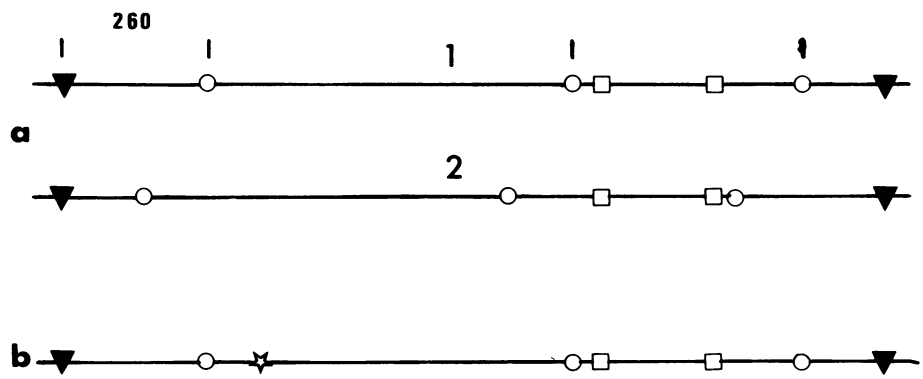

Figure 13 : The localisation of the Mbo I sites :

(a) The two possible localisations of the Mbo I sites
(b) The physical map for (▼) Eco R I, (□) Hind II, (○) Mbo I and (☆) Sac I.

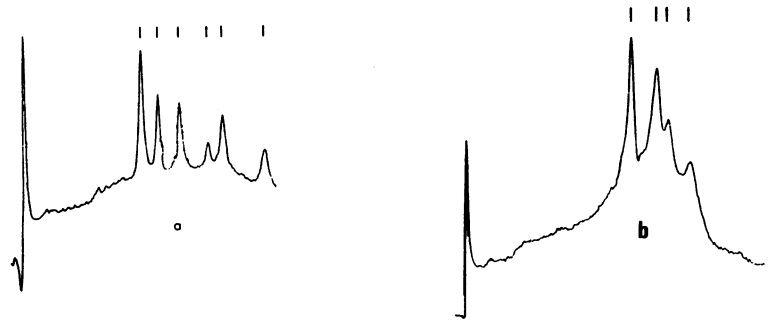

Figure 14 : Double digests of Calf satellite DNA I with Alu I and Eco RI and with Alu I and Eco R II :

The fragments generated by the double digestion of satellite DNA I with Eco R I and Alu I (a) had 1460, 1170, 935, 645,525 and 290 nucleotide pairs (the action of Alu I was not complete). The fragments generated by Eco R II alone is not significantly affected when Alu I is added (b).

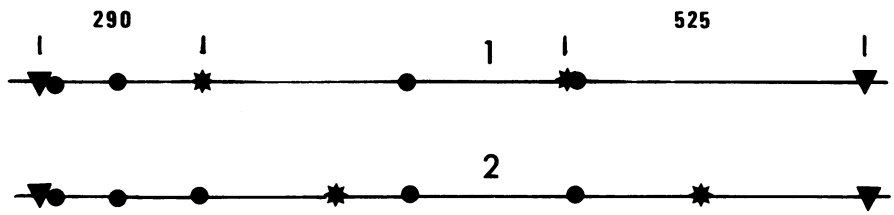

Figure 15 : The localisation of the Alu I sites :
(▼) Eco R I, (●) Eco R II, (✸) Alu I.

d) Localisation of the Alu I sites :

A double digest of satellite DNA I with Eco R I and Alu I (Fig.14a) permitted the localisation of the Alu I sites relative to the Eco R I site. Of the two possible localisations of the Alu I sites(Fig.15), the second was the correct one, since in a double digest with Alu I and Eco R II, none of the fragments of an Eco R II digest of satellite DNA I were affected.

e) Localisation of the Ava II sites :

The two fragments of a digest of satellite DNA I with Ava II (Fig. 3b) alone gave rise to fragments of 790, 380 and 290 nucleotide pairs (Fig. 16) when Eco R I was added. A double digest of Calf satellite DNA I with Ava II and Sac I shows that of the two possible localisations shown in fig.17 the first one is correct.

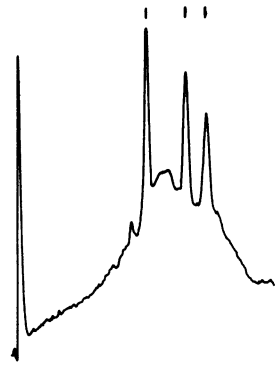

Figure 16 : Double digest of Calf satellite DNA I with Eco R I and Ava II: The fragments obtained had 790, 380 and 290 n.p.

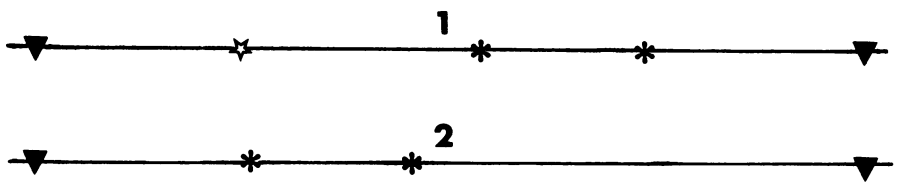

Figure 17 : The possible localisations of the Ava II sites :

(▼) Eco R I, (✱) Ava II, (✡) Sac I.

f) Localisation of the Hha I sites :

A double digest of satellite DNA I with Eco RI and Hha I (Fig.18a) permitted the localisation of the Hha I sites relative to Eco R I. Four were compatible with this result (Fig.19). Two localisations (2 and 3) corresponding to two subsatellite fractions, were compatible with the results obtained in a double digest with Eco R II and Hha I (Fig.18b), since, in both cases, the 540 nucleotide pair fragment of the Eco R II digest is converted, when the Hha I site is present, to a 420 nucleotide pair fragment, whilst neither the 360 nor the 300 nucleotide pair ones are affected.

4) The structure of Calf satellite DNA I :

The physical map of Calf satellite DNA I described here (Fig. 20) includes numerous restriction sites which have not been modified to a significant extent. One should add to this map the 14 Hae III sites which are present in the repeating unit, as well as those of Hinf I and Mbo II.

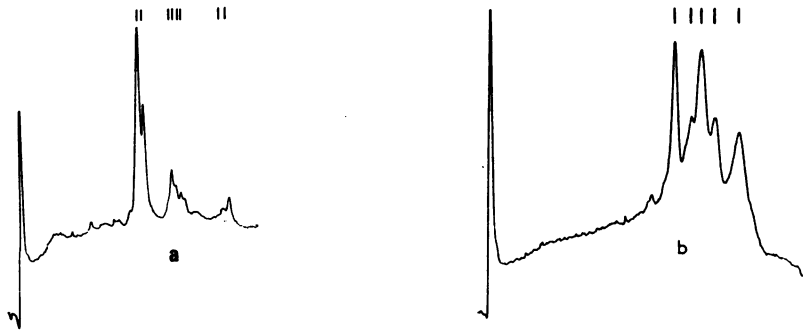

Figure 18: Double digests of Calf satellite DNA I with Hha I and Eco R I and with Hha I and Eco R II :

The fragments generated by the double digestion of satellite DNA I with Eco R I and Hha I (a) contained 1460, 1370, 995, 950, 900, 860, 510 and 465 n.p. The existence of 90 and 95 n.p. fragments which are not clearly visible here was deduced from the possible arrangements of the Hha I sites, as shown in Fig. 19. The double digestion with Eco R II and Hha I (b) generated an extra fragment of 420 n.p. which did not occur in the digest with Eco R II alone.

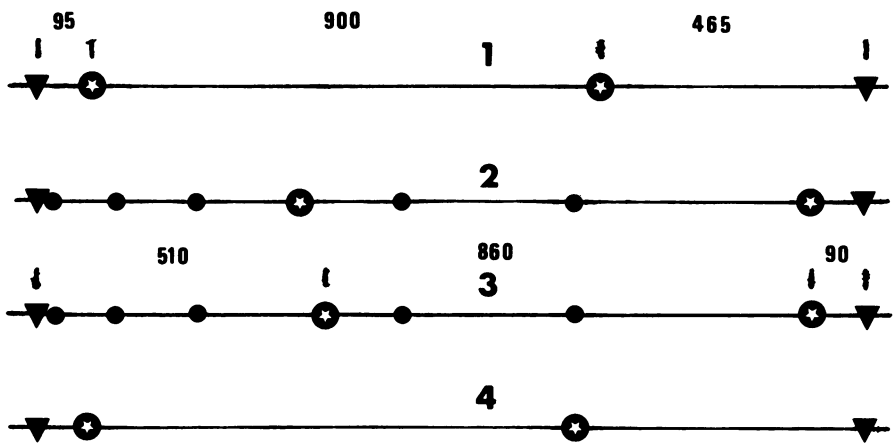

Figure 19 : The possible localisations of the Hha I sites :

(▼) Eco R I, (●) Eco R II, (✪) Hha I.

From the results described in this study, a physical map of the following restriction sites was obtained : Eco R I, Hind II, Eco R II, Mbo I, Alu I, Sac I, Ava II (Fig. 20a). On this physical map there are two different sets of Hha I sites which define two subsatellite fractions.

Except for the presence twice of the sequence Mbo I...Eco R II...Alu I in the repeating unit, there are no obvious internal repetitions. However, reassociation kinetics have shown that Calf satellite DNA I has a complexity of 200-250 nucleotide pairs (2). On the basis of this physical map

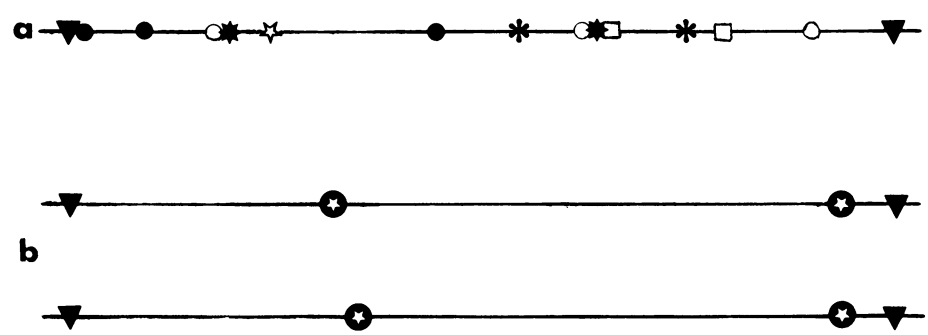

Figure 20 : The physical map of Calf satellite DNA I (a) was determined for (▼) Eco RI, (●)Eco R II, (○) Mbo I, (✸)Alu I, (☆) Sac I (✱)Ava II (□)Hind II. (b) The two subsatellite fractions determined by the two different positions of Hha I (✪).

and the fingerprints of the T1 and RNase A digests of cRNA made from satellite DNA I (S.Szala and G.Roizès, unpublished observations) it is difficult to imagine a basic repeating unit much smaller than 100 to 200 nucleotide pairs as is the case for other satellite DNAs (4-7) including Calf satellite DNA III (S.Szala and G.Roizès, unpublished observations).

Unlike the other restriction sites the Hpa II, Sac I and Hha I sites on satellite DNA I have been modified. The different distributions of these sites, as well as those of Hinf I and Mbo II, define subsatellite fractions that could have evolved together until the 1460 nucleotide pair repetition unit was formed but which from then on evolved independently.

Subsatellite fractions also occur in Calf satellite DNA III.It is composed of several subfractions as shown by the microdensitometer tracings obtained after digestion with several restriction endonucleases(Fig. 2I).

The existence of satellite DNAs in differentiated forms may be a relatively widespread phenomenon since, in the mouse, Hörz et al. (16) showed that the classical mouse satellite DNA contained a small fraction which is broken by the Hind endonuclease into a series of bands although a major fraction is resistant to the action of the same enzyme. Similar results were also obtained in the study of the mouse satellite DNA with Hae III (3) or for the guinea pig satellite DNA III with the Hind endonuclease (16). From the present study, it can be expected that many other satellite DNAs are also heterogenous.

In attempts to identify the chromosomes on which the satellite DNA sequences of different species were localised, it was always found that

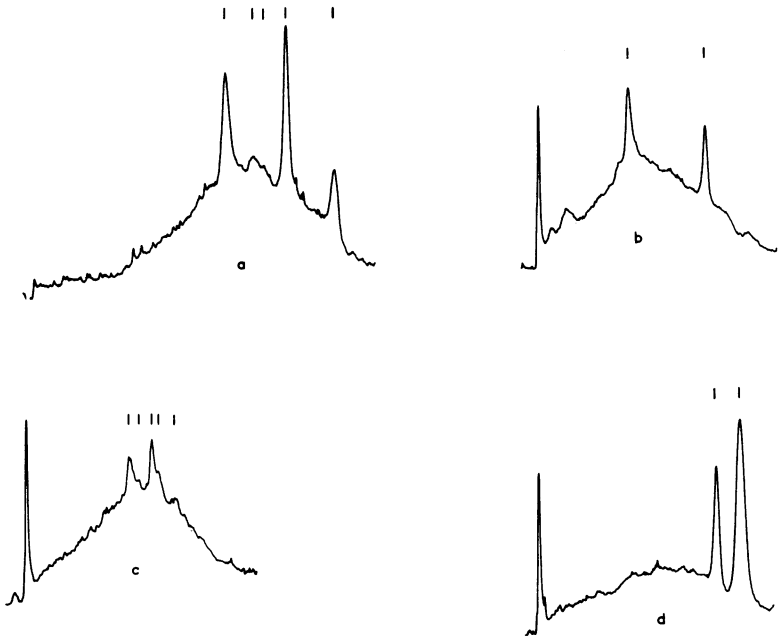

Figure 21 : Digestion of Calf satellite DNA III by several restriction endonucleases :

Satellite DNA III (ρ =1.705) was hydrolysed by a number of restriction enzymes and the fragments fractionated by gel slab electrophoresis in 1.4% agarose. It was not broken by several enzymes (Eco R I, Hpa II, Sac I, Ava I, Hpa I, Hind II, Hind III, Sac I). In some cases, most of the molecules were left intact whilst minor bands appeared(Mbo II, Alu I);with one enzyme (Hha I) although the satellite DNA III was not highly resistant to its action, no apparent banding pattern was visible. The patterns shown are those obtained with (a) Hae III, (b) Eco R II, (c) Hinf I. In these cases, bands were obtained in non isomolecular yields on an unresolved background of fragments. Finally, with (d) Mbo I, all the satellite DNA III molecules were broken into discrete fragments. These results demonstrate the highly repetitive nature of these DNA sequences and that, although they have all the same density in CsCl, the satellite DNA III molecules have not all developped the same higher order periodicities. This led me to consider that the molecules of satellite DNA III could be differentiated into several subsatellite fractions. The relative positions of the restriction sites on the different subsatellite fractions will be published elsewhere.(a) was run on a different gel than (b), (c) and (d).

almost all the chromosomes (with the exception, sometimes, of the sex chromosomes) contained these types of sequences. This is particularly true in the case of mouse (17,18). The same is true for Calf satellite DNA I since Kurnit et al. (19) showed that these DNA sequences were concentrated at the centromeres of all autosomes. However, the conditions of in situ hybridization are certainly not stringent enough to discriminate between the small

differences which one can detect in the satellite DNA I molecules with restriction endonucleases.

It is possible that not all the chromosomes have the same satellite DNA sequences and that, for instance, the different subsatellite fractions of satellite DNA I are concentrated on different chromosomes or classes of chromosomes. If this is so the satellite DNA sequences could play a role, not only on speciation as already suggested by Walker P.M.B. (20), but also in the discrimination between non homologous chromosomes during meiosis.

5) The CpG containing restriction sites in Calf satellite DNA I :

The digests of satellite DNA I obtained with Hha I, Hpa II and Sac I (minor fraction) are similar in certain aspects to that of mouse satellite DNA with Eco R II found by Southern (3). He produced evidence that unequal crossing over and random divergence were the mechanisms that led to the production of a series of bands and their intermediates (as 1/2 mer, 1 1/2 mer, 2 1/2 mer...) when this satellite DNA is digested by the restriction endonuclease Eco R II;

However it is unlikely that a similar mechanism involving unequal crossing over led to the fractionnal fragments found in Calf satellite DNA I. Reassociation kinetics have shown that Calf satellite DNA I (2) and mouse satellite DNA (21, 22) have complexities of the same order (200-250 and 150 nucleotide pairs respectively). This means that if the fractionnal units of Calf satellite DNA I were produced by a mechanism similar to that proposed by Southern (3), the unequal crossing over was restricted to a few registers among the large number possible within the repeating unit of 1460 nucleotide pairs.

A more likely explanation is that the series of bands found with Hha I, for example, are due to the modification of some of the restriction sites by either base methylation or base change. This would imply the presence of the Hha I site twice on each subsatellite fraction of Calf satellite DNA I and this gives rise to two pairs of fragments : one containing the 900 and 560 nucleotide pairs fragments and the other the 600 and 860 nucleotide pairs fragments. Modification of the Sac I site can also explain the series of bands found in the minor fraction of satellite DNA I. The same is also true with Hpa II, although this restriction site is only present once per repeating unit.

Eucaryotic DNAs are strikingly deficient in the dinucleotide CpG in comparison with the 15 other possible dinucleotides(23). Grippo

et al. (24) have shown that 5 MeC is the only minor base found in sea urchin embryo DNA and that 90% occurs in the dinucleotide CpG. Recently Bird (personal communication) has shown that the methylation of C in r DNA in X.laevis protects this DNA against cleavage with Hpa II and Hha I. It is therefore interesting to note that Hpa II and Hha I both contain CpG in their recognition sites (CCGG and GCGC). Although the recognition site sequence of Sac I is not yet known the fact that like Hpa II and Hha I it hardly degrades Calf main baund DNA (G.Roizes, unpublished observations) suggests that it also contains CpG in its recognition sequence.

All these observations suggest that the modification of Calf satellite DNA I demonstrated in this study is due to the methylation of the Cs of some of the CpG doublets.

Still, as yet the possibility that the lack of certain sites is due to base change rather than methylation cannot be excluded. If this was the case one would have to conclude that some sites diverge at high rates (Hha I, Hpa II, Sac I of the minor subfraction) whilst others (Eco R I, Eco R II, Mbo I, Ava II, Hae III, Alu I, Hind II, Sac I of the major subfraction) are quite stable. This would give rise to alternating constant and variable regions similar to those found in eucaryotic ribosomal cistrons (25-27).

The fact that the lack of restriction sites occured only in this study with CpG containing restriction enzymes makes this last hypothesis very improbable and the first one very likely.

It is hard in the present state of our knowledge to attribute a specific role to the methylation of the CpG doublet. It is interesting, however, to consider a few facts coming out from this study and others and to ask some relevant questions.

It has been thought that enzymes acting like the bacterial modification enzymes may exist in eucaryotes. These enzymes add a methyl group to a given palindromic structure and this prevents its cleavage by the corresponding restriction enzyme. The methylation in the CpG in Hpa II, Hha I and also, presumably, Sac I sites of Calf satellite DNA I are consistent with this hypothesis. Although the CpG doublet is short, it is of the restriction-modification type and because it is relatively rare, it is specific enough. However, one does not know yet if all the methylatable Cs are confined to this doublet or if all the CpGs are included in sequences of the restriction-modification type. Moreover, before such an idea can be developped further, one has to know if the methylation occurs on one or

both the DNA strands and if it is heritable during replication.

In any case, the methylation of the C in eucaryotes seems to be relatively specific since it affects the CpG doublet almost exclusively (24,28). This specificity makes the suggestion of Holliday and Pugh (29) that sites containing a methylatable C are involved in the binding of specific proteins very attractive. As far as Calf satellite DNA I is concerned, it is excluded that these sites are part of the control sequences of transcription. Also, the diversity of the sites involved makes it unlikely that they are part of the replication initiation sequences. However, before one makes any further speculations, one must know more about the methylation of the CpG doublets.

I have calculated the extent of the modification to be 0.70 for two of the four Hha I sites present in Calf satellite DNA I, but similar calculations were not possible for the other sites. Further work is therefore necessary before one can say if the distribution of the methyl groups is random or not. One must also find out if modification of the CpGs affects all kinds of eucaryotic DNA sequences, in all differentiated cells.

When such information is available, one can consider the possibility that methylation in the CpG doublet of eucaryotes has something to do with the particular structure of the genetic material in chromatin or in chromosomes.

ACKNOWLEDGEMENTS

I wish to thank M. J.MARTIN for his technical assistance, Dr B. DOD for helping to correct my English and Dr E. SOUTHERN for his helpful suggestions.

REFERENCES

1 Southern, E.M. (1970) Nature, 227, 794-798
2 Botchan, M.R. (1974) Nature, 251, 288-292
3 Southern, E.M. (1975) J.Mol.Biol., 94, 51-69
4 Endow, S.A., Polan, M.L. and Gall, J.G. (1975) J.Mol.Biol., 96,665-692
5 Gall, J.G. and Atherton, D.D. (1974) J.Mol.Biol., 85, 633-664
6 Fry, K., Poon, R., Whitcome, P., Idriss, J., Salser, W., Mazrimas, J. and Hatch, F. (1973) Proc. Nat. Acad. Sci., USA, 70, 2642-2646
7 Biro, P.A., Carr-Brown, A.C., Southern, E.M. and Walker, P.M.B., (1975) J.Mol.Biol., 94, 71-86
8 Roizès, G.P. (1974) Nucleic Acid Res., 1, 1099-1120
9 Fillipski, J., Thierry, J.P. and Bernardi, G. (1973) J.Mol.Biol., 80, 177-197.
10 Studier, F.W. (1973) J.Mol.Biol., 79, 237-248
11 Danna, K. and Nathans, D. (1971) Proc. Nat. Acad. Sci., USA, 68, 2913-2917

12 Lebowitz, P., Siegel, W. and Sklar, J. (1974) J.Mol.Biol., 88, 105-123
13 Subramanian, K.N., Pan, J., Zain, S. and Weissman, S.M. (1974) Nucleic Acid Res., 1, 727-752
14 Yoshimori, R. (1971) Ph D thesis, University of California San Francisco
15 Smith, H.O. and Wilkox, K.W. (1970) J.Mol.Biol., 51, 379-391
16 Hörz, W., Hess, I. and Zachau, G. (1974) Eur.J. Biochem., 45, 501-512
17 Pardue, M.L. and Gall, J.G. (1970) Science, 168, 1356-1358
18 Jones, K.W. (1970) Nature, 225, 912-915
19 Kurnit, D.M., Shafit, B.R. and Maio, J.J. (1973) J.Mol.Biol., 81, 273-284
20 Walker, P.M.B. (1971) Nature, 229, 306-308
21 Sutton, W.D. and Mc Callum, M. (1971) Nature, 232, 83-85
22 Hutton, J.R. and Wetmur, J.G. (1973) Biochem.Biophys.Res.Commun., 52, 1148-1155
23 Subak-Sharpe, J.H. (1969) Handbook of Molecular Cytology, 67-87 (A.Lima de Faria, ed.), North Holland, Amsterdam
24 Grippo, P. Iaccarino, M., Parisi, E. and Scarano, E. (1968) J.Mol.Biol., 36, 195-208
25 Smith, G.P. (1973) Cold Spring Harbor Symp. Quant. Biol., 38, 507-513
26 Wensink, P.C. and Brow, D.D. (1971) J.Mol.Biol., 60, 235-247
27 Miller, O.L. and Beatty, B.R. (1969) Science, 164, 955-957
28 Scarano, E. , (1971) Adv. Cytopharmacol., 1-13
29 Holliday, R. and Pugh, J.E. (1975) Science, 187, 226-232