STUDIES ON THE METHYLATION OF EUKARYOTIC DNA

by

MICHAEL J. BROWNE

A thesis presented for the degree of

Doctor of Philosophy

The University of Glasgow

October, 1976.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENTS

I am indebted to a number of persons for their help with various aspects of this work. In particular, I would like to express my gratitude to:-

Professor R.M.S. Smellie and Professor A.R. Williamson for making the facilities of the Department of Biochemistry available for this research.

Dr. R.H. Burdon for his supervision and instruction during the period of this work.

Dr. B.E.H. Maden and Dr. N.W. Fraser for advice on fingerprinting techniques.

Dr. R. Eason for assistance with sedimentation velocity analyses.

Dr. R.L.P. Adams for helpful discussions.

Dr. A. Bird (M.R.C. Mammalian Genome Unit, Edinburgh University) for advice on the Hpa II Restriction-Modification System.

Mr. D.M. Tillman for generously providing samples of SV40 DNA.

Mr. J. Turnbull for generously providing samples of DNA methylated *in vitro*.

The staff of the Wellcome Cell Culture Unit for providing stocks of cells.

Miss M. Lowdon for proof reading the Introduction.

Miss N. Golder for her patient typing of this work.

I acknowledge the receipt of a research studentship from the Science Research Council.

ABBREVIATIONS

The standard abbreviations, as recommended by the Editors of the Biochemical Journal (Biochem. J. (1973) 131, 1-20) are used with the following exceptions:

5 MeC	5-methyl cytosine.
6 MeA	N^6-methyl adenine.
Py	Pyrimidine.
Pu	Purine.
MRNA	Messenger RNA.
hnRNA	Heterogeneous nuclear RNA.
BSS	Balanced Salt Solution.
HAP	Hydroxylapatite.
DMSO	Dimethyl sulphoxide.
P.B.S.	Phosphate Buffered Saline.
SAM	S-adenosyl-L-methionine.
SDS	Sodium dodecyl sulphate.
SSC	0.15M NaCl, 0.015M Trisodium Citrate.

Nucleosides & nucleotides: the 'd' (deoxy) prefix has been omitted except in the case of possible ambiguity.

CONTENTS

Page No.

INTRODUCTION:

Page No.

EXPERIMENTAL (Contd)

LIST OF FIGURES

LIST OF TABLES

SUMMARY.

The DNA of eukaryotes has been shown to contain the 'minor' base 5-methyl cytosine. This arises by enzymic modification of cytosine already incorporated into DNA. The balance of evidence suggests that this base is involved in the control of expression of information encoded in the genome. The work presented here is an attempt to elucidate the relationship of DNA methylation to cellular differentiation using a variety of approaches.

Preliminary experiments using a model differentiating system (the Friends Erythroleukaemic cell line) showed the techniques used to measure gross changes in the DNA content of 5 MeC to be inadequate; alternative approaches were sought.

Pairs of cell lines derived from the same species, e.g. the human HeLa and Chang Liver cell lines, were used as models for intertissue comparisons and Hamster, Mouse, Human, *Xenopus* and Soya Bean Cotyledon cell lines were used in interspecies comparisons.

The characteristics of DNA methylation in these cells were determined by:

(i) DNA fingerprinting. This approach showed the distribution of 5-methyl cytosine in pyrimidine oligonucleotides to be identical in cell lines derived from the same species and to exhibit only slight variations between vertebrate species, virus transformation of cells did not alter the pattern of methylation (the plant cell line was markedly different).

(ii) Fingerprinting of DNA methylated *in vitro* by a purified DNA methylase showed that the enzyme is able to methylate DNA in a manner remarkably similar to that found *in vivo*.

This result, taken with the observed lack of variation in distribution of 5-methyl cytosine in the vertebrate cell lines,

suggests that all vertebrate cells contain only one DNA methylase and that the characteristics of the DNA methylase and nucleotide sequence methylated have been evolutionarily conserved.

(iii) Large variations in the 5 MeC content of some of the cell lines used suggest that there are 'control' elements acting upon the basic DNA methylase system mentioned in (ii).

(iv) Isopycnic centrifugation and thermal elution of DNA from hydroxylapatite showed the distribution of 5-methyl cytosine, with respect to the G+C and A+T rich sections of the genome, to be species specific. This may reflect the methylation of satellite DNAs.

(v) Rapidly reannealing DNA of HeLa cells was shown to be enriched for 5 MeC content, as has been shown for rodent cells - evolutionary conservation is implicated.

(vi) Digestion of DNA using the Hpa II restriction endonuclease, which will not cleave the sequence 5' C-C-G-G 3' when methylated, provided an alternative method for measuring the level of 5-methyl cytosine in eukaryotic DNA and furnished evidence for the clustering of unmethylated Hpa II restriction endonuclease cleavage sites.

The features of the interaction between the DNA methylase and DNA, and the control of this interaction in vertebrate cells are discussed.

INTRODUCTION

1. GENERAL DESCRIPTION OF DNA

1.1. Historical

Nucleic acids were first discovered by Meischer in 1868 but their importance was not fully realised until DNA was shown to have a genetic function. Thus Avery et al., (1944) demonstrated that the purified DNA from Pneumococcus type III could transform Pneumococcus cells from different strains to produce a type III phenotype which was maintained in the following generations. Hershey and Chase (1952) found that DNA was the only substance to enter host cells upon infection by T even bacteriophages. These observations are fundamental to much of modern biochemistry although DNA has since been deprived of its exclusive role by the demonstration that purified RNA from viruses is infective (Gierer and Schramm, 1956).

1.2. Structure

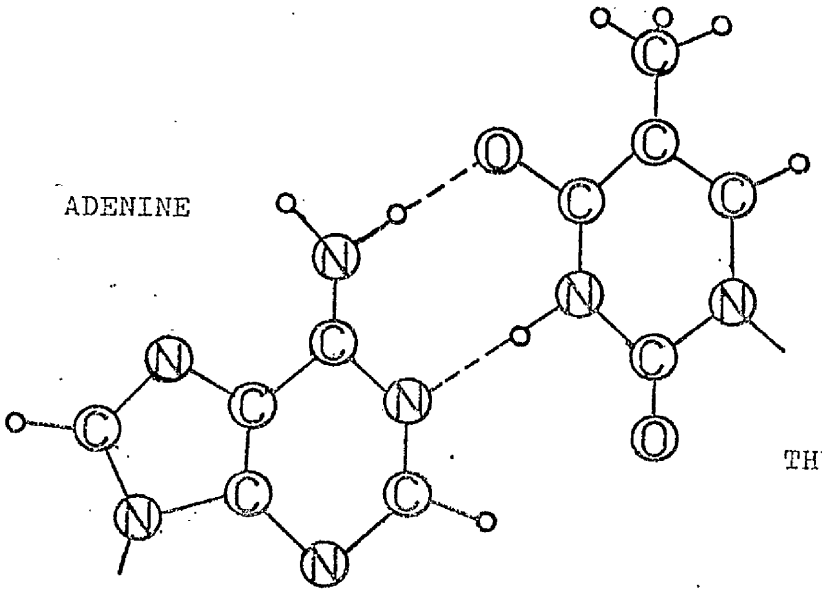
Chemical analysis has shown DNA to be a nucleotide polymer, normally two such polymers of opposite polarity are wound around the same axis and held together by hydrogen bonds between the bases (Watson and Crick, 1953), adenine in one strand always pairing with thymine in the other, and guanine always pairing with cytosine (Fig. 1).

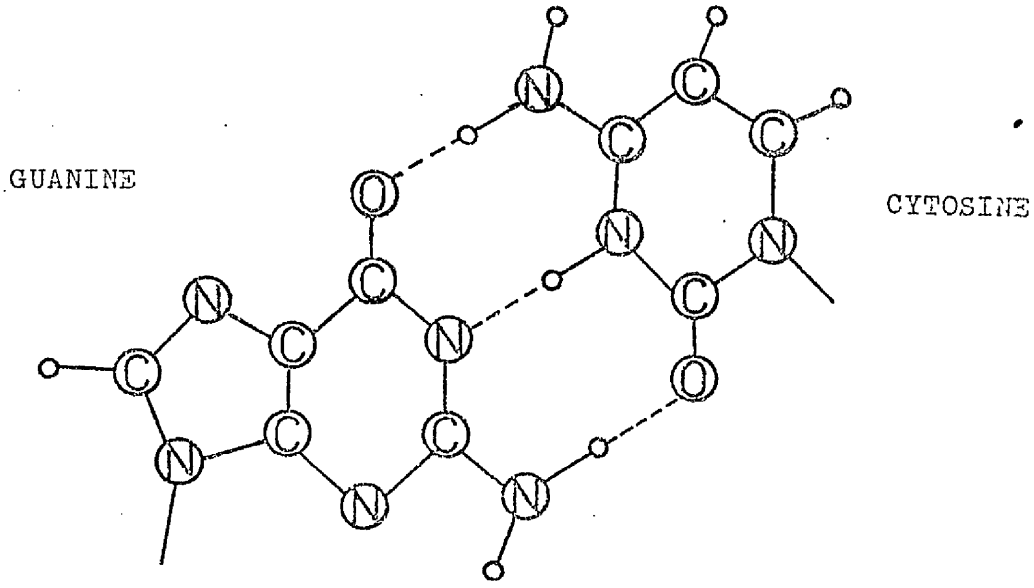
1.3. The Central Dogma

The complementary nature of the polynucleotide strands of DNA and the specific arrangement of the four nucleotides have suggested the basic mechanism by which DNA can replicate

FIG. 1

Base Pairing in DNA.





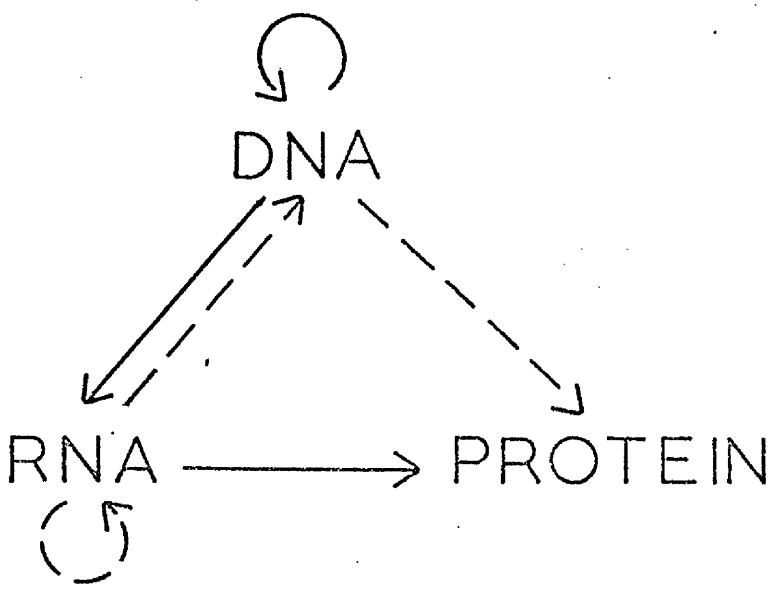


FIG.2.

"THE CENTRAL DOGMA"

The solid arrows show general transfers of sequence information dotted arrows show special transfers(after Crick,1970).

and be copied into RNA molecules of specific sequence (transcribed), these RNA molecules are in turn used as templates on which to synthesize proteins (translation), these features constitute the basis of the 'Central Dogma' (Crick, 1958) see Fig. 2. Experimental evidence for the sequence information flows indicated has been accumulated in a wide range of organisms. We are, however, unable to explain in full the manner in which flow(s) of material can be controlled in this system, nor can we explain the reason for the existence of the large amount of DNA and RNA in eukaryotes whose function remains largely speculative (Britten and Davidson, 1969; Comings, 1972; Lewin, 1975).

The following sections outline present knowledge of one particular aspect of this system, i.e. the postsynthetic methylation of bases in DNA and review its possible function(s) in eukaryotes.

2. METHYLATED BASES IN DNA

2.1. Methylated Bases in Prokaryotic DNA

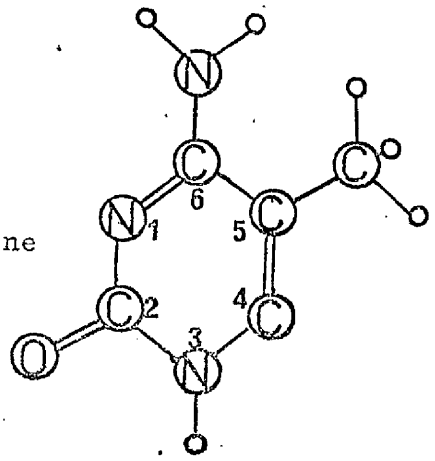
A number of postsynthetic modifications of the bases in DNA are known to occur in bacterial systems. The methylated derivative of adenine, N^6-methyl adenine (6 MeA), Fig. 3, was first discovered in bacteria by Dunn and Smith (1958), and has since been shown to occur in a large number of species (Jones and Walker, 1963; Vanyushin et al., 1965; Vanyushin, 1968). The content of 6 MeA in E. Coli and Serratia marcesens is about 0.5 to 0.7 mole %.

One other common methylated base, 5-methyl cytosine

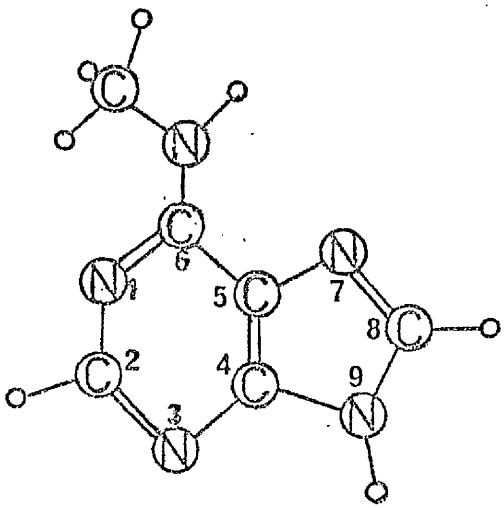
FIG.3

Methylated Bases in DNA.

5-Methyl Cytosine



N^6-Methyl Adenine



(5 MeC) Fig. 3, has been identified in a number of bacteria (Johnson and Coghill, 1925; Doskocil and Sormova, 1965a, b; Fujimoto et al., 1965; Oda and Marmur, 1966; Vanyushin, 1968). The mole % of 5 MeC is normally somewhat less than that of 6 MeA, lying in the range 0.02 to 0.2 moles % (Vanyushin, 1968). Both modifications may be present in the same strain, or one or both may be absent.

Bacteriophage DNA may also be methylated as either 6 MeA or 5 MeC. The extent of modification appears to be variable; in Xanthomonas bacteriophage XP-12 all cytosines are replaced by 5 MeC. (Ehrlich, Ehrlich and Mayo, 1975). Modifications other than the addition of methyl groups have been noted: bacteriophage T_4 contains the base 5 hydroxy methyl cytosine (Wu and Geiduschek, 1975) and bacteriophage T_6 contains the glucosylated derivative of 5 hydroxy methyl cytosine (Wyatt and Cohen, 1953).

2.2. Methylated Bases in Eukaryotic DNA.

In contrast to the prokaryotic systems the only reported modified base in higher eukaryotes is 5 MeC (Hotchkiss, 1948; Wyatt, 1951; Vanyushin, 1968). In vertebrates 5 MeC constitutes about 2 mole % of the total bases in DNA; in insects the quantity is much lower - 0.002 moles % in the locust (Wyatt, 1951) and in Drosophila 5 MeC may be negligible, i.e. only one base in 3×10^5 (I.B. Dawid, personal communication). The DNA of higher plants contains 5 to 7 moles % of 5 MeC (Wyatt, 1951; Brawerman and Chargaff, 1951; Tewari and Wildman, 1966). Fungi are reported to have no methylated bases (Guseinov et al., 1972).

The base 6 MeA was at one time thought to be confined to the prokaryotes, however, Gorovsky et al., (1973) have shown the existence of 6 MeA

as well as 5 MeC in the macronucleus of a number of Tetrahymena species (0.65 - 0.85 moles % of 6 MeA) and Cummings et al., (1974) found 2.5 moles % of 6 MeA (but no 5 MeC) in the DNA of Paramecium aurelia.

Evidence concerning the methylation of mitochondrial DNA is comparatively scant, but Nass (1973) has shown the presence of low levels of 5 MeC in mouse L cells, BHK-21/C13 cells and the polyoma transformed BHK-21/PyY cell line, and Evans and Evans (1970) found 5 MeC in the mitochondria of Physarum. Dawid (1974) was unable to detect 5 MeC in the mitochondria of HeLa and Xenopus cells. This conflict of results is probably due to the differing sensitivities of the methods used to detect 5 MeC.

2.3. Methylation of DNA as a Postsynthetic Event.

There are several lines of evidence indicating that methylation of DNA is a postsynthetic event: (i) Methylation occurs on previously synthesised DNA (Introduction, section 6.1.). (ii) DNA Methylase activities have been isolated which will catalyse the methylation of bases at the polynucleotide level in vitro (see Introduction, section 3). (iii) The formation of 5 MeCTP in vivo from dCTP or TTP and its subsequent incorporation into DNA can be excluded as Sneider and Potter showed that dCTP is not methylated in vivo, and Rubery and Newton (1973) could not detect the aminated product of TTP in DNA.

2.4. The Stability of Methylated Bases in Eukaryotic DNA.

It is important in postulating a function for the methylation of DNA to know whether the modification is transient

or not. Once cytosine is methylated there are three means of removing 5 MeC (i) deamination to T, (ii) demethylation and (iii) excision. Burdon and Adams (1969) showed that in L929 cell DNA labelled *in vivo* with (^{3}H) Thymidine and L(methyl-^{14}C)-methionine (in the presence of sodium formate to specifically label 5 MeC with ^{14}C) the ratio of label in 5 MeC to T remained constant over eight cell generations, thus ruling out *any* loss of 5 MeC. Sneider (1973) suggests that deamination of 5 MeC does not occur *in vivo* and that any low level reported is purely artefactual.

It is, of course, not possible to exclude the operation of any of the three mechanisms mentioned above, at a level below the limits of detection.

2.5.1. The Effect of Methylation on the Stability of Eukaryotic DNA.

The introduction of a bulky methyl group onto the 6 amino position of adenine in a synthetic polydeoxyribonucleotide lowers the T_M (the temperature where the polymer or DNA is half denatured) i.e. it tends to destabilise the double helix (Vanyushin, 1968). Methylation of cytosine in the 5 position, however, tends to stabilise the double helix and results in an elevated T_M (Dawid *et al.*, 1970; Gill *et al.*, 1974).

2.5.2. The Effect of Methylation on Apparent Base Composition.

As noted above (Introduction, section 2.5.1.) methylation of cytosine elevates the T_M of the DNA and thus increases its apparent G+C content. Neutral CsCl density gradient centrifugation however indicates an apparent decrease in G+C content on increasing the level of 5 MeC. Gill *et al.*, (1974)

suggest that the decreased buoyant density is due to an increase in the volume of the DNA and to altered hydration in the major groove of the DNA. Whatever may be the reason for this anomalous behaviour these observations are useful in themselves as a means for detecting the presence of 5 MeC (Dawid et al., 1970).

3. DNA METHYLASES AND THEIR SUBSTRATE SPECIFICITY

3.1. Prokaryotic Methylases.

3.1.1. The Restriction-Modification System of Bacteria. General.

Methylase enzymes responsible for the modification of bacterial and bacteriophage DNAs in vivo have been shown in many instances to be part of a Restriction-Modification (R-M) system (Nathans and Smith, 1975). The phenomenon of restriction modification is thought to be a means by which bacteria are able to eliminate foreign DNA, by endonuclease attack, meanwhile protecting homologous DNA by modification.

Restriction modification systems are characterised by (i) the ability of an enzyme activity which can catalyse double stranded breaks in DNA and (ii) the ability of an enzyme to catalyse the donation of a methyl group from S-Adenosyl methionine (SAM) onto DNA producing either 6 MeA, e.g. $ECoR_I$ methylase or 5 MeC, e.g. $ECoR_{II}$ methylase (Arber, 1974).

Modification of DNA prevents cleavage by homologous but not heterologous endonucleases suggesting that there are specific recognition sites on the DNA for each restriction modification system.

Restriction modification systems fall into two main groups, the characteristics of the restriction systems are noted in Table I.

TABLE I

TABLE OF PROPERTIES OF RESTRICTION ENZYMES

	Class I	Class II
M.W.	ca. 300,000	20,000 - 100,000
Cofactor Requirements	ATP, Mg^{2+}, SAM	Mg^{2+}
Point of Cleavage	Remote from recognition site	At recognition site
Product of cleavage	Heterogeneous fragments	Discrete fragments
Examples	ECoB, ECoK	$ECoR_1$

(from Nathan & Smith, 1975)

3.1.2. Class I R-M Systems.

Genetic studies of the R-M system of E. Coli B suggest that the system is specified by three genes (i) hsdS determining the nucleotide sequence recognition sub-unit (ii) hsdR determining the restriction enzyme and (iii) hsdM determining the modification enzyme. Mutants of the type $hsdS^+$, $hsdM^+$, $hsdR^-$ are able to catalyse the modification reaction, but mutants of the type $hsdS^+$ $hsdM^-$ $hsdR^+$ are unable to catalyse restriction. It is suggested that the gene product of hsdM is necessary for restriction as well as modification (Arber, 1974). The enzymes constituting the R-M system of E. Coli B have been isolated (Eskin and Linn, 1972; Lautenberger and Linn, 1972) and consist of three types of subunit: α, β and γ of molecular weight 135,000, 60,000 and 55,000 respectively. The β and γ subunits together are able to function as the modification system, the addition of the α subunit allows restriction to occur in the presence of SAM, omission of either β or γ prevents restriction. These observations are interesting as there is no evidence for the methylation of DNA as a step in cleavage; it has been suggested that this mechanism prevents digestion of bacterial DNA in $hsdM^-$ mutants.

3.1.3. Class II R-M Systems.

Genetic analysis has shown that the genes for $ECoR_I$ and $ECoR_{II}$ R-M systems are located on plasmids (Hattman and Cousens, 1972), there appear to be two distinct proteins in these systems which unlike Class I enzymes are able to act independently, although both restriction and modification enzymes recognise and act at the same point(s) on the DNA molecule (Dugaiczyk et al.,

1974).

3.2. Methylated Sequences in Prokaryotic DNA.

The ECoR$_I$ modification enzyme has been shown to methylate adenine bases in the following sequence

5' G A $\overset{*}{A}$ T T C 3'

3' C T T $\underset{*}{A}$ A G 5'

* = CH_3

and the homologous restriction enzyme produces the following cleavage

5' G$\downarrow$A A T T C 3'

3' C T T A A$\uparrow$G 5'

Methylation of adenine in one or both strands prevents restriction.

Many Class II recognition sequences have now been determined (Nathans and Smith, 1975). All known Class II recognition sites are 4–6 nucleotides long and possess two fold rotational symmetry, this has led to the speculation that R–M enzymes will consist of subunits arranged with two fold symmetry, and it is known that the ECoR$_I$ endonuclease is composed of two identical subunits (Greene <u>et al</u>., 1974).

Evidence suggests that the recognition sites for Class I enzymes may have similar properties (Brockes et al., 1974).

3.3. EUKARYOTIC DNA METHYLASES

3.3.1. Mammalian DNA Methylases.

The isolation of mammalian DNA methylase enzymes has been stimulated by success in characterising several bacterial systems. The fact that mammalian cells contain a DNA methylase activity was shown in a variety of relatively impure systems (Burdon et al., 1967; Kalousek and Morris, 1968; Scheid et al., 1968; Vanyushin et al., 1971; Tosi et al., 1972; Burdon and Douglas, 1974). Several authors have reported the purification of DNA methylases: from rat spleen (Kalousek and Morris, 1969a); from rat liver (Drahovsky and Morris, 1971 a,b, 1972; Morris and Pih, 1971); from HeLa cell nuclei (Roy and Weissbach, 1975); from Novikoff rat hepatoma cells (Sneider et al., 1975) and from Krebs II mouse ascites tumour cells (Turnbull and Adams, 1976).

The purified enzymes catalyse the transfer of methyl groups from SAM to the 5 position of cytosine; none of the enzymes methylates any other base. These methylases differ in some respects but all appear to be able to methylate native and denatured DNA from heterologous sources and, to a lesser extent, homologous DNA. The use of homologous DNA as a substrate in vitro indicates incomplete methylation in vivo.

Rat liver methylase (Drahovsky and Morris, 1971 a,b) has the ability to form complexes with DNA; a weak one at $0^{\circ}C$ and a tightly bound one at $37^{\circ}C$. This enzyme is believed to bind to DNA at random and then move along the length of the molecule remaining attached to the DNA between successive methylations; this idea is supported by the observation that the rate of methylation decreases as the molecular weight of the substrate DNA is decreased. The Krebs II ascites enzyme (Turnbull and

Adams, 1976) however, does not form similar complexes and there is no evidence for "walking" along the DNA. Variation of the properties of these enzymes may reflect species or tissue specificity or may be some artefact of purification, e.g. loss of subunits or aggregation. It must be noted that apart from the Krebs II ascites and the HeLa methylase none of the methylases were highly purified, and even in these latter systems they have not been shown to be single enzymes.

3.3.2. Plant DNA Methylases.

Kalousek and Morris (1969 b) showed that it was possible to prepare a crude extract of pea seedling with the ability to methylate DNA *in vitro*.

3.3.3. Eukaryotic Restriction Endonucleases.

In bacterial systems the DNA methylase enzymes are associated with a restriction endonuclease, (methylation of DNA affords protection against cleavage) however there is no evidence for analogous endonucleases in eukaryotes as methyl deficient DNA appears to be stable *in vivo*. Culp and Black (1971) effected undermethylation of mouse 3T3 cell DNA by methionine starvation but were unable to detect degradation of this DNA *in vivo*. Interpretation of this observation is difficult as only newly replicated DNA would be undermethylated under the conditions used: in bacterial systems DNA methylated in one strand only is not normally a substrate for the homologous restriction endonuclease (Introduction, section 3.1.).

A more unambiguous source of evidence is provided by the DNAs of Polyoma virus (Kaye and Winocour, 1967), Herpes Simplex virus (Low *et al.*, 1969) and the amplified ribosomal genes of

Xenopus (Dawid et al., 1970) all of which are not methylated and yet appear to be stable in vivo.

3.4. METHYLATED SEQUENCES IN EUKARYOTIC DNA.

3.4.1. Methods of DNA Sequencing.

RNases able to cleave ribonucleopolymers at a point specified by bases in the polymer, e.g. T_1 RNase which cleaves RNA only at the 3' end of guanosine residues, have proved useful in the sequence analysis of RNA (Sanger et al., 1976). Although the analysis of DNA is limited by the lack of analogous specific DNases the chemical degradation of DNA by a diphenylamine/formic acid solution allows the specific production of clusters of pyrimidine nucleotides of the general formula $_p(P_{y-p})_n$ (Burton and Petersen, 1960). These oligonucleotides can then be separated according to length into pyrimidine isostich by DEAE cellulose chromatography or by length and base composition, e.g. resolving C_{2p3}, CTp_3 and T_2p_3 by two-dimensional electrophoresis (Murray, 1970) or by electrophoresis followed by homochromatography in a second dimension (Ling, 1972). This methodology has been important in analysing the specificity of sequences methylated in DNA.

Copies of DNA molecules may be synthesised in vitro using E. Coli DNA polymerase, one of the substrate deoxyribonucleonucleoside triphosphates being α ^{32}P labelled. The newly synthesised DNA may be digested with micrococcal DNase and spleen phosphodiesterase to yield ^{32}P labelled 3' mononucleotides. The ^{32}P originally introduced by the substrate deoxynucleoside triphosphate is therefore transferred to its 'nearest neighbour' in the polynucleotide chain (Swartz et al., 1962). This technique thus allows characterisation of the arrangement of nucleotides in DNA.

A certain degree in the specificity of cleavage of DNA by DNases, e.g. DNase II (Doskocil and Sorm, 1961) allows the relatively specific production of oligonucleotides which may be separated using chromatography and further characterised.

3.4.2. Methylated Sequences in Animal DNA.

Unlike the bacterial methylase systems there is no known nucleotide sequence(s) specifying recognition by animal DNA methylases. It is known that methylation does not occur at random *in vivo*. In animals the majority of 5 MeC is found in association with monopyrimidines (Sinsheimer, 1955; Doskocil and Sorm, 1962 and Table II). In calf thymus and sea urchin DNA the most highly methylated cytosine occurs in the relatively infrequent dinucleotide C-G (Doskocil and Sorm, 1961; Grippo *et al.*, 1968; Swartz *et al.*, 1962). The majority of mammalian methylations thus occur in the sequence: Purine-5 MeC-G. However this is by no means the only sequence methylated. Table II shows that methylation also occurs in $(Py)_n$-5 MeC-G and Sneider (1972) has found substantial amounts of 5 MeC occurring at the 5' end of pyrimidine tracts and also internally.

The methylation specifying site in HeLa DNA has been designated as 5' G-5 MeC-G 3' or 5' C-5 MeC-G 3' (Roy and Weissbach, 1975), however the sequence analysis used only encompassed 40-50% of the total DNA; the fact that the artificial template poly (dG-dC·dG-dC) is saturated at a level equivalent to only 1 methyl group per 1800 bases suggests that this is due to an inhomogeneity of the polymer and indicates that the recognition and/or modification site is more complex than the partial sequence analysis showed.

TABLE II

PERCENTAGE DISTRIBUTION OF 5-METHYL CYTOSINE

IN DNA PYRIMIDINE ISOSTICHS

PYRIMIDINE LENGTH	DNA SOURCE: RAT HEPATOMA CELLS*	SEA URCHIN+
MONO	49.3	60
DI	21.5	12
TRI	13.2	10
TETRA	8.3	9
PENTA	3.9	6
HEXA	2.1	-
HEPTA	1.0	-
OCTA	0.6	-
NONO	0.3	-

* calculated from data of Sneider, 1971

\+ calculated from data of Grippo et al., 1968

All figures calculated as % of total pyrimidines in each isostich.

Fry et al., (1973) determined the sequence of the basic repeating unit in Dipodemys ordii BS-β satellite DNA to be:

```
5'     GG¦ACACAGCGGG¦ACAC   3'
       ¦              ¦
3'     CC¦TGTGTCGCCC¦TGTG   5'
```

On the basis that only the dinucleotide C-G is methylated they predicted that there should be 7.4 mole % 5 MeC in this satellite; this is in close agreement with the actual value of 6.7 ± 1.2 moles %.

Mammalian methylases are able to methylate undermethylated DNA, i.e. double stranded DNA where one strand is normally methylated and the other is 5 MeC deficient. Thus undermethylated L929 DNA produced by methionine starvation is the preferred substrate for the Krebs II ascites enzyme (Turnbull and Adams, 1976), whereas 'normally' methylated DNA isolated from stationary L929 cells is a poor substrate. In this instance mammalian methylases behave in a similar fashion to the bacterial methylases, which rapidly methylate undermethylated homologous DNA but not normally methylated DNA (Introduction, section 3.2.).

Many bacterial R-M recognition sites in DNA have been shown to possess two-fold rotational symmetry as does the 'lac' operator (Gilbert and Maxam, 1973). Two-fold rotational symmetry in the nucleotide sequence might then be supposed to be a general feature of DNA methylase/DNA interaction. However the HS-β satellite (above) does not appear to possess any two-fold rotational symmetry and it is probable that the protein nucleic acid interaction in this system must rely on some other feature of the DNA structure.

Denaturation of the DNA substrate stimulates the rate of DNA methylation thus eukaryotic methylases differ from prokaryotic methylases in this respect

(Sneider et al., 1975; Turnbull and Adams, 1976).

3.4.3. Methylated Sequences in Plant DNA.

The distribution of 5 MeC in plant DNA is somewhat different to that in animals or bacteria. Shapiro and Chargaff (1960) found that in Rye germ DNA 5 MeC occurs 70% more often in the mono-pyrimidine fraction than C, whereas in vertebrate DNA 5 MeC occurs at least twice as often as C in monopyrimidines. Doskocil and Sorm (1962) produced a more detailed analysis of Wheat germ DNA, thus 5 MeC,while occurring mainly in the nucleotide doublet 5 MeC-G, also appears in the dipyrimidines 5 MeC-T and 5 MeC-C, the % of C methylated in the dipyrimidines is almost as high as in mono-pyrimidines; other longer methylated sequences are also noted. Spencer and Chargaff (1963) found that in Rye germ DNA 31% of 5 MeC occurred as the monopyrimidine: the corresponding figure for Wheat germ was 33%.

The tendency for preferential methylation of monopyrimidines is thus less in plants than in vertebrates.

4. THE ORGANISATION OF 5 MeC IN THE EUKARYOTIC GENOME.

4.1. Methylation of Satellite DNA.

Satellite DNA was first reported in mouse DNA (Kit, 1961). Centrifugation of mouse DNA on a CsCl density gradient revealed the presence of two peaks separated due to differences in their density, the density of the main band, 1.701 g/ml, corresponds to a G+C content of 42% while the density of the satellite 1.691 g/ml corresponds to a G+C content of 34% (Flamm et al., 1967) and constitutes about 10% of the total DNA. 'Hidden'satellites have

since been demonstrated in many other species (Corneo et al., 1970, 1972). When mouse satellite DNA is isolated and denatured it is observed to reanneal rapidly (Walker & McLaren, 1965). On the basis of the kinetics of reassociation the satellite DNA is composed of units of 300-400 nucleotides length repeated 10^6 times (Waring and Britten, 1966; Flamm et al., 1969). However sequencing studies suggest that the repeating unit is only 8-13 bases long (Southern, 1970). Overestimation of the repeat length by renaturation data may be due to mismatching of bases as the repeating units, though similar, are not identical (Southern, 1970, 1971; Sutton and McCallum, 1971).

The C3H mouse embryo cell satellite DNA has a molar concentration of 5 MeC approximately twice that of main band DNA (Salomon et al., 1969). Similarly mouse L cell satellite DNA contains four times as much 5 MeC as main band DNA, and the HS-β satellite of Dipodomys ordii has a high level of 5 MeC (6.7 mole %).

The distribution of 5 MeC in pyrimidine tracts of mouse C3H cell satellite DNA is reported to be the same as main band DNA (Salomon et al., 1969) but Harbens et al., (1975), using mouse L cell DNA found that the distributions differed. The biological significance of the latter observation is unknown.

4.2. Methylation of Rapidly Reannealing DNA.

As noted above (Introduction, section 4.1.) the mouse cell satellite DNA reanneals rapidly, i.e. it has a low sequence complexity. Satellite DNAs represent one form of rapidly reannealing DNA. Britten and Kohne (1968) denonstrated that the DNA of higher organisms is composed of three main classes: highly repetitious, present many times in the haploid genome;

moderately repetitious and non-repetitious, present once or a few times in the haploid genome, the amount of each class being species specific. About 30-40% of the DNA of higher organisms consists of repeated sequences. Davidson et al., (1975) consider that 50% of the DNA of higher organisms consists of interspersed repetitious (300 ± 100 nucleotides long) and non-repetitious (800 ± 200 nucleotides long) sequences.

The rapidly reannealing DNA of Chinese hamster cells (Sneiderman and Billen, 1973) and of mouse Ehrlich ascites carcinoma DNA (Brahic and Fraser, 1971) is enriched for 5 MeC by two to three fold. Sneiderman and Billen (1973) showed that this DNA replicates early in S phase and suggested that the initiation region of each replicon contains a highly methylated rapidly reannealing DNA moiety. The involvement of 5 MeC in DNA replication is reviewed in Introduction, sections 5.1. and 5.2.

4.3. Methylation of Chromatin Fractions.

The DNA of eukaryotes in vivo exists in a complex with histone and non-histone proteins, the whole complex being known as chromatin. Comings (1972) found that late replicating, A+T rich heterochromatin in hamster DON cells was methyl deficient to an extent greater than could be accounted for by base composition alone. He proposed that as this DNA was not transcribed, the deamination of 5 MeC could occur as the evolutionary pressure to retain cytosine was reduced, thus reducing the G+C content of this DNA. This observation may not be generally applicable as mouse cultured cells have a late replicating A+T rich satellite enriched in heterochromatin which is preferentially methylated. The preferential methylation of heterochromatin in mouse strain AKR and human has been confirmed using

an antibody to 5 MeC which binds almost exclusively to hetero-chromatin (Miller et al., 1974). The proposed mechanism of 5 MeC loss could, however, explain in part the relatively low level of G+C in vertebrates and in particular the reduced frequency of the nucleotide doublet C-G. It could also explain the increased incidence of the dinucleotides T-G and C-A which would result from the production of T-G in one DNA strand and its complementation by C-A in the opposite DNA strand following replication (Swartz et al., 1962).

No evidence has yet been presented to connect the distribution of 5 MeC with the subunit structure of chromatin (Carpenter et al., 1976).

4.4. The Relationship of 5 MeC to Transcribed Regions of DNA.

The genetic information encoded in DNA must, for expression in the form of proteins or structural RNA, e.g. transfer and ribosomal RNAs, be transcribed by an RNA polymerase. Recent work using RNA:DNA hybridisation has shown that it is probable that not all nuclear DNA in higher organisms is transcribed; thus 6% of the non-repeated DNA in neonatal mice (Gelderman et al., 1971), 15% of the genome in cultured mouse cells (Grady and Campbell, 1973) and 28% of the genome in sea urchin embryos (Hough et al., 1975) is complementary to sequences in RNA from these cells, the majority of the DNA thus appears to be untranscribed.

Transcription of the genome does not yet appear to be related to any particular feature of the DNA. It is known that the major fraction of RNA is transcribed from the unique sequences in DNA. A substantial portion of RNA is, however, complementary to repetitive DNA, thus Smith et al., (1974) show that 20-25% of hnRNA molecules in sea urchins contain repetitive transcripts and Holmes

and Bonner (1974) show that rat ascites hnRNA is similar. An important feature of both papers is the demonstration that in many hnRNA molecules there are both repetitive and non-repetitive sequences covalently linked together. This implies that these molecules are transcribed from regions of DNA containing interspersed repetitive and non-repetitive sequences (Britten & Kohne, 1968).

Pagoulatos (1974) in mouse plasmocytoma cells and Cohen et al., (1973) in mouse L-cells showed the presence of RNA molecules complementary to the highly repetitious mouse satellite DNA (see section 4.1.). However both authors suggest that this may represent transcription of some inhomogeneity acquired by the satellite since the cells are no longer normal.

A simple model of cellular differentiation might postulate the transcription or nontranscription of certain genes in different ontological phases. However using an assay for the existence of globin mRNA sequences it is possible to detect low levels of these sequences in non-erythroid cells (Humphries et al., 1976). This suggests that transcriptional control may not be rigorous, involving an absolute on/off mechanism, but may be rather more subtle.

There is little evidence for a direct topographical link between methylated bases and transcribed regions of DNA. In bacteria it has been postulated that methylated bases are located in a start codon (5 MeC) or a stop codon (6 MeA) and in some way influence the activity of the transcribing RNA polymerase (Venner et al., 1974).

Experimental observation indicates that the situation in eukaryotes is distinctly different. The DNA of Polyoma virus (Kaye and Winocour, 1967); Herpes Simplex Virus (Low et al., 1969)

and the amplified ribosomal genes of Xenopus laevis (Dawid et al., 1970) is transcribed, yet there are no methylated bases present in any of these DNAs. In contrast 5 MeC has been shown to occur in both the non-transcribed spacer regions of Xenopus laevis chromosomal ribosomal genes and in the DNA sequences complementary to ribosomal RNA (Dawid et al., 1970); in addition methylation occurs on both strands of the DNA to a comparable extent. Volpe and Eremenko (1974) were able to show hybridisation of nuclear RNA to methylated regions in HeLa cell DNA whereas messenger RNA hybridised to non-methylated DNA; they suggest that the data demonstrate preferential methylation of regulatory genes. The DNA used in this experiment was methylated in vitro and may not reflect in vivo methylation as cytoplasmic and nuclear factors have been shown to influence methylation in vitro (Burdon and Douglas, 1974).

Molitor et al., (1976) have used autoradiography to detect regions of DNA in mouse L929 cells which are deficient in 5 MeC. These regions constitute ca. 10% of the total DNA. It is suggested that these regions may be involved in the transcription process.

5.1. The Relationship of Methylation to DNA Synthesis.

It has been observed that under normal conditions the methylation of DNA occurs very soon after synthesis in both prokaryotes (Billen, 1968; Lark, 1968) and eukaryotes (Burdon and Adams, 1969). However the two are not stringently linked, thus in E. Coli grown in a methionine free medium DNA synthesis continues for some time but the DNA produced is partially undermethylated. On addition of methionine to the medium DNA methylation resumes before the onset of renewed DNA synthesis (Billen, 1968; Lark, 1968).

The independence of synthesis and methylation has been demonstrated

in eukaryotes: Burdon and Adams (1969) found that in synchronised L929 cells released from aminopterin blockade there was a wave of DNA synthesis followed by a wave of DNA methylation. In this system methylation lags behind DNA synthesis by as much as one hour and Adams (1974) showed continuing DNA synthesis for up to eight hours after methionine deprivation in the same cells. Culp and Black (1971) were able to demonstrate the production of undermethylated DNA in mouse 3T3 cells: after 24 hours methionine deprivation the newly synthesised DNA was up to 40% deficient in 5 MeC. In these three eukaryotic systems DNA synthesis can proceed in the absence of methylation.

The reverse is also true. Methylation can proceed in the absence of DNA synthesis, i.e. on DNA synthesised several generations previously. Evans et al., (1973) showed that in Physarum polycephalum methylation of progeny strands of DNA synthesised in the same cell cycle was only 1% of newly incorporated cytosines whereas parental strands were methylated at a level of 5 to 7%, the difference being made up over the succeeding generations. Adams (1973) found that the DNA synthesised three to four hours after fertilisation of Paracentrotus lividus continued to be methylated for at least the next seventy hours (in this period the cell number increased one hundred-fold); by the end of this period the % methylation had increased three-fold.

Using the drug hydroxyurea in mouse L929 cells to inhibit DNA synthesis by 94%, Burdon and Adams (1969) were able to demonstrate continuing methylation; even pre-treatment with hydroxyurea for some hours before estimation of methylation only produced an inhibition of 85%.

Methylation of DNA in Physarum during interphase (G2) has been observed (Evans and Evans, 1970).

Hilliard and Sneider (1975) showed continuing methylation of parental DNA during S phase in cells synchronised by thymidine blockade. However much of this methylation was shown to be artefactual as the conditions employed to produce synchrony were inducing repair insertion of cytosines into parental DNA which were then methylated. When synchrony was achieved by sequential mitotic G1 arrest methylation of parental DNA was reduced.

The fact that synthesis and methylation of DNA can be demonstrated to occur separately probably rules out any role for methylation in synthesis and vice versa.

5.2. Relation of Methylation to Cell Cycle.

Experiments using synchronous culture show that early replicating DNA is relatively more methylated and late replicating DNA is relatively less methylated than average, (Adams, 1971; Comings, 1972; Scheiderman and Billen, 1973). This might suggest that there is a role for methylation in the initiation of DNA replication. However the observations in 6.1. make this unlikely.

6. The Relationship of Methylation to Cellular Differentiation.

A number of observations have led to the speculation that DNA methylation might be responsible for controlling cellular differentiation: Vanyushin et al., (1970; 1973) and Kappler (1971) found that the gross level of methylation is species and tissue specific; Adams (1973) showed that the level of methylation increases with the development of the sea urchin Paracentrotus lividus; Berdyshev et al., (1967) showed a decrease in methylation by 35-40% in spawning humpback salmon; human leukaemic cells have an unusually high level of methylation (Desai et al., 1971) and the polyoma

transformed BHK cell line BHK-21/PyY has an elevated level of methylation (Nass, 1973; Rubery and Newton, 1973).

The means by which methylation could influence differentiation is unknown but one possibility is that the presence of 5 MeC in some way affects the transcription of DNA. There are two main theories concerning the way in which the amount of 5 MeC might be altered:

(i) By deamination of 5 MeC.

Grippo et al., (1970) using sea urchin embryos showed that deamination of 5 MeC to T occurred at the level of the DNA polymer. They also showed that the distribution of this "minor" T in polypyrimidine tracts was non random, as might be expected if it arose from the non-randomly located 5 MeC. The presence of a novel T,it is proposed,could alter the 'programming' of the DNA and initiate or continue development; this base change would be propagated in subsequent generations. Although Grippo et al.,(1970) took great care to ensure that the 'minor' T was not a product of isolation or hydrolysis of DNA, Sneider (1973) was able to demonstrate that deamination of 5 MeC, at least in rat hepatoma cells, was in fact artefactual. The interpretation of earlier data concerning deamination is therefore difficult.

An additional difficulty is produced when invoking deamination as a control mechanism as this process has not been shown to be reversible, yet Gurdon and Laskey (1970) were able to use nuclei from differentiated cells to programme anucleated eggs. Reversal of deamination would also pose the problem of specifying those T's to be reaminated.

(ii) Increasing Methylation per se.

This is reviewed by Holliday and Pugh (1975). Essentially

they consider increasing methylation as a mechanism for switching off gene transcription, new methylation sites being occupied at every cell division thus allowing cells to precisely count the number of divisions passed through. The authors see this as more readily reversible than deamination either by a 'demethylation' system or by a dilution effect in the absence of new methylation of DNA synthesised in embryonic cells.

Riggs (1975), however, considers increasing methylation as a means of activating chromosomes. Non-methylated chromosomes would be inactive, e.g. the 'Barr' body (inactivated X chromosome).

These models of differentiation demand that the DNA should be replicated and it does seem likely that some event at the level of DNA (but not necessarily a modification of DNA) is necessary. Thus Green and Neuth (1974) using preadipose culture cells and McClintock and Papaconstantinou (1974) using Friend Erythroleukaemic cells showed an obligatory requirement for cell division before differentiation could take place.

A note of caution should be added here; it is possible that differences in the level of methylation shown in inter-tissue comparisons or during cellular differentiation are not primarily concerned with differentiation but reflect differing rates of methylation, i.e. when DNA is being synthesised rapidly the lag period before methylation may be slightly longer than normal, resulting in a decreased content of 5 MeC.

AIMS OF THE PRESENT WORK

The Introduction has outlined the present knowledge of the characteristics and possible functions of the base 5-methyl cytosine in eukaryotic DNA. Briefly, the balance of evidence suggests that DNA methylation is not connected with the events of DNA synthesis and is unimportant in maintaining the integrity of the DNA once synthesised; there is still the possibility that methylation is in some way related to the control of expression of information encoded in the genome.

The work in the following sections is an attempt to employ a variety of novel approaches to determine the characteristics of DNA methylation which vary with cellular differentiation and those which remain constant in order to assist elucidation of the relationship between DNA methylation and cellular differentiation.

1. MATERIALS

1.1. Biological.

Hamster BHK-21/C13 (Macpherson and Stoker, 1962); Hamster BHK-21/PyY (Stoker, 1964); HeLa (Gey et al., 1952); Chang liver cells (Chang, 1954); mouse L929 (Sanford et al., 1948) and a Xenopus liver cell line (originally a gift from Dr. K. Jones, Edinburgh University) were all cell lines routinely maintained in this Department.

Mouse SVT2 cells (Aaronson & Todaro, 1968) were a gift from Dr. L.V. Crawford, I.C.R.F.; Friend's Erythroleukaemic cells, clone M2 (Birnie et al., 1974) were a gift from Dr. G.D. Birnie, Royal Beatson Institute for Cancer Research; and Soya Bean Cotyledon cells were a gift from Dr. I. Henson, Glasgow University Botany Department.

1.2. Chemical.

All chemicals were, wherever possible, Analar reagents supplied by B.D.H. Chemicals Ltd., Poole, Dorset, except for the following:

Triton X-114 (Scintillation Grade)	Koch-Light Laboratories Ltd., Colnbrook, England.
Toluene (AR Grade)	"
2,5 diphenyloxazole (PPO) (scintillation grade)	"
Trichloroacetic acid (TCA)	"
Calf Serum	Bio-Cult Laboratories Ltd., Paisley, Scotland.
Amino Acids	"
Vitamins	"
Penicillin	Glaxo Pharmaceuticals, London.

Streptomycin	Glaxo Pharmaceuticals, London.
p-Bis(o-methyl-styryl) benzene (Bis-MSB) (Scintillation grade)	Kodak Ltd.
Kodirex KD 54T (35 x 43cm) X-ray film	"
DX-80 Developer	"
FX-40 X-ray liquid fixer	"
Sephadex G-25 (medium)	Pharmacia, Uppsala, Sweden.
Spleen phosphodiesterase	Worthington Biochemical Corporation via Cambrian Chemicals Ltd., London.
Hydroxylapatite (Biogel HTP) - DNA grade	Bio-Rad Laboratories, Richmond, California.
Cellulose Acetate Electrophoresis Strips (25 x 95cms)	Oxoid Ltd., London.
Agar	"
Pre-coated DEAE cellulose sheets - 40 x 20cms	Machery-Nagel and Co., Camlab, Cambridge.
Polygram cel 300 DEAE	"
Whatman DE81 Paper (46cm x 50m)	H. Reeve-Angel and Co. Ltd., London.
Whatman No. 52 and 3MM Paper 2.5cm paper discs	"
Whatman G/FC 2.5cm glass fibre discs	"
Actinomycin D (200µg/vial)	Calbiochem Ltd., Hereford, England.
Adenine, guanine, cytosine, thymine, 5-methyl cytosine	"
dCMP & d5MeCMP ($5'PO_4s$)	Sigma London Chemical Co. Ltd., Kingston-on-Thames, Surrey.
Micrococcal Nuclease	"
Pancreatic DNAase	"

Calf Thymus and Salmon Testes DNA	Sigma London Chemical Co. Ltd., Kingston-Upon-Thames, Surrey.
Agarose (for electrophoresis)	"
Ethidium Bromide	"
Hpa II Restriction Endonuclease	Miles Laboratories Ltd., Stoke Poges, England.
S_1 nuclease	"
CsCl	Hopkin and Williams Ltd., Chadwell Heath, England.

SV40 DNA (Type I and Type II) was provided by Mr. D.M. Tillman, Biochemistry Dept., Glasgow University. Samples of DNA methylated in vitro were provided by Mr. J. Turnbull, Biochemistry Dept., Glasgow University.

1.3. Radiochemicals

^{32}P-orthophosphate (^{32}Pi 10mCi/ml) carrier free (supplied in a solution of dilute HCl)	Radiochemical Centre, Amersham.
L-(methyl-^{14}C) methionine (58 C./mol)	"
L-(methyl-^{3}H) methionine (75000mCi/mmol) in aqueous solution containing 0.2% β mercapto ethanol	"
(6-^{3}H)-Thymidine (20,000-30,000mCi/mmol)	"
Deoxy(U-^{14}C) cytidine (> 405mCi/mmol)	"

TABLE III

CONSTITUENTS OF EAGLES MINIMAL ESSENTIAL MEDIUM (MEM) AS USED IN BIOCHEMISTRY DEPARTMENT, GLASGOW UNIVERSITY

M.E.M. amino acid formulation

	mg/litre
L-arginine	126.40
L-cystine	24.00
L-glutamine	292.00
L-histidine HCl	38.30
L-isoleucine	52.50
L-leucine	52.50
L-lysine	73.10
L-methionine	14.90
L-phenylalanine	33.00
L-threonine	47.60
L-tryptophan	10.20
L-tyrosine	36.20
L-valine	46.90

M.E.M. vitamins

D-calcium pentothenate	2.0
Choline chloride	2.0
folic acid	2.0
i-inositol	4.0
nicotinamide	2.0
pyridoxal HCl	2.0
riboflavin	0.20
thiamine HCl	2.0

TABLE III (Contd)

Formulation of inorganic salts and other components

	mg/litre
$CaCl_2 \cdot 6H_2O$	393.0
KCl	400.0
$MgSO_4 \cdot 7H_2O$	200.0
NaCl	6800.0
$NaH_2PO_4 \cdot 2H_2O$	140.0
glucose	4500.0
$NaHCO_3$	2240.0
phenol red	17.0
penicillin	100,000 units
streptomycin	100mg

1.4. STANDARD SOLUTIONS.

1.4.1. Cell Culture Solutions (Animal lines).

Buffered Saline Solution (BSS) (Earle, 1943). BSS consisted of 116mM NaCl, 5.4mM KCl, 1mM $MgSO_4$, 1mM NaH_2PO_4, 1.8mM $CaCl_2$ and 0.002% (w/v) phenol red. This solution was adjusted to pH 7.0 by the addition of 8.4% (w/v) $NaHCO_3$.

Minimal Essential Medium (MEM) (Busby et al., 1964). The constituents are listed in Table III.

1.4.2. Cell Culture Solutions (Plant line).

Soya bean cotyledon cells were maintained on agar plates (8% agar) in Miller's Medium (1968).

1.4.3. Cell Fractionation Solutions. (Penman, 1969).

Reticulocyte Standard Buffer (RSB). 10mM NaCl, 3mM $MgCl_2$, 10mM tris-HCl, pH 7.4.

Detergent Mixture (Magik). 6.6% v/v Tween 80 and 3.3% w/v Sodium Deoxycholate were dissolved in sterile water and stored at -10°C.

1.4.4. DNA Isolation Solutions. (Hell et al., 1972).

MUP was 8M Urea, 0.24M phosphate buffer at pH 6.8.

1.4.5. Fingerprinting and Sequencing Solutions. (Sanger et al., 1965; Brownlee and Sanger, 1969; Ling, 1972).

(a) Marker dye for all electrophoresis and homochromatography was 1% Xylene cyanol F.F (blue); 2% orange G (yellow); 1% acid fuschin (pink).

(b) Buffer for the First Dimension electrophoresis on cellulose

acetate at pH 3.5 in 7M urea was prepared by dissolving 210.2g of urea and 25ml of glacial acetic acid in distilled water and making up to 500ml after buffering to pH 3.5 by the addition of a small quantity of pyridine.

(c) Second Dimension electrophoresis on DEAE paper, 7% Formic acid was used for this purpose.

(d) Homomix C. Yeast RNA (50g) was divided in 500mls of 1M KOH and stirred at room temperature for 15 mins. The solution was then neutralised to pH 7.5 with concentrated HCl and dialysed against distilled water for ca. 3 hours; urea (420g) was added and the volume made up to 1 litre by the addition of distilled water. The homomix was stored at $-20^{\circ}C$ before use.

(e) Solutions for digestion of DNA to 3' mononucleotides. (Shatkin, 1969).

(i) DNA to be digested at a concentration of 10μmoles/ml in 0.01M tris buffer pH 8.6.

(ii) Calf spleen phosphodiesterase 15-20 units/ml in H_2O.

(iii) Micrococcal nuclease 30 units/ml in 0.1M potassium phosphate pH 7.3.

(iv) Tris buffer 0.4M, pH 8.6, 0.2M $CaCl_2$.

(v) Potassium phosphate buffer, 0.5M, pH 7.0.

(f) Diphenylamine formic acid solution. This was a freshly prepared solution of 3% diphenylamine (w/v) in 98% formic acid.

(g) pH 3.5 Acetate Buffer. This was the same as (b) but did not contain urea.

1.4.6. Gel Electrophoresis. (Hayward, 1972).

Gel Solutions.

'E' buffer:

36mM sodium phosphate buffer, 36mM tris, 0.001M EDTA, pH 7.7.

Agarose Gels:

These were on 1.5% w/v gel of agarose in E buffer.

Buffer for HpaII Restriction Endonuclease. Sharp et al., (1973) (from Haemophilus parainfluenzae).

0.01M Tris, pH 7.4, 0.01M $MgCl_2$, 0.006M KCl, 0.001M Dithiothreitol, 5µg gelatin; total volume 50µl.

1.4.7. S_1 Nuclease digestion. (Sutton, 1971).

(a) S_1 digestion buffer was 0.03M sodium acetate, 3 x 10^{-5}M $ZnSO_4$, 0.01M NaCl pH 4.5 stored at $0^{o}C$ as 10 x stock.

(b) S_1 nuclease (320 units/ml) was stored in a solution of 50% glycerol: 50% S_1 buffer at $-20^{o}C$.

(c) 'Carrier' DNA for S_1 digestion was a 2mg/ml solution of sonicated heat denatured calf thymus DNA in 0.01M tris-HCl buffer pH 8.6.

1.4.8. Scintillation Spectrometry Solutions.

(a) For most experiments a toluene based scintillation fluid was used. This consisted of 5.0g of 2,5 diphenyloxazole (PPO) per litre of toluene.

(b) Triton/toluene scintillation fluid consisted of 5g of PPO plus 0.5g of p-Bis (o-methyl-styryl) benzene (Bis MSB) in 350mls of triton X-114 and 650mls of toluene.

METHODS

1. CELL CULTURE TECHNIQUES.

1.1. Growth of Animal Cells.

HeLa, BHK-21/C13, BHK-21/PyY, L929, SVT2 and Chang liver cells were grown in the Glasgow Modification of Eagle's Minimal Essential Medium to which had been added 10% v/v calf serum (Busby et al., 1964): this medium is designated as EC10. These cell lines were grown as monolayers in Roux bottles at 37°C in an atmosphere containing 5% CO_2.

Friends Erythroleukaemic cells, clone M2, were grown in the Glasgow Modification of Eagles Minimal Essential Medium containing 2 x glutamine, Non essential amino acids, and 15% horse serum. These cells were grown in Roux bottles at 37°C in an atmosphere containing 5% CO_2.

Xenopus liver cells were grown in the same medium as HeLa cells except that the atmosphere was not made 5% CO_2.

The stock cultures were routinely examined for contamination by fungi and yeasts using Sabouraud's medium. Bacterial contamination was monitored with blood agar plates, and contamination by pleuropneumonia like organisms (PPLO) was monitored by use of PPLO agar plates.

Cells grown as monolayers were seeded at a density of approximately 5 x 10^6/Roux in 50mls of EC10 24 hours before labelling, or in the case of Xenopus cells, 48 hours.

Friends Erythroleukaemic cells were seeded at a density of 0.2 x 10^6/ml. Induction of Friends cells was as in McClintock & Papaconstantinou (1974). DMSO spectroscopic grade, non sterile, was added to a final concentration of 1.5% v/v.

Induction of haemoglobin synthesis was followed by the absorbance at 410nm. 5ml aliquots of cell suspension were pelleted at 1000g washed twice in PBS (Dulbecco & Vogt, 1954) lysed in 0.5ml of distilled water, debris was removed by centrifugation at 1000 x g and the supernatants were adjusted to 2% (w/v) in sodium dodecyl sulphate. The absorbance of the lysates was measured at 410nm. After four days growth in the presence of DMSO the A_{410} had increased to $0.4A_{410}/10^6$ cells from $0.1A_{410}/10^6$ cells at day zero, in close agreement with McClintock and Papaconstantinou (1974).

1.2. Labelling of Cells.

With $^{32}PO_4$ (Monolayer cultures).

Cells in Roux bottles were labelled 24 hours after seeding. The normal medium was poured off and fresh medium added; $^{32}PO_4$, normally 2mCi, was added and the cells incubated at 37°C for 48 hrs.

With $^{32}PO_4$ (Friend Cells).

Cells, in log phase of growth, were centrifuged at 1000g for 2 mins and the pellet resuspended in fresh medium with or without DMSO. $^{32}PO_4$ was added to a final concentration of 0.05mCi/ml.

With L(methyl-^{14}C) or L(methyl-^{3}H) methionine.

The technique of Maden et al., (1972) was used. Labelling medium contained 10^{-2}M sodium formate and 2×10^{-5}M adenosine and guanosine. Isotope was added with a sufficient quantity of "cold" methionine to give a final concentration of 5×10^{-5}M (half the normal level of EC10). The normal medium was poured off, the monolayer washed in methionine free medium and the labelling medium (50ml) added, the required amount of isotope

(125µCi of L(methyl-^{14}C) methionine or up to 750µCi L(methyl-^{3}H) methionine) was added and the medium made 5% v/v with freshly thawed dialysed calf serum. Cells were grown for 48 hrs before harvesting.

With deoxy(U-^{14}C) cytidine and (6-^{3}H)-Thymidine.

No adjustments were made to the constituents of the medium. Cells seeded 24 hrs previously were labelled by replacing the medium with fresh ECl0 and the relevant isotope at the required concentration.

1.3. Harvesting of Cells.

Monolayer cultures.

The medium was decanted and the monolayer of cells washed with ice cold BSS, the cells were then scraped off the glass into 10mls BSS using a rubber wiper. Cells were pelleted by centrifugation at 450g for 2 mins at 4°C and washed again by resuspension and centrifugation in BSS.

Friends M2.

Cells were centrifuged directly from the medium and washed twice in BSS at 4°C.

1.4. Growth, Labelling and Harvesting of Plant Cells.

Soya bean cotyledon cells were grown on 8% agar in Miller's medium (1968) in the light at 30 to 33°C. The cells were subcultured every ten days by transferring small clumps of tissue onto fresh agar plates.

L(methyl-^{3}H) methionine and $^{32}PO_4$ labelling: The concentration of isotopes was the same as used for the animal cell cultures, the isotopes were added to the hot agar, containing sodium formate at 0.01M, before pouring into petri dishes (60mm). Soya bean cells

were then subcultured onto the plates.

Harvesting was accomplished by simply removing the growing clumps of tissue from the agar.

2. EXTRACTION OF DNA.

2.1. Preparation of Nuclei.

The technique of Penman (1969) was used. Cells harvested as above were allowed to swell for 5 mins in 4 mls of RSB diluted as follows to facilitate disruption of the cells.

HeLa cells, RSB	=	1 x
Chang cells, RSB	=	½ x
L929 cells, RSB	=	1/3 x
SVT2 cells, RSB	=	1/3 x
PyY cells, RSB	=	1/6 x
C13 cells, RSB	=	1/6 x

After swelling, 0.6 mls of 'magik' detergent was added and the suspension vortexed for 30 seconds. This procedure removes the outer nuclear membrane and causes the nuclei to clump. The suspension was then centrifuged at 450g for 2 mins to recover the nuclei.

All steps were at $4^{\circ}C$.

2.2. Extraction of DNA.

DNA was prepared by the method of Hell, Birnie, Slimming and Paul (1972). Cells or nuclei were suspended in at least 10 volumes of 8M urea, 0.24M sodium phosphate, 1% sodium dodecyl sulphate, 0.01M EDTA, pH 6.8 and then sheared in a sealed filled vessel with a Waring Blender at full speed for 6 periods of 15 seconds alternating with 45 seconds cooling in ice between each

period. The lysate was then added to a thick slurry of HAP (5g per g of tissue used) suspended in MUP and the suspension was left for one hour at room temperature, with occasional mixing. The suspension was then poured into a sintered glass funnel (Gallenkamp Sinta No. 3) and the liquid drawn off under low suction. The HAP was washed with MUP with frequent stirring, to prevent channelling, until no further material was eluted, as judged by A_{260} and A_{280} measurements. Washing was continued with 0.014Msodium phosphate buffer, the removal of urea being monitored by measurement of refractive index. DNA was then eluted with 0.4M phosphate buffer and dialysed against 10-20 volumes of distilled water and precipitated by the addition of NaCl, final concentration 0.2M, and absolute alcohol 2½ volumes. The solution was left overnight at -20°C and the DNA recovered as a precipitate by centrifugation at 10^4g for 20 mins at -20°C.

The DNA product had an A_{260}:A_{280} ratio of 1.8-2.0 and an average single-stranded molecular weight of 1.1 x 10^6 daltons.

2.3. Extraction of Plant Cell DNA.

Plant cell DNA was extracted as for animal cells except that a further stage of purification was used. Ethanol precipitated DNA was dissolved in 1 x SSC and sodium chlorate added to a final concentration of 1M. The solution was extracted with an equal volume of chloroform-isoamyl alcohol (24:1, V/V) for 30 mins. The two phases were separated by centrifugation at 10,000g for 10 mins and the upper aqueous phase removed. DNA was precipitated by addition of two volumes of absolute alcohol (Rubery and Newton, 1973).

3. SHEARING OF DNA AND ESTIMATION OF MOLECULAR WEIGHT.

(Shenkin & Burdon, 1974).

DNA preparations in 0.1 x SSC were sheared by sonication using a Dawe Soniprobe. The glass vial containing 5 mls of DNA solution was held in a precooled metal container, surrounded by melting ice. Twelve successive cycles of sonication for 15 secs at 7 amps followed by 45 secs cooling were performed. The average single stranded molecular weight was determined by one of two methods.

(i) Sedimentation through 0.9M NaCl - 0.1M NaOH in the Model E analytical ultracentrifuge. The observed S value was corrected to $S^{o}_{20,w}$ (the value expected with DNA at infinite dilution in a solution having the density and viscosity of water at $20^{o}C$) by multiplying by 1.16 (Studier, 1965). The molecular weight (M) can then be estimated using the relationship of Studier (1965):-

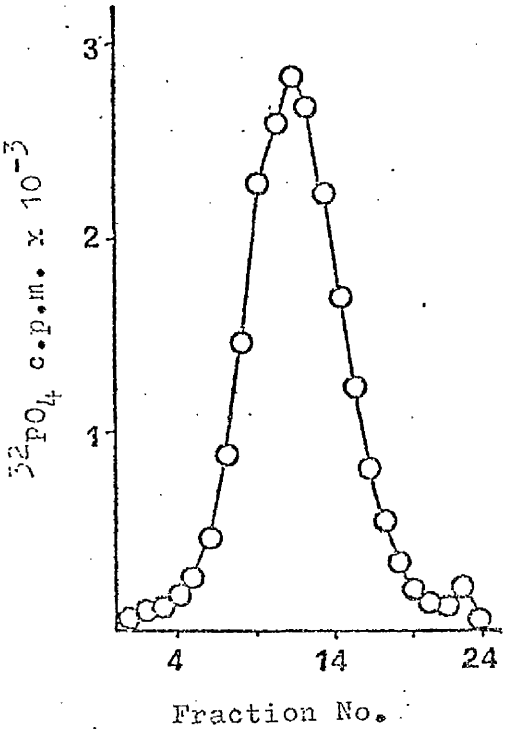
$$S^{o}_{20,w} = 0.0528 . M^{0.400}$$

The validity of this relationship however has not been proven for molecular weights below 10^{6}.

The average single-stranded molecular weight of DNA after sonication is about 6×10^{4} daltons.

(ii) DNA labelled to high specific activity could not be sized on the Model E ultracentrifuge. Instead the use of 5%-20% sucrose gradients was adopted (Burgi and Hershey, 1963). 3.5ml linear gradients of 5% - 20% sucrose in 0.1M NaCl, 0.1M NaOH in polyallomer centrifuge tubes were centrifuged in the SW56 rotor of a Beckman preparative ultracentrifuge at $4^{o}C$ at 45,000 rpm for 21 hours. Fractions were collected dropwise (by piercing the bottom of the

FIG.4



SEDIMENTATION OF DNA ON A 5-20% SUCROSE GRADIENT

$^{32}PO_4$ labelled HeLa cell nuclear DNA (Methods 1.2 & 2.)was sonicated to a single-stranded molecular weight of $6x10^4$ and then centrifuged through a 3.5ml. sucrose gradient(5-20%) at 45K r.p.m. for 21h. at 4^oC in the SW56 rotor of a Beckman ultracentrifuge.After centrifugation fractions were collected dropwise from the bottom of the centrifuge tube and the amount of acid precipitable $^{32}PO_4$ label in each fraction estimated as in Methods section 3.

tube using a syringe needle) onto Whatman 3MM discs; the discs were washed twice in ice cold 5% TCA, twice in methylated spirit, once in ether and then dried. The discs were placed in scintillation vials and the radioactive material solubilised by the addition of 0.5mls hyamine hydroxide (1M solution in methanol) and incubated at 60°C for 20 mins. Toluene PPO scintillation fluid (5mls) was added and the position of the DNA on the gradient was estimated by scintillation counting. The S value and molecular weight of the DNA were then calculated. Fig. 4 shows a 5% - 20% alkaline sucrose gradient of $^{32}PO_4$ labelled, sonicated HeLa DNA.

4. FRACTIONATION OF DNA ON CsCl GRADIENTS.

(a) In the presence of Actinomycin D:

This follows the procedure of Birnstiel et al., (1974). A solution of CsCl (Refractive Index 1.3900 at 3°C) in 0.05M sodium borate pH 9.0 containing 200µg Actinomycin D, total volume 5mls, containing 5 to 15µg of labelled DNA was centrifuged to equilibrium in a T;50 rotor at 35,000rpm for 72 hrs at 2 to 5°C. Fractions were collected, dropwise from the bottom of the tube, onto Whatman 3MM discs. These were washed in ice cold 5% TCA (twice), methylated spirit (twice), ether (once), dried in air: the radioactivity was solubilised with hyamine hydroxide at 60°C, 5.0mls of Toluene PPO added and the radioactivity estimated by scintillation counting.

(b) In the absence of Actinomycin D.

This is essentially a control for the above gradient. The procedure is the same except that the refractive index measurement (1.400) and centrifugation were performed at room temperature.

5. THERMAL DISSOCIATION OF DNA ON HYDROXYLAPATITE. (Bernardi, 1965).

DNA in 0.12M phosphate buffer pH 6.8 was adsorbed onto hydroxylapatite pre-equilibrated with buffer (at least 1g HAP/100µg DNA) at 60°C in a jacketed column. The slurry was stirred occasionally over the course of 2 hours to ensure complete adsorption. The temperature of the column was then increased stepwise and 30 volumes of 0.12M phosphate buffer washed through with stirring at each step. Each fraction was made 10% v/v TCA, 100µg salmon testes DNA added and acid precipitable DNA recovered on Whatman GF/C filters. The filters were washed in 5% TCA, alcohol and ether, dried, solubilised with 0.5mls hyamine hydroxide at 60°C for 20 mins. 5.0mls of Toluene PPO scintillation fluid was added and the radioactivity estimated by scintillation counting. Recovery of DNA was 95 - 100%.

6. DEPURINATION OF DNA. (Ling, 1972).

Ethanol precipitated DNA, 5 - 20µg, prepared as above was redissolved in a small volume of distilled water and desalted by passing over a small (10mls) Sephadex G-25 column. The void volume was collected (ca. 1 - 2mls) in a siliconised test tube. The DNA was lyophilised and then taken up in 20λ of distilled water, 40µl of a fresh solution of diphenylamine 3% w/v in formic acid added. The solution was sealed in a capillary and digested for 18 hrs at 30°C in the dark. Formic acid and water were removed under vacuum. Diphenylamine was removed by extraction with ether.

7. FINGERPRINTING OF DNA. (Sanger et al., 1965 ; Brownlee and Sanger, 1969; Ling, 1972).

7.1. Oligonucleotides from depurinated DNA (prepared as in Methods .6.) were separated by electrophoresis in the first dimension on cellulose acetate at pH 3.5. 2.5 x 95cm strips of cellulose acetate were soaked in 7M urea, pH 3.5 buffer and excess buffer removed from the application area by blotting with pads of tissue paper. The depurinated DNA was applied as a spot 10 - 15cm from one end of the strip and a spot of marker dye (Materials 1.4.5.) was applied at either side. The rest of the cellulose acetate strip was wiped clean of excess buffer before the strip was placed in position in the electrophoresis tank. Electrophoresis was carried out at 4.5KV for ca. 60 mins. After electrophoresis the cellulose acetate strip was allowed to drip to remove excess white spirit. The section 3 - 4cms behind the blue marker dye and 3 - 4cms in front of the yellow dye was then placed 4cms from the short side of a DEAE cellulose thin layer. Three strips of Whatman 3MM paper 2.5cms x 30cms soaked in distilled water were laid over the cellulose acetate strip. Pressure was then applied,using a glass plate, to blot the oligonucleotides from the first dimension onto the second dimension; where they were bound to the DEAE groups of the thin layer. After 20 - 25 mins most of the radioactivity was transferred to the thin layer. The DEAE thin layer sheet was briefly chromatographed with distilled water till the solvent front had travelled about 5cms; the thin layer was transferred to a chromatography tank containing 100ml of homomix (at 60^{o}C). Chromatography was for 16 to 30 hours (depending on the batch of thin layer plates). The ascending solvent front was soaked up by a wad of 2.5cm Whatman 3MM strips attached to the top of the thin layer. Brownlee and Sanger (1969) suggested the use of three different homomixes for separating oligonucleotides of different size ranges. Homomix C (5%) was

FIG.5.

FINGERPRINT OF DEPURINATED $^{32}PO_4$ LABELLED HELA NUCLEAR DNA.

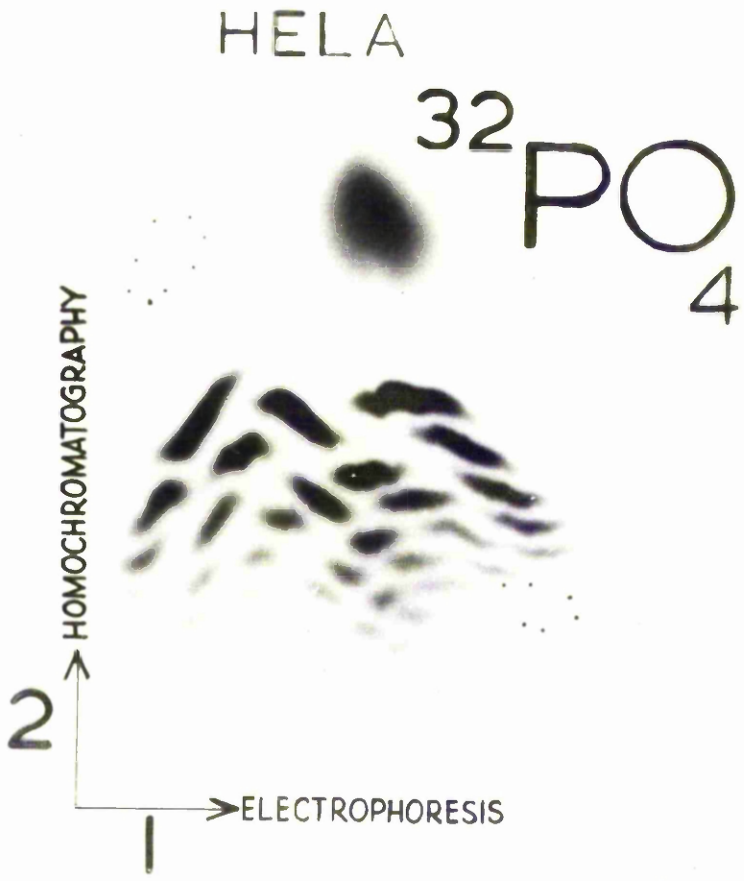
HeLa nuclear DNA labelled with $^{32}PO_4$ for 48h. (Methods 1.2 & 2.)was depurinated (Methods 6.) and fingerprinted by electrophoresis at pH3.5 ($1^{st.}$dimension) and Homochromatography at $60^{o}C$($2^{nd.}$dimension)as in Methods section 7.

(A.) Autoradiograph of depurinated HeLa DNA fingerprint.

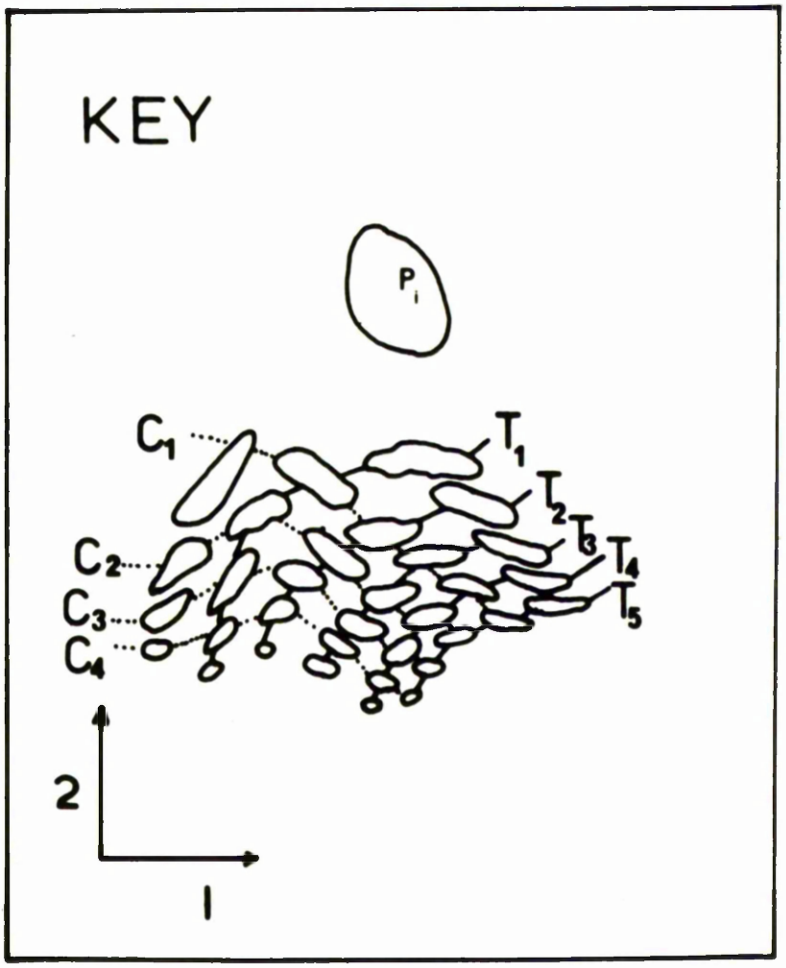
(B.) Key: identification of pyrimidine oligonucleotides in autoradiograph(after Ling,1972).

FIG. 5.

A.


HELA
$^{32}PO_4$
HOMOCHROMATOGRAPHY
2
1
ELECTROPHORESIS

B.


KEY
P_1
C_1
C_2
C_3
C_4
T_1
T_2
T_3
T_4
T_5
2
1

used in this study.

7.2. Autoradiography.

When dry the DEAE thin layer was marked for identification with ^{35}S-sulphate ink and then placed over a sheet X-ray film in a lead-lined folder. These folders were stored (5 to 10 days for $^{32}PO_4$ labelled material and 2 to 4 weeks for ^{14}C labelled material) before development. The films were developed for 4 - 6 mins., washed briefly in water and fixed. Processing was performed in a Kodak P3 X-ray film processing unit. An autoradiograph of a $^{32}PO_4$ labelled HeLa DNA fingerprint is shown in Fig. 5.

7.3. Estimation of Radioactivity.

The appropriate areas of the DEAE thin layers were identified by reference to their autoradiographs and cut out. In the case of $^{32}PO_4$ labelled material these were placed directly into scintillation vials with 10mls of toluene PPO scintillation fluid. For fingerprints employing ^{14}C label or ^{3}H and $^{32}PO_4$ the DEAE thin layer was solubilised as in Methods section 3. Vials were then counted in a scintillation spectrometer using appropriate window settings. In the case of double labelled material (i.e. ^{3}H and $^{32}PO_4$) allowance was made for spillover between windows.

7.4.

7.4.1. CALCULATION OF RELATIVE FREQUENCY OF METHYLATION.

This is simply calculated as c.p.m. (^{3}H or ^{14}C) in an oligonucleotide divided by the c.p.m. in the cytosine monopyrimidine fraction.

7.4.2. CALCULATION OF RELATIVE FREQUENCY OF CYTOSINE IN AN OLIGONUCLEOTIDE.

The depurinated products of DNA have 5' and 3' terminal phosphates. Thus in an oligonucleotide there are N+1 labelled phosphates (in $^{32}PO_4$ labelled DNA), where N = no. of bases in the oligonucleotide, which may be estimated from the position on the homochromatogram (Ling, 1972). As the base composition and length of the oligonucleotide can be obtained from the fingerprint the $^{32}PO_4$ label in cytosine containing mononucleotides is thus $\frac{N_{cyt}}{N_T+1} \times {}^{32}PO_4$ cpm (where N_{cyt} is the number of cytosines in the oligonucleotide of total base number N_T). As it is difficult to estimate the total amount of radiolabel on a fingerprint the relative molar frequencies of cytosine, in any particular isostich, have been calculated taking the cytosine monopyrimidine as 100.

7.4.3. CALCULATION OF RELATIVE FREQUENCY OF OLIGONUCLEOTIDES.

The relative frequency of a $^{32}PO_4$ labelled pyrimidine oligonucleotide may be calculated once allowance for terminal phosphate labelling has been made.

8.1. RESTRICTION OF DNA BY HpaII.

Class II restriction endonucleases cleave DNA at specific sites dictated by the nucleotide sequence. HpaII cleaves DNA at the sequence:

5' C C G↓G 3'
3' G↑G C C 5'

(Nathans and Smith, 1975).

DNA (1µg) was dissolved in the digestion buffer and one unit

of HpaII restriction endonuclease added. This was incubated at 37°C for four hours. (One unit is defined as that amount of enzyme required to digest 1µg of phage λ DNA in one hour). The reaction was stopped by freezing in solid CO_2.

The activity of the enzyme was routinely assayed by digesting a known quantity of SV40 DNA containing both form I (supercoiled) and form II (relaxed circular molecules). HpaII produces a single break in the SV40 DNA converting the form I and form II molecules to form III linear molecules. The mobilities of the three molecules in 1.5% agarose gels are different; Fig. 6.

Gel nos. 3 and 4 contain a mixture of digested and nondigested SV40 DNA to demonstrate the relative mobilities of the different forms of SV40.

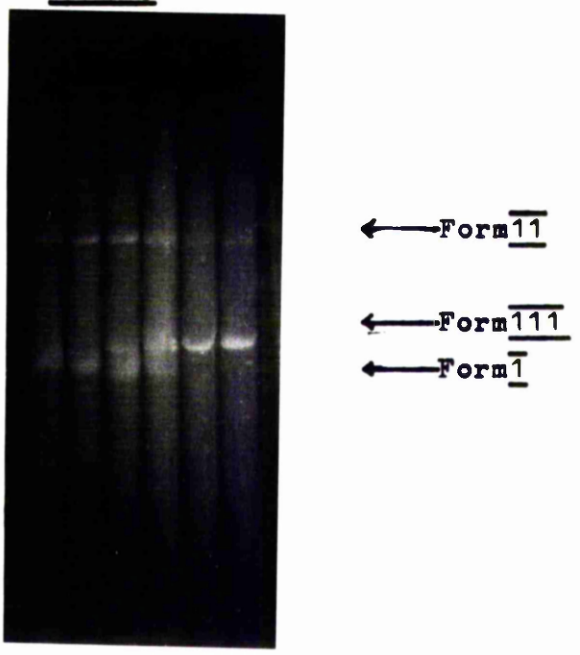
8.2.

Gel Electrophoresis of Restricted DNA:

DNA restricted with HpaII endonuclease, 50µl, was made 1M with respect to sucrose, a small amount of bromophenol blue added, and layered onto 0.6cm x 14cm, 1.5% cylindrical agarose gels in glass tubes. The gels were subjected to electrophoresis at 10 amps per gel for two hours. The gels were removed from the tubes and sliced into 2mm slices using a Mickle Gel Slicer; the slices were placed in scintillation vials and 0.3mls of a 100% H_2O_2 solution added. The vials were then incubated at 200°C for several hours until the gel slices were fully solubilised. Triton/toluene scintillation fluid (10 vols) was then added and the vials counted for radioactivity.

In the case of gels where there is sufficient material the position of the DNA can be visualised by soaking the gels in a

FIG.6.



GEL No. 1 2 3 4 5 6

AGAROSE GEL ELECTROPHORESIS OF SV40 DNA

The photograph shows the separation of the 3 forms of SV40 DNA by agarose gel electrophoresis, the bands of DNA were visualised by staining with ethidium bromide (Methods 8.2).

Gels 1 & 2 contain DNA not restricted by Hpa II restriction endonuclease, i.e. Forms I & II only.

Gels 3&4 contain a mixture of restricted and unrestricted DNA to demonstrate the differing mobilities of SV40 DNA Forms I &III

Gels 5 & 6 contain restricted DNA only(a small amount of Form II is still visible indicating incomplete digestion to Form III in this instance).

Conditions of restriction and electrophoresis are described in Methods 8.1 & 8.2 .

solution of 0.5µg/ml ethidium bromide (in gel buffer) for 30 mins. The DNA can then be visualised using a U.V. lamp. Using this procedure as little as 0.5µg of DNA may be detected (Sharp et al., 1973).

9. BASE COMPOSITION ANALYSIS OF DNA.

9.1. By Enzymatic Digestion (Shatkin, 1969).

(a) Digestion.

30µl of labelled DNA, 5µl of $CaCl_2$-tris buffer, and 5µl of micrococcal nuclease were mixed in a siliconised 3ml test tube: this was incubated for two hours at 37°C. At the end of this period 20µl of potassium phosphate buffer was added to decrease the pH to 7.0. The digestion was continued by adding 10µl of spleen phosphodiesterase and incubating for 1 hour at 37°C, another 10µl of spleen phosphodiesterase was added and the solution was reincubated. This was repeated (i.e. a total of 3 x 10µl of spleen phosphodiesterase was added). Further aliquots of spleen phosphodiesterase were sometimes necessary to complete digestion.

(b) Separation of nucleotides.

At the end of the incubation the DNA digest was applied to sheets of Whatman No. 52 paper (43 x 57cm) about 10cm from one end, in 1.5cm bands. The paper was wetted with pH 3.5 acetate buffer (Materials section 1.4.5.(b)), placed in an electrophoresis tank and electrophoresed at 4.5KV for 45 mins. The sheet was then removed and dried in air. After completely removing all traces of pyridine the sheet was autoradiographed and processed as in Methods section 7.

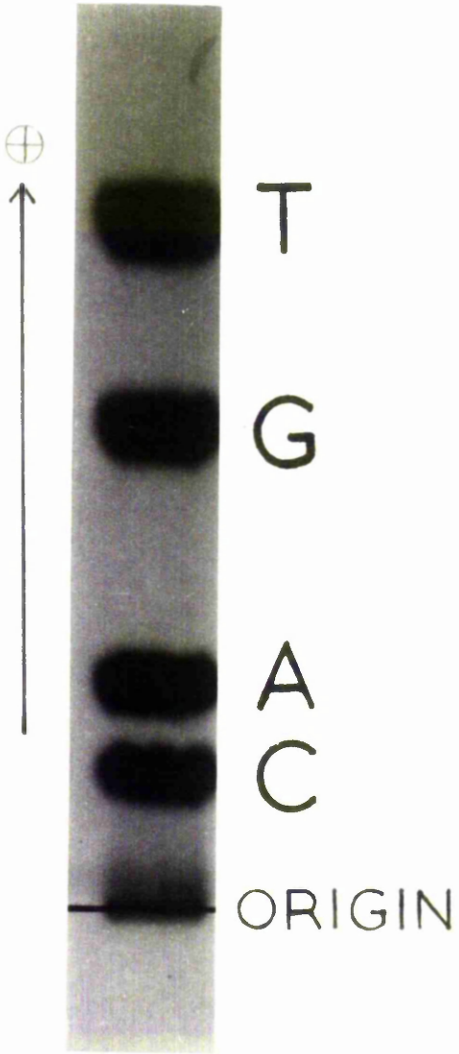


FIG.7

The autoradiograph shows the electrophoretic separation of $^{32}PO_4$ labelled mononucleotides (see text of Methods 9.1.)

Base Composition of BHK-21/PyY DNA.

Base	% of Total $^{32}PO_4$ Label/mononucleotide
T	28·9 ± 1·7
G	19·5 ± 2·4
A	30·5 ± 1·7
C	20·9 ± 2·3

Figure 7 shows an autoradiograph of the separated $^{32}PO_4$ labelled nucleotides from nuclear DNA of BHK-21/PyY cells (labelled as Methods 1.2.) The table alongside gives the nucleotide composition of the nuclear DNA from these cells.

9.2. By Acid Hydrolysis.

This was performed by the method of Adams (1973). Lyophilised DNA samples were taken up in 12N $HCLO_4$ and after incubation at 100°C for 60 mins the perchlorate was removed as a precipitate of $HCLO_4$, the supernatant was neutralised and portions applied as spots to sheets of Whatman No. 1 chromatography paper (thymidine, cytosine, 5 methyl-cytosine, adenine and guanosine, 50µg of each, were applied as markers) and subjected to descending chromatography in n-butanol-water-NH_3 (87:13:1) or isopropanol-HCl-H_2O (68:16.4:15.6) (Rubery and Newton, 1973) the position of the bases being determined by use of a U.V. lamp.

10.S_1 NUCLEASE DIGESTION OF DNA.(Sutton, 1971)

Quench cooled heat denatured radiolabelled DNA in 0.1 x SSC (0.2mls) was digested as follows:

0.2mls heat denatured DNA

0.005mls 2mg/ml carrier DNA

0.050mls S_1 x 10 buffer (stock)

0.120mls H_2O

0.125mls S_1 nuclease

0.5mls total volume

The digestion was performed at 37°C. After completion, 0.1mls of a 50% solution of trichloroacetic acid (TCA) was added plus 0.1mls of a 2mg/ml solution of salmon testes DNA; acid precipitable DNA was

retained on Whatman GF/C filters and the radioactivity estimated as in Methods, section 3.

EXPERIMENTAL

1. DNA METHYLATION IN A SYSTEM DIFFERENTIATING IN VITRO

Section 5 of the Introduction outlined the case for involvement of DNA methylation in the control of differentiation; in brief, there appears to be a correlation between methylation and development. It was of interest then to examine the methylation of DNA in a well defined differentiating system, in particular one which would allow isotopic labelling experiments. The Friend virus transformed leukaemic mouse cell line clone M2 (Birnie et al., 1974) was considered an ideal candidate for such an investigation. The induction of differentiation by DMSO is easily controllable and requires two cell generations before induction of haemoglobin synthesis can be detected (McClintock and Papaconstantinou, 1974). This suggests that some event at the level of DNA is involved. The possibility that this 'event' is a change in the level of 5 MeC in these cells was investigated as below.

1.1. Methylation of Parental DNA Strands.

Friend M2 cells in mid log phase of growth were subcultured into fresh medium at a density of 0.2×10^6 cells/ml. Deoxy (U-^{14}C) cytidine was added to a final concentration of 1μCi/ml and the cells allowed to grow for 24 hrs (approximately one generation). Each culture was then split into 3 aliquots: Aliquot A; the cells were pelleted by centrifugation (Methods, section 1.3) and stored at -20°C; Aliquot B, the cells were pelleted by centrifugation, the medium poured off and the cells resuspended in fresh unlabelled medium at a density of 0.2×10^6 cells/ml; Aliquot C was treated as for B except DMSO was added to the medium at a final concentration of 1.5% (v/v). Aliquots B and C were incubated for 48 hrs.

After this period the cell density of aliquot B had increased to 1.2 x 10^6/ml (i.e. 2 to 3 generations had elapsed), the cell density of aliquot C had increased to 0.8 x 10^6/ml (2 generations), Fig. 8. The cells were pelleted by centrifugation (Methds, 1.3) and total cellular DNA was extracted from all three aliquots, hydrolysed in 12N perchloric acid, and chromatographed to estimate the relative amount of 5 MeC in each (Methods, 9.2). Fig. 9 shows the separation of 5 MeC from C, Table IV (a) lists the values obtained for each aliquot.

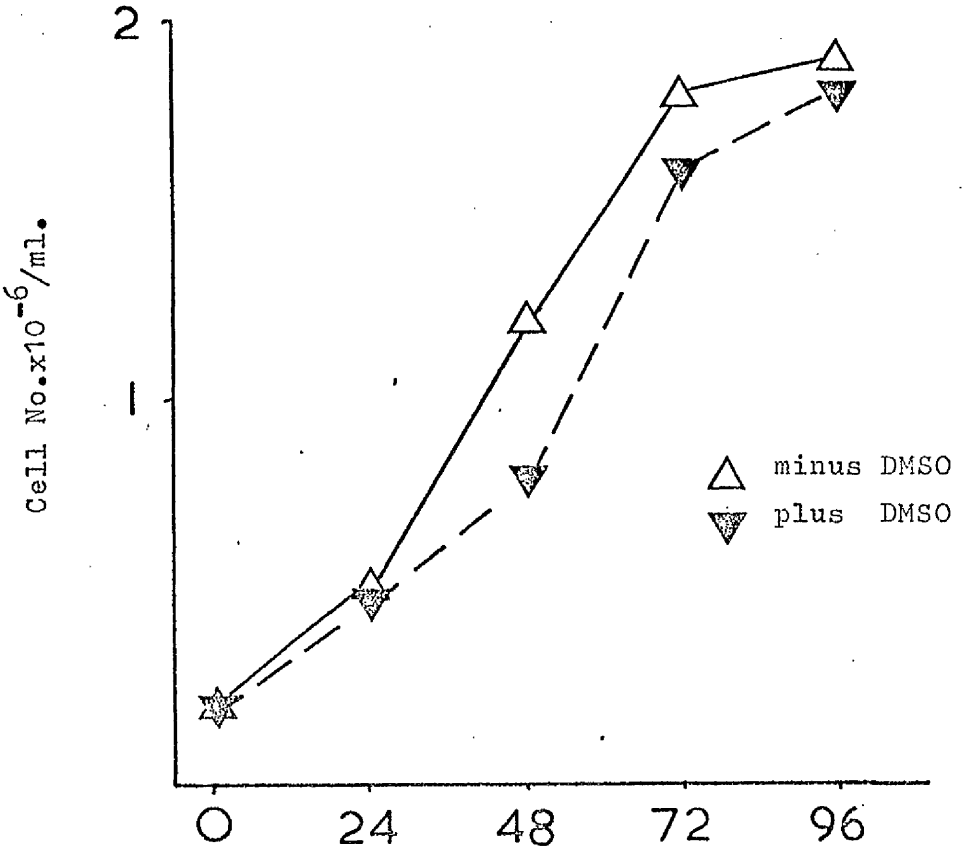
^{14}C label was observed in areas of the chromatogram which did not correspond to C or 5 MeC U.V. markers. No attempt was made to further characterise them; similar results have been obtained before (Rubery and Newton, 1973).

1.2. Methylation of Progeny DNA Strands.

The results displayed in Table IV (a) illustrate one acute problem associated with the determination of the level of 5 MeC in DNA, i.e. that of accuracy and reproducibility. Obviously in an experiment of this kind accuracy is of the utmost importance. Another chromatographic system was sought: Tamm et al., (1953) report the complete separation of d5MeCMP from dCMP (mobilities: d5MeCMP, 1.26; dCMP, 0.74: relative to dTMP as 1.00) using a solvent of isobutyric acid (10 vols) NH_3 0.5N (6 vols) pH 3.6-3.7. Use of nucledides in the analysis rather than nucleosides or bases permits labelling of d5MeCMP with $^{32}PO_4$, this allows a higher specific activity to be attained more conveniently.

The chromatographic system was examined using dCMP and d5MeCMP (50µg of each) as markers, their position being determined (after running in the isobutyric acid/NH_3 solvent on Whatman No. 1 paper

FIG.8



Growth Curve of Friend (clone M2) Cells.

Cells were seeded at a density of 0.2x10^{6}/ml. in medium either with or without DMSO(1.5%, w/v.)and the cell density monitored(duplicate samples)using a haemocytometer(Methods 1.1)

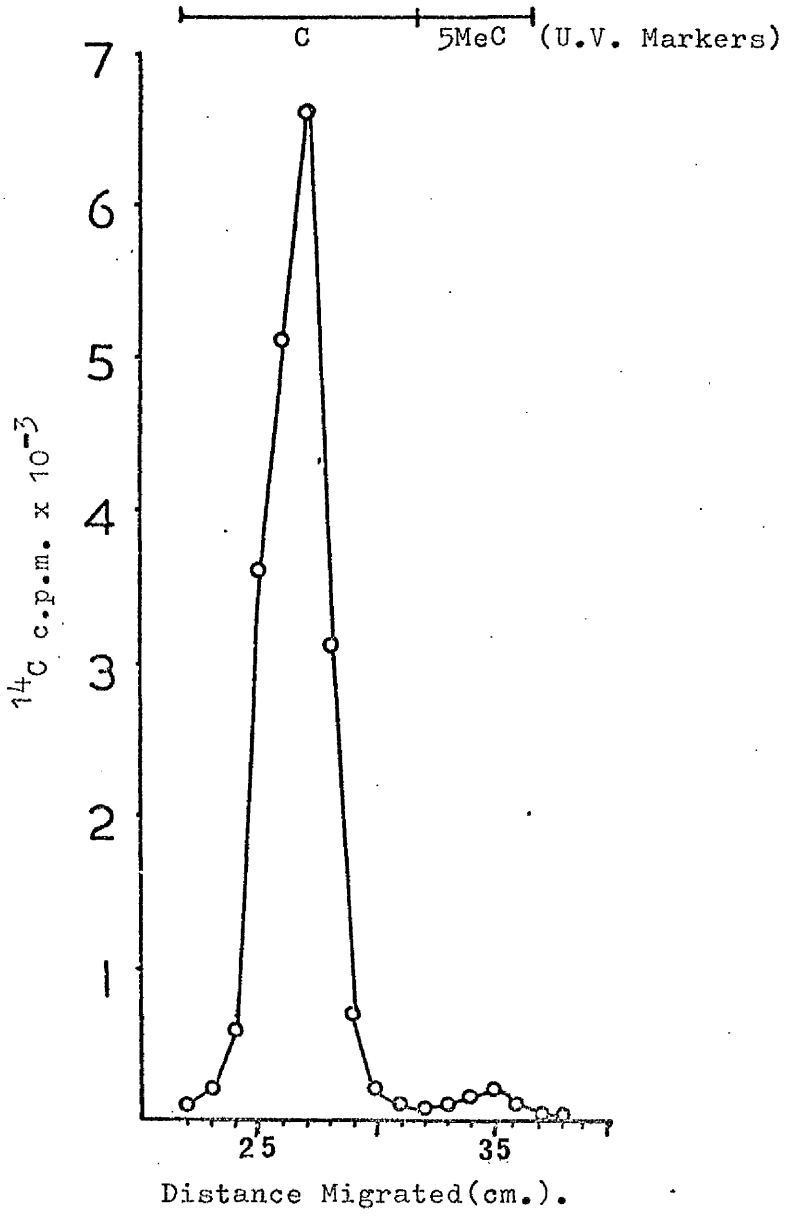


FIG.9

THE SEPARATION OF CYTOSINE & 5-METHYL CYTOSINE.

Friend cell(clone M2) DNA,labelled with deoxy(U-^{14}C) cytidine for 24h.,was hydrolysed to the bases(Methods 9.2) and the bases chromatographed on Whatman No.1 chromatography paper .50µg of each of the bases(thymine,cytosine, guanine, adenine & 5-methyl cytosine) were used as markers,their position being determined by use of a U.V. lamp.The chromatogram was then cut into 1cm. strips and the total amount of radioactivity in each strip estimated as in Methods 7.3

for 24 hrs) by U.V. absorbtion. Under the conditions used the actual R_F of the d5MeCMP was only some 10% greater than dCMP itself. Resolution of d5MeCMP could be improved by rerunning the chromatogram after turning through 90^{o}. Figure 10 shows the separation of $^{32}PO_4$ labelled d5MeCMP from dCMP. (These nucleotides were prepared as in Methods, section 9.1).

The identity of the minor spot was further confirmed by chromatographing $^{32}PO_4$ labelled dCMP and $^{32}PO_4$ and L(methyl-^{3}H) methionine labelled d5MeCMP obtained from an electrophoretic separation of the nucleotides (Methods, section 9.1.) in one dimension only. The chromatogram was cut into 1cm strips, solubilised with hyamine, toluene PPO added (Methods, 7.3.) and the radioactivity estimates by scintillation counting. The distribution of ^{3}H and ^{32}P is shown in Figure 11. The separation is similar to that of the U.V. markers (above).

The deoxy (U-^{14}C) cytidine labelling experiment is not suitable for determining the level of 5 MeC in progeny DNA strands as the ^{14}C label remains in the parental DNA. To rule out the possibility that methylation is modulated in progeny strands alone Friends M2 cells grown for 2 generations with or without DMSO (1.5% v/v) and in mod log phase of growth were subcultured into fresh medium at a density of 0.8×10^{6} cells/ml, $^{32}PO_4$ was added to a final concentration of 0.05mCi/ml and the cells incubated for 24 hours. The cells were harvested by centrifugation (Methods, section 1.3.), the total cellular DNA extracted, digested to mononucleotides with micrococcal nuclease and spleen phosphodiesterase (Materials and Methods, section 9.1.) and the relative amount of 5 MeC determined by chromatography in isobutyric acid/NH_3. The results are shown in Table IV (b).

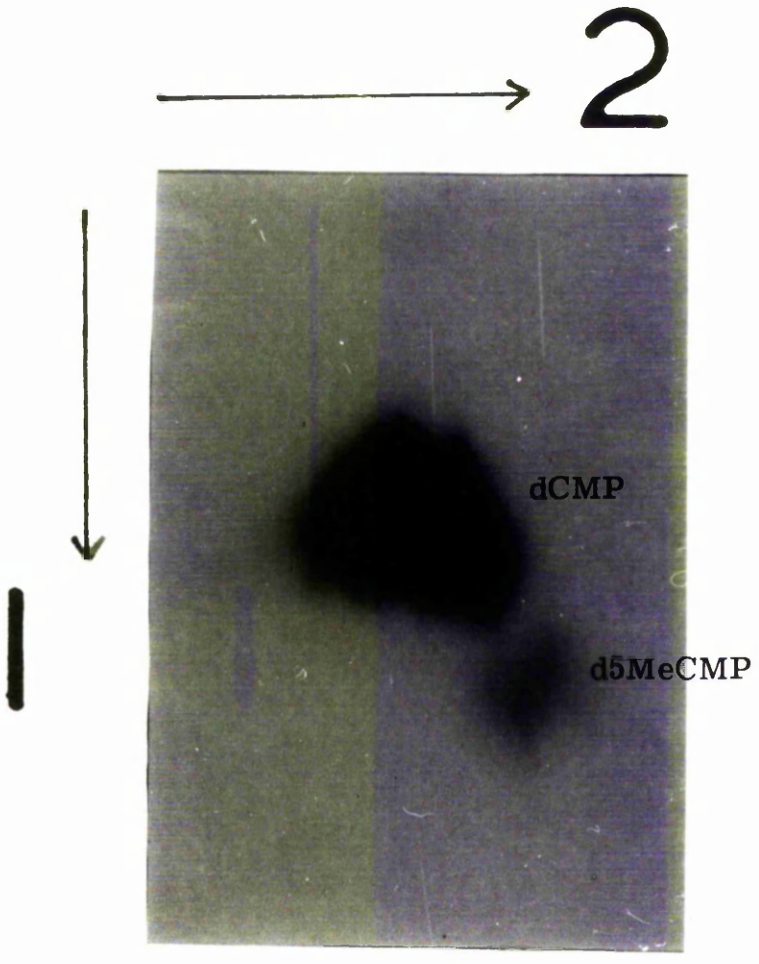
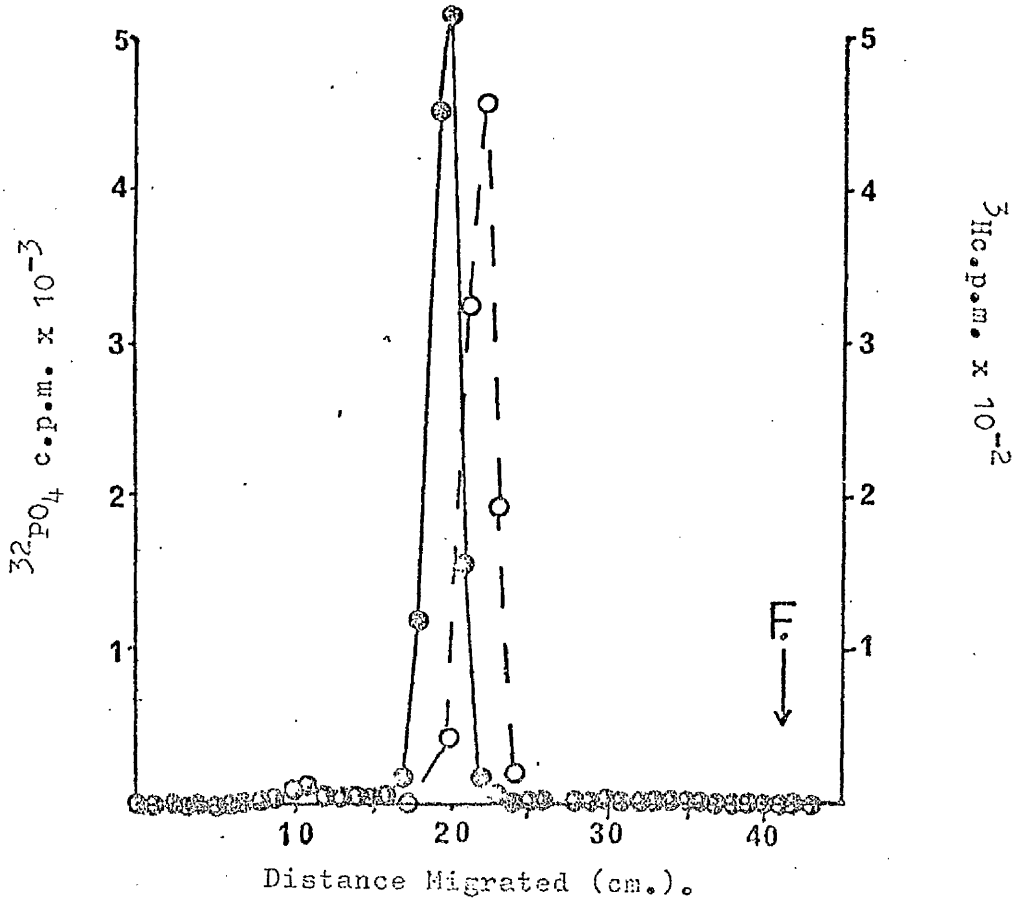


FIG.10.

THE SEPARATION OF $^{32}PO_4$ LABELLED dCMP & d5 MeCMP (AUTORADIOGRAPH).

$^{32}PO_4$ labelled mononucleotides, obtained by enzymatic digestion of DNA(Methods 9.1), were electrophoresed at pH3.5 and the area corresponding to dCMP, FIG.7, excised and sewn to a sheet of Whatman No.1 chromatography paper (46 x 57 cm.). Chromatography, in a solvent of isobutyric acid/0.5M NH_3 (10/6, v./v.) was for 30h. in both dimensions. Autoradiography was performed as in Methods 7.2.

FIG.11.



The Separation of $^{32}PO_4$ labelled dCMP & L-(methyl-^{3}H) labelled d5 MeCMP.

Samples of $^{32}PO_4$ & L-(methyl-^{3}H)methionine labelled DNA were combined & digested to the mononucleotides, after electrophoresis at pH3.5(Methods 9.1) the area corresponding to dCMP was excised & sewn to a sheet of Whatman No.1 chromatography paper (46x57cm.). Chromatography in a solvent of isobutyric acid/0.5M NH_3(10/6,v/v)was for 20h.(to ensure the front remained on the sheet).Routine separations e.g. FIG.10 used runs of ca.30h.,thus improving resolution.

O——O L-(methyl-^{3}H)

O——O $^{32}PO_4$

F.=Solvent front.

The Level of 5 MeC in Induced and Uninduced Friend Cells.

TABLE IV (a)

Deoxy (U-^{14}C) Cytidine Labelling.

Cell State	% 5 MeC/C + 5 MeC	No. Determinations
Aliquot A, 24 hrs. growth in deoxy (U-^{14}C) cytidine	2.3 ± 0.6	(3)
Aliquot B, 48 hrs, growth after labelling, no DMSO	2.3 ± 0.3	(3)
Aliquot C, 48 hrs. growth after labelling, with DMSO	2.4 ± 1.5	(3)

TABLE IV (b)

$^{32}PO_4$ Labelling.

Cell State	% 5 MeC/C + 5 MeC	No. Determinations
Cells grown for 2 generations without DMSO, 24 hrs. in $^{32}PO_4$ labelling medium	1.84 ± 0.2	(7)
Cells grown for 2 generations with DMSO, 24 hrs. in $^{32}PO_4$ labelling medium	1.85 ± 0.5	(6)

In both experiments, above, no detectable change was noted in the level of DNA methylation either in the presence or absence of DMSO. The absence of change in methylation of parental DNA, measured by the ^{14}C label in d 5 MeC, after 2 generations indicates that (i) all methylation is completed soon after incorporation of the deoxycytidine into DNA (ii) addition of DMSO and subsequent "differentiation" has no apparent effect on the level of 5 MeC. The fact that the level of methylation in the progeny DNA strands of induced or uninduced cells is the same rules out the possibility that changed in methylation are manifested in progeny DNA. We can therefore conclude that these experiments rule out any connection between methylation and differentiation and as a possible corollary, transcription. This conclusion may be facile because (i) the experiments do not exclude the possibility that methyl groups are rearranged or removed and replaced by new methyl groups (ii) changes in methylation may well be too small to be detected by the systems used.

Acknowledging that the possibility of changes below the level of detectability are occurring (the errors shown in Table IV are considerable) an alternative and more sensitive analytical procedure(s) is required. The following sections describe other approaches to the problem of defining the role of DNA methylation.

2. ANALYSIS OF THE TISSUE AND SPECIES SPECIFICITY OF 5 MeC DISTRIBUTION IN PYRIMIDINE TRACTS METHYLATED IN VIVO.

The specific production of pyrimidine tracts from DNA has been invaluable in sequencing short regions of DNA (Southern, 1970) and for characterising features of the genome such as the distribution of methylated bases (Introduction, section 3.4.1.).

Most previous work on the distribution of 5 MeC has either used a one dimensional separation, to resolve isostichs only, or separation by base composition has also been employed, and has necessitated the use of time-consuming cumbersome techniques. A combination of depurination and the 2 dimensional resolution of the products by electrophoresis (Murray, 1970) or electrophoresis and homochromatography (Ling, 1972) seemed to offer the most favourable approach.

The lack of resolution amongst the longer oligonucleotides using 2D electrophoresis makes analysis difficult. For this reason it was decided to use the methodology of electrophoresis and homochromatography.

A preliminary study using HeLa nuclear DNA labelled with $^{32}PO_4$ or specifically in 5 MeC using L(methyl-^{14}C) methionine (Methods, 1.2.) was performed.

2.1. Homochromatography of Depurinated HeLa DNA.

Autoradiographs of homochromatograms of $^{32}PO_4$ and L-(methyl-^{14}C) methionine labelled HeLa nuclear DNA, labelled, extracted and fingerprinted as in Methods, 1.2.2. and 7.1. are shown in Figs. 5 and 12. The amount of label appearing in the fingerprints after losses due to digestion, extraction and transfer was about 50% - 60% of the starting material.

Fig. 12 shows the pattern of labelled obtained by using DNA labelled in 5 MeC only. It should be noted that there is little or no label in the area corresponding to Tp_2, this indicates that the labelling procedure is highly specific. Further confirmation of the specificity of labelling was provided by base analysis of L(methyl-^{14}C) methionine labelled HeLa DNA using perchloric acid

FIG.12.

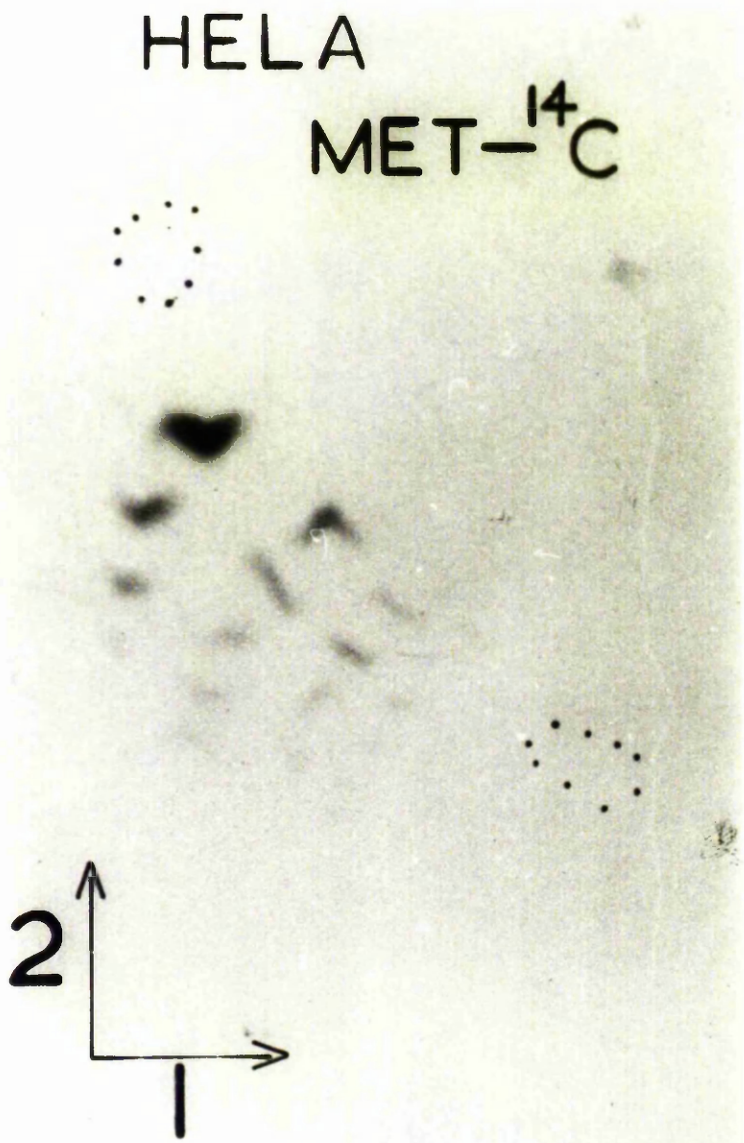
FINGERPRINTS OF L-(METHYL-^{14}C)METHIONINE LABELLED DNA.

Nuclear DNA ,labelled with L-(methyl-^{14}C) methionine,under conditions where 95% of(^{14}C) appears in 5 MeC(Methods 1.2 &2.;& Table 5),was depurinated(Methods 6) and fingerprinted by electrophoresis at pH3.5($1^{st.}$dimension)and Homochromatography at 60^{o}C ($2^{nd.}$dimension) as in Methods section 7.

(A.) Fingerprint of HeLa DNA produced as above.

(B.) Fingerprint of BHK-21/C13 DNA produced as above.

(C.) Key to HeLa DNA fingerprint,identification of oligonucleotides was by reference to Ling(1972) and to the results obtained by fingerprinting DNA labelled with L-(methyl-^{3}H)methionine & $^{32}PO_4$ (Tables V1 & V11) .



A.

B.

FIG.12 CONT'D.

Key to HeLa L(methyl-^{14}C)
Fingerprint.

C.

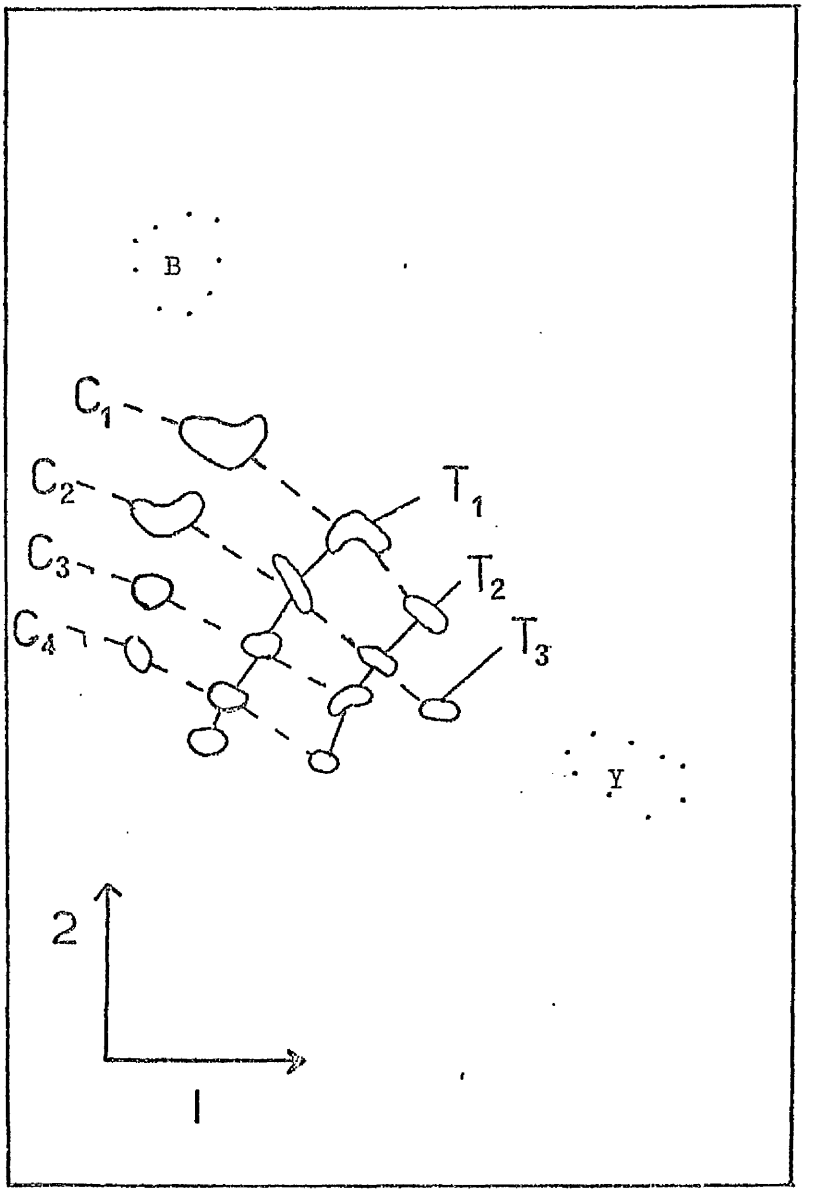


TABLE V.

SPECIFICITY OF METHYL LABELLING OF 5 MeC.

Chromatography Solvent:	Base	% of total ^{14}C in base
n-butanol-H_2O-NH_3	G	2.6
	C	1.6
	5 MeC	92.7
	A	1.6
	T	1.5
label in C + 5 MeC = 94.3%		
isopropanol-HCl-H_2O	G	0.9
	C	8.6
	5 MeC	88.0
	A	2.1
	T	0.5
label in C + 5 MeC = 96.6%		

digestion followed by chromatography (Methods, section 9.2.). The distribution of label is shown in Table V. Ca. 95% of the ^{14}C label appears in C + 5 MeC.

Fig. 12 shows that there are no very highly methylated sequences present in low molar amounts, e.g. C_{10} T_{10}. This is important in the design of most of the following experiments.

There are two main disadvantages in using L-(methyl-^{14}C) methionine to label 5 MeC, namely cost and low specific activity. L-(methyl-^{3}H) methionine on the other hand is much cheaper and of far higher specific activity: however the presence of ^{3}H cannot easily be detected by autoradiography; fluorography (Harbers et al., 1975), could be used but a combination of $^{32}PO_4$ and L-(methyl-^{3}H) methionine labelled DNA allows detection of oligonucleotides not containing 5 MeC; provides a check for preferential loss of oligonucleotides during ether extraction or transfer from the 1st to the 2nd dimension of fingerprinting and makes a separate set of experiments to determine the relative occurence of the oligonucleotides unnecessary. This procedure would not, however, detect long oligonucleotides of low molar frequency which were highly methylated. Most of the following experiments use nuclear DNA extracted from cells labelled with $^{32}PO_4$ and L-(methyl-^{3}H) methionine.

The results of 5 separate HeLa nuclear DNA preparations are shown in Table VI. Variation in results between DNA preparations was no different to the variation between chromatograms from the same preparation; differences in growth conditions and cell density thus have no noticeable effect on the pattern of methylation.

The use of this technique allows us to examine the tissue and species specificity of DNA methylation in a far more subtle way than measuring the mol % of 5 MeC alone. In the process of cellular

TABLE VI (Contd)

Random distribution calculated from Peterson (1962).

The results for HeLa nuclear DNA are based on experiments using 2 L(methyl-^{14}C) methionine labelled cultures and 3 L(methyl-^{3}H) methionine plus $^{32}PO_4$ labelled cultures. The results for Chang liver cells are based on one L(methyl-^{3}H) methionine plus $^{32}PO_4$ labelled culture.

Figures in parenthesis are values for standard deviations.

TABLE VI

DEPURINATED HUMAN NUCLEAR DNA FINGERPRINT DATA

Oligo-Nucleotide Base Composition	HELA Rd.Freq.of * Methylation	CHANG Rd.Freq.of Methylation	CYTOSINE Rel.Molar† Frequency	CYTOSINE Random Distribution
C	100 (-)	100	100 (-)	100
C_2	27.6 (2.9)	27.7	44.3 (6.6)	42.8
C_3	9.3 (2.1)	8.2	15.3 (4.2)	13.8
C_4	2,4 (0.8)	2.3	5.4 (1.3)	3.9
CT	22.6 (3.0)	23	73.3 (7.0)	56.6
C_2T	10.2 (3.2)	10	50.0 (1.0)	36.4
C_3T	6.8 (2.6)	6.5	27.8 (2.6)	15.6
C_4T	2.5 (0.7)	2.6	7.8 (4.0)	-
C_5T	1.8 (0.3)	1.3	5.6 (1.0)	-
CT_2	5.3 (1.6)	5.6	30.2 (4.1)	24.0
C_2T_2	6.4 (2.1)	8.4	30.7 (4.4)	20.6
C_3T_2	2.1 (0.7)	2.8	16.2 (1.0)	11.0
C_4T_2	1.0 (0.5)	1.3	8.3 (0.5)	4.7
CT_3	2.3 (1.9)	1.8	14.2 (1.5)	9.1
C_2T_3	1.7 (0.5)	1.4	14.9 (1.0)	9.7
C_3T_3	1.5 (0.1)	0.9	9.9 (0.5)	6.3
C_4T_3	1.1 (0.1)	0.6	6.4 (1.0)	3.1
CT_4	0.7 (0.1)	0.6	6.1 (1.0)	11.4
C_2T_4	0.8 (0.1)	-	7.0 (1.0)	4.1
T	0.6 (0.5)	0.2	-	-

* and † calculated as in Methods, section 7.4.

differentiation one could imagine the operation of a number of DNA methylases (Rubery and Newton, 1973; Holliday and Pugh, 1975; Riggs, 1975) with differing substrate specificities. Their activity can potentially be detected by differences in the methylation pattern of a fingerprint.

2.2. Tissue Specificity of DNA Methylation.

In performing this comparison it was not possible to use DNAs derived from different tissues. Instead a model system consisting of two cell lines was adopted: HeLa, a human cervical epithelial cell line and Chang, a human liver cell line. It was assumed that the two cell types would be representative of their tissue of origin. Nuclear DNA labelled with $^{32}PO_4$ and L-(methyl-^{3}H) methionine (Methods, section 1.2.) was depurinated and fingerprinted (Methods, 7.2.); the data are presented in Table VI. It is evident that the Chang 5 MeC distribution is idential to that of HeLa.

During the course of this experiment it was learned that the Chang liver cell line possesses many of the characteristics of the HeLa cell line (Nelson-Rees and Flandermeyer, 1976); thus it is not certain whether the Chang liver cells are in fact of a different origin and the value of this model inter-tissue comparison is difficult to assess. As a result, use of this system was discontinued. On the other hand, Chang liver cells are morphologically distinct from HeLa cells and can therefore be considered to have 'differentiated' to some extent. Assuming this to be the case the fact that there is no difference in methylation pattern suggests that the distribution of 5 MeC is not related to differentiation.

2.3. SPECIES SPECIFICITY OF 5 MeC DISTRIBUTION IN PYRIMIDINE TRACTS

The 5 MeC content of DNA is tissue and species specific (Introduction, section 6). The reported mole % variation between species is on the whole larger than that between tissues (Kappler, 1971; Vanyushin et al., 1970; 1973) and it would seem likely that if any variation in 5 MeC distribution does occur it would be detected more easily using an interspecific comparison.

The following study was therefore performed.

2.3.1. Comparison between Mammalian Cell Lines.

Two cell lines were used: mouse L929 fibroblast cells and hamster BHK-21/C13 cells. L929 cells were labelled for 48 hours with $^{32}PO_4$ and L-(methyl-^{3}H) methionine, and the BHK-21/C13 cells with $^{32}PO_4$ and L-(methyl-^{3}H) methionine or with L-(methyl-^{14}C) methionine alone (Methods, section 1.2.). The nuclear DNA was extracted, depurinated and fingerprinted as before. The results are displayed in Table VII and Fig. 12(b).

The data show a very similar distribution of 5 MeC throughout the oligonucleotides. There do appear to be some differences between species, e.g. in mouse L929 DNA the C_2, C_2T and C_2T_2 oligonucleotides are relatively less methylated than the corresponding oligonucleotides in HeLa, Chang or BHK-21/C13 DNA. This undermethylation of some oligonucleotides in L929 DNA appears to be consistent with the results of Harbers et al., (1975) who found that these oligonucleotides were not methylated in L929 satellite DNA. However the sample size and the errors involved make unambiguous interpretation of the results difficult. The fact that the 5 MeC distributions are so similar is somewhat surprising in view of the reported variation in level of 5 MeC in these cells (Adams, 1973; Rubery and Newton, 1973; Dawid, 1974).

TABLE VII (Contd)

The results for the L929 nuclear DNA are based on 2 L-(methyl-^{3}H) methionine and $^{32}PO_4$ labelled cultures, those for BHK-21/C13 are based on 3 L-(methyl-^{3}H) methionine and $^{32}PO_4$ labelled cultures and one L-(methyl-^{14}C) methionine, only, labelled culture.

R.M.F. (Relative Molar Frequency).

Calculations as in Table 6.

TABLE VII

DEPURINATED L929 AND BHK-21/C13 CELL NUCLEAR DNA

FINGERPRINT DATA.

Oligo-nucleotide Base Composition	L929		BHK-21/C13	
	Relative Frequency of Methylation	Cytosine (R.M.F.)	Relative Frequency of Methylation	Cytosine (R.M.F.)
C	100 (-)	100 (-)	100 (-)	100 (-)
C_2	14.8 (3.3)	42.4 (7.4)	28.4 (6.0)	49.6 (11)
C_3	2.7 (0.6)	12.5 (2.1)	6.9 (3.1)	22.7 (7.8)
C_4	1.1 (0.3)	3.7 (2.2)	3.3 (1.8)	13.4 (7.0)
CT	16.5 (3.0)	61.3 (9.5)	25.6 (3.1)	73.6 (6.4)
C_2T	6.5 (0.7)	38.0 (5.3)	13.0 (0.8)	39.1 (6.2)
C_3T	3.0 (0.5)	13.2 (3.9)	7.0 (4.7)	25.9 (8.4)
C_4T	1.1 (0.4)	5.0 (0.7)	-	-
C_5T	0.6 (-)	1.3 (-)	-	-
CT_2	4.6 (1.8)	26.5 (5.0)	8.4 (2.4)	29.2 (4.3)
C_2T_2	3.3 (1.0)	24.4 (6.4)	7.5 (3.1)	34.9 (9.0)
C_3T_2	1.9 (0.9)	12.5 (2.3)	2.8 (1.8)	20.1 (3.3)
C_4T_2	0.6 (0.2)	4.2 (0.2)	-	-
CT_3	4.2 (0.4)	12.9 (1.4)	2.6 (0.4)	18.4 (-)
C_2T_3	2.1 (0.6)	12.7 (1.4)	3.4 (0.6)	13.7 (-)
C_3T_3	1.3 (0.6)	5.2 (3.5)	2.1 (-)	-
CT_4	0.1 (0.1)	5.9 (0.6)	-	-
C_2T_4	0.2 (0.2)	9.4 (0.2)	-	-
T	0.6 (0.4)	-	1.3 (1.3)	-

2.3.2. Comparison with an Amphibian Cell Line.

Analysis of the 5 MeC distribution in a number of mammalian cell lines has shown the distribution to be very similar. As a result of this finding it was decided to examine the distribution of 5 MeC in an amphibian cell line. As amphibians are evolutionarily distant from mammals such a comparison should reveal whether the pattern of methylation has been conserved or has only recently evolved.

Cultured Xenopus liver cells were labelled with $^{32}PO_4$ and L-(methyl-^{3}H) methionine and the total cellular DNA extracted (as difficulty was experienced in trying to prepare nuclei). The results of fingerprints from two cultures are shown in Table VIII.

The results, while showing some degree of variation from the mammalian 5 MeC distribution are remarkably similar to the mammalian pattern.

2.4. THE EFFECT OF VIRUS TRANSFORMATION ON THE DISTRIBUTION OF 5 MeC IN PYRIMIDINE TRACTS.

Nass (1973) and Rùbery and Newton (1973) observed an increase in the 5 MeC content of DNA in the polyoma transformed cell line BHK-21/PyY compared to the parent strain BHK-21/C13. (Our own cell stocks of the same cell lines also have this characteristic, Table XI). They suggested that this increase may be due to the presence of a new methylase(s) either coded by the virus or induced by the presence of the viral genome. Comparison between 'normal' and virally 'transformed' cell lines thus provides an excellent opportunity to search for new methylases and is particularly interesting in the context of the possible relationship between methylation and transcriptional control.

Grady and Campbell (1973) found that the proportion of the

TABLE VIII

DEPURINATED XENOPUS DNA FINGERPRINT DATA.

Oligonucleotide Base Composition	Relative Frequency of Methylation	Cytosine (R.M.F.
C	100 (-)	100 (-)
C_2	25 (2.5)	41.4 (3.5)
C_3	4.9 (0.7)	12.0 (1.3)
CT	34.1 (2.6)	70.3 (5.5)
C_2T	11.3 (2.4)	36.9 (1.2)
C_3T	4.8 (1.5)	16.6 (0.2)
C_4T	1.0 (0.2)	3.3 (1.3)
CT_2	18.2 (0.9)	29.2 (3.2)
C_2T_2	7.8 (2.0)	20.0 (2.7)
CT_3	3.7 (0.2)	13.1 (1.9)
C_2T_3	2.5 (0.5)	11.5 (3.2)
C_3T_3	1.6 (0.1)	6.8 (3.9)
C_4T_3	1.2 (0.1)	4.8 (0.6)
CT_4	2.4 (0.8)	5.4 (1.0)
C_2T_4	0.6 (0.6)	6.5 (1.9)
T	1.2 (-)	-

These results are based on 2 cultures labelled with L-(methyl-^{3}H) methionine and $^{32}PO_4$, the frequencies were calculated as previously (Table VI.)

R.M.F. Relative Molar Frequency.

genome transcribed in cultured mouse cells was increased when the cells had been transformed by integration of viral sequences into the host DNA; thus the polyoma transformed line PY AL/N transcribes 30.9% of the genome while the AL/N cells themselves only transcribe 18.9%. Similarly the SV40 transformed Balb 3T3 line, SVT2, transcribes 30% of the genome and Balb 3T3 cells transcribed 19.5%. By challenging DNA with a mixture of RNA from transformed and untransformed cells it was possible to show that some sequences present in the RNA of nontransformed cells were not present in transformed cells, although transformed cells transcribe more of the genome. It is evident that transformation causes a profound change in the organisation of transcription.

It is possible that a new methylase with a different substrate specificity would produce a novel pattern of methylated oligonucleotides when the DNA is subjected to fingerprinting. In the case of the BHK-21/C13 and BHK-21/PyY cells the 70% increase in methylation would make any change in distribution easily detectable. The possibility was explored as below.

Two pairs of cell lines were used in this study: (1) the BHK-21/C13 line and its polyoma transformed derivative BHK-21/PyY (2) L929 mouse cells and the SV40 transformed mouse cell line SVT2 (N.B. SVT2 is not derived from L929 but is a transformed mouse cell line).

$^{32}PO_4$ and L-(methyl-^{3}H) methionine labelled DNA samples were prepared from BHK-21/PyY and SVT2 cells. In addition DNA was also prepared from BHK-21/PyY labelled with (^{14}C-methyl) methionine alone (Methods 1.2.). The nuclear DNA was extracted and fingerprinted as before.

The results obtained from DNA fingerprints are shown in

TABLE IX (Contd)

These results are based on 3 L-(^{3}H-methyl) methionine and $^{32}PO_4$ labelled cultures of SVT2 and BHK-21/PyY cells and one L-(^{14}C-methyl) methionine labelled BHK-21/PyY culture.

Results were calculated as previously (Table VI).

TABLE IX

5 MeC DISTRIBUTION IN THE DNA OF 2 TRANSFORMED CELL LINES.

Oligo-nucleotide Base Composition	SVT2 Relative Frequency of Methylation	SVT2 Cytosine (R.M.F.)	BHK-21/PyY Relative Frequency of Methylation	BHK-21/PyY Cytosine (R.M.F.)
C	100 (-)	100 (-)	100 (-)	100 (-)
C_2	19.2 (3.6)	44.3 (2.3)	31.7 (4.9)	45.0 (4.6)
C_3	4.3 (1.7)	14.4 (3.3)	7.0 (2.2)	17.0 (2.7)
C_4	1.2 (0.3)	5.0 (0.9)	2.1 (0.2)	5.4 (0.7)
CT	21.7 (3.7)	67.9 (4.1)	25.0 (9.3)	73.7 (4.5)
C_2T	7.8 (0.7)	42.5 (1.4)	13.7 (3.2)	47.6 (5.1)
C_3T	4.3 (1.2)	20.4 (1.2)	7.3 (1.6)	24.8 (2.7)
C_4T	1.0 (0.4)	6.5 (0.5)	2.1 (0.2)	8.1 (1.4)
CT_2	4.6 (0.2)	26.0 (1.4)	6.6 (2.0)	31.6 (2.1)
C_2T_2	3.5 (0.4)	30.0 (1.3)	6.9 (1.4)	31.4 (5.1)
C_3T_2	1.6 (-)	14.0 (1.8)	3.3 (0.7)	17.8 (1.7)
C_4T_2	0.6 (0.1)	6.9 (0.4)	1.3 (0.1)	7.7 (1.0)
CT_3	3.9 (0.1)	11.4 (1.0)	1.6 (1.2)	15.3 (4.2)
C_2T_3	1.4 (0.2)	10.4 (3.8)	1.9 (0.7)	18.2 (1.9)
C_3T_3	1.0 (0.9)	7.4 (3.2)	1.8 (0.5)	12.9 (5.7)
C_4T_3	0.2 (0.2)	4.0 (1.8)	1.3 (0.1)	9.7 (3.2)
CT_4	3.2 (0.5)	5.5 (0.4)	0.9 (-)	7.7 (3.6)
T	< 0.01		0.2 (0.2)	

Table IX. The results from the L-(methyl-^{14}C) methionine labelled BHK-21/PyY DNA have been included with the rest of the data as there was no detectable difference between the ^{14}C and ^{3}H distribution. The main purpose of this additional L-(methyl-^{14}C) methionine labelled DNA experiment was to rule out the presence of a new methylase which preferentially methylated the long (15 - 50) oligonucleotides. This would not have been detected using the relatively light $^{32}PO_4$ label employed in these 'double label' experiments.

Comparison between the homologous transformed and untransformed lines on the basis of the 5 MeC distribution in pyrimidine tracts reveals little difference. Thus as far as it is possible to assess at this stage there are no new methylases present in untransformed cells.

It should be noted that the C_2, C_2T and C_2T_2 oligonucleotides in mouse SVT2 cell DNA appear to be methyl deficient compared to BHK-21/PyY DNA; it will be recalled that mouse L929 DNA also appeared to be deficient in the same manner (Table VII).

2.5. 5 MeC DISTRIBUTION IN PYRIMIDINE TRACTS OF PLANT DNA.

The previous sections have demonstrated the basic similarity of the distribution of 5 MeC in a number of vertebrate cell lines. The specific distribution and its apparent conservation throughout the vertebrates suggests there has been strong evolutionary pressure for the retention of 5 MeC in its present distribution. One question that can now be asked is: how far back in evolution does this pattern extend ?

The most practical way to answer this question is to compare the vertebrate 5 MeC distribution with that of some distantly

related eukaryote. In view of the low level of 5 MeC reported in insects and fungi (Introduction, section 2.2.) it was decided to use a comparison with plant tissue, i.e. a soya bean cotyledon cell line (Materials, section 1.1.).

The soya bean culture was maintained on 8% agar containing Millers medium (Methods, section 1.4.). The cells were routinely subcultured every 10 days. To specifically label 5 MeC with L-(methyl-^{3}H) methionine the medium was made 0.01M with respect to sodium formate, L-(methyl-^{3}H) methionine 5μCi/ml and $^{32}PO_4$ 5μCi/ml were added to the gel before cooling and setting, small, 5 day old, colonies of the culture were transplanted onto the labelling medium (10ml volume in 60mm petri dishes) and allowed to grow for 48 hrs before harvesting.

Total cellular DNA was extracted by a two stage process, first, using the same procedure as for the animal cell lines followed by a chloroform/isoamyl alcohol extraction (Methods, 2.2.), (there appeared to be a large amount of white fibrous material remaining after the MUP extraction.) The DNA was then fingerprinted as before. The results are displayed in Table X.

Because of poor resolution in the first dimension (electrophoretic) separation the results have been expressed only in terms of distribution amongst pyrimidines of differing length. Analysis using separation in the first dimension was not attempted. It was, however, possible to identify the T spot to estimate the specificity of the (methyl-^{3}H) labelling. (Table X).

For comparison the data from L929 DNA fingerprints calculated in the same manner is shown alongside. The distribution of 5 MeC appears substantially different in soya bean DNA to that in the

TABLE X.

DISTRIBUTION OF 5 MeC IN SOYA BEAN COTYLEDON DNA.

DISTRIBUTION OF ISOSTICHS

% OF TOTAL PYRIMIDINES IN ISOSTICH

Isostich Number	Soya	σ	L929	σ
1	39.3	(3.2)	43.3	(2.5)
2	41.2	(2.5)	32.2	(4.1)
3	19.6	(1.7)	24.4	(0.9)

DISTRIBUTION OF 5 MeC IN ISOSTICHS

% OF (METHYL-^{3}H) IN ISOSTICH

Isostich Number	Soya	σ	L929	σ
1	35	(10.0)	69.4	(3.0)
2	46	(2.0)	21.2	(2.9)
3	25	(4.0)	9.5	(1.8)

Values calculated as % of total $^{32}PO_4$ (after subtraction of 3' terminal $^{32}PO_4$) or ^{3}H in the first 3 isostichs (i.e. mono, di and tripyrimidine). Values for Soya bean DNA taken from 3 parallel cultures, those for L929 from 2 separate cultures.

(Methyl-^{3}H) label in monopyrimidine T was only detected in one culture at a level of 6.4% of that in monopyrimidine C, none was detected in the other 2.

σ, standard deviation

L929 DNA. The relative abundance of the mono, di and tri pyrimidines in soya bean cotyledon differs to that reported for wheat and rye germ DNA (Spencer and Chargaff, 1963). (This may be due to uneven transfer from the 1st to the 2nd dimension of the fingerprints). However it should be noted that the relative methylation of each of the isostichs in soya bean DNA is markedly different to the relative methylation of the isostichs in vertebrate DNA (exemplified by the L929 DNA fingerprinting results). These results are in general agreement with the observation of Diskocil and Sorm (1962) that the replacement of C by 5 MeC in wheat germ DNA is approximately the same in mono and di pyrimidines, and Spencer and Chargaff (1963) noted that only 31% and 33% of the total 5 MeC in rye germ DNA and wheat germ DNA, respectively, occurred as the monopyrimidine.

It is unlikely that any contribution to the 5 MeC distribution is made by chloroplast DNA as Tewari and Wildman (1966) demonstrated that chloroplast DNA contains no 5 MeC.

These results indicate that the methylation pattern of plant nuclear DNA is considerably different to that in vertebrate and suggests that the particular features of vertebrate DNA methylation have only recently evolved.

2.6. DETERMINATION OF THE LEVELS OF 5 MeC IN CULTURED CELLS.

Nelson-Rees and Flandermeyer (1976) have presented evidence which suggests that many cell lines in culture are contaminated by HeLa cells; this may be the cause of the similarity of results obtained by fingerprinting. The degree of methylation in a number of cell lines was therefore measured.

Nuclear $^{32}PO_4$ labelled DNA was prepared from a number of cell lines (total cellular DNA from Xenopus cells) and analysed for 5 MeC

content as in Experimental, section 1.2. The results are shown in Table XI.

The data indicate that there is a wide and significant variation in the levels of 5 MeC between cell lines. The levels of 5 MeC detected are consistent with previously reported figures (Adams, 1974; Nass, 1973; Rubery and Newton, 1973; Dawid et al., 1970). The level of 5 MeC in HeLa cells is lower than that determined by Dawid (1974), but approximately the same as that reported by Lawley et al., (1972).

It is unlikely that any of these cell lines was contaminated by HeLa cells as they were clearly all morphologically distinct: the variation in 5 MeC contents does, however, exclude the possibility that the distribution of 5 MeC remains constant because the 5 MeC content of all the cells is the same.

3. OLIGONUCLEOTIDE DISTRIBUTONS IN VERTEBRATE DNA.

The results obtained from fingerprint analysis of DNA appear to be useful in terms of a comparative analysis of the distribution of 5 MeC in vertebrate cell lines. There are however, several points in the process of digestion and fingerprinting of DNA where preferential losses of oligonucleotides could occur; Birnboim et al., (1973) reported the preferential loss of high molecular weight isostichs when formic acid and diphenylamine were removed by ether extraction, and transfer between the first and second dimensions of the fingerprinting procedure is not completely quantitative (Brownlee and Sanger, 1969). Potential losses of material make it difficult to relate these fingerprinting studies to others in the literature which employed different techniques, or to interpret

TABLE XI.

THE 5 MeC CONTENT OF DNA IN CULTURED CELLS.

Cell Line	% 5 MeC/C + 5 MeC	No. of Determinations
HeLa	1.3 ± 0.2	(3)
L929	2.9 ± 0.4	(3)
BHK-21/Cl3	1.0 ± 0.1	(3)
BHK-21/PyY	1.7 ± 0.2	(3)
Xenopus	4.1 ± 0.2	(3)

All data are for nuclear DNA except Xenopus cells. (total cellular DNA).

($^{32}PO_4$ labelling was for 48 hours in all cases).

the data in 'absolute' terms. It was necessary, then, to establish whether the oligonucleotide frequencies obtained by $^{32}PO_4$ labelling and fingerprinting were in agreement with reported frequencies obtained by other techniques.

(i) Monopyrimidine Analysis.

The relative molar amounts of T and C in the monopyrimidine fraction have been determined for calf thymus DNA by use of paper chromatography and DEAE cellulose chromatography (Petersen, 1962; Spencer and Chargaff, 1963). Table XII presents these values and the values obtained by fingerprinting for the cell lines used.

(ii) Dipyrimidine Analysis.

Swartz et al., (1962) analysed the DNA of a number of organisms in terms of the frequency of occurrence of nucleotide doublets (e.g. CpG, CpC, ApT). In order to allow a comparison with the fingerprint data, which does not resolve the nucleotide doublets TpC and CpT, the data of Swartz et al., (1962) have been recalculated to express the amount of TpC + CpT relative to CpC. Data for calf thymus DNA obtained by pyrimidine tract analysis have also been included (Table XII).

(iii) Isostich Analysis.

Spencer and Chargaff (1963) using calf thymus DNA and Sneider (1971) using rat hepatoma DNA noted the non random distribution of pyrimidines in isostichs. In particular there are fewer di and tri pyrimidines and more monopyrimidines than expected on the basis of a random distribution of nucleotides in DNA. Data from $^{32}PO_4$ labelled fingerprints of HeLa, BHK-21/PyY, mouse L929 and Xenopus liver cell DNA were analysed in terms of the distribution of pyrimidines in the mono to tetra pyrimidine isostichs (Fig. XII).

TABLE XII.

THE RELATIVE OCCURRENCE OF MONO AND DIPYRIMIDINES IN DNA.

(i) Fingerprint Data (this work)		RELATIVE MOLAR FREQUENCY OF NUCLEOTIDE		
		$^{*}T/C$	$^{*}CT/C_2$	$T_2/^{*}C_2$
No. Expts.	Source of DNA			
(3)	HeLa cells	2.12 (0.1)	3.29 (0.3)	2.32 (0.3)
(1)	Chang "	2.02 (-)	2.94 (-)	2.29 (-)
(3)	BHK-21/PyY cells	2.09 (0.3)	3.17 (0.1)	1.83 (.45)
(3)	BHK-21/C13 cells	2.14 (0.1)	3.01 (0.6)	1.98 (-)
(2)	Mouse L929 cells	2.21 (0.2)	2.91 (0.3)	1.98 (.27)
(3)	Mouse SVT2 cells	1.80 (0.3)	3.08 (0.3)	1.74 (0.1)
(2)	Xenopus liver cells	2.30 (0.1)	3.38 (0.1)	2.7 (0.1)

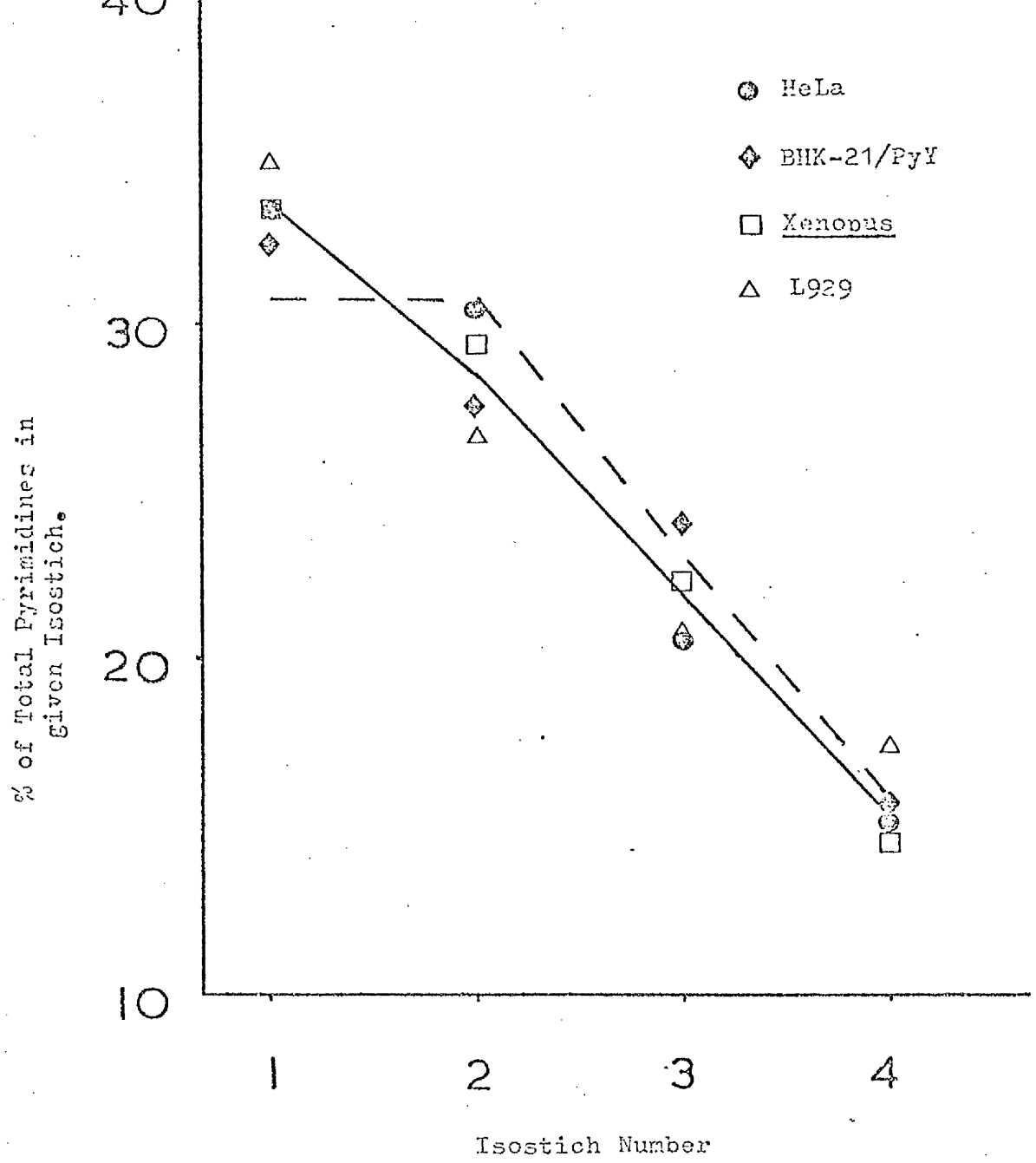
(ii) Nearest Neighbour Data. (calculated from Swartz et al., (1962)).

Mouse lymphoma DNA	-	2.74	1.84
Mouse ascites tumour DNA	-	3.00	1.92
Human spleen DNA	-	2.72	2.06
Rabbit liver DNA	-	2.78	2.04
Bovine thymus DNA	-	2.43	1.41

(iii) Pyrimidine Oligonucleotide Separation Data.

Calf thymus (Petersen, 1962) DNA	1.6	2.6	1.3
Calf thymus DNA (Spencer & Chargaff, 1963).	2.4	3.4	1.8

* Calculated as in Methods, section 7.4.



ISOSTICH DISTRIBUTIONS IN VERTEBRATE DNA.

The distributions shown represent the % of total pyrimidines occuring in each isostich(calculated as a % of the total pyrimidines in the 4 isostichs,allowance has been made for the terminal $^{32}PO_4$;Methods 7.4).Each data point is the mean value obtained from 2 $^{32}PO_4$labelled fingerprints for each of the species.

———,shows the mean distribution(calculated as the mean distribution of pyrimidines for all 4 species).

- - -,shows the predicted random distribution recalculated from the values of Sneider(1971).

The data presented in Table XII and Fig. XII show that the oligonucleotide frequencies observed in these experiments are very similar to the oligonucleotide frequencies obtained by other techniques, the preferential loss of short C rich oligonucleotides, in the course of fingerprinting, noted by Harbers et al., (1975) could not be detected.

(iv) The Distribution of Cytosine.

Tables VI to IX show the relative occurrence of 5 MeC and C in pyrimidine tracts of a number of species. The expected random distribution of C + 5 MeC calculated from the predicted values for calf thymus DNA (Petersen, 1962) are provided for comparison in Table V. Examination of the data for the distribution of $^{32}PO_4$ labelled C_p reveals that nearly all the oligonucleotides contain $^{32}PO_4$ labelled C_p at a frequency greater than random: Tables VI to IX. This apparently anomalous result is explained by the fact that all frequencies of Cp have been expressed relative to the monopyrimidine Cp as 100. However Spencer and Chargaff (1963) showed that in calf thymus DNA only 22.4% of the total C + 5 MeC occurred as the monopyrimidine, not 25% as would be expected, and Petersen (1962) found that C occurred in the monopyrimidine fraction of calf thymus DNA at a level 27% lower than predicted.

4. THE DISTRIBUTION OF 5 MeC IN DNA METHYLATED IN VITRO

Several DNA methylases have been extracted from animal and plant cells (Introduction, 3.3.1. and 3.3.2.). It is not yet known whether each cell contains a number of methylases, nor how they are controlled or how they recognise and act upon their substrates to produce a specific distribution of 5 MeC (Introduction, section 3). A DNA methylase isolated from Krebs II

ascites cells (Turnbull & Adams, 1976) was available in this laboratory; some of its properties were examined employing a modification of the technique used to analyse the distribution of 5 MeC *in vivo* in pyrimidine tracts.

The procedure used by Turnbull and Adams (1976) to methylate DNA *in vitro* was adapted to produce quantities of DNA suitable for fingerprinting (ca. 50µg or less).

The assay mixture (140µl) contained 40µg of DNA (duplicate samples of *E. Coli* denatured and native DNA, and native Calf Thymus DNA were used); 3.3µCi S-adenosyl-L-(methyl-^{3}H) methionine and 100µl of buffered enzyme solution (final concentrations were: 7.15×10^{-4}M dithiothreitol and EDTA, 7.2% glycerol; 3.6×10^{-3}M tris HCl (pH 7.8) and 2.36×10^{-4}M S-adenosyl methionine). After a one hour incubation at 37°C the reaction was stopped by adding 2mls of a solution of 1% SDS, 2mM EDTA, 3% 4-amino-salicyclic acid, 5% butanol, 0.5M NaCl. The mixture was extracted with phenol 88%/M-cresol, 12%/8 hydroxy-guinoline 0.1%, and then alkali (0.1M NaOH) treated for 48 hours at room temperature. The product was neutralised with HCl and made 0.1M with respect to phosphate buffer pH 6.8. This was then desalted through G-25 Sephadex. 70,000cpm of $^{32}PO_4$ labelled BHK-21/PyY DNA (ca. 5µg, prepared as in Methods 1.2 and 2.2) was added and the DNA depurinated and fingerprinted as before. (The addition of $^{32}PO_4$ labelled DNA allows location of the oligonucleotides by autoradiography). The amount of (methyl-^{3}H) label in each spot on the fingerprints was calculated as before (Methods, section 7.4.), results shown in Table XIII.

During the course of this work it was discovered that methylation of native *E. Coli* DNA *in vitro* occurred in *N. crassa* endonuclease sensitive regions (Turnbull and Adams, 1976), i.e. that methylation was occurring in single stranded regions. The pattern of methylation

of 'native' and 'denatured' E. Coli DNA is basically similar, however there do appear to be some differences, the significance of which is unclear as in both bases the DNA being methylated is single stranded. Further experiments are necessary to substantiate the differences.

The partial characterisation of the 5 MeC distribution of DNA methylated in vitro allows us to make several suggestions about the properties of the DNA methylase:

(a) The data are consistent with the purification extracting an enzyme able to methylate pyrimidine tracts in DNA in a manner similar to that performed in vivo. This in turn suggests that the purification procedure has not been selecting one particular DNA methylase from a spectrum of enzymes present in the cell (selective purification of one enzyme at the expense of others might be expected to produce a novel distribution of 5 MeC when DNA methylated in vitro is fingerprinted) i.e. the purification procedure used is isolating the only enzyme present in the cell.

(b) Whether the DNA is single (both E. Coli substrates) or double stranded (Native Calf Thymus DNA) the enzyme is able to methylate DNA and the pattern of methylation is not greatly affected by the secondary structure of the DNA.

This section is discussed more fully later (Discussion).

5. VARIATION OF THE 5 MeC DISTRIBUTION WITH RESPECT TO THE G+C CONTENT OF DNA.

Previous sections have detailed analyses of DNA which characterise the distribution of 5 MeC in short sections of DNA. It was considered to be of interest to examine the distribution of 5 MeC in longer sequences. One possible approach is the use of equilibrium density ultra-centrifugation techniques to isolate satellite DNAs (Kit, 1961) to detect possible preferential association of 5 MeC.

The isolation of satellites, however, tells us little about the organisation of 5 MeC in DNA other than the satellites; Comings

A = Native E. Coli DNA.

B = Denatured E. Coli DNA.

C = Native Calf Thymus DNA.

(-) indicates oligonucleotide methylation only detected in one fingerprint.

* figures indicate occurrence of 5 MeC relative to 5 MeC as monopyrimidine.

Results obtained from fingerprints of duplicate samples of DNA methylated in vitro in each case.

TABLE XIII

THE DISTRIBUTION OF 5 MeC IN PYRIMIDINE TRACTS OF DNA METHYLATED IN VITRO

Oligonucleotide Base Composition	Substrate					
	A		B		C	
C	* 100		100		100	
C_2	27.2	(3.0)	18.5	(1.4)	17.5	(1.5)
C_3	5.5	(0.5)	4.3	(0.7)	5.7	(0.3)
C_4	2.1	(1.4)	1.0	(-)	1.5	(0.5)
CT	36.5	(3.5)	21.4	(3.7)	42.0	(1.0)
C_2T	10.5	(1.5)	9.0	(1.0)	13.8	(1.8)
C_3T	3.5	(0.5)	3.0	(1.0)	5.7	(0.7)
C_4T	3.8	(-)	2.3	(-)	1.6	(-)
CT_2	8.5	(2.5)	6.3	(0.3)	8.3	(3.3)
C_2T_2	7.3	(2.7)	7.5	(1.6)	7.5	(0.5)
C_3T_2	6.9	(0.2)	6.6	(0.1)	1.9	(-)
C_4T_2	5.5	(0.3)			2.4	(-)
CT_3	3.0	(-)	2.8	(-)	3.0	(1.0)
C_2T_3	3.7	(-)	4.8	(-)	2.9	(-)
C_3T_3	5.3	(-)	4.5	(-)	2.8	(-)
C_4T_3	5.3	(-)			1.8	(-)
T	4.0	(-)	6.0	(-)	2.0	(-)

(1972) was able to demonstrate preferential methylation of G+C rich DNA in Chinese hamster DON cells but this work necessitated synchronising the cells.

A modification of the basic equilibrium density gradient technique was suggested by the work of Birnstiel et al., (1974) who were able to isolate sea urchin histone genes by using Actinomycin D/CsCl equilibrium gradients; the Actinomycin D intercalates with G+C rich DNA (Sobel, 1974) and preferentially alters its density. In the following experiments DNA of a relatively low molecular weight ca. 5.0×10^6 was used as this was considered more likely to produce separation of sequences less clustered than, for example, mouse satellite DNA.

In a preliminary experiment L929 nuclear DNA labelled with L(methyl-^{3}H)-methionine and $^{32}PO_4$ (Methods, section 1.2.) was centrifuged to equilibrium on CsCl gradients in the absence of Actinomycin D (Fig. 14) (Methods, section 4). The majority of the (methyl-^{3}H) label is coincident with the $^{32}PO_4$ label, i.e. there is no evident bias towards A+T or G+C rich DNA. There is a small amount of material on the A+T rich side of the main band DNA which is probably satellite DNA (although no attempt was made to characterise it further); It does appear to be labelled with (methyl-^{3}H) to a higher level than main band DNA, as might be expected for mouse satellite DNA (Introduction, section 4).

Addition of Actinomycin D (200µg/5ml gradient) causes a profound shift in the relative distributions of (methyl-^{3}H) and $^{32}PO_4$ labels, the (methyl-^{3}H) distribution is heavily biased towards the A+T rich section of the gradient (Fig. 14(b)).

An alternative method for separating DNA fractions with respect to base composition is by thermal elution from hydroxylapatite. The

FIG.14.

ISOPYCNIC CENTRIFUGATION OF L929 DNA.

(a) L-(methyl-^{3}H)methionine & $^{32}PO_4$ labelled nuclear DNA from mouse L929 cells (Methods 1.2.&2.),of average single-stranded molecular weight 1.1×10^{6}, was centrifuged to equilibrium in a CsCl gradient (average Refractive Index 1.400), fractionated dropwise onto Whatman 3MM discs and the radio-activity estimated as in Methods 4.

O——O L-(methyl-^{3}H)

●——● $^{32}PO_4$.

Density increases from right to left.

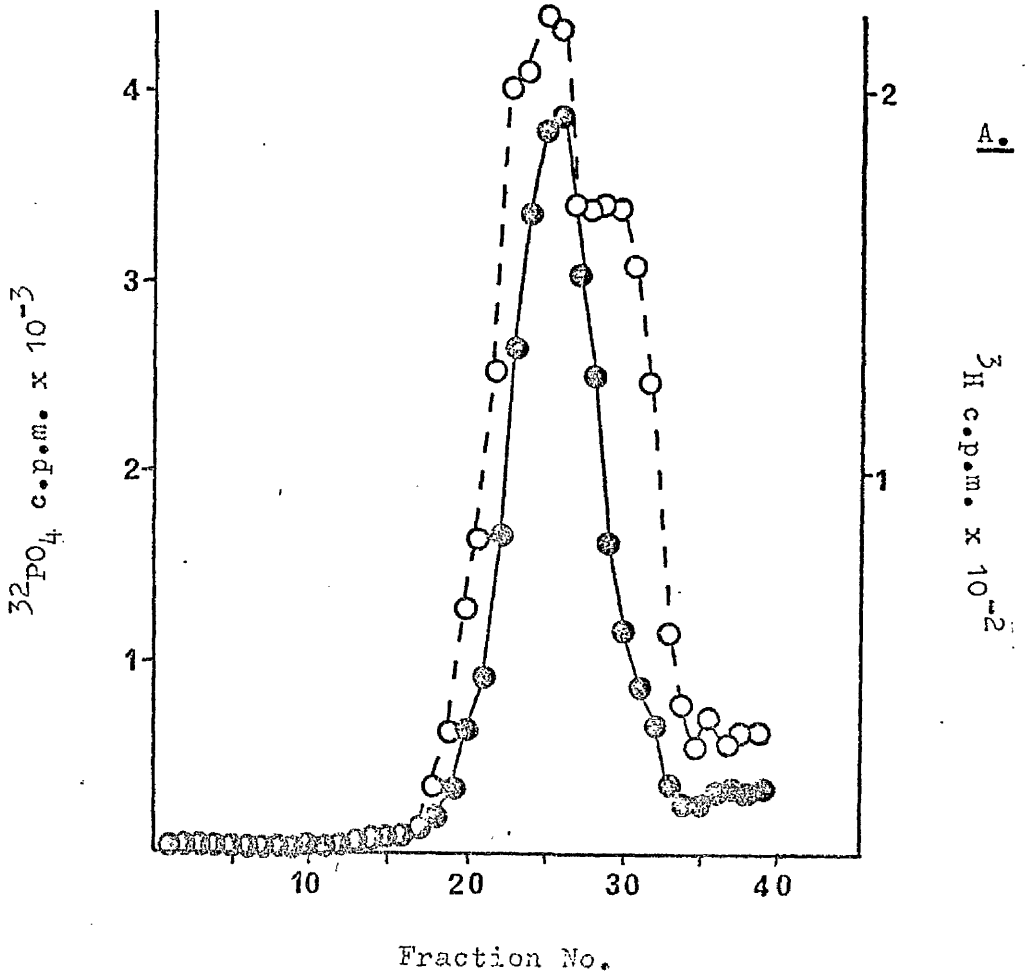
(b) DNA prepared as in (a) was centrifuged to equilibrium in a CsCl gradient (average Refractive Index 1.3900 at 3°C) containing actinomycin D at a concentration of 40 µg/ml(Methods 4). The gradient was harvested as in (a).

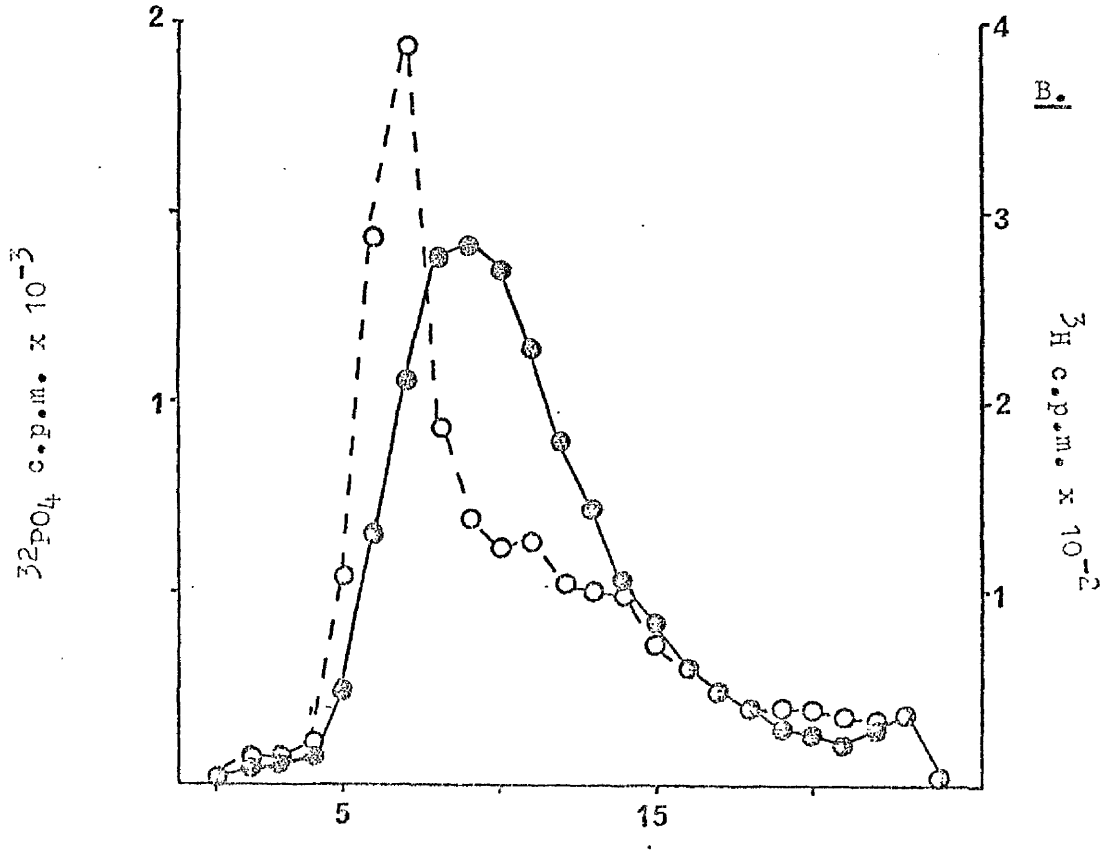
O——O L-(methyl-^{3}H)

●——● $^{32}PO_4$.

Density increases from right to left.

Radioactivity estimated as acid precipitable c.p.m.


A.
$^{32}PO_4$ c.p.m. x 10^{-3}
^{3}H c.p.m. x 10^{-2}
Fraction No.


B.
$^{32}PO_4$ c.p.m. x 10^{-3}
^{3}H c.p.m. x 10^{-2}

separation relies on the fact that the double helix in G+C rich regions of DNA is more stable than in A+T rich regions and is thus retained on a hydroxylapatite column at higher temperatures.

L929 nuclear DNA labelled as above and either sonicated or non-sonicated (Methods, section 3) was eluted from a hydroxylapatite column using 30 volumes of 0.12M sodium phosphate buffer pH 6.8 at each temperature interval indicated (Fig. 15) (Methods, section 5).

Whether the L929 DNA was sonicated or not the T_M of the (methyl-^{3}H) label was considerably lower than that of the $^{32}PO_4$ label, i.e. 5 MeC is associated with A+T rich regions of the DNA.

The elevated T_M apparent in the non-sonicated thermal elution compared to the sonicated thermal elution is probably due to the retention of denatured DNA by virtue of its covalent attachment to non-denatured G+C rich regions of DNA, as a decrease in molecular weight of the DNA depresses the T_M (Fig. 15(a) and (b)).

This result independently confirms the observed qualitative behaviour of (methyl-^{3}H) label on Actinomycin D/CsCl gradients and supports the validity of using these gradients in this analysis.

COMPARATIVE ANALYSES USING ACTINOMYCIN D GRADIENTS.

The distribution of 5 MeC in pyrimidine clusters was found to be highly conserved in vertebrates. The question now posed is whether the 5 MeC distribution seen in L929 DNA on Actinomycin D/CsCl gradients is also conserved. An intraspecific and interspecific comparison was performed.

For the intraspecific comparison nuclear DNA from mouse SVT2 cells was prepared and analysed on Actinomycin D/CsCl gradients as for L929 DNA; the gradient profile is shown in Fig. 16. The isotope distributions of both mouse cell lines are very similar,

FIG.15.

THERMAL ELUTION OF L929 DNA FROM HAP.

(A.) Mouse L929 cell nuclear DNA labelled with L-(methyl) methionine & $^{32}PO_4$ (Methods 1.2 & 2.), single-stranded molecular weight 1.1×10^6, was adsorbed onto a column of hydroxylapatite; this was washed with 3x10 column volumes of 0.12M phosphate buffer at the temperatures indicated(Methods ,5).

(B.) Mouse L929 DNA labelled and extracted as in (A.) and sonicated to single-stranded molecular weight 6×10^4 (Methods 3) was adsorbed onto hydoxylapatite and eluted as in (A.).

In both cases radioactivity was measured as acid precipitable c.p.m.(Methods,5)

O——O L-(methyl-^{3}H).

●——○ $^{32}PO_4$.

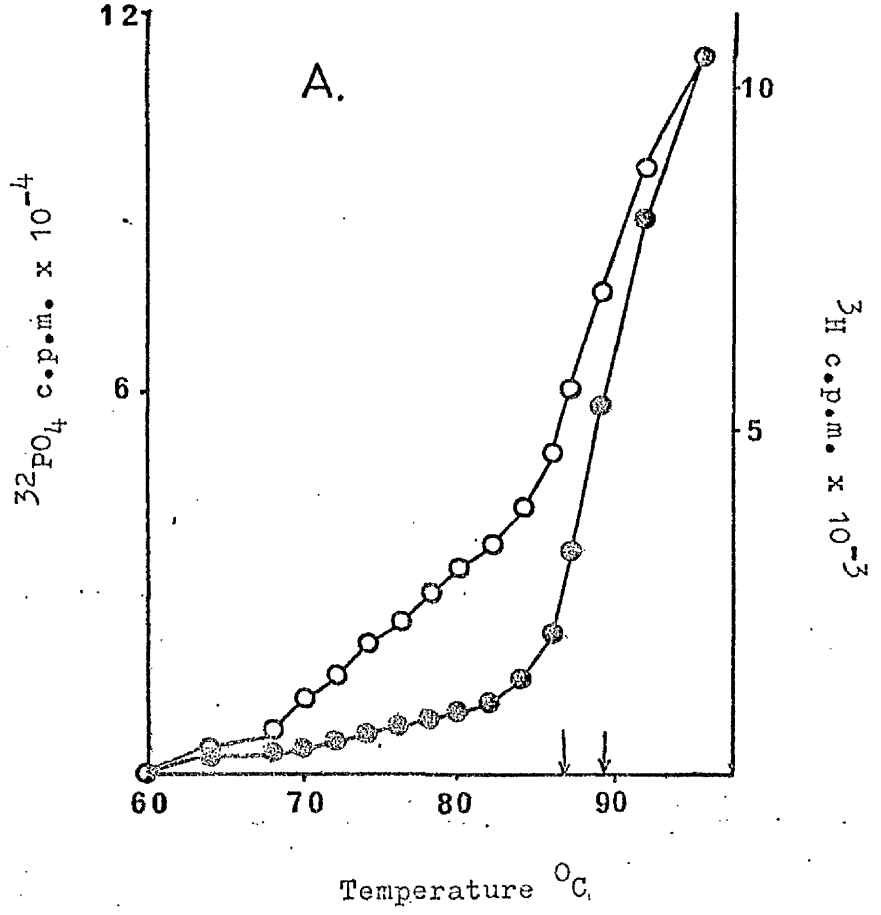
Arrows indicate Tm.

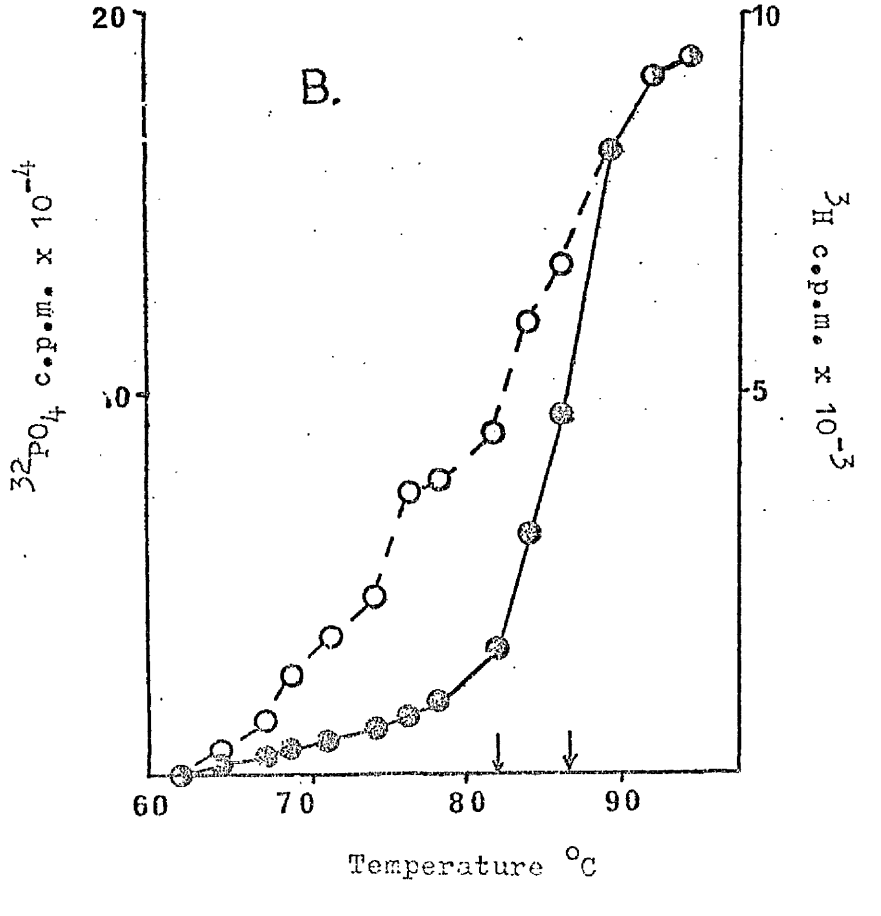
Tm. $^{32}PO_4$ labelled unsonicated DNA=89°C

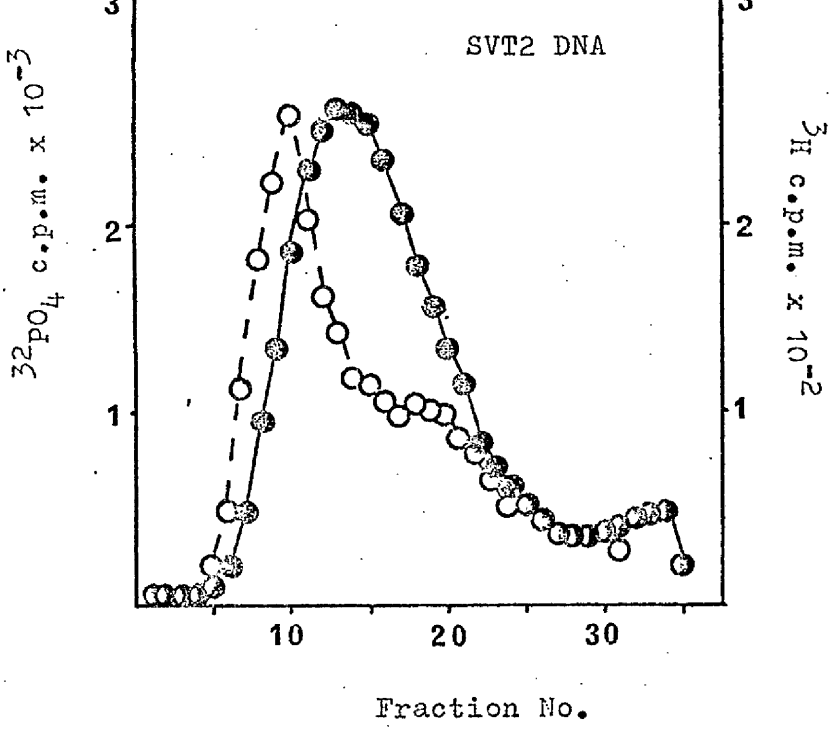
" " " sonicated " =87°C

Tm. (methyl-^{3}H) " unsonicated DNA=87°C

" " " sonicated DNA=82°C


A.
12
6
10
5
60
70
80
90
$^{32}PO_4$ c.p.m. x 10^{-4}
^{3}H c.p.m. x 10^{-3}
Temperature °C


B.
20
10
10
5
60
70
80
90
$^{32}PO_4$ c.p.m. x 10^{-4}
^{3}H c.p.m. x 10^{-3}
Temperature °C


SVT2 DNA
$^{32}PO_4$ c.p.m. x 10^{-3}
^{3}H c.p.m. x 10^{-2}
1
2
3
10
20
30
Fraction No.

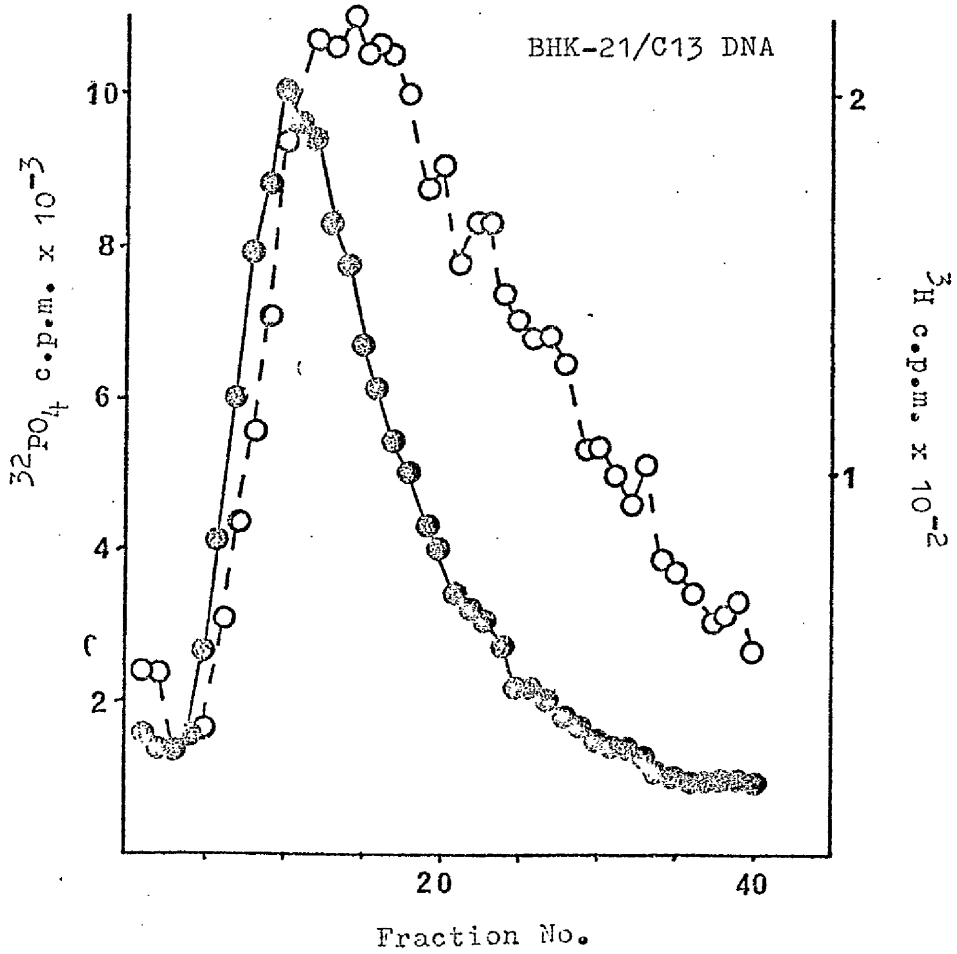

BHK-21/C13 DNA
$^{32}PO_4$ c.p.m. x 10^{-3}
^{3}H c.p.m. x 10^{-2}
2
4
6
8
10
1
2
20
40
Fraction No.

FIG.16.

ISOPYCNIC CENTRIFUGATION OF SVT2 DNA.

Mouse SVT2 cell nuclear DNA,labelled with L-(methyl-^{3}H) ethionine & $^{32}PO_4$ (Methods,1.2 & 2.) was centrifuged to quilibrium on a CsCl gradient containing actinomycin D, onditions were as in legend to FIG.14(b).

O——O L-(methyl-^{3}H).

●——● $^{32}PO_4$.

Density increases from right to left.

Radioactivity estimated as acid precipitable c.p.m.

FIG.17.

ISOPYCNIC CENTRIFUGATION OF BHK-21/C13 DNA.

Hamster BHK-21/C13 cell nuclear DNA,labelled with L-(methyl-^{3}H)methionine & $^{32}PO_4$ (Methods,1.2 & 2.),was centrifuged to equilibrium on a CsCl gradient containing actinomycin D, conditions were as in the legend to FIG.14(b).

O——O L-(methyl-^{3}H).

●——● $^{32}PO_4$.

Density increases from right to left.

Radioactivity estimated as acid precipitable c.p.m.

FIG.18.

ISOPYCNIC CENTRIFUGATION OF HELA DNA.

HeLa nuclear DNA labelled with L-(methyl-^{3}H)methionine and $^{32}PO_4$(Methods 1.2 & 2.)was analysed on CsCl gradients containing actinomycin D,conditions as in legend to FIG.14(b.)

O——O L-(methyl-^{3}H)

●——● $^{32}PO_4$

Density increases from right to left. Radioactivity estimated as acid precipitable c.p.m.

FIG.19 .

THERMAL ELUTION OF HELA DNA FROM HAP.

HeLa nuclear DNA labelled with $^{32}PO_4$ or L-(methyl-^{14}C) methionine(Methods 1.2 & 2.) of single-stranded molecular weight 1.1×10^6 was adsorbed onto a column of hydroxylapatite and eluted ,at the temperature steps indicated,by3x10 column volume washes of 0.12M phosphate buffer (Methods 5.).

O——O L-(methyl-^{14}C)

●——● $^{32}PO_4$

Radioactivity estimated as acid precipitable c.p.m.

FIG.18

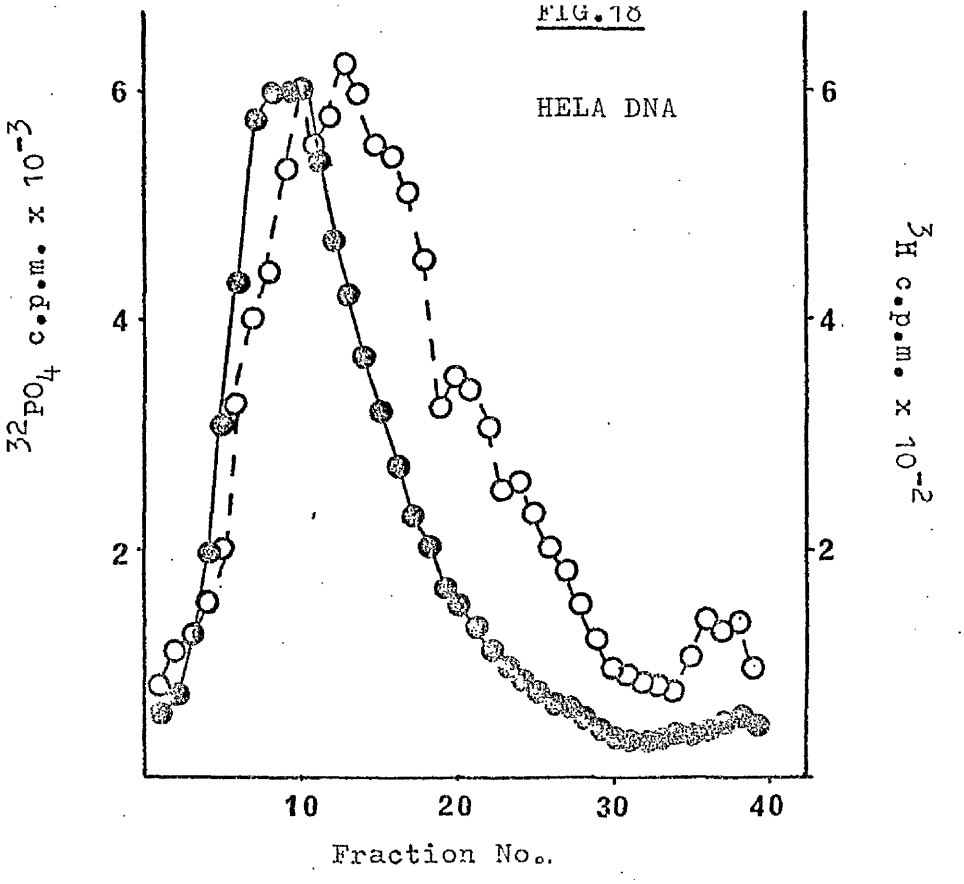

HELA DNA
$^{32}PO_4$ c.p.m. x 10^{-3}
^{3}H c.p.m. x 10^{-2}
Fraction No.
2
4
6
10
20
30
40

FIG.19

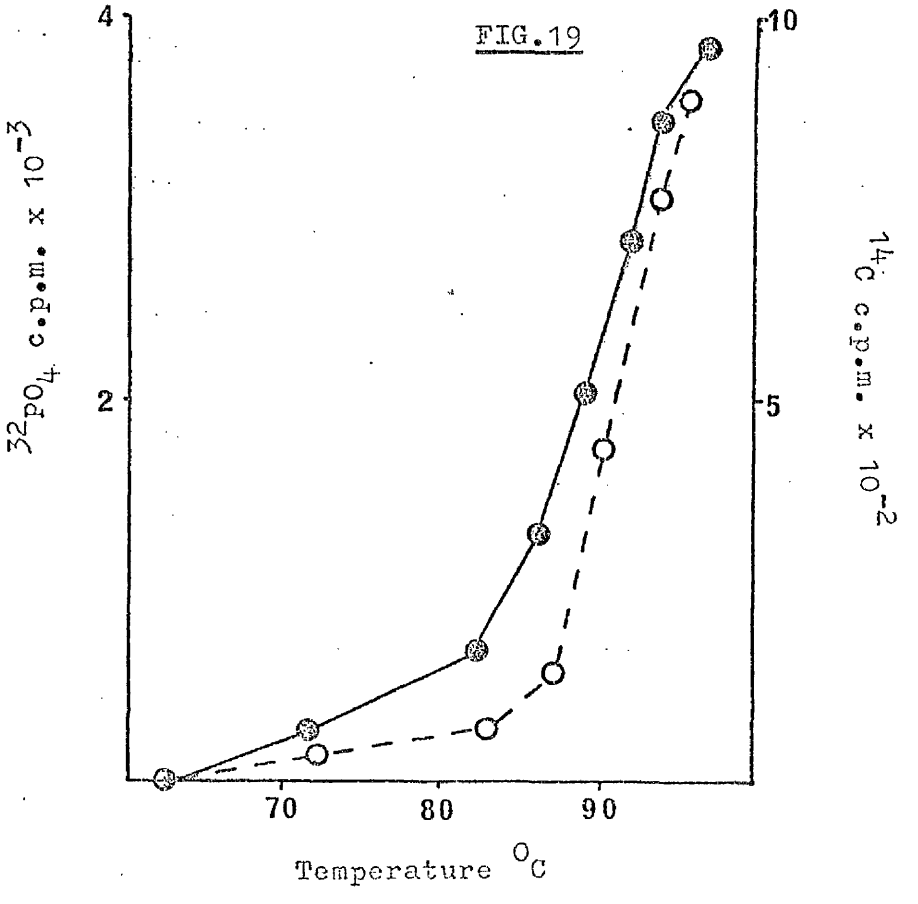
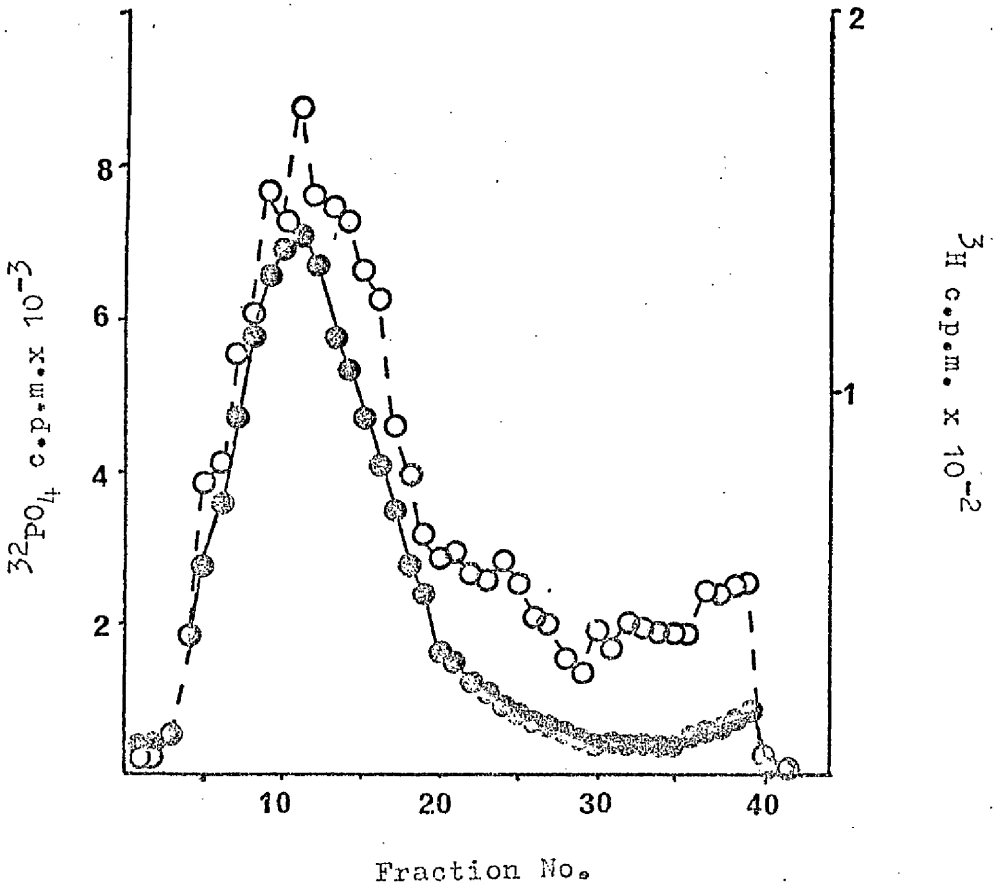

$^{32}PO_4$ c.p.m. x 10^{-3}
^{14}C c.p.m. x 10^{-2}
Temperature ^{O}C
2
4
5
10
70
80
90

FIG.20



ISOPYCNIC CENTRIFUGATION OF XENOPUS DNA.

Xenopus DNA labelled with L-(methyl-^{3}H)methionine and $^{32}PO_4$ (Methods 1.2 & 2.) was centrifuged to equilibrium in a CsCl gradient containing actinomycin D ,conditions as in Methods 4.

O——O L-(methyl-^{3}H)

●——● $^{32}PO_4$

Density increases from right to left.Radioactivity estimated as acid precipitable c.p.m.

if not identical.

Nuclear DNA from BHK-21/C13 and HeLa cells was similarly analysed. In both these cell lines the 5 MeC distribution was biased towards G+C rich DNA (Fig. 17 and Fig. 18). The result for the BHK-21/C13 cell DNA is in close agreement with the distribution of 5 MeC in the Chinese hamster DON cell line (Comings, 1972).

As a control samples of unsonicated $^{32}PO_4$ labelled and L(methyl-^{14}C) methionine labelled HeLa nuclear DNA (Methods, section 1.2) were also subjected to thermal elution from hydroxylapatite (Fig. 19). The T_M of the ^{14}C label is somewhat higher than the $^{32}PO_4$ label and is consistent with the Actinomycin D/CsCl gradient data.

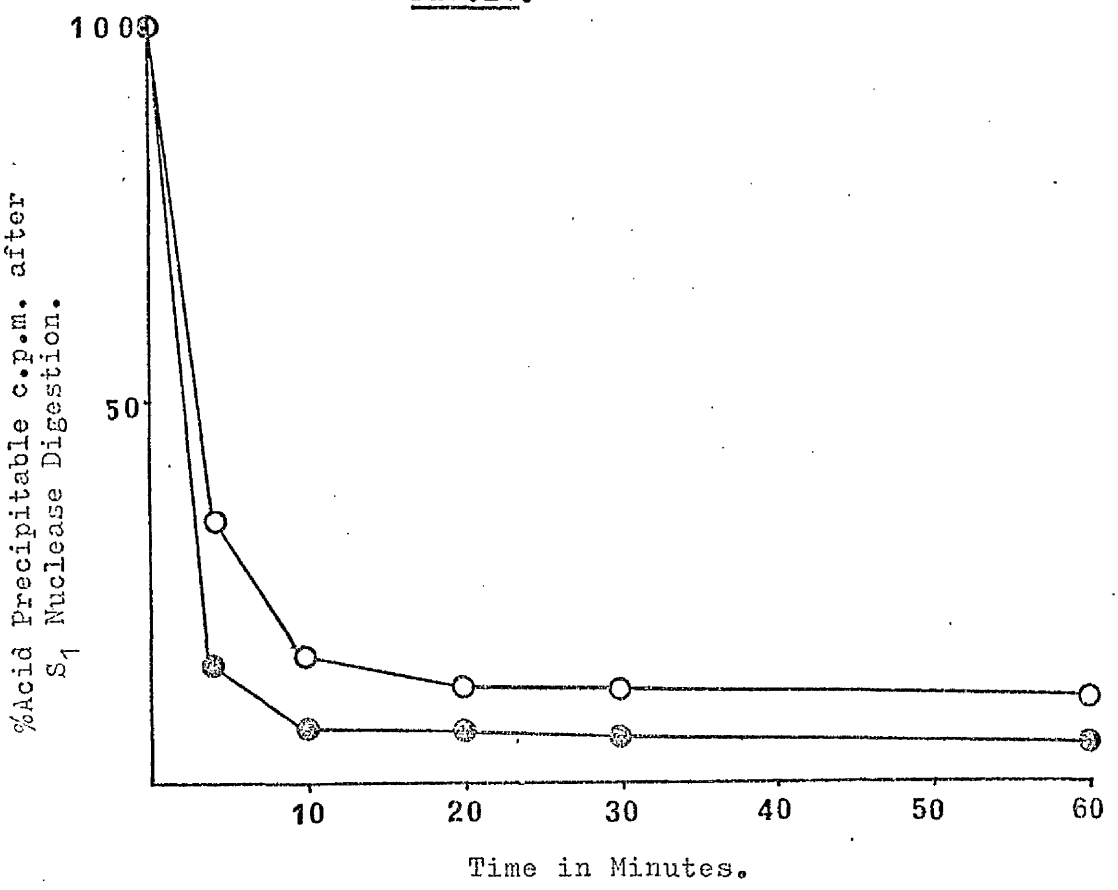
Figure 20 shows the distribution of (methyl-^{3}H) and $^{32}PO_4$ labels obtained on Actinomycin D/CsCl gradients of Xenopus (whole cell) DNA labelled with L(methyl-^{3}H)methionine and $^{32}PO_4$. In this instance the L(methyl-^{3}H) distribution seems to be largely coincident with the $^{32}PO_4$ label but with a substantial portion associated with the G+C rich fraction of DNA.

The results obtained from isopycnic ultracentrifugation and thermal elution of DNA suggest that there is a certain degree of species specificity of 5 MeC distribution within the genome as a whole.

6. METHYLATION OF RAPIDLY REANNEALING DNA IN HeLa CELLS.

Brahic and Fraser (1971) described a rapidly reannealing fraction of mouse DNA, consisting ca. 10% of the total genome, which was enriched in 5 MeC; this fraction was slightly A+T rich. Schneiderman and Billen (1973) isolated a rapidly reannealing fraction of DNA from Chinese hamster cell DNA which was also highly methylated. The possibility that all mammalian DNA has a 5 MeC rich rapidly reannealing component was examined using DNA derived from a non-rodent source,

FIG.21.



TIME COURSE OF S_1 NUCLEASE DIGESTION OF HELA NUCLEAR DNA

L-(methyl-^{14}C)methionine labelled HeLa nuclear DNA 15µg (700c.p.m./µg) and 9µg of $^{32}PO_4$ labelled HeLa nuclear DNA (3000c.p.m./µg)(Methods 1.2 & 2.)were dissolved together in 2ml. of 0.1xSSC and boiled at 100°C for 10min.,the solution was rapidly cooled on ice and 200µl aliquots were digested for up to 60 min. with S_1 Nuclease using heat denatured sonicated calf thymus DNA as carrier;the final concentration of total DNA was 40µg/ml. and NaCl was 0.1M.Radioactivity was estimated as acid precipitable c.p.m.,allowance was made for 'spillover' between counting channels.

O——O L-(methyl-^{14}C)

●——● $^{32}PO_4$

i.e. HeLa cell DNA.

HeLa nuclear DNA labelled with L(methyl-^{14}C) methionine was mixed with $^{32}PO_4$ labelled HeLa nuclear DNA (Methods, 1.2.), heat denatured (Legend to Fig. 21), and digested with S_1 nuclease for varying periods (Methods, 10). Fig. 21 shows the ^{14}C label to be relatively more resistant to digestion than $^{32}PO_4$, indicating that the rapidly reannealing fraction is 5 MeC rich. To ensure that this represented an increased % methylation of C, a similar time course using (6-^{3}H) thymidine labelled DNA (Methods, section 1.2.) plus L(^{14}C-methyl) methionine labelled DNA was performed. The relative resistance of the radiolabels to digestion are shown in Table XIV.

Double stranded HeLa DNA ($^{32}PO_4$ labelled) was resistant to S_1 digestion, i.e. <5% of the label became acid soluble after digestion for 60 mins.

The precise amount of S_1 resistance was found to vary slightly between experiments, however the ratio of labels in the resistant fractions remained roughly constant.

Examination of Table XIV shows that there is a small rapidly reannealing component of HeLa DNA enriched for 5 MeC. The increase in 5 MeC is not due to an elevated C + 5 MeC content, but to an increased % methylation of C. (An elevated C + 5 MeC content would mean a marked reduction in the % of thymidine in this fraction).

TABLE XIV

METHYLATION OF RAPIDLY REANNEALING DNA IN HeLa CELLS.

Source of Label	% S_1 Resistant (Label)	Ratio
1. $^{32}PO_4$ and (methyl-^{14}C) methionine	10.8 (^{14}C) 5.4 ($^{32}PO_4$)	2.0 ($^{14}C/^{32}PO_4$)
2. (6-^{3}H) Thymidine and (methyl-^{14}C) methionine	14.1 (^{14}C) 8.2 (^{3}H)	1.7 ($^{14}C/^{3}H$)

Data from duplicate time courses.

7. DIGESTION OF MAMMALIAN NUCLEAR DNA BY THE Hpa II RESTRICTION ENDONUCLEASE.

The bacterial restriction endonuclease Hpa II cleaves unmodified double stranded DNA as follows:

5' C↓C-G-G 3'

3' G-G-C-C 5'
↑

(Nathans and Smith, 1975).

Methylation of the 3' cytosine in the cytosine doublet prevents restriction (A. Bird, personal communication). This enzyme is particularly interesting in the context of vertebrate DNA methylation as the cleavage site contains the nucleotide doublet C-G which is known to be methylated (Introduction, section 3.4.). The Hpa II restriction endonuclease can accordingly be exploited as a probe for detecting methylation of this tetranucleotide in DNA either for comparison of variation in levels of methylation between tissues of one organism or for detection of unmethylated regions in DNA.

The two cell lines BHK-21/Cl3 and its polyoma transformed derivative BHK-21/PyY have been shown to have a different 5 MeC content (Introduction, section 5). It was considered to be of interest to attempt to confirm this observation using an independent method, i.e. by using Hpa II restriction endonuclease.

Samples of $^{32}PO_4$ labelled DNA prepared from BHK-21/Cl3 and $^{32}PO_4$ and L(methyl-^{3}H) methionine labelled BHK-21/PyY DNA (Methods, sections 1.2 and 2) were digested with Hpa II restriction endonuclease (4µg of DNA with 4 units of enzyme) and then electrophoresed on 1.5% agarose gels for 2 hours at 10 amps per gel (Methods, section 8). The digestions were assumed to have reached completion as no further degradation was apparent when the incubation time was increased

from 4 to 16 hrs. The distribution of $^{32}PO_4$ label is shown in Fig. 22.

In DNA of random base composition each possible tetranucleotide should occur once every 256 base pairs; taking into account the CpG doublet deficiency in vertebrate DNA (Swartz et al., 1962) and the low C+G content (40% of total bases) the tetranucleotide 5' C-C-G-G 3' should occur approximately once every 256 x 4 x 1.25 base pairs, i.e. cleavage of totally unmethylated vertebrate DNA should produce fragments of average molecular weight 8×10^5. Examination of the gel profiles in Fig. 22 reveals a large amount of material in both the restriction products of BHK-21/C13 and BHK12k/PyY DNA which has not been reduced in molecular weight, i.e. remains at 5×10^6 (Methods 2.0.) The heterogeneous nature of the material on the gels makes precise measurement of Hpa II sensitivity difficult, however an estimate of material moving faster than the main band on the gels shows BHK-21/C13 DNA to be 40% 'digested', i.e. reduced in molecular weight and moves faster on gels, whereas BHK-21/PyY DNA is only 20% 'digested'. This is in qualitative agreement with the observed differences in 5 MeC content of the cells, however without a reasonably precise estimate of the average size of the 'digested' DNA it is not possible to quantitate the relative resistance of DNA to Hpa II cleavage.

If the non-methylated, and therefore cleaved Hpa II sites had been distributed randomly the 'digested' DNA should have shown a completely random molecular weight distribution. This is not the case, indeed the 'digested' DNA appears to contain a large amount of material of fairly discrete molecular weight implying that the unmethylated Hpa II cleavage sites are separated by a short relatively specific distance on the DNA molecule.

Examination of the distribution of L(methyl-^{3}H) label in

FIG.22.

AGAROSE GEL ELECTROPHORESIS OF BHK-21/C13 & BHK-21/PyY DNA.

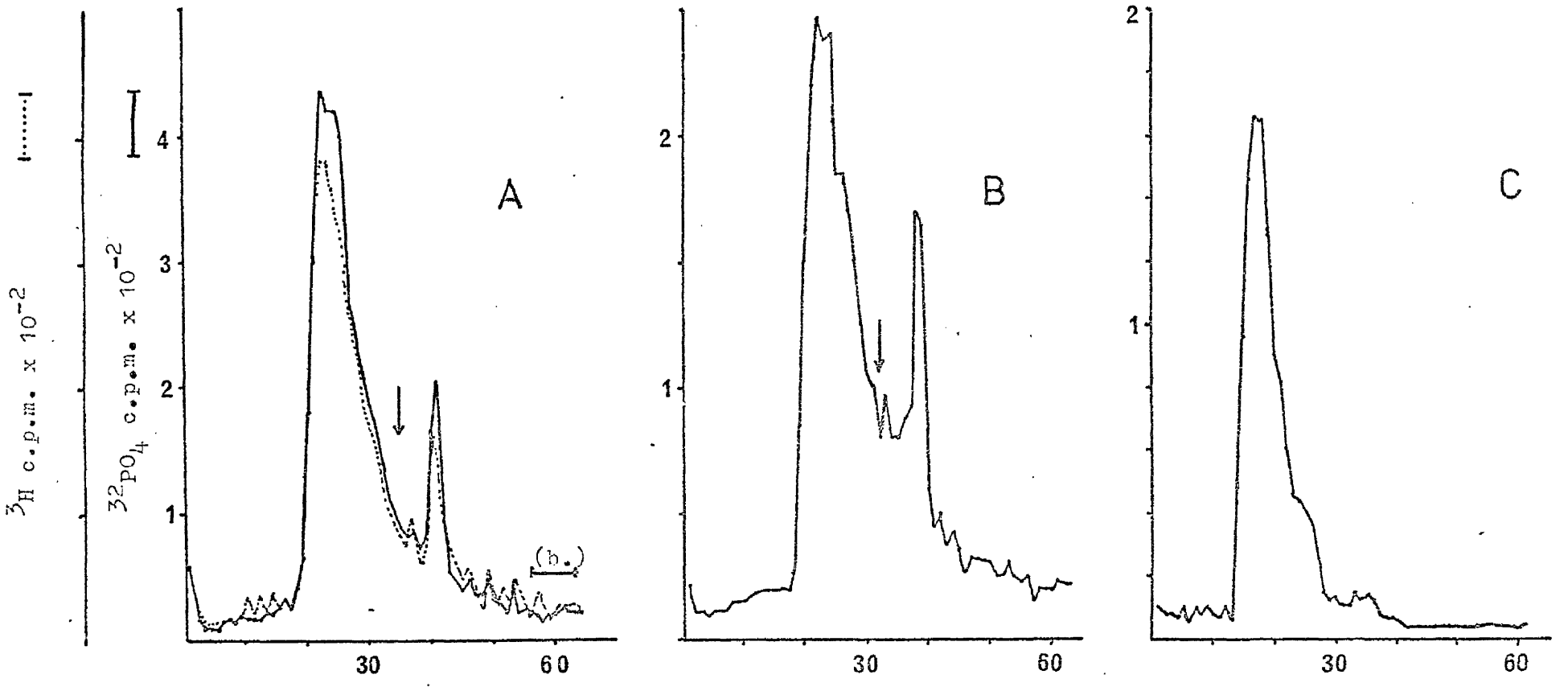
(A.) Gel profile of BHK-21/PyY nuclear DNA labelled with L-(methyl-^{3}H)methionine & $^{32}PO_4$ (Methods 1.2 & 2.).The DNA was restricted with the Hpa$\overline{11}$ restriction endonuclease and electrophoresed through a 1.5% agarose gel, after electrophoresis the gel was sliced(2mm.), solublised & the total radioactivity in each slice estimated(Methods 8.1 &8.2).

(B.) Gel profile of BHK-21/C13 nuclear DNA, labelled with $^{32}PO_4$ alone(Methods 1.2 &2.),conditions of restriction and electrophoresis as for (A.)

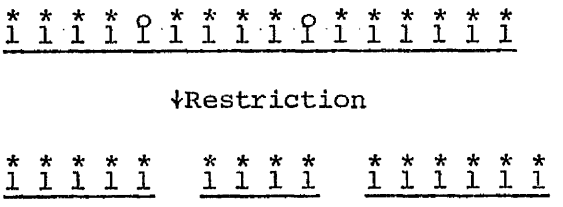
(C.) Control. Gel profile of unrestricted BHK-21/C13 nuclear DNA (profile for BHK-21/PyY nuclear DNA identical) Conditions as in (A.) except that Hpa$\overline{11}$ restriction endonuclease was not added to the incubations.

Arrows indicate point of division of "undigested" & "digested" DNA(determined arbitrarilly).

(b.)= Position of bromophenol blue.


^{3}H c.p.m. x 10^{-2}
$^{32}PO_4$ c.p.m. x 10^{-2}
A
B
C
(b.)
30
60
Gel slice No.
Gel slice No.
Gel slice No.
Direction of Migration.

gel A of Figure 22 shows that the 'digested' DNA is methylated and consideration of the ratio of $^{32}PO_4$ to L(methyl-^{3}H) label in the 'undigested' and 'digested' DNA shows the 'digested' DNA to be normally methylated. Restriction of BHK-21/PyY DNA may then be considered to have occurred as follows:-



*1 = Methylated regions of DNA

♀I = unmethylated Hpa II cleavage sites

\- = double stranded DNA.

These results thus suggest that unmethylated 5' C-C-G-G 3' sequences are distributed throughout the genome in a highly specific manner.

DISCUSSION

1. SPECIFICITY OF LABELLING 5 MeC.

One of the problems associated with specifically labelling 5 MeC is that of labelling other bases by addition of labelled methyl groups, particularly thymine. The fact that the labelling procedure adopted here is highly specific is indicated by:

(i) Acid hydrolysis and chromatographic analysis of DNA (Table V).

(ii) The lack of visible spots on autoradiographs, of ^{14}C labelled fingerprints, which correspond to T_1, T_2 or T_3 (Fig. 12).

(iii) The relative amount of L(methyl-^{3}H) label in the T_1 spot of $^{32}PO_4$ and L(methyl-^{3}H) methionine labelled fingerprints is very low (Tables VI to X).

Deamination of 5 MeC has been suggested to have a role in control of differentiation (Introduction, section 6), although recent evidence suggests that earlier reports of deamination *in vivo* are the result of artefacts (Sneider, 1973). The specificity of 5 MeC labelling shows that little (methyl-^{3}H) or (methyl-^{14}C) label appears in T. It is not possible to assess the significance of this low level of labelling as no rigorous precautions were taken to prevent entry of (methyl-^{3}H) or (methyl-^{14}C) label into the C_1 carbon pool, nor to prevent deamination of 5 MeC during isolation or fingerprinting of DNA.

2. THE EVOLUTIONARY CONSERVATION OF METHYLATED SEQUENCES IN VERTEBRATE DNA.

Previous work in this field has partially characterised the nucleotide sequences methylated in DNA (Introduction, section

3.4.2.). The purpose of the series of experiments in Experimental section 2 was to use the distribution of 5 MeC in pyrimidine tracts as a sensitive probe for tissue and species specificity of methylation and in a search for new virus coded or virus induced methylases (presumably able to methylate DNA in new sequences).

(i) TISSUE SPECIFICITY:

Comparison of the fingerprinting data obtained from pairs of cells derived from the same species (HeLa and Chang; BHK-21/C13 and BHK-21/PyY; mouse L929 and SVT2) reveals a high degree of similarity in the pattern of methylation (Experimental, section 2).

(ii) SPECIES SPECIFICITY:

Comparison of the fingerprinting data obtained for the human hamster, mouse and Xenopus cell lines shows little difference in the distribution of 5 MeC (Experimental, section 2). The sample size used in these experiments is too small to be confident about the quantitation of minor variations which may in fact represent a degree of species specificity, e.g. the undermethylation of some oligonucleotides in mouse cell DNA (Tables VII and IX).

In view of the large range in the 5 MeC content of the nuclear DNA of the cells used (Table V) the differences in 5 MeC distribution are very small, variations in 5 MeC distribution may then represent deviations from a basic pattern.

The fact that the vertebrate pattern of 5 MeC distribution differs from that of the higher plants (Experimental, section 2.6) and bacteria (Introduction, section 3.2) suggests that the pattern has evolved fairly recently and has since been highly conserved.

The data in Tables VI to IX are largely consistent with the possibility that 5 MeC occurs in the sequence 5' Py-5MeC-G 3' (Table XV). However without a dinucleotide or terminal

nucleotide analysis of pyrimidine tracts this cannot be confirmed. (The amount of label used in the fingerprints was insufficient for this purpose).

3. THE EXISTENCE OF VIRUS CODED OR VIRUS INDUCED METHYLASES.

Rubery and Newton (1973) suggested that the elevated 5 MeC content of BHK-21/PyY cells might be due to the presence of a new DNA methylase(s). If this is so one might expect the action of this enzyme to produce an alteration in the distribution of 5 MeC; a 70% increase in the total 5 MeC content occurring in only one or two pyrimidine tracts would be easily detectable. The data in Table IX show no difference in the 5 MeC distribution in BHK-21/C13 and BHK-21/PyY DNA, nor any difference between mouse L929 and SVT2 DNA. The 'new' methylations are arranged in the normal manner and the methylase(s) specificity is therefore unchanged: thus as far as this test is applicable there is apparently no new DNA methylase present after virus transformation of a cell.

This suggestion, however, assumes that labelling DNA with L(methyl-^{3}H) or L-(methyl-^{14}C) methionine would reveal all the 5 MeC on a fingerprint. This assumption may be erroneous as Ehrlich et al., (1975) were able to identify a mechanism in Xanthomonas phage XP-12 which utilised serine in place of SAM to methylate C. This phenomenon, however, has not yet been observed in eukaryotes.

The increased content of 5 MeC in virally transformed cells poses the problem of accommodation of the new bases. There are two possible mechanisms: (i) Previously unmethylated DNA sequences become methylated or (ii) Sequences normally methylated in DNA are amplified.

TABLE XV

COMPARISON OF THE DISTRIBUTION OF 5 MeC WITH THE DISTRIBUTION OF 5'Py-C-Pup 3' IN BHK-21/Cl3 DNA.

Oligonucleotide Base Composition	Rel. Frequency of Methylation	Relative Frequency of 5'Py-C-Pup 3'
C	100	100
C_2	28.4	24.7
C_3	6.9	7.6
C_4	3.3	3.4
CT	25.6	32.3
C_2T	13.0	13.0
C_3T	7.0	6.5
CT_2	8.4	9.7
C_2T_2	7.5	8.7
C_3T_2	2.8	4.0
CT_3	2.6	4.6
C_2T_2	3.4	2.7

The relative frequency of distribution of 5' Py-C-Pu 3' in pyrimidine tracts has been calculated from the data in Table VII. Values for the quantity of 3' terminal cytosines are obtained by dividing the relative frequency of cytosines in an oligonucleotide by the length of the oligonucleotide.

Allowance for the inequality of occurrence of T-C and C-T (Swartz et al., 1962) has been made in the case of the CT pyrimidine tract.

(i) $\overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \underline{1}\ \underline{1}\ \underline{1} \rightarrow \overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \underline{1}$

(ii) $\overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \underline{1}\ \underline{1}\ \underline{1} \rightarrow \overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \overset{*}{\underline{1}}\ \underline{1}\ \underline{1}\ \underline{1}$

* = CH_3

1 = Methylation Site

→ = Viral Transformation

Possibility (ii) could operate by the amplification of (say) 5 MeC rich satellite DNA as a whole, or by amplification of relatively short (4 - 10 nucleotides long) sections of DNA.

The two mechanisms could be distinguished by precise quantitation of the 5 MeC content of the DNA and the number of Hpa II cleavage sites unmethylated per genome. Alternatively the chromosomes of each cell line could be examined for the presence of 5 MeC rich satellites (Miao & Schildkraut, 1968; Miller, 1974).

4. DNA SEQUENCES METHYLATED IN VITRO.

The distribution of 5 MeC in pyrimidine tracts methylated *in vitro* whether on single stranded *E. Coli* DNA or double stranded calf thymus DNA was basically the same as *in vivo* (although there is some variation depending on the template used (Table XIII).

The data suggest the enzyme purification has isolated a 'bona fide' DNA methylase and has not merely selected one activity from a spectrum of enzymes. The fact that the 5 MeC distribution is similar whether the substrate is native or denatured argues against a role for secondary structural features (except perhaps transient ones) in determining the position of methylation and implies that nucleotide sequence alone is important.

The distribution of 5 MeC amongst pyrimidine tracts gives

incomplete information about the nucleotides neighbouring 5 MeC, and it is possible that all the 5 MeC is 3' terminal on oligo-pyrimidine tracts and only sequences 3' to C are relevant in specificity (Table XV). There may then be a number of methylation sites each with a corresponding methylase, i.e. there will be a number of different methylases which will be indistinguishable by pyrimidine tract analysis. The ability of the Krebs II ascites methylase to completely compensate for the under-methylation of L929 DNA produced by methionine starvation (Turnbull and Adams, 1976) and the fact that the methylase used is highly purified (by salt extraction, separation by size on an Ultrogel column, elution from phosphocellulose and DEAE cellulose) and is still able to methylate all the pyrimidine tracts in the same manner as in vivo suggests that this enzyme activity represents the only DNA methylase in these cells. The possibility that there are a number of methylases which have co-purified (and are thus very similar) cannot be excluded. The presence of other methylases which contribute only a small percentage of the total 5 MeC of the cell's DNA is also possible.

If there is only one DNA methylase in the Krebs II ascites cells, bearing in mind the relative lack of tissue or species specificity of 5 MeC distribution in pyrimidine tracts (Tables VI to IX), it is possible that there is only one type of DNA methylase in all vertebrate cells.

If there is only one DNA methylase enzyme present in each vertebrate cell which is able to accomplish all the necessary methylations then additional factors must be employed to accommodate variation of 5 MeC levels in development (Holliday and Pugh, 1975; Riggs, 1975). The enzyme isolated from Krebs II ascites cells

(Turnbull and Adams, 1976), preferentially methylates semi-methylated DNA and would therefore be able to 'maintain' the level of 5 MeC in progeny DNA after cell division but would be unlikely to methylate new sites as it is unable to methylate 'normally' methylated L929 DNA to any great extent. To allow methylation of new sites a new enzyme (either the 'maintenance' enzyme ± a regulatory subunit(s) or a completely new enzyme) would be needed. Its presence would however be only transient; after development only the 'maintenance' enzyme would be detectable. Thus in the cultured cells used in these studies (including the Krebs II ascites cells) one would not expect to find a number of different DNA methylases. A more likely source of regulatory elements and/or enzymes would be an embryonic tissue, e.g. Xenopus embryos.

5. THE NATURE OF THE METHYLATION SPECIFYING SITE IN VERTEBRATE DNA.

The fingerprinting data obtained from analysis of DNA methylated in vivo and in vitro (Experimental, sections 2 and 4) suggest that the features specifying the cytosines to be methylated have been highly conserved in vertebrates. If only one DNA methylase is involved (at least for the great majority of the methylations) one might expect some feature of the polynucleotide sequence to emerge.

For the purposes of the discussion the methylation recognition site will be considered to consist of an unknown nucleotide sequence which is essential for methylation to occur, e.g. the methylase binding site on DNA. The methylation site itself is considered to be the methylated cytosine plus a small number of nucleotides both 5' and 3' to the cytosine. The two may or may not be identical.

Much of the 5 MeC present in vertebrate DNA has been found in association with the C-G doublet (Introduction, section 3). This has led to speculation that the C-G doublet might constitute the recognition and methylation site. It is unlikely that the C-G doublet is the only determinant of methylation as (a) methylation occurs as frequently in the C-T doublet as the T-C doublet and occurs at a substantial level internally and 5' to many other pyrimidine tracts (Sneider, 1972) (b) purified methylases are able to methylate only a very few of the cytosines in a substrate such as $(dC\text{-}dG)_n$. $(dC\text{-}dG)_n$ or $(dG,dC)_n$ (Roy and Weissbach, 1975; Sneider et al., 1975) (c) the 5 MeC content of DNA is tissue specific and thus requires some selection of methylation other than that determined by sequence alone, i.e. a regulatory element(s) of some kind. (d) Data from nearest neighbour analysis of DNA indicate that in mouse liver DNA 2.2% of total C occurs in the C-G nucleotide doublet (Swartz et al., 1962), whereas the total 5 MeC content of mouse liver cells is 1.0 mole % i.e. 5% of Cs are methylated (Vanyushin et al., 1970).

Roy and Weisbach (1975) suggested that the methylation site nucleotide sequence is either 5' C-C-G 3' or 5' G-C-G 3'. This model can be applied to the HS-β satellite DNA (Fry et al., 1973), however saturation studies on artificial templates in vitro, e.g. $(dC\text{-}dG)_n$. $(dC\text{-}dG)_n$ and $(dG,dC)_n$ indicate that the recognition site, at least, is more complex. If the sequences 5' C-C-G 3' or 5' G-C-G 3' were the only determinants of methylation neither of the nucleotide doublets in the TC pyrimidine isostich could be methylated; this has in fact been shown to be methylated in vivo (Experimental, section 2 and Sneider, 1971) and in vitro (Experimental, section 4).

Turnbull and Adams (1976) showed that the Krebs II ascites methylase is most active on native undermethylated L929 DNA; similarly Sneider et al., (1975) noted a stimulation of the rate of methylation when they used undermethylated rat hepatoma DNA. The undermethylation of both these DNAs was achieved by growing synchronised cells in methionine free medium for a short time, the newly replicated DNA is then deficient in 5 MeC, thus the DNA substrate used in these methylations in vitro is probably methylated in one strand only (parental), the opposite (progeny) strand being largely unmethylated. It appears that mammalian methylases may be similar to prokaryotic methylases in that the preferred substrate; has a methyl group on one DNA strand but not the other, i.e. is semi-methylated (Arber, 1974).

The methylation site could then consist of two C-G nucleotide doublets in opposing strands of the DNA

5' ——————*C-G——————— 3'

3' ——————G-C*——————— 5'

If sequences immediately 3' to a C-G doublet were not involved in specifying methylation, sequences 5' to C-G in the opposite strand would not be involved (as a result of base pairing). It may well be that the bases immediately adjacent to a cytosine are unimportant or at least not stringently required in defining the methylation site. There are precedents for both these situations in nucleic acid/protein interactions in prokaryotic systems. The Hind II restriction modification enzyme recognition site is degenerate, the nucleotide sequence is GTPyPuAC, thus several similar sequences may specify recognition (Nathans and Smith, 1975) and the endonuclease from Haemophylus parahaemolyticus cleaves DNA at an AT or TA base

pair but the recognition site is remote from the site of cleavage (Kleid et al., 1976).

Thus far the only known complete methylated sequence in mammalian DNA is the HS-β satellite (Fry et al., 1973):

5' ACACAG*CGGG 3'

3' TGTGTCG*CCC 5'

* = probable position of CH_3 group.

The basic repeat (above) is 10 nucleotides long; i.e. one turn of the DNA helix. However more than 50% of the satellite consists of variants of this basic unit which are 11 or 12 nucleotides long. Thus a model recognition site remote, e.g. one helical turn, from the actual methylated base would have to allow some flexibility.

The authors did not identify where 5 MeC occurred in this sequence but proposed that those cytosines marked with an asterisk (above) were methylated. Examination of the nucleotide sequences adjacent to these cytosines in each strand reveal that each C*G doublet is preceded by a repeated nucleotide doublet, viz:

5' ACAC CA C G 3'

5' GTGT CC C G 3'

which can be generalised to:

5' Pu Py Pu Py N_n CG 3'

n = 2,3 or 4 and represents the area of inhomogeneity in the satellite sequence. A preliminary search for such sequences in mammalian DNA complementary to mammalian tRNA and 5sRNA species (Barrell and Clark, 1974); mouse immunoglobulin mRNA (Milstein et al., 1974) and globin mRNA (Proudfoot and Brownlee, 1974) revealed no special feature to distinguish the nucleotide sequence 8 nucleotides 3' or 5' to the C-G doublets in DNA (nor was there any specificity in the 2

nucleotides one turn of the DNA helix, 5' or 3' remote from a C-G doublet): these sequances may of course not be representative of the whole genome. In addition this system would not allow methylation of the dinucleotides C-A or C-Py, which have been shown to be methylated.

Alternatively the recognition site could consist of a mono-pyrimidine C opposite a 3' terminal C on an oligopyrimidine, viz:

5' N G $\overset{*}{C}$ G G N 3'

3' N C G C C N 5'
↑

This is consistent with the approximate figure of 50% for 5 MeC found in monopyrimidines, but does not allow methylation of single stranded DNA in vitro unless the specificity is altered, nor methylation of the Hpa II restriction endonuclease cleavage site (Experimental, section 7).

The nature of the characteristic(s) determining the site of methylation of vertebrate DNA is thus as yet unknown: it is possible to consider that during the course of evolution an original remote recognition site plus an otherwise randomly placed cytosine has been evolving into a recognition site plus a non randomly placed cytosine (i.e. most often 5' to a G) with a number of possibilities still available.

Using the values obtained for the saturation of DNA in vitro by Roy and Weissbach (1975) it is possible to estimate the probable number of bases specifying the DNA methylase recognition site. The single stranded M. luteus DNA is methylated in vitro once per 116 bases and E. Coli DNA once per 230 bases; assuming a random base distribution the site should consist of approximately 4 bases. (If these bases are remote from the actual methylated base the problem of their identification will be increased).

If the DNA methylase is able to recognise and methylate one or

a small number of DNA sequences it should be possible to determine the general characteristics of these sequences by fingerprinting the products of a partial digestion of L(methyl-^{14}C) methionine labelled DNA(either limiting the amount of DNase in the DNA digests or modifying the DNA using a carbodiimide before digestion Ho & Gilham, 1974). One might expect to find a number of spots belonging to a homologous series if the recognition and methylation sites are coincident; a random distribution of 5 MeC might indicate that the recognition site is remote from the methylation site.

6. THE DISTRIBUTION OF 5 MeC IN THE VERTEBRATE GENOME,

(i) Base Composition.

Data from Actinomycin D/CsCl gradients and thermal elution from hydroxylapatite (Experimental, section 5) indicate that there is a degree of species specificity in the distribution of 5 MeC with respect to A+T and G+C rich fractions of the genome. This information cannot easily be related to the distribution of 5 MeC amongst pyrimidine tracts (Experimental, section 2) as the size of DNA being considered is different.

The distribution of 5 MeC in HeLa and BHK-21/C13 DNA is similar to the distribution in Chinese hamster DON cell line (Comings, 1972), i.e. 5 MeC is biased towards the G+C rich section of DNA. The fact that the 5 MeC distribution in the mouse L929 and SVT2 cells is not biased towards G+C rich DNA implies that interpretation of 5 MeC distribution in gross evolutionary terms (Comings, 1972) may be misleading.

Mouse DNA is known to contain an A+T rich satellite (Introduction, section 4.1.) which is highly methylated. This could account at least in part for the bias of 5 MeC toward A+T

rich DNA. Corneo et al., (1970, 1972) have shown human DNA to contain a number of DNA satellites either G+C or A+T rich. With this in mind it is possible that a biased distribution of 5 MeC reflects the methylation of satellite DNA's, or species specific repetitive sequences, imposed upon an otherwise random distribution of 5 MeC.

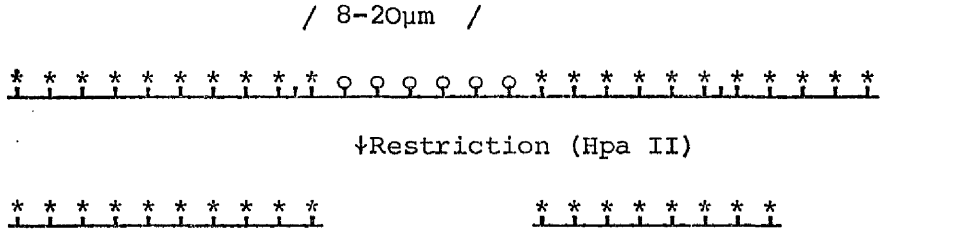
(ii) Rapidly Reannealing DNA.

Rapidly reannealing DNA from HeLa cells was shown to be enriched for 5 MeC (Experimental, section 6). HeLa DNA thus shares this characteristic with mouse and Chinese hamster DNA (Brahic and Fraser, 1971; Schneiderman and Billen, 1973) and suggests that this feature has been conserved during the course of evolution. It is not possible to assign a role for this fraction of DNA, However,its evolutionary conservation suggests it must have some important function (possibilities have been outlined in Sections 4 and 5 of the Introduction).

(iii) Undermethylated Regions in DNA.

The digestion of BHK-21/C13 and BHK-21/PyY DNA by the Hpa II restriction endonuclease reveals considerable resistance to cleavage (Experimental, section 7). This suggests that the majority of Hpa II restriction endonuclease cleavage sites are methylated in vivo but that certain cleavage sites,located in a highly specific manner, are unmethylated. Molitor et al., (1976) have detected regions of DNA (8 - 20μm long) in L929 cells which are thought to be unmethylated. If we consider the sequence 5' C-C-G-G 3' to be randomly distributed throughout these unmethylated regions the DNA digested' by the Hpa II restriction

endonuclease should be unmethylated:



← unmethylated fragments.

♀ = unmethylated 5' C-C-G-G 3'

* = methylated regions of DNA

__ = double stranded DNA

The fact that the short fragments of BHK-21/PyY DNA are normally methylated suggests that they are not related to the long unmethylated regions of Molitor et al., (1976).

The existence of specific unmethylated sequences in DNA implies rigorous control of the operation of the DNA methylase system in the cell to enable some Hpa II restriction endonuclease cleavage sites to be methylated while others are not.

CONCLUSION

The results obtained by fingerprinting vertebrate DNA methylated *in vivo* and *in vitro* suggest the existence of only one type of DNA methylase in vertebrate cells which is able to methylate DNA at specific sites, the nature of which has been highly conserved throughout the evolution of the vertebrates: 'Control' elements are necessary to allow variation in the distribution, quantitative or qualitative, of 5 MeC. Species specificity in the distribution of 5 MeC in longer sequences (> 200 bases) detected by thermal elution from hydroxylapatite, Actinomycin D/CsCl gradients and Hpa II endonuclease digestion implies that the short methylation sites are specifically interspersed within longer sequences. This specific distribution of 5 MeC may be related to some function of these sequences, however without understanding the role of satellite, rapidly reannealing or Hpa II restriction endonuclease sensitive regions it is not yet possible to determine the reasons for the distribution of 5 MeC.

REFERENCES

Aaronson, S.A. & Todaro, G.J. (1968). Science 162, 1024.

Adams, R.L.P. (1973). Nature 244, 27.

Adams, R.L.P. (1974). Biochim. Biophys. Acta 335, 365.

Arber, W. (1974). Prog. Nucl. Acid Res. & Mol. Biol. 14, 1.

Avery, O.T., MacLeod, C.M. & McCarthy, M. (1944). J. Exp. Med. 79, 137.

Barrell, B.G. & Clark, B.F.C. (1974). In Handbook of Nucleic Acid Sequences, Joynson-Bruvvers Ltd., Oxford.

Berdyshev, G.D., Korotaev, G.K., Boyarskikh, G.V. & Vanyushin, B.F. (1967). Biokhimiya 32, 988.

Bernardi, G. (1965). Nature 206, 709.

Billen, D. (1968). J. Mol. Biol. 31, 477.

Birnboim, H.C., Mitchel, R.E.J. & Strauss, N.A. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 2189.

Birnie, G.D., MacPhail, E., Young, B.D., Getz, M.J. & Paul, J. (1974). Cell Differentiation 3, 221.

Birnstiel, M., Telford, J., Weinberg, E. & Stafford, D. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 2900.

Brahic, M. & Fraser, M.J. (1971). Biochim. Biophys. Acta 240, 23.

Brawwerman, G. & Chargaff, E. (1951). J. Amer. Chem. Soc. 73, 4052.

Britten, R.J. & Davidson, E.H. (1969). Science 165, 349.

Britten, R.J. & Kohne, D.E. (1968). Science 161, 529.

Brockes, J.P., Brown, P.R. & Murray, K. (1974). J. Mol. Biol. 88, 437.

Brownlee, G.G. & Sanger, F, (1969). Europ. J. Biochem. 11, 395.

Burdon, R.H. & Adams, R.L.P. (1969). Biochim. Biophys. Acta 174, 332.

Burdon, R.H. & Douglas, J.T. (1974). Nucl. Acid Res. 1, 97.

Burdon, R.H., Martin, B.T. & Lal, B.M. (1967). J. Mol. Biol. 28, 357.

Burgi, E. & Hershey, A.D. (1963). Biophys. J. 3, 309.

Burton, K. & Petersen, G.B. (1960). Biochem. J. 75, 17.

Busby, D.W.G., House, W. & McDonald, J.R. (1964). In Virological Techniques, p.218, Churchill, London.

Chang, R.S. (1954). Proc. Soc. Exp. Biol. Med. 87, 440.

Cohen, A.K., Huh, T.Y. & Helleiner, C.W. (1973). Canad. J. Biochem. 51, 529.

Comings, D.E. (1972). Exp. Cell Res. 74, 383.

Corneo, G., Ginelli, E. & Polli, E. (1970). J. Mol. Biol. 48, 319.

Corneo, G., Zardi, L. & Polli, E. (1972). Biochim. Biophys. Acta 269, 201.

Crick, F.H. (1970). Nature 227, 561.

Culp, L.A. & Black, P.H. (1971). Biochim. Biophys. Acta 247, 220.

Cummings, D.J., Tait, A. & Goddard, J. (1974). Biochim. Biophys. Acta 374, 1.

Davidson, E.H., Hough, B.R., Amenson, C.S. & Britten, R.J. (1973). J. Mol. Biol. 77, 1.

Dawid, I.B. (1974). Science 184, 80.

Dawid, I.B., Brown, D.D. & Reeder, R.H. (1970). J. Mol. Biol. 51, 341.

Desai, L.S., Wulff, U.C. & Foley, G.E. (1971). Exp. Cell Res. 65, 260.

Doskocil, J. & Sorm, F. (1961). Coll. Czech. Chem. Comm. 26, 2739.

Doskocil, J. & Sorm, F. (1962). Biochim. Biophys. Acta 55, 953.

Doskocil, J. & Sormova, Z. (1965a). Biochim. Biophys. Acta 95, 513.

Doskocil, J. & Sormova, Z. (1965b). Biochim. Biophys. Res. Comm. 20, 334.

Drahovsky, D. & Morris, N.R. (1971a). J. Mol. Biol. 57, 475.

Drahovsky, D. & Morris, N.R. (1971b). J. Mol. Biol. 61, 343.

Drahovsky, D. & Morris N.R. (1972). Biochim, Biophys. Acta 277, 245.

Dugaiczyk, A., Hedgpeth, J., Boyer, H.W. & Goodman, H.M. (1974). Biochemistry, 13, 501.

Dulbecco, R. & Vogt, M. (1954). J. Exp. Med. 99, 167.

Dunn, D.B. & Smith, J.D. (1958). Biochem. J. 68, 627.

Earle, W.R. (1943). J. Nat. Cancer Inst. 4, 165.

Ehrlich, M., Ehrlich, K. & Mayo, J.A. (1975). Biochim. Biophys. Acta 395, 109.

Eskin, B. & Linn, S. (1972). J. Biol. Chem. 247, 6183.

Evans, H.H. & Evans, T.E. (1970). J. Biol. Chem. 245, 6436.

Evans, H.H., Evans, T.E. & Littman, S. (1973). J. Mol. Biol. 74, 563.

Flamm, W.G., McCallum, M. & Walker, P.M.B. (1967). Proc. Nat. Acad. Sci. U.S.A. 57, 1729.

Flamm, W.G., Walker, P.M.B. & McCallum, M. (1969). J. Mol. Biol. 40, 423.

Fry, K., Poon, R., Whitcombe, P., Idriss, J., Salser, W., Mazrimas, J. & Hatch, F. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 2642.

Fujimoto, D., Srinivasan, P.R. & Borek, E. (1965). Biochemistry 4, 2849.

Gelderman, A., Rake, A.V. & Britten, R.J. (1971). Proc. Nat. Acad. Sci. U.S.A. 68, 172.

Gey, G.O., Coffman, W.D. & Kubicek, M.T. (1951). Cancer Res. 12, 264.

Gierer, A. & Schramm, G. (1956). Nature 177, 702.

Gilbert, W. & Maxam, A. (1973). Proc. Nat. Acad. Sci. U.S.A. 70, 3581.

Gill, J.E., Mazrimas, J.A. & Bishop, C.C. (1974). Biochim. Biophys. Acta 335, 330.

Gorovsky, M.A., Hattman, S. & Pleger, G.L. (1973). J. Cell Biol. 56, 697.

Grady, L.J. & Campbell, W.P. (1973). Nature New Biol. 243, 195.

Green, H. & Neuth, M. (1974). Cell 3, 127.

Greene, P.J. Betlach, M.C. & Boyer, H.W. (1974). Methods Mol. Biol. 7, 87.

Grippo, P., Iaccarino, M., Parisi, E. & Scarano, E. (1968). J. Mol. Biol. 36, 195.

Grippo, P., Parisi, E., Larcstia, C. & Scarano, E. (1970). Biochem. 9, 2605.

Gurdon, J.B. & Laskey, R.A. (1970). J. Embryo. Exp. Morph. 24, 227.

Guseinov, V.A., Mazin, A.L., Vanyushin, B.F. & Belozerskii, A.N. (1972). Biokimiya 37, 312.

Harbers, K., Harbers, B. & Spencer, J.H. (1975). Biochem. Biophys. Res. Comm. 66, 738.

Hayward, G. (1972). Virology 49, 342.

Hattman, S. & Cousens, L. (1972). J. Bact. 112, 1428.

Hell, A., Birnie, G.D., Slimming, T.K. & Paul, J. (1972). Anal. Biochem. 48, 369.

Hershey, A.D. & Chase, M. (1952). J. Gen. Physiol. 36, 39.

Hilliard, J.K. & Sneider, T.W. (1975). Nucl. Acid. Res. 2, 809.

Holliday, R. & Pugh, J.E. (1975). Science 187, 226.

Hotchkiss, R.D. (1948). J. Biol. Chem. 175, 315.

Holmes, D.S. & Bonner, J. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 1103.

Hough, B.R., Smith, M.J., Britten, R.J. & Davidson, E.H. (1975). Cell 5, 291.

Ho, N.W.Y. & Gilham, P.T. (1974). Biochemistry 13, 1082.

Humphries, S., Windass, J. & Williamson, R. (1976). Cell 7 267.

Johnson, T.B. & Coghill, R.D. (1925). J. Amer. Chem. Soc. 47, 2838.

Jones, A.S., Stacey, M. & Watson, B.E. (1957). J. Chem. Soc. p. 2454.

Johes, A.S. & Walker, R.T. (1963). J. Gen. Microbiol. 31, 187.

Kalousek, F. & Morris, N.R. (1968). J. Biol. Chem. 243, 2440.

Kalousek, F. & Morris, N.R. (1969a). J. Biol. Chem. 244, 1157.

Kalousek, F. & Morris, N.R. (1969b). Science 164, 721.

Kappler, J.W. (1971). J. Cell. Physiol. 78, 33.

Kaye, H. & Wincour, E. (1967). J. Mol. Biol. 24, 475.

Kit, S. (1961). J. Mol. Biol. 3, 711.

Kleid, D., Humayun, Z., Jeffrey, A. & Ptashne, M. (1976). Proc. Nat. Acad. Sci. U.S.A. 73, 293.

Lark, C. (1968). J. Mol. Biol. 31, 389.

Lautenberger, J.A. & Linn, S. (1972). J. Biol. Chem. 247, 6176.

Lawley, P.D., Carthorn, A.R., Shah, S.A. & Smith, B. (1972). Biochem. J. 128, 133.

Lewin, B. (1975). Cell 4, 77.

Ling, V. (1972). J. Mol. Biol. 64, 87.

Low, M., Hay, J. & Keir, H.M. (1969). J. Mol.Biol. 46, 205.

McClintock, P.R. & Papaconstantinou, J. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 4551.

MacPherson, I.A. & Stoker, M.G.P. (1962). Virology 16, 147.

Maden, B.E.H., Salim, M. & Summers, D.F. (1972). Nature New Biol. 237, 5.

Maio, J.J. & Schildkraut, C.L. (1968). J. Mol.Biol. 40, 203.

Meisher, F. (1868). Page 1. In The Nucleic Acids Vol. 1 (1955). J.N. Davidson & E. Chargaff (Eds. E. Chargaff & J.N. Davidson) Acad. Press, New York.

Miller, C.O. (1968). In Biochemistry and Physiology of Plant Growth Substances, 33pp (Eds. Wightman, F. & Setterfield, G.) Runge Press, Ottawa.

Miller, O., Schnedl, W., Allen, J. & Erlanger, B.F. (1974). Nature 251, 636.

Milstein, C., Brownlee, G.G., Cartwright, E.M., Jarviss, J.M. & Proudfoot, N.J. (1974). Nature 252, 354.

Morris, N.R. & Pih, K.D. (1971). Cancer Res. 31, 433.

Murray, K. (1970). Biochem J. 118, 831.

Nass, M. (1973). J. Mol. Biol. 80, 155.

Nathans, D. & Smith, H.O. (1975). Ann. Rev. Biochem. 44, 273.

Nelson-Rees, W.A. & Flandermeyer, R.R. (1976). Science 191, 96.

Oda, K. & Marmur, J. (1966). Biochemistry 5, 761.

Pagoulatos, G.N. (1974). Proc. Nat. Acad. Sci. U.S.A. 71, 3664.

Penman, S. (1969). In Fundamental Techniques in Virology (Eds. Habel, K. & Salzman, N.P.) Ch.5, p35. Acad. Press, New York.

Proudfoot, N.J. & Brownlee, G.G. (1974). Nature 252, 359.

Riggs, A.D. (1975). Cytogenet. Cell Genet. 14, 9.

Roy, P.H. & Weissbach, A. (1972). Nucl. Acid Res. 2, 1669.

Rubery, E.D. & Newton, A.A. (1973). Biochim. Biophys. Acta 324, 24.

Salomon, R., Kaye, A.M. & Herzberg, M. (1969). J. Mol. Biol. 43, 581.

Sanford, K.K. Earle, W.R. & Likely, G.D. (1948). J. Nat. Cancer Inst. 9, 229.

Sanger, F., Brownlee, G.G. & Barrell, B.G. (1965). J. Mol. Biol. 13, 373.

Scheid, B. , Srinivasan, P.R. & Borek, E. (1968). Biochemistry 7, 280.

Schneiderman, M.H. & Billen, D. (1973). Biochim. Biophys. Acta 308, 352.

Shapiro, H.S. & Chargaff, E. (1960). Biochim, Biophys. Acta 39, 68.

Sharp, P.A., Sugden, B. & Sambrook, J. (1973). Biochemistry, 12, 3055.

Shatin, A. (1969). In Fundamental Techniques in Virology (Habel, K. & Salzman, N.P. eds.) p.496. Acad. Press, New York.

Shenkin, A. & Burdon, R.H. (1974). J. Mol. Biol. 85, 19.

Sinsheimer, R.L. (1955). J. Biol. Chem. 215, 579.

Smith, M.J., Hough, B.R., Chamberlain, H.E., Davison, E.H. (1974). J. Mol. Biol. 85, 103.

Sneider, T.W. (1971). J. Biol. Chem. 246, 4774.

Sneider, T.W. (1972). J. Biol. Chem. 247, 2872.

Sneider, T.W. (1973). J. Mol. Biol. 79, 731.

Sneider, T.W. & Potter, V.R. (1969). J. Mol. Biol. 42, 271.

Sneider, T.W., Teague, W.M. & Rogachevsky, L.M. (1975). Nucl. Acid Res. 2, 1685.

Sobel, H.M. (1974). Scientific American 231, (2), 82.

Southern, E.M. (1970). Nature 227, 794.

Southern, E.M. (1971). Nature New Biol. 232, 82.

Spencer, J.H. & Chargaff, E. (1963). Biochim. Biophys. Acta 68, 18.

Stoker, M. (1964). Virology 24, 165.

Studier, F.W. (1965). J. Mol. Biol. 11, 373.

Sutton, W.D. (1971). Biochim. Biophys. Acta 240, 522.

Sutton, W.D. & McCallum, M. (1971). Nature New Biol. 232 83.

Swartz, M.N., Trautner, T.A. & Kornberg,A. (1962). J. Biol. Chem. 237, 1961.

Tamm, C., Shapiro, H.S., Lipshitz, R. & Chargaff, E. (1953). J. Biol. Chem. 203, 673.

Tewari, K.K. & Wildman, S.G. (1966). Science 153, 1269.

Tosi, L., Granieri, A. & Scarano, E. (1972). Exp. Cell Res. 72, 257.

Turnbull, J. & Adams, R.L.P. (1976). Nucl. Acid Res. 3, 667.

Vanyushin, B.F. (1968). Uspekhi Sovremennoi Biologii 65, 163.

Vanyushin, B.F., Kiryanov, G.I., Kurdryashova, I.B. & Belozersky, A.N. (1971). F.E.B.S. Lettr. 15, 313.

Vanyushin, B.F., Kökurina, N.A. & Belozersky, A.N. (1965). Dokl. An. S.S.S.R. 161, 1451.

Vanyushin, B.F., Mazin, A.L., Vasilyev, V.K. & Belozersky, A.N. (1973). Biochim. Biophys. Acta 299, 397.

Vanyushin, B.F., Tkacheva, S.G. & Belozersky, A.N. (1970). Nature 225, 948.

Walker, P.M.B. & McLaren, A. (1965). Nature 208, 1175.

Waring, M. & Britten, R.J. (1966). Science 154, 791.

Watson, J.D. & Crick, F.H. (1953). Nature 171, 737.

Wu, R. & Geiduschek, P. (1975). J. Mol. Biol. 96, 513.

Wyatt, G.R. (1951). Biochem. J. 48, 584.

Wyatt, G.R. & Cohen, S.S. (1953). Biochem, J. 55, 774.