



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>  
[research-enlighten@glasgow.ac.uk](mailto:research-enlighten@glasgow.ac.uk)

CHROMATIN STRUCTURE AND DNA METHYLATION

By

TERENCE DAVIS

Thesis submitted to the University of Glasgow  
for the degree of Doctor of Philosophy

Department of Biochemistry,  
The University of Glasgow,  
Glasgow G12 8QQ.

August, 1984.

ProQuest Number: 10391224

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10391224

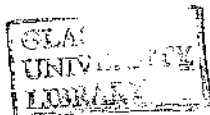
Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

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

Thesis  
7000  
copy 2



ACKNOWLEDGMENTS

I would like to express my sincere thanks to the following people:

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.L.P. Adams for his guidance, criticism and patience throughout the course of this work.

Dr R. Eason for his help with the computer analyses.

All members of Lab. A-20 for helpful suggestions.

The staff of the Wellcome Cell Culture unit and the Animal House for their cooperation and help.

Mrs L. Peedle for typing this manuscript.

I acknowledge the financial support of the Medical Research Council.

ABBREVIATIONS

The abbreviations used in this thesis are in agreement with the recommendations of the editors of the Biochemical Journal (Biochem. J. (1983) 209, 1-27), except the following.

T	Thymine
A	Adenine
C	Cytosine
G	Guanine
m <sup>5</sup> C	5-methylcytosine
m <sup>4</sup> C	N <sup>4</sup> -methylcytosine
m <sup>6</sup> A(6MeA)	N <sup>6</sup> -methylcytosine
m <sup>7</sup> G	7-methylguanine
Py	Pyrimidine
Pu	Purine
L929	Mouse fibroblast cells
Laza	Mouse fibroblasts treated with 5-azadeoxycytidine
SDS	Sodiumdodecylsulphate
PMSF	Phenylmethylsulphonylfluoride
TCA	Trichloroacetic acid
AdoMet(SAM)	S-adenosylmethionine
MPE	Methidiumpropyl-EDTA.iron(II)

<u>LIST OF CONTENTS</u>	<u>Page</u>
Title	i
Acknowledgements	ii
Abbreviations	iii
Contents	iv
List of Figures	x
List of Tables	xiii
Summary	xv
I. INTRODUCTION	1
I.1. The presence of methylated bases in DNA	2
I.1.1. Methylated bases in bacterial DNA	2
I.1.2. Methylated bases in viral DNA	2
I.1.3. Methylated bases in eukaryotic DNA	4
I.1.4. DNA methylation as a post-synthetic event	5
I.1.5. The stability of methylated bases in DNA	5
I.1.6. The effect of methylation on DNA stability	6
I.1.7. The effect of methylation on apparent base composition	8
I.1.8. The distribution of $m^5C$ in the eukaryotic genome	9
I.2. DNA methylases	10
I.2.1. Prokaryotic DNA methylases	10
I.2.1.1. Methylases involved in restriction- modification	10
I.2.1.2. The <u>dam</u> and <u>dcm</u> methylases	12
I.2.2. Eukaryotic DNA methylases	13
I.2.2.1. Occurrence and properties	13
I.2.2.2. The sequence specificity of eukaryotic DNA methylases	15
I.3. Possible functions for DNA methylation in prokaryotes	16
I.3.1. A role for methylation in DNA replication	16
I.3.2. DNA methylation and repair	17
I.3.3. DNA methylation and recombination	18
I.3.4. Methylation and the Mu-Mom gene	18
I.4. Possible functions of DNA methylation in eukaryotes	19

	<u>Page</u>
I.4.1. DNA methylation and gene expression	19
I.4.1.1. General introduction	19
I.4.1.2. The use of inhibitors of DNA methylation	20
I.4.1.3. The effects of <u>in vitro</u> methylation on gene expression	21
I.4.2. DNA methylation and cellular differentiation	22
I.4.3. DNA methylation and X-chromosome inactivation	23
I.4.4. The effect of methylation on viral expression	24
I.4.5. The maternal inheritance of DNA in <u>Chlamydomonas</u> chloroplasts	25
I.4.6. DNA methylation and carcinogenesis	25
I.4.7. Aging	26
I.5. Chromatin structure	26
I.5.1. The nucleosome concept	26
I.5.2. The nucleosome 'core'	27
I.5.3. The internal structure of the 'core'	28
I.5.4. The role of histone H1	28
I.5.5. Nucleosome phasing	29
I.5.6. The structure of transcriptionally active chromatin	31
I.5.7. Chromosomal scaffold	33
I.5.8. Nuclear matrix	33
I.5.8.1. DNA sequences associated with the nuclear matrix	34
I.6. Chromatin structure and DNA methylation	35
<u>Aims</u>	36
II. MATERIALS AND METHODS	37
II.1. MATERIALS	38
II.1.1. Biological	38
II.1.2. Chemical	38
II.1.3. Radiochemicals	40
II.1.4. Cell culture solutions	40
II.1.5. Stock buffers	40
II.2. METHODS	44



	<u>Page</u>
II.2.1. Cell culture techniques	44
II.2.1.1. Growth of cells	44
II.2.1.2. Contamination checks	44
II.2.1.3. Labelling of cells	45
II.2.1.4. Harvesting of cells	45
II.2.1.5. Preparation of nuclei	46
II.2.1.6. Preparation of a cell lysate	46
II.2.2. <u>In vitro</u> methylation	46
II.2.2.1. Methylation of nuclei	46
II.2.2.2. Methylation of chromatin and DNA	46
II.2.3. Digestion of nuclei using nucleases	46
II.2.3.1. Micrococcal nuclease	46
II.2.3.2. Pancreatic deoxyribonuclease (DNase I)	47
II.2.4. Purification of DNA for scintillation counting	47
II.2.5. Double label technique	48
II.2.6. Sanders' extraction procedure	48
II.2.7. Digestion of nuclei using methidium- propyl-EDTA.iron (II)	50
II.2.8. Preparation of nuclear non-histone proteins	50
II.2.9. Purification of DNA	51
II.2.10. Nick-translation of DNA	51
II.2.11. Base composition analysis of DNA	51
II.2.11.1. By enzymatic digestion	51
II.2.11.2. By acid pyrolysis	52
II.2.12. Agarose gel electrophoresis	52
II.2.13. Acrylamide gel electrophoresis	52
II.2.14. Computer analyses	53
III. RESULTS	54
III.1. <u>In vivo</u> methylation of DNA in L929 cells relative to nucleosomes	55
III.1.1. The distribution of $m^5C$ in chromatin	55
III.1.2. Core DNA of L929 chromatin is enriched in $m^5C$	55
III.1.3. Calculation of the base composition of DNA	58

	<u>Page</u>
III.1.4. Digestion of DNA with micrococcal nuclease	59
III.1.5. Digestion of chromatin using methidiumpropyl-EDTA.iron(II)	60
III.1.6. Reconstitution of chromatin	62
III.1.7. Models predicting the locations of nucleosomes on DNA	64
III.1.7.1. Application of the model of Zhurkin to specific gene sequences	64
III.1.7.2. The variation in the C + G content along mouse genes	65
III.2. <u>In vitro</u> methylation of chromatin	71
III.2.1. Optimization of conditions for methylation	71
III.2.2. The sole product of methylation of DNA <u>in vitro</u> is m <sup>5</sup> C	75
III.2.3. Methylation of mosquito nuclei	75
III.2.4. Methylation of nuclei from stationary phase L929 cells	75
III.2.5. Methylation of nuclei from log phase L929 cells	81
III.2.6. Comparison of the methylation of chromatin and DNA	81
III.2.6.1. Core particles are poor acceptors of methyl groups	81
III.2.6.2. DNA is a better substrate than chromatin	89
III.2.7. Histones inhibit DNA methylation <u>in vitro</u>	89
III.2.8. Nuclear non-histone proteins do not inhibit DNA methylation <u>in vitro</u>	94
III.3. <u>In vivo</u> methylation of DNA in transcriptionally active chromatin of L929 cells	94
III.3.1. Use of DNase I to study transcribing chromatin	94
III.3.2. Sanders' extraction procedure for isolation of transcribing chromatin	96

	<u>Page</u>
III.3.3. <u>In vitro</u> transcription of nuclei from L929 cells	98
III.3.4. Active regions of chromatin are deficient in m <sup>5</sup> C	98
III.3.5. Location of DNA less than one hour old	104
III.4. Transcribing chromatin is preferentially methylated <u>in vitro</u>	105
III.4.1. Log phase L929 cells	105
III.4.2. Stationary phase L929 cells	107
III.4.3. Mosquito cells	108
III.4.4. DNA methylase is associated preferentially with SO	110
III.5. Delayed methylation	113
III.5.1. DNA methylation is not completed for several hours after DNA synthesis	113
III.5.2. Delayed methylation occurs predominantly in core DNA	114
III.5.3. Effect of salt extraction of nuclei on DNA methylation	116
III.5.4. DNA methylation is associated with the nuclear matrix	116
III.5.5. Effect of 5-azadeoxycytidine on <u>in vitro</u> methylation	118
III.5.6. The methylation in Laza nuclei is preferentially in core DNA	119
IV. DISCUSSION	122
IV.1. Base composition of nucleosomal core DNA	123
IV.1.1. Core DNA is enriched in C + G residues	123
IV.1.2. Problems inherent in these studies	123
IV.1.3. Reconstitution of chromatin	125
IV.1.4. Application of DNA flexibility models	125
IV.1.5. Implications of these studies	129
IV.1.6. Alternative explanations	131
IV.2. Methylation of chromatin and DNA <u>in vitro</u>	131
IV.2.1. Effect of chromatin structure on DNA methylation	131

	<u>Page</u>
IV.2.2. Effects of proteins on DNA methylation <u>in vitro</u>	134
IV.2.3. Conclusions	135
IV.3. Methylation of DNA in transcriptionally active chromatin	136
IV.3.1. Isolation of transcribing chromatin	136
IV.3.2. Evidence that the nucleosomes in S0 and S50 are derived from transcribing chromatin	136
IV.3.3. Characterization of the DNA in fractions S0 and S50	137
IV.3.4. Methylation of transcribing chromatin <u>in vitro</u>	138
IV.4. Studies on delayed methylation	140
IV.4.1. DNA methylation can be delayed	140
IV.4.2. Effect of salt extraction	141
IV.4.3. Methylation occurs in the nuclear matrix	142
IV.4.4. Studies using 5-azadeoxycytidine	143
IV.5. Conclusions and further work	144
REFERENCES	146

LIST OF FIGURES

	<u>Page</u>
Figure 1. Methylated bases in DNA	3
Figure 2a. The rotational isomers of N <sup>6</sup> -methyl-adenine	7
Figure 2b. Prediction of kinks in poly(dA.dT). poly(dA.dT)	7
Figure 3. Mechanism of action of DNA methylase	14
Figure 4. Calibration curve for the determination of the amount of <sup>14</sup> C in the <sup>3</sup> H channel of a liquid scintillation spectrometer	49
Figure 5. Agarose gel of DNA from total and nucleosomal core chromatin from L929 cells	56
Figure 6. High voltage electrophoresis of 3'-mono-nucleotides	57
Figure 7. Agarose gel of DNA from total and methidiumpropyl-EDTA.iron(II) digested chromatin	61
Figure 8. Agarose gel of DNA from reconstituted chromatin digested using micrococcal nuclease	63
Figure 9. Predicted flexibility of the $\beta$ 1-globin gene	66
Figure 9a. bases 1200 to 1500	67
Figure 9b. bases 121 to 421	67
Figure 9c. bases 50 to 1550	68
Figure 10. Plot of the variation in C + G content along mouse genes	69
Figure 11. Time course of methylation of mosquito nuclei using ascites methylase	72
Figure 12. Dependence of methylation on the concentration of AdoMet	73
Figure 13. Dependence of methylation on the concentration of DNA	73
Figure 14. Dependence of methylation on the amount of DNA methylase	73
Figure 15. The sole product of <u>in vitro</u> methylation using ascites methylase is m <sup>5</sup> C	76

	<u>Page</u>
Figure 16. Digestion of <u>in vitro</u> methylated nuclei of mosquito cells using micrococcal nuclease	78
Figure 17. Agarose gel of DNA from mosquito chromatin digested using micrococcal nuclease	80
Figure 18. Digestion of <u>in vitro</u> methylated nuclei from stationary phase L929 cells	82
Figure 19. Digestion of nuclei from log phase L929 cells after methylation using the endogenous methylase	84
Figure 20. Digestion of <u>in vitro</u> methylated nuclei from log phase L929 cells (methylated using ascites methylase)	85
Figure 21. Comparison of the methyl group acceptor ability of various chromatins from L929 cells	87
Figure 22. Linweaver-Burke plot of Figure 21	87
Figure 23. Methylation of DNA of core and total chromatin from mosquito cells	90
Figure 24. Methylation of DNA purified from total and core chromatin from L929 cells	91
Figure 25. Methylation of DNA purified from total and core chromatin of mosquito cells	92
Figure 26. Inhibition of DNA methylation by histones	93
Figure 27. Effect of nuclear non-histone proteins on DNA methylation <u>in vitro</u>	95
Figure 28. Digestion of <u>in vitro</u> methylated nuclei from mosquito cells using DNase I	97
Figure 29. Acrylamide gels of the proteins in the fractions of a Sanders' extraction	100

	<u>Page</u>
Figure 30. Agarose gel of the DNA present in the fractions of a Sanders' extraction	102
Figure 31. Agarose gel of the DNA present in the fractions of a Sanders' extraction using mosquito nuclei	109
Figure 32. Glycerol gradients of the fractions of a Sanders' extraction of nuclei from L929 cells	111
Figure 33. Delayed methylation occurs preferentially in nucleosomal core DNA	115
Figure 34. Effect of Salt extraction of nuclei on <u>in vitro</u> methylation	117
Figure 35. Digestion of <u>in vitro</u> methylated nuclei of L929 cells prelabelled with 5-azadeoxycytidine	120
Figure 36. Predicted flexibility of the <u>E.coli</u> <u>ilv</u> locus ( <u>ilvL</u> - <u>ilvE</u> )	127

<u>LIST OF TABLES</u>		<u>Page</u>
Table 1.	Properties of restriction-modification systems	11
Table 2.	Constituents of Eagles Minimal Essential Medium (MEM) as used in the Biochemistry Department, University of Glasgow	41
Table 3.	Constituents of Mitsuhashi and Maramorosch Basal medium	42
Table 4.	Non-essential amino acids	42
Table 5.	Base composition of the DNA from total and nucleosomal core chromatin from L929 cells	58
Table 6.	5-methylcytosine content of DNA at several times of digestion using micrococcal nuclease	59
Table 7.	5-methylcytosine content of the DNA in chromatin and methidiumpropyl-EDTA.iron(II) treated chromatin	60
Table 8.	5-methylcytosine content of DNA from reconstituted chromatin at several times of digestion using micrococcal nuclease	62
Table 9.	C + G content of the DNA in predicted nucleosome positions on mouse DNA sequences	70
Table 10.	Kinetic parameters for ascites DNA methylase using chromatin from L929 cells as substrate	86
Table 11.	<u>In vitro</u> transcription in nuclei from log phase L929 cells	99
Table 12.	Transcribing chromatin is deficient in $m^5C$	103
Table 13.	Location of DNA less than 1 hour old in a Sanders' extraction	104
Table 14.	Methylation of nuclei from log phase L929 cells using endogenous DNA methylase followed by Sanders' extraction	105



	<u>Page</u>
Table 15. Methylation of a cell lysate of log phase L929 cells using endogenous methylase followed by Sanders' extraction	106
Table 16. Methylation of nuclei from stationary phase L929 cells using added ascites methylase followed by Sanders' extraction	107
Table 17. Methylation of a cell lysate from stationary phase L929 cells using added ascites methylase followed by Sanders' extraction	107
Table 18. Methylation of nuclei from mosquito cells followed by Sanders' extraction	108
Table 19. New DNA is less methylated than old DNA	114
Table 20. Methylation occurs preferentially in DNA associated with the nuclear matrix	118
Table 21. Methylation of DNA in nuclei from L929 cells treated with 5-azadeoxycytidine	119

## SUMMARY

The DNA of eukaryotes has been shown to contain the minor base 5-methylcytosine. This arises by the enzymic modification of cytosine already incorporated into DNA. The distribution of methylation in chromatin is not random but there are distinct methylated and unmethylated domains, one such unmethylated domain consists of transcriptionally active DNA.

The work undertaken during the tenure of this award has been to examine the effects of chromatin structure and chromatin constituents upon DNA methylation, with the aim of finding a specific inhibitor of methylation which may be responsible for the undermethylation of transcribing regions of chromatin.

When the distribution of methyl groups in chromatin is examined, it is found that nucleosomal core DNA is enriched in these groups when compared to linker DNA. This DNA is also enriched in the bases cytosine and guanine.

When nuclei from log phase cells are methylated in vitro with the endogenous methylase, the methyl groups are added predominantly to core DNA. As DNA synthesis is not continuing in vitro, and most of the methylation has already occurred in vivo, the methylation in these nuclei is "delayed methylation". Evidence is presented that this methylation is occurring in the nuclear matrix, and the methylation pattern observed simply reflects the pattern in vivo.

But when mouse ascites methylase is added to these nuclei, the methyl groups go on to linker DNA, showing that this DNA is more susceptible to methylation than core DNA. These two observations suggest that the methylation pattern found in vivo is not a function of chromatin structure, and it is thought that nucleosomes bind to DNA enriched in cytosine and guanine, and that the enrichment of methylcytosine is a consequence of this.

This idea is reinforced by several lines of evidence:

i) Histones are very effective inhibitors of DNA methylation in vitro when added to a methylase assay.

ii) Chromatin is a poor acceptor of methyl groups compared to DNA. Core particles, which have a higher histone to DNA ratio, are not as good acceptors as chromatin.

iii) When native mouse DNA and histones are reconstituted into chromatin, the nucleosomes are again found on methyl rich and cytosine and guanine rich DNA.

When transcriptionally competent chromatin is examined, the DNA is found to be deficient in 5-methylcytosine compared to total chromatin. When nuclei from log phase cells are incubated in vitro the methylation occurs predominantly in the fraction of chromatin containing transcribing regions. As this fraction is also thought to contain DNA associated with the nuclear matrix it is not certain which DNA is being methylated. As the methylation in isolated nuclei is not affected by extraction with 0.2 M NaCl, which removes all of the soluble DNA methylase, this methylation is a product of a bound DNA methylase which is thought to be in the nuclear matrix.

Adding soluble DNA methylase to nuclei in vitro results in an elevated methylation of DNA in this fraction, and again the methyl groups are being added to transcribing (matrix) chromatin.

When the effects of nuclear non-histone proteins on DNA methylation in vitro are examined no inhibition is observed by either high mobility group proteins or nuclear non-histone proteins in general.

These results suggest that there are no specific proteins in transcribing regions which inhibit DNA methylation in vivo.

1. INTRODUCTION

## I.1. The presence of methylated bases in DNA

### I.1.1. Methylated bases in bacterial DNA

The first observation of a minor base in bacterial DNA was the detection of 5-methylcytosine ( $m^5C$ , Fig. 1a) in tubercle bacilli (Johnson & Coghill, 1925). This discovery was not confirmed until 1965 (Doskočil & Šormova, 1965) and  $m^5C$  has since been found in many bacteria (Oda & Marmur, 1966; Vanyushin et al., 1968). The amount of  $m^5C$  present in the DNA is usually low, ranging from 0.02 - 0.2 mole %.

In 1958 a second minor base was discovered in the DNA of the bacterium Escherichia coli 15T<sup>-</sup> (Dunn & Smith, 1958) and was found to be N<sup>6</sup>-methyladenine ( $m^6A$ , Fig. 1b). This modification occurs at a level which is normally higher than  $m^5C$ , e.g. 0.5 - 0.7 mole % for E.coli and Serratia marcescens (Jones & Walker, 1963; Vanyushin et al., 1965) and has been found in a wide range of bacterial species.

These modifications are also found in the DNA of mycoplasmas (Razin & Razin, 1980).

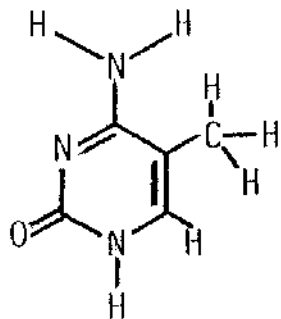
Both these modifications may be present in the same species or one or both may be absent.

Recently a third modified base was discovered, the base N<sup>4</sup>-methylcytosine ( $m^4C$ , Fig. 1c), in the DNA of Bacillus centrosporus (Janulaitis et al., 1983).

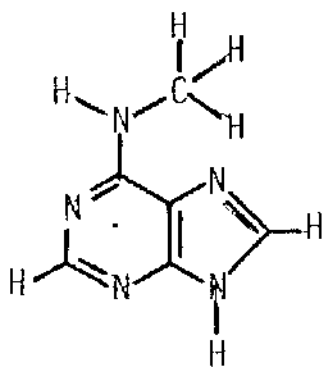
### I.1.2. Methylated bases in viral DNA

Methylated bases have also been found in the DNAs of many viruses; e.g. in the Xanthomonas bacteriophage XP-12 all the cytosines are present as  $m^5C$  (Kuo et al., 1968), and the bacteriophages T<sub>2</sub> and C of E.coli and Salmonella spp contain  $m^6A$  (Dunn & Smith, 1955). The T<sub>even</sub> phages have the related base 5-hydroxymethylcytosine substituted for cytosine (Wyatt & Cohen, 1953).

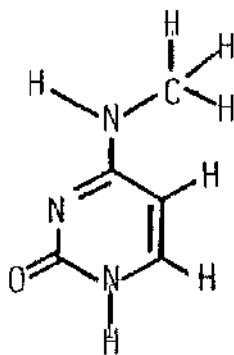
$m^5C$  has also been found in the DNA of several animal viruses, e.g. Herpesvirus saimiri (Desroisiers et al., 1979), adenovirus type 12 (Sutter & Doerfler, 1980), mouse mammary tumor virus (Cohen, 1980) but only when the viral DNA is integrated into the host's DNA and  $m^5C$  has not been found in mature virions (Low et al., 1969; Gunthert et al., 1976;



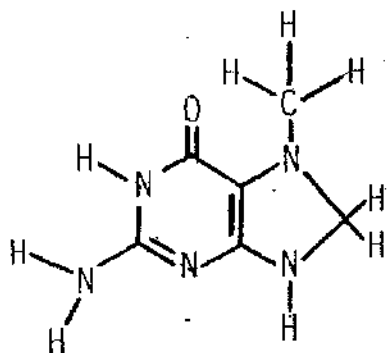
1a. 5-methylcytosine ( $m^5C$ )



1b.  $N^6$ -methyladenine ( $m^6A$ )



1c.  $N^4$ -methylcytosine ( $m^4C$ )



1d. 7-methylguanine ( $m^7G$ )

Figure 1. Methylated Bases in DNA.

von Acken et al., 1979) with the exception of frog virus 3 (Willis & Granoff, 1980).

Other modifications found in bacteriophages include the glycosylated derivative of 5-hydroxymethylcytosine in phage T<sub>6</sub> (Wyatt & Cohen, 1953) and the base 7-methylguanine (Fig. 1d) in the phage DDVI of Shigella dysenteriae (Nikolskaya et al., 1976).

### I.1.3. Methylated bases in eukaryotic DNA

In contrast to prokaryotes the only detected modified base in the DNA of higher eukaryotes (excluding insects) is m<sup>5</sup>C (Hotchkiss, 1948; Wyatt, 1951; Vanyushin et al., 1970). The level of modification varies greatly between species, the amount of m<sup>5</sup>C is very low in insects, 0.002 mole % in locust (Wyatt, 1951), 0.03 mole % in mosquito (Adams et al., 1979) and very high in plants, varying between 4 and 8 mole % (Wagner & Capesius, 1981; Shmookler et al., 1981). The level in most other species lies between these two extremes at 1-2 mole % (Wyatt, 1951; Tewari & Wildman, 1966; Vanyushin et al., 1973a).

The minor base m<sup>6</sup>A was once thought to be present only in the DNA of prokaryotes but has since been found in several species of protozoa, e.g. Tetrahymena spp (Gorovsky et al., 1973; Hattman et al., 1978a) and m<sup>6</sup>A but no m<sup>5</sup>C has been found in Paramecium aurelia (Cummings et al., 1974). M<sup>6</sup>A has also been found in Chlamydomonas reinhardi (Hattman et al., 1978b) and even in higher eukaryotes as trace amounts have been found in the mosquito Aedes albopictus (Adams et al., 1979). Deobagkar et al. (1982) have reported finding substantial amounts of the bases m<sup>5</sup>C, m<sup>6</sup>A and even m<sup>7</sup>G in the DNA of the mealy bug (Planococcus spp), but this result may be artifactual due to the presence of phenolic oxidases which these authors state could lead to the production of these modifications.

The only eukaryotes (except possibly some insects) not known to have methylated bases are the fungi (Guseinov et al., 1972).

In eukaryotes DNA is not limited to the nuclear region but is also found in organelles, and  $m^5C$  has been found in chloroplast DNA (Burton et al., 1979).  $m^5C$  was also reported in the DNA of mitochondria (Nass, 1973) but this is strongly contested (Dawid, 1974; Cummings et al., 1974; Groot & Kroon, 1979).

#### 1.1.4. DNA methylation as a post-synthetic event

When does DNA methylation occur? Does it occur before or after DNA synthesis?

It has been shown that there is no in vivo methylation of deoxycytidine triphosphate (Rubery & Newton, 1973) and no amination of deoxythymidine triphosphate (Sneider & Potter, 1969). This indicates that the methyl groups are not incorporated as deoxy-5-methylcytidine during DNA synthesis. The exception to this is provided by Xanthomonas bacteria when undergoing infection by phage XP-12 (Wang et al., 1982).

The observation that cytosine can be methylated after incorporation into DNA (Burdon & Adams, 1969) and the discovery of specific DNA methylases (Kalousek & Morris, 1968; Drahovsky & Morris, 1971; Turnbull & Adams, 1976) indicate that DNA methylation occurs at the polymer level, i.e. after DNA synthesis.

#### 1.1.5. The stability of methylated bases in DNA

The presence of methylated bases in DNA suggests that these modifications either possess or have possessed a defined function, therefore they should be stable from generation to generation.

There are several mechanisms by which methylated bases could be removed from the DNA: i) the action of specific demethylases, ii) deamination of  $m^5C$  to give thymine, iii) an excision repair mechanism and iv) replication of DNA without subsequent methylation.

Burdon & Adams (1969) found no net loss of  $m^5C$  as the ratio of  $m^5C/T$  of DNA in cultured mouse cells remained constant over several cell generations.



Deamination of  $m^5C$  to give T has been shown to occur in the E.coli lac I gene (Coulondre et al., 1978) and is thought to be responsible for the 'mutational hotspots' found in this gene. A high abundance of G,C  $\rightarrow$  A,T transition mutations also occur in the 5S RNA genes of some plants (Goldsborough et al., 1982) and are thought to occur by this method. In vertebrate DNA the very low frequency of the dinucleotide CpG may also be explained by the deamination of  $m^5C$  (Bird, 1980). Support for this idea is provided by the high frequencies of the dinucleotides TpG and CpA, and the observation that CpG is highly methylated in vertebrate DNA (Section I.2.3).

Two mechanisms exist by which alkylated bases can be removed from DNA, i) an excision repair mechanism which removes bases that do not occur naturally in DNA (Warren et al., 1979; Marghison & Pegg, 1981) and ii) a specific demethylase which removes the methyl group from  $O^6$ -methyl-guanine (Pegg & Hui, 1978). Neither of these mechanisms removes  $m^5C$  but a demethylase activity has been reported in crude nuclear extracts of murine erythroleukaemia cells (Gjerset & Martin, 1982) though the loss of  $m^5C$  observed may have been due to a nuclease activity.

That  $m^5C$  is stable in DNA has been shown by Wigler (1981) who showed that the in vitro methylation of  $\phi$ X174 RFDNA is stable over 25 cell generations when it is reintroduced into cultured mouse cells.

Specific demethylations actually occur in vivo in response to hormones (Section I.4.1) but these events require the replication of DNA (Compere & Palmiter, 1981; Colgan et al., 1982; Meijlink et al., 1983), and so occur by replication of DNA without subsequent methylation.

#### I.1.6. The effect of methylation on DNA stability

The base  $m^6A$  can exist as two rotational isomers, cis and trans (Fig. 2a), the preferred configuration being N1-cis- $N^6$ -methyladenine. In this configuration the nitrogen attached to carbon six cannot participate in the formation

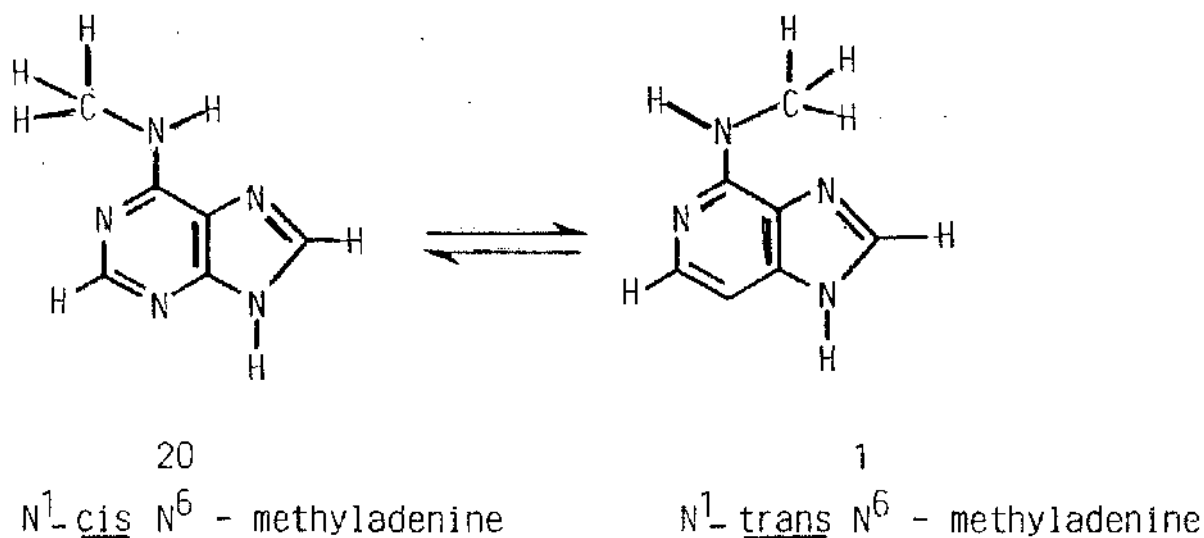


Figure 2a) The rotational isomers of N<sup>6</sup> - methyladenine

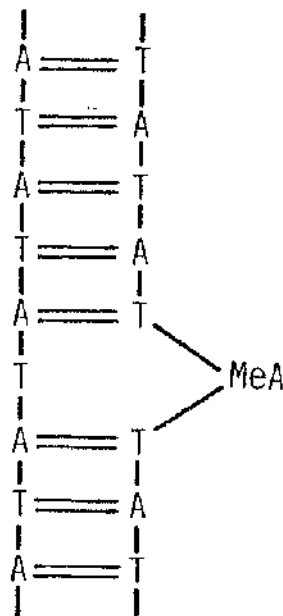


Figure 2b) Prediction of 'kinks' in poly(dA.dT). poly (dA.dT).  
Data from Engel and von Hippel, 1978.

of hydrogen bonds with thymine, thus causing a local destabilization of the double helix, subsequently lowering the melting temperature ( $T_m$ ) (Vanyushin, 1968; Engel & von Hippel, 1978). When  $m^6A$  is present in a poly (dA.dT), poly (dA.dT) copolymer distinct kinks in the structure are thought to occur (Fig. 2b). When the configuration is N1-trans  $N^6$ -methyladenine the methyl group lies in the major groove of the double helix and does not affect the base pairing.

The methyl group of  $m^5C$  lies in the major groove and actually stabilizes the double helix, thus raising the  $T_m$  (Gill et al., 1974).

Recent evidence has shown that DNA can exist as a left handed double helix (Z-DNA) at high concentrations of NaCl (> 2.0 M) and is the preferred structure for DNA having alternate residues of C, or more particularly  $m^5C$ , and G (Behe & Felsenfeld, 1981; van Lier et al., 1983). Z-DNA has been detected in the polytene chromosomes of Drosophila (Nordheim et al., 1981). The low frequency of CpG in vertebrate DNA would tend to exclude the possibility of Z-DNA formation, but Z-DNA has been detected in the metaphase chromosomes of primates, including man (Viegas-Pequignot et al., 1983), though some evidence that these results are due to artifacts caused by acid fixation has been found (Hill & Stollar, 1983).

#### I.1.7. The effect of methylation on apparent base composition

The melting temperature ( $T_m$ ) of DNA is dependent upon the base composition, i.e. DNA with a high C + G content has a high  $T_m$ . As  $m^5C$  stabilizes the double helix a high level of  $m^5C$  would give a high apparent C + G e.g. in the bacteriophage XP-12 the  $T_m$  is  $6.1^\circ C$  higher than animal DNA of the same C + G content (Ehrlich et al., 1975). When DNA is centrifuged on a neutral caesium chloride gradient its buoyant density (i.e. apparent C + G content) decreases as the level of  $m^5C$  increases (Gill et al., 1974). This anomalous behaviour is useful for detecting the presence of  $m^5C$  (Dawid et al., 1970).

Similarly poly(m<sup>6</sup>A)poly(U) copolymers have a T<sub>m</sub> of about 15°C (Ikeda et al., 1970) which compares to a T<sub>m</sub> of 56°C for poly(A). poly(U) copolymers.

1.1.8. The distribution of m<sup>5</sup>C in the eukaryotic genome

The m<sup>5</sup>C residues in eukaryotic chromosomes are not distributed at random throughout the genome but occur in long tracts of DNA interspersed with undermethylated regions (Molitor et al., 1976). The methyl deficient DNA makes up about 10% of the total DNA in mouse L929 cells (Molitor et al., 1976) and about 60% in sea urchin (Bird et al., 1979). In the nuclear DNA of Physarum polycephalum this methyl deficient DNA makes up about 80% of the genome (Whittaker & Hardman, 1980). Thus there is a wide variation in the amount of the methyl deficient regions. (See also Bird, 1984).

This methyl deficient DNA is highly enriched in unique and lowly repetitive sequences (Browne et al., 1978; Boehm & Drahovsky, 1979; Drahovsky et al., 1979) including the transcriptionally active gene sequences (Davie & Saunders, 1981; Naveh-Manly & Cedar, 1981). This DNA is also enriched in transcriptionally active repetitive sequences such as the histone genes (Bird & Taggart, 1980) though the distribution of ribosomal RNA genes between these domains depends upon the organism (Bird, 1984).

The methyl rich DNA consists of highly repetitive sequences such as satellite DNA (Brahic & Fraser, 1971; Schneidermann & Billen, 1973); foldback and interspersed repetitive sequences (Brahic & Fraser, 1971; Drahovsky et al., 1979; Whittaker et al., 1981) and the inactive amplified ribosomal RNA genes (Tantravahi et al., 1981). Satellite DNA is however undermethylated in sperm (Sturm & Taylor, 1981; Adams et al., 1983a).

The functions of this highly methylated satellite DNA are unknown but these sequences have been found clustered at the heterochromatic centromeres and telomeres of chromosomes in mammals (Lubit et al., 1974; Schnedl et al., 1975; Lubit

et al., 1976; Schreck et al., 1977), and may play a role in chromosome alignment at meiosis (Peacock et al., 1975).

## I.2. DNA methylases

### I.2.1. Prokaryotic DNA methylases

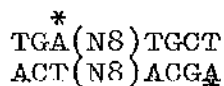
#### I.2.1.1. Methylases involved in restriction-modification

The restriction-modification systems of prokaryotes are characterised by, i) endonucleases which produce double stranded breaks in DNA, and ii), methylases, which modify the DNA at specific sites. Methylation of the DNA prevents cleavage by the corresponding nuclease.

There are three classes of restriction systems and their properties are summarized in Table 1.

Class I systems are high molecular weight, multisubunit enzymes which show DNA methylating as well as endonuclease activity and require ATP and AdoMet for activity (Lautenberger & Linn, 1972; Eskin & Linn, 1972a; 1972b). Class II systems consist of separate methylases and endonucleases of relatively low molecular weight (Roberts, 1976; Roberts, 1980; Modrich, 1982). Class III systems are similar to class I systems but consist of only two subunits (Reiser & Yuan, 1977; Bachi et al., 1979).

All three systems have specific recognition sites, but differ in their modes of cleavage. The recognition sites for class I systems are complex, e.g. for EcoB the sequence is



where N8 is an eight base variable sequence. The site of methylation is marked by an asterisk (Ravetch et al., 1978). The sites of cleavage are usually found at least 1000 bp from the recognition site (Yuan, 1981).

The class II enzymes recognize simple four to six base pair palindromic sequences, e.g. for EcoRI the sequence is

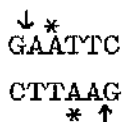


Table 1. Properties of restriction-modification systems.

	Type I	Type II	Type III
(1) Activities	Single multifunctional enzyme	Separate methylase and nuclease	Single multifunctional enzyme
(2) Protein structure	3 different sub-units	Simple	2 different sub-units
(3) Cofactor requirements	AdoMet, ATP, Mg <sup>2+</sup>	Mg <sup>2+</sup>	ATP, Mg <sup>2+</sup> ; (AdoMet)
(4) Recognition sequence	Complex: eg. EcoB; TGA(N8) TGCT	2 fold symmetry eg. EcoRI, GAATTC	Simple: eg. EcoPI, AGAUC
(5) Site of cleavage	Possibly random. At least 1000 bp from recognition sequence	Usually at recognition sequence	24-26 bp on the 3' side of the recognition sequence
(6) Methyl group donor	AdoMet	AdoMet	AdoMet
(7) Site of methylation	Recognition site	Recognition site	Recognition site

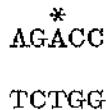
Data from Yuan, 1981.

and the Hpa II enzymes recognize



where the asterisk indicates the base methylated and the arrow the site cleaved by the nuclease (Mann & Smith, 1978; Roberts, 1980).

Only three class III systems are known (Yuan, 1981). These are EcoPI, EcoPI5 and HinF III. Their recognition sites are distinct from the other classes in that methylation occurs on only one strand, e.g. for EcoPI the site is



and the cleavage occurs about 24-26 base pairs away on the 3' side of this site (Bachi et al., 1979; Yuan, 1981).

An important lesson to be learned from these restriction-modification systems is that enzymes can recognize and interact with short DNA sequences and that this recognition is profoundly affected by the presence of methyl groups.

#### 1.2.1.2. The dam and dcm methylases

The DNA of E.coli has many more methylated bases than can be accounted for by the restriction-modification methylases (Vanyushin et al., 1968; Mamelak & Boyer, 1970). The level of m<sup>6</sup>A in wild type E.coli is 0.5 mole % and mutants have been isolated which have only 0.08 mole%, all of which can be accounted for by the EcoB methylase (Marinus & Morris, 1974). These mutants (dam<sup>-</sup> mutants) lack an enzyme which methylates the sequence GATC; and in wild type E.coli all these sequences are methylated (Hattman et al., 1978a; Geier & Modrich, 1979; Razin et al., 1980).

Similarly for E.coli C most of the methylation (m<sup>5</sup>C residues) can be accounted for by the action of the dcm protein which methylates all of the CC(A/T)GG sequences at the internal C (Schlagman et al., 1976; Razin et al., 1980). Other sequences must also be methylated in E.coli C as the CC(A/T)CG sites can only account for 50% of the m<sup>5</sup>C residues (Razin & Friedman, 1981).

The possible functions of this methylation are discussed in Section I.3.

## I.2.2. Eukaryotic DNA methylases

### I.2.2.1. Occurrence and properties

Eukaryotic DNA methylases have been purified from many sources, e.g.; rat spleen (Kalousek & Morris, 1968), rat liver (Drahovsky & Morris, 1971), HeLa cells (Roy & Weisbach, 1975), bovine thymus (Sano et al., 1983) and human placenta (Pfieffer et al., 1983).

All the methylases so far studied have similar properties, high molecular weight (130,000 - 180,000 Daltons), all are inhibited by high salt concentrations and all have the ability to methylate DNA from the same source as the enzyme.

These enzymes can also methylate both hemimethylated DNA (maintenance methylation) and unmethylated DNA (de novo methylation) though the latter occurs at a slower rate than the former (Gruenbaum et al., 1982). This de novo activity can be stimulated by brief digestion of the methylase using trypsin (Adams et al., 1983b).

These enzymes all catalyze the transfer of a methyl group from S-adenosylmethionine to the 5 position of cytosine in DNA, no other product has been found and no other cofactor is required for activity.

The actual mechanism of action is unknown but studies by Drahovsky and Morris (1971) suggested that the enzyme forms a complex with the DNA and then moves along the DNA remaining attached to the molecule between successive methylations. This mechanism of action has also been proposed by Simon et al. (1978), though the ascites enzyme is not thought to work in this way (Turnbull & Adams, 1976). Santi et al., (1983) suggest that the enzyme binds covalently to cytosine at position 6 producing a carbanion at carbon 5 which nucleophilically attacks the methyl donor AdoMet (Fig. 3).

Most of the methylation in eukaryotes occurs after a lag of about two minutes after DNA synthesis (Burdon & Adams, 1969; Kappler, 1970; Hotta & Hecht, 1971; Adams, 1974) but much of the methylation is still not complete after several hours (Burdon & Adams, 1969; Adams, 1971; Evans et al., 1973; Adams, 1974; Geraci et al., 1974; Kiryanov et al., 1980).



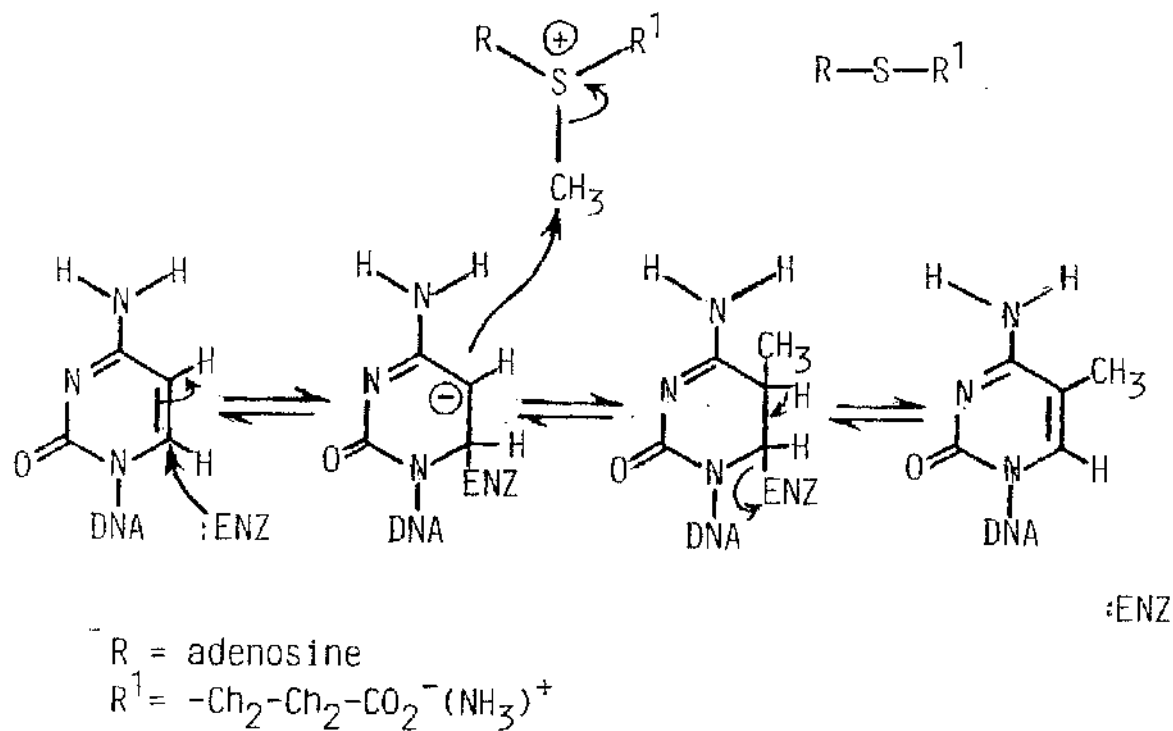


Figure 3. Mechanism of action of DNA methylase, according to Santi et al, 1983.

Although initiation of methylation occurs shortly after DNA synthesis it does appear that Okazaki fragments are ligated prior to methylation (Hotta & Hecht, 1971; Adams, 1974; Drahovsky & Walker, 1975). In certain instances it appears that methylation occurs well beyond the end of S-phase (Adams, 1973; Evans et al., 1973) but this methylation does not continue into the next cell cycle as the methylation which occurs in isolated nuclei is on newly synthesized DNA (Adams & Hogarth, 1973; Bird, 1978).

There is now evidence for two types of DNA methylases in the nuclei of mouse L929 cells, i) a 'soluble' form which can be extracted using 0.4 M NaCl, and ii) a 'bound' form which cannot be extracted even using 2 M NaCl, and which is thought to be bound to the nuclear matrix (Qureshi et al., 1982). An interesting hypothesis is that the 'bound' methylase may be responsible for the methylation which occurs at DNA replication and the 'soluble' methylase responsible for the delayed methylation, though there is no evidence to support this idea.

There are no restriction enzymes in higher eukaryotes comparable to those found in prokaryotes and unmethylated DNA is stable in vivo, e.g. the DNAs of polyoma virus (Kaye & Winocour, 1967) and Herpes simplex virus (Low et al., 1969).

#### I.2.2.2. The sequence specificity of eukaryotic DNA methylases

Unlike prokaryotic DNA methylases, the methylases of eukaryotes do not have specific multibase recognition sequences but it is known that methylation does not occur at random in vivo.

Using isotich analysis (a method involving depurination of DNA followed by two dimensional chromatography of the pyrimidine tracts) most of the  $m^5C$  in mammalian DNA is found associated with monopyrimidines, i.e. with cytosine (Sneider, 1972; Browne & Burdon, 1977; Browne et al., 1977). A considerable amount of  $m^5C$  is also found at the 5' ends of CT isotichs (Sneider, 1972) though this method involves distinguishing between CT and TC using nucleases which is

not too specific. In general the sequences which are methylated are found to be CpPu, PypCpPu, and PupCpTpPu.

Sinsheimer (1955) found that  $m^5C$  occurs predominantly at CpG dinucleotides and this has been confirmed by the use of restriction nucleases which will not act if the C of the CpG is methylated (Waalwijk & Flavell, 1978; Bird et al., 1979; McGhee & Ginder, 1979; Rae & Steele, 1979) though some CpC is also methylated (Sneider, 1980).

When purified DNA methylases are used to methylate DNA in vitro the predominant sequence methylated is also CpG (Browne et al., 1977; Simon et al., 1978; Sano et al., 1981; 1983) though Simon et al. (1983) have found considerable methylation of CpA and CpT.

The distribution in plants is somewhat different as there is more  $m^5C$ : up to 30% of the cytosines can be  $m^5C$ . This  $m^5C$  is found in the sequences CpG, CpA and CpT and even in CpC (Gruenbaum et al., 1981). These authors suggest that the methylation in non CpG sequences must be in the general sequence CpNpG so that these sites can be inherited (they are palindromic) and they found that over 80% of these sites are methylated in vivo.

### I.3. Possible functions for DNA methylation in prokaryotes

It was discussed in Section I.2.1.1 that some of the methylation in bacteria has a specific function in restriction-modification systems. Only a small proportion of the methylation is involved in these systems and the majority is a product of the dam and dcm methylases (Section I.2.1.2). What are the possible functions of this methylation?

#### I.3.1. A role for methylation in DNA replication

If DNA methylation is inhibited in phage  $\phi X174$  infected E.coli C cells by nicotinamide, few viable phage particles are produced (Razin et al., 1975) suggesting a major role of DNA methylation in phage maturation. The DNA of phage  $\phi X174$  has one  $m^5C$  residue in its 5500 nucleotides, this methyl group being located at the origin of replication very close to a major promoter (Razin et al., 1970; 1975) and is the

product of a phage induced methylase (Razin, 1973). Further work showed that nicotinamide inhibition resulted in a build up of multigenome length single stranded tails at the replication complex (Freidman & Razin, 1976), i.e. the circular <sup>double</sup> stranded form of  $\phi$ X174 could not be produced. These authors therefore suggested that the methyl group acts as the recognition site for a specific endonuclease which could be the gene A product. The methyl group does not function as part of a host restriction system as E.coli C is devoid of these systems (Arber & Linn, 1969). Nicotinamide derived cofactors are involved in many cellular processes and it is thus possible that the nicotinamide is not only inhibiting DNA methylation but also the nuclease, and therefore its use presents problems in the interpretation of the results of Razin et al. (1975).

### 1.3.2. DNA methylation and repair

By the use of dam methylase deficient mutants of E.coli K12 (dam-3) Marinus & Morris (1974; 1975) find that the DNA of these strains shows an increased sensitivity to both chemical mutagens and U.V. radiation. The DNA is also found to have many single stranded breaks. These authors suggest that the presence of  $m^6A$  protects the DNA from nucleolytic degradation but failed to find the nuclease responsible for this action. The observation that the phage  $\phi$ d grown on dam-3 mutants is not degraded when grown on wild type E.coli K12 would tend to exclude this possibility.

Dam<sup>-</sup> mutants are also hypermutable by the base analogs 2-aminopurine (an analog of G) and 5-bromouracil (an analog of T). These analogs are incorporated into DNA and result in base pairing errors as 2-AP can base pair with T and 5-BU can base pair with C (Radman et al., 1978). These data suggest that dam<sup>-</sup> mutants have lost the ability to correct mismatched bases in their DNA and has prompted the idea that the  $m^6A$  residues serve to discriminate between the two strands during DNA repair and also provides an explanation of the observations of Marinus & Morris (1974; 1975). Evidence

for such a system is provided by Glickman & Radman (1980), Radman (1981) and Pukkila et al. (1983).

### I.3.3. DNA methylation and recombination

When the phage  $\lambda$  is grown on E.coli arl<sup>-</sup> mutants (arl stands for accumulation of recombinogenic lesions) they undergo a much elevated frequency of recombination events (Korba & Hays, 1982a). A high frequency of such events has also been noted in plasmids from arl<sup>-</sup> mutants and is associated with partially deficient cytosine methylation at CC(A/T)GG sequences (Korba & Hays, 1982b). These results suggest that DNA methylation may play a role in recombination. As the DNA from arl<sup>-</sup> mutants is still partially resistant to the restriction nuclease EcoRII, it has been suggested that the increase in recombination is due to the presence of hemimethylated sites (Korba & Hays, 1982a) and it has been shown that hemimethylated duplexes do indeed recombine more frequently than DNA from arl<sup>+</sup> bacteria.

Hyper-recombination events also occur in dam<sup>-</sup> mutants of E.coli K12 (Marinus & Konrad, 1976).

### I.3.4. Methylation and the Mu-Mom gene

The temperate bacteriophage Mu encodes an interesting modification function termed Mom (modification of Mu) (Toussaint, 1976). This modification function is interesting in two respects: first, the modification is an acetoamidation (adenine is modified to N<sup>6</sup>-(1-acetoamido)-adenine) (Swinton et al. cited by Kahman, 1984); and second, the host's dam modification methylase is required for Mom expression (Toussaint, 1977; Khartoon et al. cited by Kahman, 1984). This suggests that methylation is required for Mom expression. Expression of this gene also requires a transactivating function which is thought to be a product of the C gene of Mu (Westmaas et al., 1976; Plasterk et al., 1983). When regions of the Mu genome at the 5' side of Mom are deleted the requirement for an active dam function ceases (Kahman, 1983; Plasterk, 1983) and when the Mom gene is transcribed from a foreign promoter the need for the transactivator also

ceases (Plasterk et al., 1983). Thus a model involving both methylation and the transactivating function which act at a sequence (or sequences) at the 5' side of the Mom gene, by which the Mom gene is regulated has been proposed (Kahman 1984). The actual mechanism is unknown but methylation of the promoter is thought unlikely (Kahman, 1984). This is the only known case of methylation regulating gene expression in prokaryotes.

#### I.4. Possible functions of DNA methylation in eukaryotes

##### I.4.1. DNA methylation and gene expression

###### I.4.1.1. General introduction

Many studies have been carried out to examine the methylation of specific genes and to correlate this with transcription. By the use of restriction nucleases (e.g. the isoschizomers MspI and HpaII which recognize the same sequence but HpaII is sensitive to methylation) and specific gene probes, it has been shown that the globin genes in tissues expressing these genes are undermethylated with respect to non-expressing tissues (Waalwijk & Flavell, 1978; McGhee & Ginder, 1979; Shen & Maniatis, 1980; van der Ploeg & Flavell, 1980). This negative correlation between gene expression and undermethylation is also seen in the ovalbumin, ovotransferrin and ovomucoid genes of chicken (Mandel & Chambon, 1979), the mouse immunoglobulin genes (Rogers & Wall, 1981; Yagi & Koshland, 1981); the mouse  $\alpha$ -fetoprotein genes (Andrews et al., 1982) and the human growth hormone gene (Hjelle et al., 1982).

These observations suggest that methylation may play a role in 'switching off' genes as the inactive genes are more highly methylated.

There are several complications in this simple scheme. First, there are certain genes whose methylation is the same whether the genes are active or inactive, e.g. the lens  $\delta$ -crystallin genes of chicken and the  $\alpha 2(1)$ -collagen gene of chicken (McKeon et al., 1982; Bower et al., 1983). Second, there are some genes, e.g. the Morris hepatoma  $\alpha 2u$ -globulin gene, which are undermethylated when they are inactive (Nakhasi et al., 1982) though all these genes are in primary cancer cells (Section I.4.6) which are hypomethylated when compared to normal cells. Third, there are large differences in the amount of methylation in organisms from different

taxonomic phyla (Section I.1.3), how can methylation regulate genes in insects (very little methylation) and in plants (up to 30% of the cytosines methylated)?

One possible answer is that not all of the methyl groups are involved in gene expression. Perhaps the loss of a single methyl group is sufficient to 'switch on' a gene. Such a loss would be difficult to detect.

Such demethylations actually occur in the response of several genes to the steroid hormone estrogen in roosters, e.g. the avian apoprotein II (Colgan et al., 1982) and the vitellogenin genes (Wachsmith & Jost, 1976; Wilks et al., 1982; 1984; Meijlink et al., 1983) the demethylation occurring in the 5' flanking sequences. One problem is that the demethylation event occurs subsequently to the onset of transcription, so it seems that this is a consequence rather than a cause of transcriptional activity.

In contrast to the above genes this demethylation event does not occur in the estrogen inducible vitellogenin genes of Xenopus which are expressed when fully methylated (Gerber-Huber et al., 1983).

#### I.4.1.2. The use of inhibitors of DNA methylation

Very good evidence that methylation plays a role in gene expression comes from the use of the DNA methylation inhibitors ethionine and 5-azacytidine. Christman et al. (1977) found that treatment of Friend erythroleukaemia cells with ethionine results in the expression of the globin genes. This expression is associated with hypomethylation of these genes. Ethionine is methionine with an S-ethyl terminal group instead of an S-methyl group and gets incorporated with adenosine to form S-adenosylethionine. This molecule will inhibit all methyl transferase enzymes which use AdoMet and thus will inhibit many more reactions than DNA methylation. This presents problems in interpreting data resulting from its use.

5-azacytidine on the other hand only gets incorporated into DNA and RNA. This base cannot be methylated due to the nitrogen group at position 5 and less problems are inherent

in its use.

In mouse thymoma cells the metallothionein-1 gene (MT-1) becomes sensitive to induction by cadmium and glucocorticoids after only 5 hours exposure to 5-azacytidine. This response requires the presence of DNA synthesis, i.e. the 5-azacytidine needs to be incorporated into the DNA, thus suggesting that demethylation is responsible for the inducibility of this gene (Compere & Palmiter, 1981).

Treatment with 5-azacytidine also results in the induction of endogenous viruses, e.g. the mouse type C virus (Niwa & Sugahara, 1981) and the chicken eV-1 virus (Groudinc et al., 1981).

#### I.4.1.3. The effects of in vitro methylation on gene expression

The effects of in vitro methylation of specific DNAs can be studied using either the oocyte injection system or transfection assays.

These studies have been done mainly with viral DNAs methylated with either the HpaII or the EcoRI methylases.

Using polyoma virus, methylation has no effect upon the transcription of the gene for the large T antigen (Subramanion, 1982). Similar results are obtained for the SV40 tumour antigen, though EcoRI methylation of the Herpes simplex virus thymidine kinase gene inhibits transcription (Waechter & Baserga, 1982). In the latter case an adenine residue is modified (Roberts, 1980) so an artificial situation is produced and this may produce invalid results. In contrast when the adenovirus type 2 specific binding protein gene is methylated and microinjected into oocytes no expression of this gene is observed whereas the unmethylated gene is expressed (Vardimon et al., 1982).

The problem with this type of experiment is that the bacterial methylase only methylates specific sites and the gene control region may lack this site. This problem is demonstrated by the observation that methylation of a mouse retrovirus genome in vitro using HpaII methylase has no effect on infectivity, but when a rat liver enzyme is used the virus is no longer infective (Simon et al., 1983).

Other studies involve specific eukaryotic genes, e.g. the expression of a transfected thymidine kinase gene is



dependent upon its level of methylation (Christy & Scangos, 1982).

A very significant result comes from an experiment by MacLeod & Bird (1983) who microinjected fully methylated ribosomal DNA from Xenopus sperm into Xenopus oocytes and compared its transcriptional activity with cloned ribosomal DNA. No difference is found, both DNAs are efficiently transcribed. Thus DNA methylation has no effect on transcription of ribosomal DNA. These authors do stress however that ribosomal DNA is transcribed by RNA polymerase I (not RNA polymerase II) and that this polymerase may be insensitive to DNA methylation.

#### I.4.2. DNA methylation and cellular differentiation

The distribution of  $m^5C$  in various organisms has been found to be both species and tissue specific (Kappler, 1971; Vanyushin et al., 1973a; 1973b; Ehrlich et al., 1982) and it has thus been suggested that methylation may play a role in cellular differentiation. The fact that in general sperm DNA has less  $m^5C$  than somatic cell DNA (Vanyushin et al., 1970; 1973a; Kaput & Sneider, 1979; Sturm & Taylor, 1981) prompts the idea that differentiation involves an increase in DNA methylation. Kaput & Sneider (1979) suggested that sperm DNA is strand asymmetrically methylated, so that when a cell differentiates those genes not required would have the other strand methylated, and those required would be demethylated due to cell division.

In vivo studies have on the other hand shown a selective loss of methylation during development. There is a 35% loss of methylation during differentiation of rabbit embryos (Manes & Manzel, 1981), and loss of methylation in the ribosomal genes accompanies the onset of gene activity in the early development of Xenopus laevis (Bird et al., 1981). This loss is not found in sea urchin (Pollock et al., 1978; Bird et al., 1979) and no changes in methylation are found in the early development of mouse (Singer et al., 1979).

These observations apparently contradict the fact that sperm DNA is less methylated than somatic cell DNA.

Two lines of evidence provide an answer to this contradiction. First, it has been shown that the level of methylation

of specific genes in sperm which are not being expressed, e.g. the rabbit  $\beta$ -globin gene (Waalwijk & Flavell, 1978), and the chicken ovalbumin genes (Mandel & Chambon, 1979), is higher than in the somatic cells in which these genes are expressed, whereas two genes for housekeeping functions, adenine phosphoribosyltransferase and dihydrofolate reductase are methylated to the same extent in both sperm and somatic cells (Stein et al., 1983). Second, the increase in methylation in somatic cells is due to the methylation of satellite DNA which is undermethylated in sperm (Sturm & Taylor, 1981; Adams et al., 1983a; Sanford et al., 1984).

Thus differentiation is associated with a selective loss of methylation in specific genes accompanied by an overall increase in methylation levels.

It is unknown at present whether these changes in methylation are a result or a cause of cellular differentiation, but they are not related to the process of meiosis (Rahe et al., 1983).

Further evidence for a role of methylation in differentiation comes from the use of 5-azacytidine. Treatment of cultured mouse embryo cells with this drug results in marked changes in the differentiated state of these cells (the drug induces the production of muscle cells) (Jones & Taylor, 1980). Differentiation can also be induced in Friend erythroleukaemia cells (Creusot et al., 1982).

#### I.4.3. DNA methylation and X-chromosome inactivation

The mammalian female karyotype differs from the male in that it has two X-chromosomes instead of an X and a Y chromosome. It has been known for many years that one of these X-chromosomes is inactive (Ohno et al., 1959; Lyon, 1961) and it appears in the nucleus in a highly condensed form ('barr' body) as part of the heterochromatin. The reasons for this inactivity are unknown but it has been suggested that DNA modification might be involved (Riggs, 1975; Sager & Kitchin 1975). Riggs proposed that each X-chromosome has an inactivation centre and that during embryogenesis this centre becomes inactive by methylation, thus

'switching off' the entire chromosome. Evidence that inactivation might depend upon an intrinsic property of the chromosome comes from studies showing that deactivation does not involve regulatory proteins and does not involve chromosome-chromosome interactions (Kahan & deMars, 1980). That this intrinsic property may be methylation has been shown by Mohandas et al. (1981). These authors treated mouse-human hybrid cells deficient in the enzyme hypoxanthine: guanine phosphoribosyltransferase (HPRT<sup>-</sup>) with 5-azacytidine (Section 1.4.1.2 for action) and obtained HPRT<sup>+</sup> revertants.

Many similar experiments have been carried out with the same result (Graves, 1982; Jones et al., 1982; Shapiro & Mohandas, 1982).

Other supportive evidence comes from transformation experiments; when DNA from 5-azacytidine treated cells is used to transform HPRT<sup>-</sup> cells, many HPRT<sup>+</sup> revertants are obtained (Mohandas et al., 1981; Venolia et al., 1982).

One problem with this simplistic scheme is that the expression of other X-linked loci, such as the phosphoglycerokinase and glucose-6-phosphate dehydrogenase genes, is not affected by 5-azacytidine (Stein et al., 1983).

#### 1.4.4. The effect of methylation on viral expression

When a virus infects a cell the viral DNA can integrate into the host's chromosomes and become a provirus. If this provirus is inactive the viral DNA is found to be extensively methylated (Desroisiers et al., 1979; Sutter & Doerfler, 1980; Hynes et al., 1981; Fradin et al., 1982). Thus methylation can be correlated with non-activity of the viral DNA.

Part of the integrated viral DNA can also be in an active form as in viral induced tumourigenesis and in this form the DNA is found to be undermethylated, e.g. murine mammary tumour provirus in mouse thymomas (Mermod et al. 1983) and avian sarcoma provirus in permissive chicken cells (Guntaka et al. 1980).

Treatment of cells containing proviruses with 5-azacytidine results in deintegration of the viral DNA and production of virus particles of the mouse type C virus (Niwa & Sugahara,

1981) and results in expression of the structural genes of the chicken eV-1 virus (Groudine et al., 1981).

In contrast to integrated viral DNA no methylation can be detected in mature viral particles (but see Section I.1.3) (Low et al., 1969; Desroisiers et al., 1979; Sutter & Doerfler, 1980; Youssoufian et al., 1982).

#### I.4.5. The maternal inheritance of DNA in Chlamydomonas chloroplasts

In the chloroplast of the green alga Chlamydomonas reinhardtii the genetic material is inherited from the female gamete (Sager, 1954). The DNA in the female gamete undergoes a density shift consistent with DNA methylation before zygote formation (Sager & Lane, 1972) and the DNA of male origin is degraded shortly after zygote formation. These observations led to the proposal that the maternal DNA becomes modified which protects it from nucleolytic degradation (Sager & Lane, 1972; Sager & Ramanis, 1973).

The discovery that maternal DNA becomes extensively methylated (Burton et al., 1979; Sager et al., 1981) and the discovery of a site specific endonucleolytic activity (Burton et al., 1977) led to the proposal that a restriction-modification system exists in Chlamydomonas which resembles the type I systems described in Section I.2.1.1. This nucleolytic activity has not yet been extensively purified and characterized though it seems to produce single strand nicks in DNA and excision of single stranded regions, forming 'gaps' in the DNA (Sager et al., 1984).

#### I.4.6. DNA methylation and carcinogenesis

Treatment of eukaryotic cells with chemical carcinogens leads to a pronounced demethylation of the DNA in these cells (LaPeyre & Becker, 1979; Wilson & Jones, 1983) suggesting a role for DNA methylation in the prevention of cancer. Further evidence comes from the observations that the DNA in cancers is generally hypomethylated (Feinberg & Vogelstein, 1983a) and that hypomethylation accompanies 5-azacytidine induced tumourigenesis (Harrison et al., 1983).

What this role may be is unknown but the discovery of specific cellular oncogenes (Stehelin et al., 1976; Spector et al., 1978) and the observations that DNA may play a role in gene expression (Section I.4.1) suggests that methylation may help to keep these genes inactive (or at a low level of activity). An inverse correlation between the activity of oncogenes and methylation has been found in human cells (Feinberg & Vogelstein, 1983b) and in mouse cells (Gattoni et al., 1983).

A similar pattern of methylation and activity is found in tumourigenic viruses (Section I.4.4).

#### I.4.7. Aging

In general the DNA in tissues from old animals has less  $m^5C$  than DNA from young animals (Vanyushin et al., 1973b; Romanov & Vanyushin, 1981) which suggests a possible role for DNA methylation in aging. Scarano (1969) proposed that  $m^5C$  may be deaminated to form thymine and that if this were to happen in a programmed manner this mechanism could be used as a timing device (Scarano, 1969; Tosi et al., 1972). Such deaminations do occur in vivo but there is no evidence for such a programme (Section I.1.5).

Holliday & Pugh (1975) proposed that methylation could be used to measure the number of generations a cell had undergone. At each cell division de novo methylation would occur in a specific DNA sequence, thus the number of methyl groups incorporated into this sequence would equal the number of cell divisions. So far no evidence for such a system has been found.

### I.5. Chromatin structure

#### I.5.1. The nucleosome concept

When chromatin from interphase rat thymus nuclei is analysed using an electron microscope, a linear array of spherical bodies  $70 \text{ \AA}$  in diameter connected by  $15 \text{ \AA}$  long strands is observed (Olins & Olins, 1974). This same structure is seen in chromatin depleted of histone H1 (Oudet et al., 1975).

These observations have been supplemented with studies using nucleases, e.g. the action of the endogenous nuclear  $\text{Ca}^{2+}/\text{Mg}^{2+}$  dependent rat liver nuclease on chromatin produces a series of bands when the DNA is electrophoresed on agarose gels (Hewish & Bourgoyne 1973). This pattern is also seen using micrococcal nuclease (Noll, 1974a). These bands form a regular repeating pattern and are multiples of 180-230 base pairs.

Cross-linking studies of histone aggregates in solution (Kornberg & Thomas, 1974) led to the proposal that chromatin is a regular repeating structure in which the basic unit is a histone octamer associated with about 200 base pairs of DNA, termed a nucleosome (Kornberg, 1974).

#### I.5.2. The nucleosome 'core'

Studies with nucleases has shown that the DNA repeat length of nucleosomes is 185 base pairs in rat liver (Axel et al., 1975) and in duck erythrocytes (Sollner-Webb & Felsenfeld, 1975). This repeat length is not constant but varies between different species (Compton et al., 1976; Noll, 1976; Lohr et al., 1977), different tissues (Spadafora et al., 1976; Thomas & Thompson, 1977) and even in single cell types (Kornberg, 1977).

Further digestion of nucleosomes produces a relatively nuclease resistant particle with 140 base pairs of DNA (Sollner-Webb & Felsenfeld, 1975). This 'core' particle has been found to be invariant between different species (Morris, 1976a; Thomas & Thompson, 1977; Lohr et al., 1977). More recent studies have shown that the length of DNA in the core particle is in fact 146 base pairs (Lutter, 1979; Bryan et al., 1979).

Neutron and X-ray scattering patterns of isolated core particles in solution show that cores are oblate structures 50 Å in height and 110 Å in diameter with the DNA confined to two rings on the outside (Pardon et al., 1977; Richards et al., 1977) and contain two each of the histones H2A, H2B, H3 and H4 (Bently et al., 1981; Finch et al., 1981).

### I.5.3. The internal structure of the 'core'

Limit digests of core particles with nucleases have given a lot of information about the arrangement of DNA in the core. With staphylococcal nuclease digestion, fragments of DNA from 40 to 140 base pairs at 10 base pair intervals are produced (Camerini-Otero, 1976). DNase I, DNase II and the endogenous  $\text{Ca}^{2+}/\text{Mg}^{2+}$  rat liver nuclease all make single stranded nicks at 10.5 base pair intervals (Noll, 1974b; Sollner-Webb et al., 1976; Bryan et al., 1979; Lutter, 1979).

The same pattern of nicking by DNase I is seen when protein free DNA is adsorbed to calcium phosphate (Liu & Wang, 1978). These results show that the pattern of cutting is related to the repeat of the DNA duplex, i.e. the DNA is only exposed on one side. The idea that the DNA in nucleosomes had 'kinks' at every 10.5 base pairs was also suggested to explain these results though this has not been proved (Crick & Klug, 1975; Sobell et al., 1976).

When the DNase I susceptible sites are mapped they are found to be symmetrically distributed around a dyad axis in the nucleosome (Simpson & Whitlock, 1976) and this observation led to the proposal that nucleosomes could divide into two equal halves during DNA replication or gene transcription (Weintraub et al., 1976).

### I.5.4. The role of histone H1

The nuclear content of histone H1 is about one molecule per nucleosomal core particle of chromatin (Rall et al., 1977; Albright et al., 1979) and this histone has several subtypes (Smerdon & Isenberg, 1976). These subtypes vary considerably between different tissues of the same organism (Seyedin & Kistler, 1979; von Holt et al., 1979) and have been related to the different linker lengths (Morris, 1976b).

Studies on nucleosome structure using micrococcal nuclease have shown that histone H1 is removed from the nucleosome when the DNA length is decreased from 200 to 140 base pairs (Noll & Kornberg, 1977; Simpson, 1978). If histone H1 is

removed with 0.35 M NaCl prior to digestion the linker DNA is digested much more rapidly than before (Whitlock & Simpson, 1976; Varshavsky et al., 1976; Noll & Kornberg, 1977). These studies suggest that histone H1 is in contact with the linker DNA. Physical studies on histone H1 show that the molecule has a globular central part with two peptide 'tails' (Chapman et al., 1978) and it has been shown that the histone spans the nucleosome core touching the linker on either side (Cole et al., 1977; Lawrence & Goeltz, 1981; Frado et al., 1983).

Chromatin isolated from nuclei in the presence of  $Mg^{2+}$  ions appears as a supercoiled solenoidal structure with a diameter of 300 Å (Finch & Klug, 1976; Thoma et al., 1979). When histone H1 is not present this solenoid opens out into 100 Å diameter filaments (Thoma & Koller, 1977). Cross-linking studies have shown that histone H1 polymers can be produced which are made up of eight or more histone H1 subunits (Olins & Wright, 1973; Chalkley, 1975) so the nucleosomal cores must be in close proximity and this suggests that histone H1 is involved in higher order packing of nucleosomes.

#### I.5.5. Nucleosome phasing

Since the basic structure of chromatin was elucidated (Section I.5.1) many studies have been undertaken to determine whether there is any relationship between the location of nucleosomes and the underlying DNA sequence (Zachau & Igo-Kemenes, 1981). These studies have usually involved relating the observed micrococcal nuclease cutting sites in linker DNA with restriction nuclease recognition sites, in specific DNA sequences.

The first suggestion that nucleosome locations may be fixed came from studies on the  $\alpha$ -satellite DNA of African green monkey (Musich et al., 1977) though this is strongly contested (Fittler & Zachau, 1979; Singer, 1979). Recently Zhang et al. (1983) has shown that there are at least eight different phases of nucleosomes on this DNA.



Since 1977 many examples of nucleosome phasing have been demonstrated, e.g. the tRNA genes of chicken (Wittig & Wittig, 1979), the 5SRNA genes of Drosophila and Xenopus (Louis et al., 1980; Gottesfeld & Bloomer, 1980), the histone genes of Drosophila (Samal et al., 1981) and the J<sub>k</sub> and C<sub>k</sub> immunoglobulin genes of mouse liver (Weischet et al., 1983).

The concept of nucleosome phasing is attractive as in one phase a transcriptional control region (or any other control region) may be covered by a nucleosome, whereas in another phase this region is not covered, so allowing access to the RNA polymerase complex. This idea is supported by evidence from the 5SRNA genes of Drosophila (Louis et al., 1980) and in the tRNA genes of Xenopus laevis (Bryan et al., 1981).

Many criticisms of these experiments have been voiced. First, micrococcal nuclease is sequence specific in its cleavage of DNA, it preferentially cuts DNA at adenine and thymine rich regions (Nelson et al., 1979a; Horz & Altenberger, 1981; Dingwall et al., 1981). Second, micrococcal nuclease will cleave DNA in the core particles and does so at adenine and thymine rich regions, so many of the potential 'phases' will be lost as only the 140 base pairs obtained from agarose gels is examined (McGhee & Felsenfeld, 1983). Third, the study of repetitive sequences may be misleading due to the presence of regularly spaced micrococcal nuclease sensitive sites (Horz et al., 1983). Fourth, the observation that the nucleosomes in total chromatin are not phased (Strauss & Prunell, 1980).

These results do not mean that nucleosomes are not fixed relative to the DNA but only that different methods must be employed. The nucleosomes may be positioned relative to structural aspects of the DNA. Hogan et al. (1983) has shown that poly(dG).poly(dC) homopolymers are less flexible than poly(dA).poly(dT) homopolymers suggesting that the former would be less likely to be found in nucleosomes, though in practice neither homopolymer is found in nucleosomes (Simpson & Kunzler, 1979).

Similarly models of DNA flexibility based upon the regular spacing of certain dinucleotides have been proposed

(Trifonov, 1980; Zhurkin, 1983). These models have been applied to DNA sequences which have been studied for nucleosome phasing and the predicted locations of nucleosomes agree quite well with the locations determined by nuclease studies (Mengeritsky & Trifonov, 1983; Zhurkin, 1983).

#### 1.5.6. The structure of transcriptionally active chromatin

Two approaches have been used to determine whether transcriptionally active chromatin is organized into nucleosomes: electron microscopy and digestion with nucleases. The use of the electron microscope has revealed that the very active ribosomal RNA genes do not have nucleosomal-like particles between the RNA polymerase molecules in Oncopeltus fasciatus (Foc et al., 1976), Xenopus laevis oocytes (Scheer, 1978; Pruitt & Grainger, 1981), Tetrahymena pyriformis (Borkhardt & Nielson, 1981; Borchsenius et al., 1981), or Physarum polycephalum (Johnson et al., 1978). These genes are only non-nucleosomal when they are being expressed and the non-transcribed spacers are always nucleosomal.

Studies with micrococcal nuclease on the other hand have shown that the rRNA chromatin in Physarum and Tetrahymena has a repeat length of 180 base pairs suggestive of nucleosomes (Mathis & Gorovsky, 1976; Piper et al., 1976; Grainger & Ogle, 1978), though Grainger & Ogle interpreted this as protection of the DNA due to the binding of RNA polymerase. With Xenopus laevis a nucleosomal ladder can barely be detected after micrococcal nuclease digestion and the ribosomal RNA genes are preserved in a high molecular weight form (Spadafora & Crippa, 1984) showing that proteins are protecting the DNA from digestion but that these are not nucleosomes. Recently it has been shown that in Tetrahymena the non-transcribed spacers of the ribosomal RNA genes are packaged into nucleosomes but the coding regions are not. (Palen & Cech, 1984).

Histones have been found in the nucleolar regions of Physarum polycephalum (Johnson et al., 1978) but not in Xenopus laevis (Labhardt & Koller, 1982). As little or no compaction of the ribosomal RNA genes is observed in the electron microscope then if nucleosomes are present they must

be in a radically altered form (Woodcock et al., 1976; Scheer, 1978).

Other highly active genes such as the putative transcriptional unit of the silk fibroin gene of Drosophila melanogaster are also devoid of nucleosomes (McKnight et al., 1976).

Analysis of the DNA sequences represented in the mononucleosome fraction of a micrococcal nuclease digest has shown that virtually all sequences are represented in the same proportion as in the total DNA (Lacy & Axel, 1975; Kuo et al., 1976; Gottesfeld & Butler, 1977; Bloom & Anderson, 1978) suggesting that all sequences are nucleosomal. Studies on individual genes have shown that the active globin genes (Axel, 1976), ovalbumin genes (Garell & Axel, 1976) and the integrated active leukaemia virus DNA (Panet & Cedar, 1977) are all nucleosomal. The active histone genes in sea urchin show a blurred micrococcal nuclease digestion ladder showing an altered nucleosomal structure (Spinelli et al., 1982) and so are similar to the ribosomal RNA genes.

Other characteristics of transcribing chromatin are: 1) an enrichment in the high mobility group proteins 14 and 17 (Teng et al., 1979; Kuehl et al., 1980), 2) an enrichment in highly acetylated histones (Ruiz-Carillo et al., 1975; Nelson et al., 1979b) and 3) a deficiency in histone H1 (Faber et al., 1981; Gabrielli et al., 1981). When mononucleosomes from a micrococcal nuclease digest are electrophoresed on gels three distinct subtypes are found (Levy-Wilson & Dixon, 1979; Miki et al., 1980). One of these subtypes has all of the above properties and is enriched in active DNA sequences (Levy-Wilson & Dixon, 1979; Miki et al., 1980; Egan & Levy-Wilson, 1981). The other subtypes have little or no high mobility group proteins and are enriched in histone H1 and non-transcriptionally active DNA sequences. Other methods, e.g. DNase digestion followed by  $Mg^{2+}$  solubility of transcribing chromatin have shown that transcriptionally active DNA is associated with nucleosomes (Gottesfeld & Partington, 1977; Davie & Saunders, 1981).

### 1.5.7. Chromosomal scaffold

In addition to the repeating structure (nucleosomes) chromatin possesses a wide range of non-histone proteins. Of these some have been shown to be responsible for the maintenance of chromosome structure and are organized into a structural entity known as the 'chromosomal scaffold' (Adolph et al., 1977a; Paulson & Laemmli, 1977). This structure can be isolated by treating chromatin or metaphase chromosomes with 2M NaCl or polyanions (dextran sulphate/heparin) (Adolph et al., 1977a). This treatment removes all of the histones and about 70% of the non-histones. The remaining non-histones are organized into a scaffold structure which is the same shape as the original metaphase chromosome (Paulson & Laemmli, 1977; Adolph et al., 1977b). Observations with the electron microscope have shown that the DNA is attached to this scaffold in loops of between 30 and 90 kilobases in length (Marsden & Laemmli, 1979). Transverse sections through swollen but unextracted chromosomes show a star like arrangement of the chromatin fibre consistent with a radial loop model (Marsden & Laemmli, 1979; Adolph, 1980). A central 'core' or scaffold has also been observed in mitotic chromosomes by using silver staining and light microscopy (Howell & Hsu, 1979).

It has recently been shown that the scaffold requires a specific metalloprotein for stability. The histone depleted structures were specifically and reversibly stabilized with  $\text{Cu}^{2+}$  ions and less specifically with  $\text{Ca}^{2+}$  ions (Lewis & Laemmli, 1982). Two proteins were shown to be involved in this stabilization, termed  $\text{Sc}_1$  and  $\text{Sc}_2$  (170,000 and 135,000 Daltons). More recently it has been shown using microscopy that the scaffold is a defined structure and not a non-specific aggregate (Earnshaw & Laemmli, 1983).

### 1.5.8. Nuclear matrix

When nuclei are extracted using 2M NaCl and treated with low amounts of deoxyribonucleases a residual structure remains which is termed the nuclear matrix (Berezney & Coffey, 1974; 1977; Mitchelson et al., 1979). This structure accounts for about 10-20% of the total nuclear protein and about 2-10% of the DNA (Berezney and Coffey, 1977).

Electron microscopic analysis of the nuclear matrix shows a peripheral shell derived from the pore lamina elements of the nuclear envelope (Aaronson & Blobel, 1975; Gerace & Blobel, 1980) together with an internal framework to which the DNA is attached. (Berezney, 1980; Fisher et al., 1982).

Further treatment using nucleases (DNases/RNases) leaves a structure termed the nuclear protein matrix which has less than 0.1% of the total DNA (Berezney & Coffey, 1977).

This nuclear protein matrix has also been shown to be stabilized by  $\text{Ca}^{2+}$  (Wunderlich & Herlan; 1977) and  $\text{Cu}^{2+}$  or  $\text{Ca}^{2+}$  (Lebkowski & Laemmli, 1982). It has been suggested that metals play a role in the long range compaction of DNA in both histone depleted metaphase chromosomes and in inter-phase nuclei.

#### I.5.8.1. DNA sequences associated with the nuclear matrix

Matrix DNA from HeLa cells is highly enriched in transcriptionally active DNA sequences, as demonstrated by hybridization of the total DNA and matrix DNA to an excess of nuclear RNA (Jackson et al., 1981). Similarly the active ovalbumin and conalbumin gene sequences of hen oviduct are found in matrix DNA, whereas this is not seen in hen brain where these genes are inactive (Robinson et al., 1982; 1983). This is also seen with the chicken  $\beta$ -globin genes in erythrocytes (Hentzen et al., 1984). These observations suggest an important role for the nuclear matrix in gene expression.

Pulse labelling experiments using  $[^3\text{H}]$ -thymidine have shown that matrix DNA is highly enriched in newly synthesized DNA (Berezney & Coffey, 1975; Keller & Riley, 1976; Miller et al. 1978; Hunt & Vogelstein, 1981; Aelen et al., 1983). One minute after rats are injected with  $[^3\text{H}]$ -thymidine more than 90% of the tritium is found in the DNA which remains associated with the nuclear matrix. The same result is observed with 3T3 cells in culture (Vogelstein et al., 1980).

Recently it has been shown that a form of DNA methylase is tightly associated with the nuclear matrix (Qureshi et al.,

1982; Qureshi, 1983). This is significant as DNA methylation is coupled with DNA synthesis (Section I.2.2.1) and so the methylase would be expected to be located at or near the site of DNA replication.

Other sequences which show an enrichment in matrix DNA are middle-repetitive sequences in Chinese hamster, mouse and chicken (Jeppeson & Bankier, 1979; Razin et al., 1979; Kuo, 1982).

#### I.6. Chromatin structure and DNA methylation

Both chromatin structure and DNA methylation have been reported to play a role in the transcriptional activity of genes (Sections I.4 and I.5.6) so it is of interest to study the relationship between chromatin structure and DNA methylation.

When the distribution of  $m^5C$  in the chromatin from chicken cells is examined it is found that nucleosomal core DNA is enriched in this base (Razin & Cedar, 1977). This distribution was lost when the chromatin was reconstituted. Adams et al. (1977) found no such enrichment in the core DNA in Chinese hamster ovary chromatin but have recently found an enrichment in mouse L929 chromatin (Adams et al., 1984).

Chromatin structure can be probed by the use of DNA methylases and when nuclei from chicken erythrocytes are treated using the E.coli methylase over 95% of the methyl groups are found in nucleosomal linker DNA (Bloch & Cedar, 1976). This suggests that the linker DNA is more susceptible to enzymatic action than core DNA and is in apparent contradiction with the noted distribution of  $m^5C$ . Similarly the methylase in Friend erythroleukaemia cells has been located in the linker DNA (Crousot & Christman, 1981) though this conclusion is based on the solubilization of the methylase using nucleases to digest the DNA in chromatin and may be misleading.

These results suggest that chromatin structure can effect DNA methylation and shows that further study along these lines is potentially interesting.

## AIMS

Nucleosomes are the basic structural units of eukaryotic chromatin and the simplistic model of a nucleosome depicted as 146 base pairs of DNA wrapped around an octamer of histones has been useful in elucidating nucleosome structure. However, the discovery of histone modifications and high mobility group proteins associated with nucleosomes in transcriptionally active chromatin, has led to the idea that nucleosomes are not just packaging vehicles, but the structural basis for the functional state of genes.

As DNA methylation has also been implicated in the regulation of gene activity, particularly those  $m^5C$  residues located in the 5' regions of genes, it is of interest to determine whether there is any relationship between chromatin structure and DNA methylation.

The aims of this work have been:

- (1) To investigate the distribution of  $m^5C$  in chromatin.
- (2) To attempt to methylate chromatin of mouse and insect cells in vitro and to determine where the methyl groups are incorporated.
- (3) To determine the effects of histones and nuclear non-histone proteins on methylation in vitro.
- (4) To establish whether the undermethylation of transcribed regions could be a result of inhibition of methylation by proteins associated with these regions, and if so which proteins.

II. MATERIALS AND METHODS



II.1. MATERIALS

II.1.1. Biological

Mouse L929 cells (Sanford et al., 1948) are a cell line routinely maintained in this department.

Mosquito cells (Aedes albopictus) (Singh, 1967) were obtained from Flow Laboratories.

II.1.2. Chemical

All chemicals were analytical grade reagents except the following:

Foetal calf serum	Bio-Cult Laboratories Ltd., Paisley, Scotland.
Newborn calf serum	"
Non-essential amino acids	"
Mitsubishi and Maramorosch Basal medium	Flow Laboratories Ltd., Irvine, Ayrshire, Scotland.
Triton X-100 (scintillation grade)	Koch-Light Laboratories Ltd. Colnbrook, England.
2,5-diphenyloxazole (PPO) (scintillation grade)	"
Dithiothreitol	"
DX-80 developer	Kodak Ltd., Manchester
FX-40 X-ray liquid fixer	"
Micrococcal nuclease	Boehringer Mannheim House, Lewes, E. Sussex.
PMSF (phenylmethylsulfonyl- fluoride)	"
Calf thymus histone	"
<u>Neurospora crassa</u> endonuclease	"
<u>E.coli</u> polymerase I	"
Bathophenanthroline- disulfonate	Sigma London Chemical Co. Ltd., Poole, Dorset, England.
S-adenosylmethionine (chloride salt)	"
Ethidium bromide	"
Tween 80 (polyoxyethylene sorbitan mono-oleate)	"
Salmon testis DNA	"

EGTA (ethyleneglycol-bis-( $\beta$ -aminoethyl ether) N,N' tetra-acetic acid)	Sigma London Chemical Co. Ltd., Poole, Dorset, England
8-Hydroxyquinoline	"
Agarose (type II)	"
Deoxyribonucleotides	"
Ribonucleotides	"
Thymidine	"
Hypoxanthine	"
Aminopterin	"
Hydroxyurea	"
m-Cresol	"
Ribonuclease A	"
Pancreatic deoxyribonuclease (DNase I)	"
Whatman No.1 paper and 3MM 2.5 cm paper discs	H. Reeve-Angel and Co. Ltd., London.
Sephadex G-50	Pharmacia, Uppsala, Sweden
Hyamine hydroxide (1M solution in methanol)	Fisons Scientific Apparatus, Loughborough, England.
Fuji Rx X-ray Film	Fuji Photo Film Co. Ltd., Tokyo 106, Japan.
Aminex A 6 Resin (sodium form)	BIO-RAD Laboratories Ltd. Caxton Way, Watford, England.
Silver stain reagents	"
Exonuclease III	Bethesda Research Laboratories (UK) Ltd., Science Park, Cambridge, England.
5'-azadeoxycytidine was kindly donated by Dr A. Piskala, Czechoslovak Academy of Sciences.	
Methidiumpropyl-EDTA, iron (II) was kindly donated by Professor P.B. Dervan, Division of Chemistry, California Institute of Technology.	
Ascites DNA methylase (purified according to Turnbull & Adams, 1976) was kindly provided by Dr R.L.P. Adams.	

II.1.3. Radiochemicals

Deoxyguanosine 5'- [ $\alpha$ - $^{32}\text{P}$ ]- triphosphate ( $\sim 3000$ Ci/mmole)	Radiochemical Centre, Amersham.
Deoxy- [ $^{\text{u}}\text{-}^{14}\text{C}$ ]-cytidine (480 mCi/mmole)	"
[ 6 - $^3\text{H}$ ]-uridine (25 Ci/mmole)	"
[ 6 - $^3\text{H}$ ]-thymidine (23 Ci/mmole)	"
L- [methyl - $^3\text{H}$ ]-methionine (15 Ci/mmole)	"
S-adenosyl-L- [methyl - $^3\text{H}$ ]- methionine (15 Ci/mmole)	"
[ 5 - $^3\text{H}$ ]-uridine 5'-triphosphate (1.0 Ci/mmole)	"
[U - $^{14}\text{C}$ ]-toluene (250 nCi/ml)	Packard Instrument Co. Inc., Downers Grove, Illinois.

II.1.4. Cell culture solutions

Glasgow modification of Eagles' Minimal Essential Medium (MEM) (Busby *et al.*, 1964). The constituents are listed in Table 2. This medium is supplemented with 10% (v/v) new born calf serum and designated EC<sub>10</sub>.

Basal medium for the culture of mosquito cells (Mitsuhashi & Maramorosch, 1964). The constituents are listed in Table 3. This medium is supplemented with 10% (v/v) foetal calf serum, 1% (v/v) 200 mM glutamine and 1% (v/v) non-essential amino acids. The pH was adjusted to 8.0 with 1.0 N NaOH. This medium is designated as MM.

The constituents of non-essential amino acids are listed in Table 4.

II.1.5. Stock Buffers

Buffered Saline Solution (BSS) (Earle, 1943)

NaCl	116.0 mM
KCl	5.4 mM
MgSO <sub>4</sub>	1.0 mM
NaH <sub>2</sub> PO <sub>4</sub>	1.0 mM
CaCl <sub>2</sub>	1.8 mM
phenol red	0.002% (w/v)

Table 2. Constituents of Eagles Minimal Essential Medium (MEM)  
as used in the Biochemistry Department, Glasgow University.

	<u>mg/litre</u>
D-glucose	4500.0
L-arginine HCl	126.9
L-cystine	28.5
L-glutamine	584.6
L-histidine HCl	21.0
L-isoleucine	52.5
L-leucine	52.5
L-lysine HCl	73.1
L-methionine	14.9
L-phenylalanine	33.0
L-threonine	47.6
L-tryptophan	8.2
L-tyrosine	36.2
L-valine	46.9
D-calcium pantothenate	2.0
choline chloride	2.0
folic acid	2.0
i-inositol	4.0
nicotinamide	2.0
pyridoxal HCl	2.0
riboflavin	0.2
thiamine HCl	2.0
MgSO <sub>4</sub> ·7H <sub>2</sub> O	200.0
KCl	400.0
NaCl	6400.0
NaH <sub>2</sub> PO <sub>4</sub> ·2H <sub>2</sub> O	140.0
CaCl <sub>2</sub> ·6H <sub>2</sub> O	264.9
phenol red	17.0

Table 3. Constituents of Mitsubishi and Maramorosch Basal Medium.

	<u>mg/litre</u>
Lactalbumin hydrolysate	8125.0
Yeastolate	6250.0
CaCl <sub>2</sub> ·2H <sub>2</sub> O	250.0
KCl	250.0
MgCl <sub>2</sub> ·6H <sub>2</sub> O	125.0
NaCl	8750.0
NaHCO <sub>3</sub>	150.0
NaH <sub>2</sub> PO <sub>4</sub> ·2H <sub>2</sub> O	282.6
D-glucose	5000.0

Table 4. Non-essential amino acids.

	<u>mg/litre</u>
L-alanine	8.9
L-asparagine·H <sub>2</sub> O	15.0
L-aspartic acid	13.3
glycine	7.5
L-glutamic acid	14.7
L-proline	11.5
L-serine	10.5

Buffer M

Tris-HCl pH 7.8	50.0 mM
EDTA	1.0 mM
dithiothreitol	1.0 mM
sodium butyrate	5.0 mM
glycerol	50.0% (v/v)

Buffer M<sup>+</sup> has PMSF in ethanol at 6 mg/ml added at 1% (v/v) immediately before use.

Stop Solution

sodium dodecyl sulphate	1.0% (w/v)
EDTA	2.0 mM
butan-1-ol	5.0% (v/v)
4-aminosalicylic acid	3.0% (w/v)

The 4-aminosalicylic acid is added immediately before use.

Sanders' buffers

Buffer A

Tris-HCl pH 7.0	0.1 mM
CaCl <sub>2</sub>	1.0 mM
MgCl <sub>2</sub>	2.0 mM
2-mercaptoethanol	15.0 mM
sodium butyrate (pH 7.0)	30.0 mM
glycerol	10.0% (v/v)

Buffer B

Tris-HCl pH 7.0	50.0 mM
MgCl <sub>2</sub>	2.0 mM
KCl	25.0 mM
2-mercaptoethanol	15.0 mM
sodium butyrate	30.0 mM
sucrose	250.0 mM

Transcription Buffer (TC) (Hay et al., 1982)

Hepes-NaOH pH 8.0	30.0 mM
ammonium sulphate	100.0 mM
MnCl <sub>2</sub>	1.5 mM
KCl	6.0 mM
dithiothreitol	1.25 mM
glycerol	12.5% (v/v)

MPE buffer

Tris-HCl pH 7.4	15.0 mM
NaCl	15.0 mM
KCl	60.0 mM
sucrose	0.5 M

Agarose gel buffers

Sample buffer

glycerol	25.0% (v/v)
sodium dodecyl sul- phate	2.0% (w/v)
bromophenol blue	0.025% (w/v)

Running buffer

Tris-HCl pH 7.7	36.0 mM
NaH <sub>2</sub> PO <sub>4</sub>	30.0 mM
EDTA	1.0 mM

Scintillation solutions

- toluene/PPO. 5.0 g of PPO per litre of toluene.
- toluene/triton/PPO. 5.0 g of PPO in 350 ml of triton X-100 and 650 ml of toluene.

II.2. METHODS

II.2.1. Cell culture techniques

II.2.1.1. Growth of cells

Mouse L929 cells were seeded at a density of approximately  $5 \times 10^6$  per Roux bottle in 50 ml of EC<sub>10</sub> and grown as monolayers at 37°C in an atmosphere containing 5% CO<sub>2</sub>.

Mosquito cells were seeded at a density of approximately  $5 \times 10^6$  per Roux bottle in 50 ml of MM and grown as monolayers at 27°C.

II.2.1.2. Contamination checks

All media and passaged cells were checked regularly for bacterial, fungal or PPLO infection as follows:

- a. Bacterial contamination: aliquots were added to blood agar plates and brain-heart infusion broth and were incubated at 37°C. Results were considered to be negative if no growth is seen after 7 days.

b. Fungal contamination: aliquots were added to Sabouraud's medium and cultured at 32°C. No growth after 7 days is assumed to indicate the absence of fungal contamination.

c. PPLO infection: agar plates were seeded with passaged cells by piercing the agar surface with a charged pasteur pipette. The plates were grown in an atmosphere of 95% N<sub>2</sub> and 5% CO<sub>2</sub> at 37°C. Infected cells resulted in the occurrence of the characteristic "fried egg" appearance of PPLO colonies on examination of the plates under the microscope.

#### II.2.1.3. Labelling of cells

With deoxy [U-<sup>14</sup>C]-cytidine, [6-<sup>3</sup>H]-uridine and [6-<sup>3</sup>H]-thymidine:

These isotopes were added to Roux bottles 24 hours after seeding at 0.1 µCi, 50 µCi and 50 µCi respectively. No alterations were made to the medium.

With L- [methyl-<sup>3</sup>H]-methionine:

L929 cells were grown for 2 days as in II.2.1.1. above. The medium was then replaced with 15 ml of EC<sub>10</sub> containing 3 mM hydroxyurea (to inhibit DNA synthesis), 60 µM hypoxanthine, 10 µM thymidine, 2 µM aminopterin and 20 mM sodium formate. After a further 1 hour at 37°C, 250 µCi of L- [methyl-<sup>3</sup>H]-methionine was added and the cells were harvested 3 hours later.

With 5'-azadeoxycytidine (Qureshi, 1983):

L929 cells at stationary phase were subcultured at a density of approximately 10 x 10<sup>6</sup> per Roux bottle and incubated at 37°C for 10 hours. 5'-azadeoxycytidine was then added to a final concentration of 1 µM and the incubation was continued for a further 10 hours after which the cells were harvested.

#### II.2.1.4. Harvesting of cells

The medium was decanted and the monolayer of cells washed in ice cold BSS, the cells were then scraped off the glass into 10 ml of BSS using a rubber wiper. Cells were pelleted by centrifugation at 850 x g for 10 minutes at 4°C.



#### II.2.1.5. Preparation of nuclei

Cells harvested as above were resuspended in 1% (v/v) Tween 80. These cells were homogenized with six strokes of a Potter homogenizer and nuclei recovered by centrifugation at 850 x g for 10 minutes at 4°C. Nuclear preparations were checked using a phase contrast microscope.

#### II.2.1.6. Preparation of a cell lysate

Cells harvested as above (II.2.1.4) were resuspended in buffer M minus glycerol and allowed to swell for 10 minutes. The cells were then gently pipetted up and down in a 1.0 ml tip (approximately 0.5 mm diameter) using a Gilson G-200 pipette. The preparation was checked using a phase contrast microscope.

### II.2.2. In vitro methylation

#### II.2.2.1. Methylation of nuclei

Nuclei prepared as in Section II.2.1.5 were resuspended in buffer M at a concentration of approximately  $5 \times 10^7$  per ml. 50  $\mu$ Ci of S-adenosyl-L- [methyl-<sup>3</sup>H]-methionine (15 Ci/mole) was added and the nuclei incubated for 3 hours at 37°C. In some experiments 40 units of a partially purified ascites DNA methylase were added prior to methylation.

#### II.2.2.2. Methylation of chromatin and DNA

The standard assay mixture (200  $\mu$ l) contained 0-20  $\mu$ g of DNA, 5  $\mu$ Ci of S-adenosyl-L- [methyl-<sup>3</sup>H]-methionine (15 Ci/mole), 3 nmole of non radioactive AdoMet and 40 units of ascites DNA methylase in buffer M. After incubation for 3 hours at 37°C the reaction was stopped by the addition of 2.0 ml of stop solution.

### II.2.3. Digestion of nuclei using nucleases

#### II.2.3.1. Micrococcal nuclease

Nuclei in buffer M containing  $\text{CaCl}_2$  at a concentration of 2 mM were treated with micrococcal nuclease at 500 units per ml at 37°C. Samples were taken at several times up to 30 minutes to which 1.5 ml of stop solution was added. Samples were also

taken at these times for electrophoresis on 1.5% (w/v) agarose gels.

For the production of 'core' chromatin particles, nuclei were treated as above until 50% of the DNA was rendered acid soluble.

#### II.2.3.2. Pancreatic deoxyribonuclease (DNase I)

This digestion procedure was the same as for micrococcal nuclease (Section II.2.3.1) except  $MgCl_2$  was used instead of  $CaCl_2$  and an enzyme concentration of 10 units per ml.

#### II.2.4. Purification of DNA for scintillation counting

DNA was purified from nuclei suspended in stop solution by the following procedure.

The suspensions were heated at 60°C for 20 minutes to dissolve the nuclei. 100 µg of salmon testis DNA was added as a carrier DNA for the precipitation step. Protein was removed from the mixture by extraction with a solution of phenol (88%); m-cresol (12%) and 8-hydroxyquinoline (0.1%). After centrifugation at 850xg for 10 minutes the upper aqueous layer (containing the DNA) was removed, leaving protein in the interphase material.

DNA was precipitated by mixing vigorously with 2 volumes of absolute ethanol. After centrifugation at 850xg for 10 minutes at -10°C this precipitate was redissolved in 100 µl of 0.5 N NaOH. The solution was incubated at 37°C for 1 hour to digest any RNA in the mixture. The solution was then spotted onto Whatman 3MM filter circles (2.5 cm diameter), washed four times in ice-cold 5% (w/v) trichloroacetic acid (10 ml per filter) and dried using methylated spirits and ether. The DNA was dissolved by heating in 0.3 ml of 1M hyamine hydroxide at 60°C for 20 minutes. 5 ml of toluene/PPO scintillator was added to the vial and radioactivity assayed using a liquid scintillation spectrometer.

### II.2.5. Double label technique

Most cells used in this study have had their DNA labelled with deoxy  $[U-^{14}C]$ -cytidine so that in each experiment the amount of  $^{14}C$  dpm is directly proportional to the amount of DNA, so different samples can be easily compared. The methylation is carried out using S-adenosyl-L-  $[methyl-^3H]$ -methionine, therefore the amount of methyl groups incorporated is directly proportional to the amount of  $^3H$  dpm. Therefore the relative methylation of different samples in one experiment can be easily compared.

The radioactivity of both  $^3H$  and  $^{14}C$  can be determined provided that the amount of  $^{14}C$  in the  $^3H$  channel of the scintillation counter can be determined.

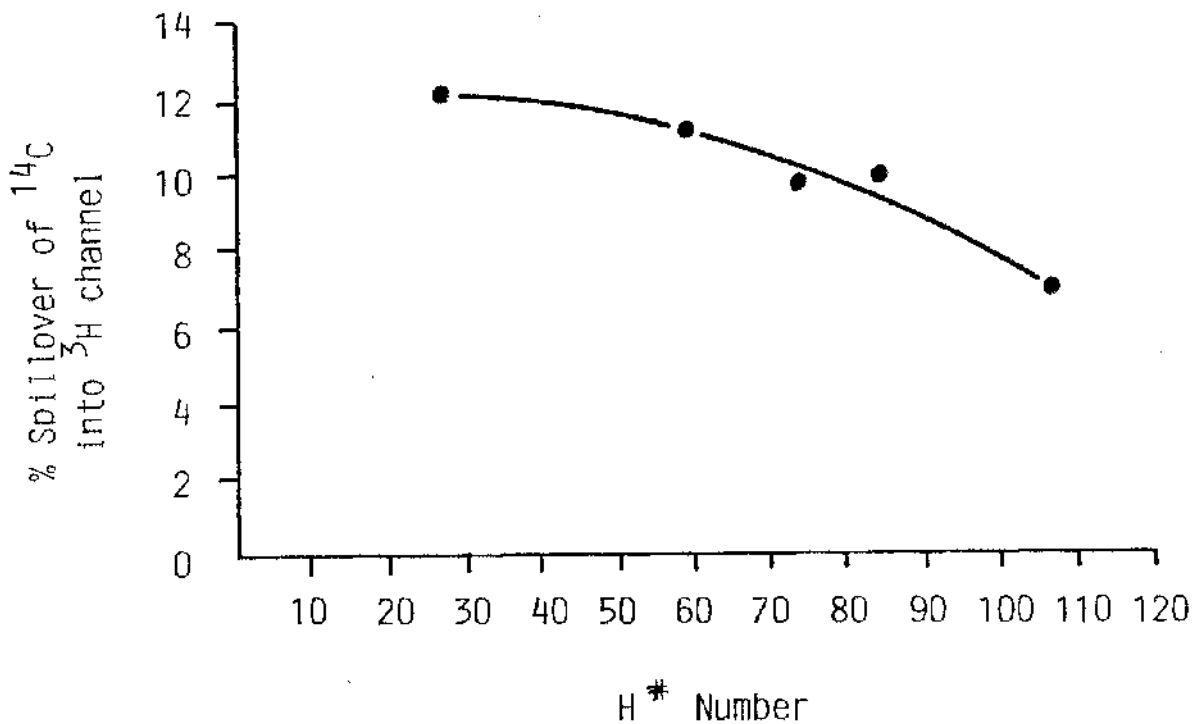
Six samples of 100  $\mu$ l  $[U-^{14}C]$ -toluene (250 nCi/ml) in 10 ml of toluene/PPO scintillation fluid were prepared. Chloroform was added to each sample to 0, 50, 100, 200, 300, and 400  $\mu$ l. These samples were counted on the dual label program of a scintillation counter. A calibration curve of percentage spillover into the  $^3H$  channel against H number was constructed (Fig. 4) so the amount of  $^3H$  can now be determined for any sample.

### II.2.6. Sanders' extraction procedure

This procedure is a modification of the original procedure of Sanders (1978) as used by Davie & Saunders (1981).

Nuclei were prepared as in II.2.1.5 and resuspended in buffer A at a concentration of approximately  $5 \times 10^7$  per ml. These nuclei were digested with micrococcal nuclease at 200 units per ml for 2 minutes at  $37^\circ C$ . The reaction was stopped by the addition of EGTA to a final concentration of 2 mM. The nuclei were centrifuged at  $850 \times g$  for 5 minutes and the supernatant (S0) was removed. The nuclei were resuspended in 500  $\mu$ l of buffer B containing 50 mM NaCl, and recentrifuged and the supernatant removed (S 50). This procedure was repeated using buffer B containing 100 mM, 200 mM, and 400 mM NaCl to yield supernatants S 100, S 200, and S 400 respectively. The final pellet was resuspended in 500  $\mu$ l of buffer B. To each fraction 1.5 ml of stop solution was added. Samples were taken at each

Figure 4. Calibration curve for the determination of the amount of  $^{14}\text{C}$  cpm in the  $^3\text{H}$  channel of a liquid scintillation spectrometer.



Six samples of 100  $\mu\text{l}$   $[\text{U-}^{14}\text{C}]$ -toluene (250 nCi/ml) in 10 ml of toluene/PPO scintillation fluid, containing 0, 50, 100, 200, 300 and 400  $\mu\text{l}$  of chloroform respectively, were prepared. These samples were counted on the dual label program of a scintillation counter. The % spillover of  $^{14}\text{C}$  cpm into the  $^3\text{H}$  channel was determined and plotted against H number.

The H number method measures the response produced in any liquid scintillation sample by the same electron energy and is thus a measure of the amount of quench in the sample. This method is used with Beckman LS-8000 series scintillation spectrometers.

stage in the procedure for electrophoresis on 1.5% agarose gels and for scintillation counting.

When nuclei were methylated prior to this extraction procedure they were washed once in buffer A prior to treatment with micrococcal nuclease.

#### II.2.7. Digestion of nuclei using methidiumpropyl-EDTA.iron (II)

Nuclei were prepared as in II.2.1.5 and resuspended in MPE buffer at a concentration of approximately  $5 \times 10^7$  per ml.

An equal volume of a mixture containing 100  $\mu$ M methidiumpropyl-EDTA, 10 mM dithiothreitol and 0.1 mM ferrous ammonium sulphate was added.

Then EDTA and hydrogen peroxide (30% v/v) were added to final concentrations of 1 mM and 0.5 mM respectively. The mixture was incubated for 16 hours at 20°C and the cleavage reaction was stopped by the addition of 0.1 volume of 50 mM bathophenanthrolinedisulfonate.  $MgCl_2$  was added to a final concentration of 20 mM and 1 unit of exonuclease III was added. After incubation for 10 minutes at 37°C, 2 volumes of stop solution was added.

DNA was purified as in Section II.2.9 and redissolved in 50 mM tris-HCl pH 8.0 containing 200 mM NaCl and 10 mM  $MgCl_2$ . The DNA was digested with Neurospora crassa nuclease at 20  $\mu$ g per ml for 10 minutes at 37°C. The reaction was stopped with 2 volumes of stop solution and the DNA repurified. The base composition of the DNA was then determined as in Section II.2.11.2.

#### II.2.8. Preparation of nuclear non-histone proteins

Nuclei from ascites cells in buffer  $M^+$  were extracted with buffer  $M^+$  containing 0.35 M NaCl and centrifuged at 850xg for 10 minutes. The supernatant was recovered and dialyzed against buffer  $M^+$  containing 50% (v/v) glycerol for 24 hours.

To prepare high mobility group (HMG) proteins the undialysed 0.35 M NaCl extract was made 2.5% (w/v) with respect to trichloroacetic acid and centrifuged at 10,000xg for 10 minutes. The supernatant was neutralized and dialyzed against buffer  $M^+$  containing 50% (v/v) glycerol for 24 hours.

### II.2.9. Purification of DNA

Nuclei in buffer M were incubated with ribonuclease A (boiled for 10 minutes before use) at 10 units per ml for 1 hour at 37°C. Two volumes of stop solution were added and protein was removed as previously described. The aqueous layer was extracted with a mixture of chloroform (95%), isoamylalcohol (5%) and centrifuged at 850 x g for 10 minutes. The aqueous layer was removed and DNA precipitated as previously described. The DNA was washed twice with absolute ethanol and dried.

DNA was assayed by the method of Burton (1956).

### II.2.10. Nick-translation of DNA (Maniatis et al., 1975).

50  $\mu$ Ci of deoxyguanosine 5'-[ $\alpha$ -<sup>32</sup>P]-triphosphate was put into a siliconized Eppendorf tube and lyophilized. To this tube was added 5  $\mu$ l of a DNA solution (1 mg/ml), 6  $\mu$ l of a solution of dATP, dTTP and dCTP (16 mM with respect to each); 3  $\mu$ l of 500 mM tris-HCl pH 7.5 containing 50 mM MgCl<sub>2</sub>, 1  $\mu$ l of 2% (v/v) 2-mercaptoethanol and 5 units of *E.coli* polymerase I. Water was added to give a final volume of 30  $\mu$ l.

The tube was incubated for 1 hour at 12°C and the reaction was stopped by the addition of 2  $\mu$ l of 0.5 M EDTA and incubation for 5 minutes at 60°C.

The reaction mixture was applied to a 5 cm x 0.5 cm column of sephadex G-50 and eluted with 10 mM tris-HCl pH 7.5 containing 0.1 mM EDTA. The DNA peak was collected and NaCl added to a final concentration of 0.2 M. The DNA was precipitated by the addition of 3 volumes of absolute ethanol and storage at -20°C overnight followed by centrifugation at 10,000xg for 1 hour.

### II.2.11. Base composition analysis of DNA

#### II.2.11.1. By enzymatic digestion (Shatkin, 1969)

50  $\mu$ l of labelled DNA; 8  $\mu$ l of 0.4 M tris-HCl pH 8.6 containing 0.2 M CaCl<sub>2</sub> and 30 units of micrococcal nuclease were incubated for 2 hours at 37°C. 34  $\mu$ l of 0.5 M potassium phosphate pH 7.0 buffer was then added to adjust the pH to 7.0. The digestion was continued by the addition of 2 units of spleen phosphodiesterase and incubating for 1 hour at 37°C.

A further 2 units of spleen phosphodiesterase was added and the mixture reincubated. This was repeated once more (i.e. a total of 6 units of spleen phosphodiesterase was used).

At the end of the incubation, samples of the DNA digest were applied as single spots to 45 cm x 57 cm sheets of Whatman No.1 paper, prewetted with pH 3.5 acetate buffer (acetic acid 3.3% (v/v), pyridine 0.33% (v/v), EDTA 0.38% (w/v)), about 10 cm from one end. The paper was electrophoresed at 3.0 kV for 75 minutes, dried in air and autoradiographed. The appropriate radioactively labelled areas of the sheets were cut out and placed in vials containing 10 ml of water and the radioactivity assayed using the Cerenkov program of a liquid scintillation counter.

#### II.2.11.2. By acid pyrolysis. (Adams et al., 1979)

Samples of DNA were dissolved in 150  $\mu$ l of 98% formic acid and sealed in thick walled pyrex tubes. The tubes were heated at 170°C for 2 hours. The tubes were cooled, opened and the excess formic acid evaporated. The pyrolysed material was taken up in 100  $\mu$ l of 20 mM ammonium carbonate (adjusted to pH 10.0 with ammonia) and applied under pressure to a 20 cm x 1 cm column of Aminex A6 (equilibrated at 50°C) at a flow rate of 1.0 ml per minute. 1.0 ml samples were collected to which was added 10 ml of triton/toluene/PPO scintillator and the radioactivity was assayed.

#### II.2.12. Agarose gel electrophoresis (Hayward, 1972)

Samples of DNA were dissolved in agarose gel sample buffer and layered onto 10 cm x 8 cm, 1.5% (w/v) agarose slab gels. The gels were subjected to electrophoresis at 50 mA for 3 hours. Gels were stained with ethidium bromide at 0.5  $\mu$ g per ml for 10 minutes and DNA was visualized using ultraviolet light.

#### II.2.13. Acrylamide gel electrophoresis

Protein samples were dissolved in a mixture of glycerol (20% v/v), sodium dodecylsulphate (1% w/v), 2-mercaptoethanol (1% v/v) and bromophenol blue (0.04% w/v) and electrophoresed

on a 10 cm x 10 cm 7% (w/v) slab acrylamide gel at 10 mA for 10 hours. The running buffer was 0.12 M sodium phosphate pH 6.5 containing 1% (w/v) sodium dodecylsulphate. Gels were stained using the silver stain technique (BIO-RAD manual).

#### II.2.14. Computer analyses

A number of different mouse gene sequences were selected from the EMBL Nucleotide Sequence Data Library (Release 2) on a PDP11/34 computer.

The C + G content of 100 nucleotide segments were derived by scoring their frequency of occurrence starting successively at nucleotides 1, 2, ..., N-99 in the N nucleotides being analyzed. The C + G content of 148 nucleotide segments was also determined.

The bendability function values for 148 nucleotide segments were derived by scoring the number of RY and YR dinucleotides at intervals of 5-6 base pairs starting successively at nucleotides 1, 2, ..., N-147 in the N nucleotides being analyzed (148 nucleotides being chosen because this is the 'core' repeat length of nucleosomes) (Zhurkin, 1983).



III. RESULTS

Note: The experiments described in Sections III.2.7, III.2.8, III.4.4, III.5.3 and III.5.4 were performed in collaboration with Dr R.L.P. Adams.

III.1. In vivo methylation of DNA in L929 cells relative to nucleosomes

III.1.1. The distribution of  $m^5C$  in chromatin

In Section I.6 it was mentioned that the nucleosomal core DNA of chicken and calf thymus chromatin is enriched in  $m^5C$  (Razin & Cedar, 1977) though a similar enrichment was not observed in Chinese hamster ovary cells (Adams *et al.*, 1977).

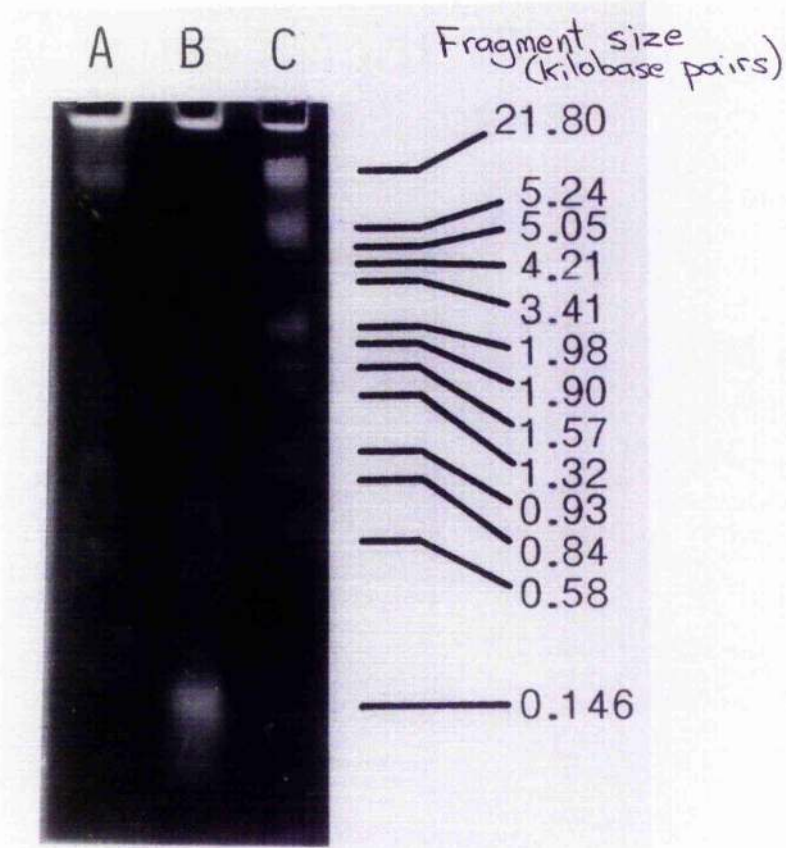
Subsequently I have investigated the distribution of  $m^5C$  in the chromatin of mouse L929 cells and related this to the predicted flexibility of DNA as determined by the method of Zhurkin (1983).

III.1.2. Core DNA of L929 chromatin is enriched in  $m^5C$

In the present work nuclei were prepared from [ $6-^3H$ ]-uridine prelabelled L929 cells (Section II.2.1.5) and these were treated with micrococcal nuclease until about 50% of the DNA was rendered acid soluble (the DNA which remains acid insoluble is about 140 base pairs in length, Fig. 5) (Section II.2.3.1). The residual DNA was purified (Section II.2.9) and pyrolysed using formic acid to the individual bases which were separated on a column of Aminex A6 (Section II.2.11.2). The cytosine and methylcytosine fractions were collected and the amount of  $m^5C$  determined (Section III.1.3). The amount of the dinucleotide CpG in the DNA was determined by nick-translation of DNA in the presence of [ $\alpha-^{32}P$ ]-dGTP (Section II.2.10), hydrolysis using nucleases to 3'-monophosphates (Section II.2.11.1) followed by separation by high voltage electrophoresis (Fig. 6). The C + G content was also determined.

The results are given in Table 5. As can be seen the nuclease resistant DNA is enriched in  $m^5C$  by 25% compared to total DNA. This DNA is also enriched in cytosine and guanine and in the dinucleotide CpG (this is the DNA methylase

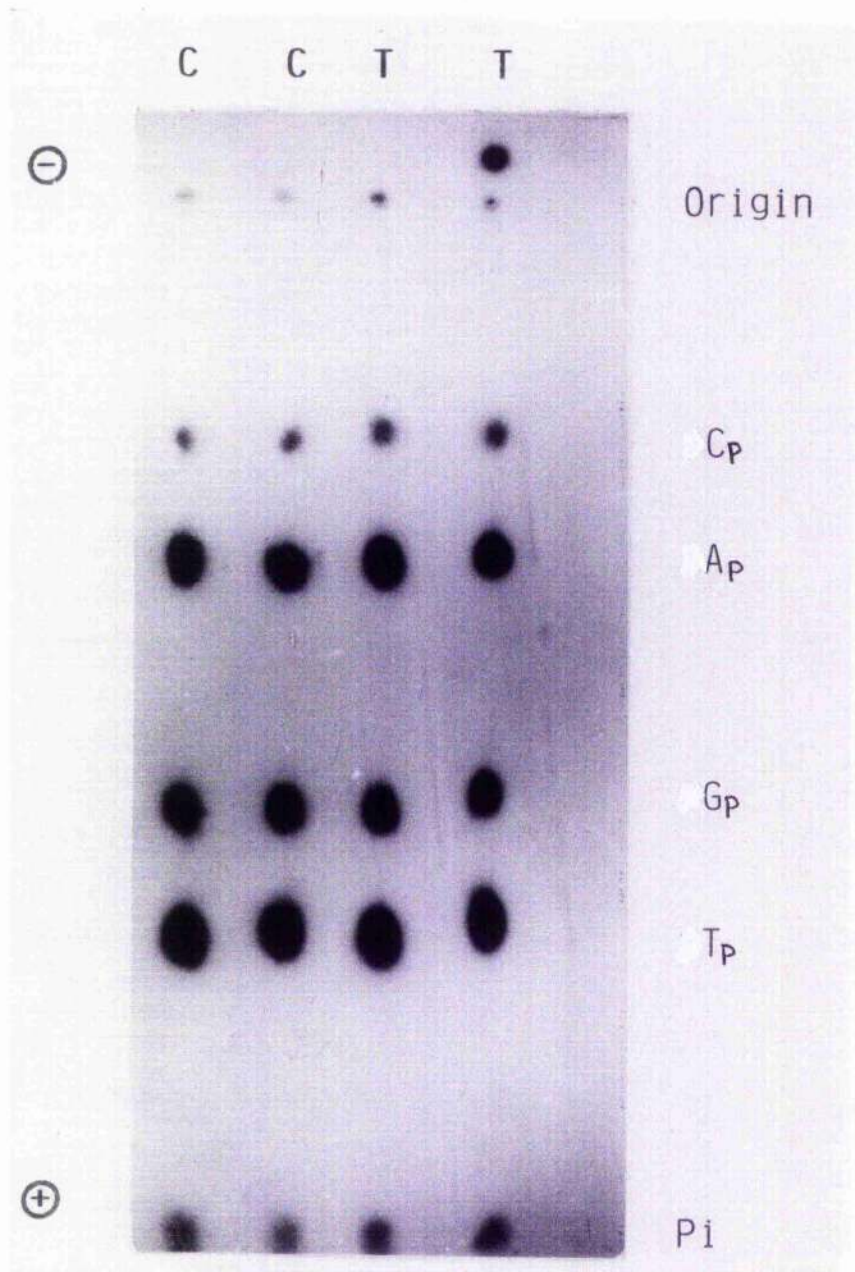
Figure 5. Agarose gel of DNA from total and nucleosomal core chromatin from L929 cells.



Lanes: A: DNA from total chromatin  
B: DNA from core chromatin  
C: Marker.  $\lambda$  DNA digested with EcoRI and Hind III.

Nuclei were digested with micrococcal nuclease until about 50% of the DNA was rendered acid soluble. Samples of DNA remaining acid insoluble were electrophoresed on a 1.5% agarose gel (Section II.2.12). DNA was visualized using ethidium bromide.

Figure 6. High voltage electrophoresis of 3'-mononucleotides.



Lanes: C: nucleotides from core DNA  
T: nucleotides from total DNA

DNA was nick-translated in the presence of deoxyguanosine 5'-[ $\alpha$ - $^{32}\text{P}$ ]-triphosphate (Section II.2.10) and digested to give 3'-mononucleotides which were separated by high voltage electrophoresis (Section II.2.11.1).

recognition site) and it appears possible that the enrichment in  $m^5C$  is a result of this enrichment (See also Adams & Eason, 1984).

Table 5. Base composition of the DNA from total and nucleosomal core chromatin from L929 cells.

	Total chromatin	Core chromatin	<sup>†</sup> Statistical significance
* $m^5C$	3.69 ± 0.20 (5)	4.79 ± 0.30 (5)	p < 0.001
%C + G	41.67 ± 1.53 (11)	45.24 ± 1.18 (9)	p < 0.001
*CpG	3.37 ± 0.36 (9)	4.23 ± 0.36 (7)	p < 0.001

\* See Section III.1.3.

† The statistical test used is the Student's t test based on the standard error of the difference between the means. The values are the mean ± standard deviation. The number of samples is given in brackets.

### III.1.3. Calculation of the base composition of DNA

As the DNA was labelled using [6-<sup>3</sup>H]-uridine both cytosine and  $m^5C$  will be labelled and the percent  $m^5C$  is simply

$$\% m^5C = \frac{\text{dpm in } m^5C \times 100}{\text{dpm in } (m^5C + C)}$$

Similarly for the CpG. The amount of <sup>32</sup>P in each 3'-monophosphate is proportional to the amount of the NpG in the original DNA. Therefore the amount of CpG is expressed by

$$\% CpG = \frac{\text{dpm in dCMP} \times 100}{\text{dpm in } (dCMP + dGMP + dAMP + dTMP)}$$

The percent C + G is obtained by measuring the area under the peaks in the Aminex A6 separation profile (by computer integration) and dividing by the appropriate  $m^M$  extinction coefficient at pH 10.0 for each base. These are: T, 3.7; A, 10.45; G, 6.4; C, 5.55;  $m^5C$ , 4.45. This gives the relative proportions of each base (Figure 15, page 76 shows an Aminex A6 separation profile).

#### III.1.4. Digestion of DNA with micrococcal nuclease

The type of experiment described in Section III.1.2 only gives a valid result if the nuclease used cleaves the DNA in a random manner. Micrococcal nuclease has two known specificities. The primary cutting sites are in adenine and thymine rich regions of DNA (Nelson et al., 1979a; Dingwall et al., 1981; Horz & Altenberger, 1981). Micrococcal nuclease has a high preference for cleavage in the linker regions of chromatin (Rill & van Holde, 1973; Noll, 1974a) but will also cleave in the core regions at adenine and thymine rich sites (McGhee & Felsenfeld, 1983). It is therefore necessary to determine that these specificities are not the cause of the result observed in Section III.1.2.

DNA was purified from [6-<sup>3</sup>H]-uridine prelabelled L929 cells (Section II.2.9) and digested using micrococcal nuclease. Samples were taken at several times during the digestion and the amount of m<sup>5</sup>C was determined (Table 6).

Table 6. 5-methylcytosine content of DNA at several times of digestion using micrococcal nuclease.

Percent DNA rendered acid soluble	*m <sup>5</sup> C
0	3.62 ± 0.03 (3)
10	3.89 ± 0.22 (3)
23	3.70 ± 0.14 (3)
45	3.59 ± 0.16 (3)

\*See Section III.1.3.

The figures are the mean ± standard deviation. The number of samples is given in brackets.

This result shows that the level of m<sup>5</sup>C does not change over the first 45% of the DNA made acid soluble. This suggests that the observed sequence specificity of micrococcal nuclease applies to the initial endonucleolytic activity but

not to the subsequent exonucleolytic activity.

III.1.5. Digestion of chromatin using methidiumpropyl-EDTA.iron (II)

The problem due to cleavage in the core DNA can be countered by the use of an agent which is not base specific in its cleavage and has a preference for linker DNA. Such an agent is the chemical reagent methidiumpropyl-EDTA.iron (II). This reagent can intercalate into DNA and in the presence of oxygen will efficiently cleave DNA (Hertzberg & Dervan, 1982). This chemical has no observable base specificity (van Dyke & Dervan, 1982; 1983), it has a strong preference for linker DNA and it can be used in the presence of low amounts of EDTA (Cartwright et al., 1983).

When the DNA in chromatin from L929 cells is digested using this reagent followed by digestion using exonuclease III and Neurospora crassa nuclease (Section II.2.7) a DNA fragment which is about 140 base pairs is produced (Fig. 7). This DNA is again found to be enriched in m<sup>5</sup>C (Table 7).

Table 7. 5-methylcytosine content of the DNA in chromatin and methidiumpropyl-EDTA.iron(II) treated chromatin.

	Total chromatin	Digested chromatin	<sup>+</sup> Statistical significance
*m <sup>5</sup> C	3.60 ± 0.09 (5)	4.46 ± 0.23 (4)	p < 0.001

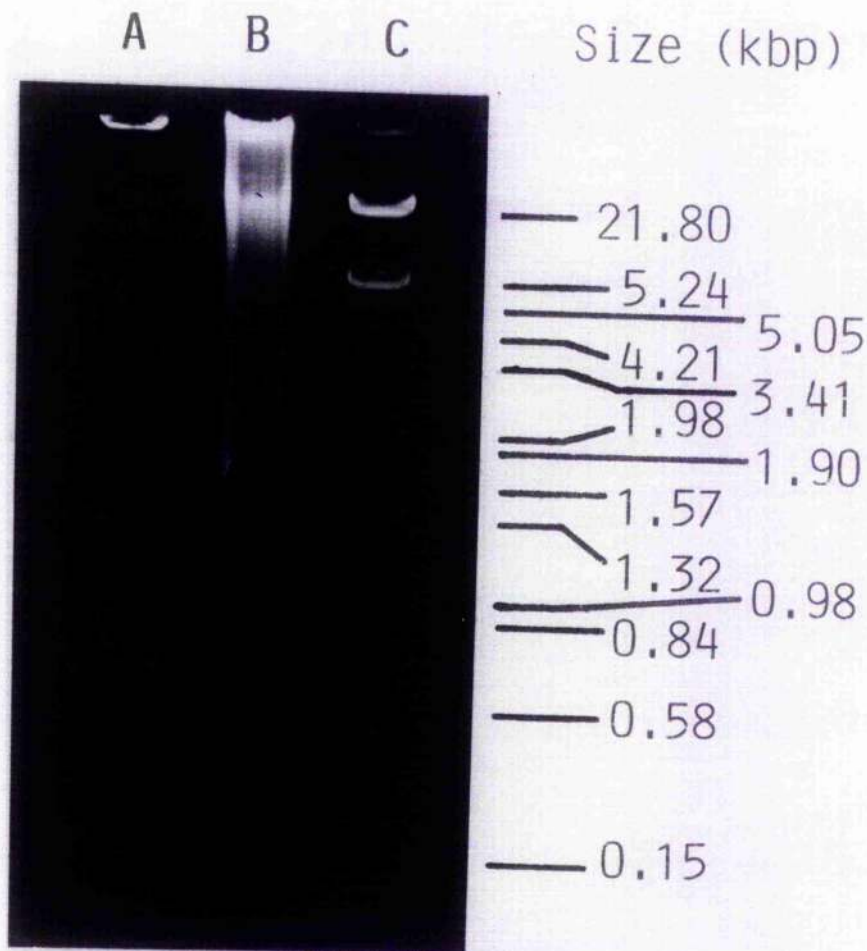
\*See Section III.1.3.

+See Table 5 (page 58).

The figures are the mean ± standard deviation. The number of samples is given in brackets.

These two results taken together effectively eliminate artifacts due to specificities of micrococcal nuclease digestion. The problem of internal cleavage in core DNA is also countered by the fact that all the DNA remaining acid insoluble was used and not simply the 140 base pair fragment.

Figure 7. Agarose gel of DNA from total and methidiumpropyl-EDTA.iron (II) digested chromatin.



Lanes; A: DNA from methidiumpropyl-EDTA.iron (II) treated chromatin.

B: DNA from total chromatin

C: Marker,  $\lambda$  DNA digested with EcoRI and Hind III.

Nuclei were treated with the methidium reagent, exonuclease III and Neurospora crassa nuclease as described in Section II.2.7. Samples of DNA from total and methidiumpropyl-EDTA.iron (II) treated chromatin were electrophoresed on a 1.5% agarose gel (Section II.2.12). DNA was visualized using ethidium bromide.



III.1.6. Reconstitution of chromatin

The experiments above have shown that nucleosomes are associated with C + G and m<sup>5</sup>C rich DNA. This can be explained in one of three ways: 1) nucleosomes associate preferentially with this DNA following DNA replication, 2) DNA methylase has a preference for nucleosomal core DNA and 3) it is possible that nucleosomes located on specific sequences decrease the propensity of methylcytosines in these sequences to undergo deamination to thymine.

DNA purified from [6-<sup>3</sup>H]-uridine prelabelled cells was mixed with an equal mass of total calf thymus histones in buffer M containing 2.0 M NaCl. This mixture was dialysed against 500 volumes of buffer M containing 0.6 M NaCl for 2 days followed by dialysis against buffer M alone for 2 days at 4°C. The reconstituted chromatin was digested using micrococcal nuclease and the m<sup>5</sup>C content of the DNA was determined at several times (Table 8 and Fig. 8).

Table 8. 5-methylcytosine content of DNA from reconstituted chromatin at several times of digestion using micrococcal nuclease.

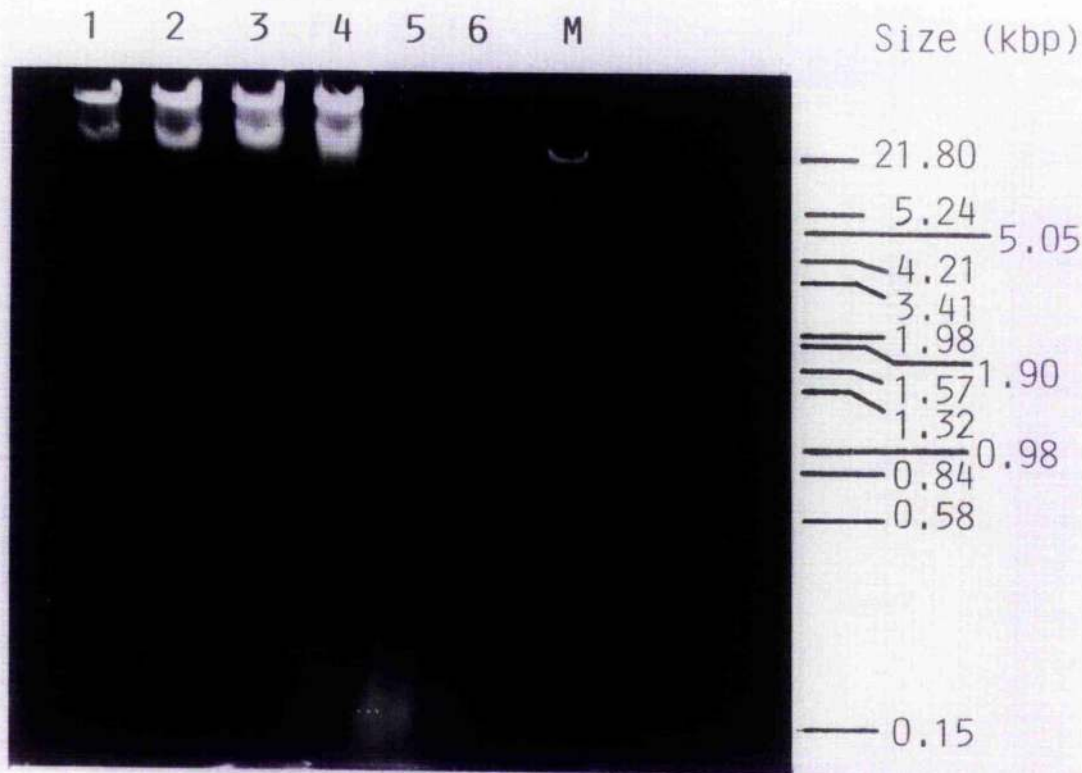
Percent DNA rendered acid soluble	*m <sup>5</sup> C
0	3.76 ± 0.10 (4)
10	4.30 ± 0.15 (3)
28	4.87 ± 0.25 (3)
42	4.96 ± 0.14 (3)
53	5.05 ± 0.23 (3)

\*See Section III.1.3.

The figures are the mean ± standard deviation. The number of samples is given in brackets.

As before the DNA in the chromatin which is resistant to nuclease digestion is enriched in m<sup>5</sup>C, thus suggesting that nucleosomes reassociate preferentially with m<sup>5</sup>C (C + G) rich DNA. This result is in contrast to that of Razin & Cedar (1977)

Figure 8. Agarose gel of DNA from reconstituted chromatin digested using micrococcal nuclease.



Lanes; 1-6: Times of digestion, 0, 1/2, 2, 6, 15 & 30 mins.

M: Marker,  $\lambda$  DNA digested with EcoRI and Hind III.

Chromatin was reconstituted as described in Section III.1.6 and digested using micrococcal nuclease. Samples were taken at the above times and electrophoresed on a 1.5% agarose gel (Section II.2.12). DNA was visualized using ethidium bromide.

who found that reconstitution of chicken chromatin resulted in a random reassociation of nucleosomes.

As can be seen from Table 8 most of the increase in  $m^5C$  content occurs during the first 30% of DNA made acid soluble. As linker DNA makes up about 30% of the total DNA, this reinforces the idea that it is indeed the linker DNA that is undermethylated.

### III.1.7. Models predicting the locations of nucleosomes on DNA

The DNA of the core particle in chromatin is wrapped around an octamer of histones (Section I.5.2). It is therefore reasonable to assume that this DNA has a higher degree of flexibility than linker DNA.

Two models have been proposed to determine the degree of flexibility of a particular DNA sequence. The model of Trifonov (1980) depends upon the concept that certain base pairs in DNA form 'wedges' which 'kink' the DNA in a certain direction and so will facilitate the formation of nucleosomes. The base pairs which have the greatest tendency to form these wedges are made up of adenine and thymine residues. This model has two drawbacks: 1) it does not take into account the 10.5 base pair periodicity inherent in the DNA of the core particle and 2) it assumes that all base pairs in a sequence are equally important. Further work by Zhurkin (1981) has shown that the periodicities observed by Trifonov occur largely in protein coding regions of DNA and are a consequence of the base sequence required to code for proteins. Zhurkin argues that only those residues at intervals of about 10.5 base pairs are important in the folding of DNA in nucleosomes. He predicts that the presence of alternating purine-pyrimidine and pyrimidine-purine residues at positions 1, 5, 11, 16 etc. in the nucleosome will enhance the flexibility of the DNA and so facilitate the formation of nucleosomes (Zhurkin, 1983).

#### III.1.7.1 Application of the model of Zhurkin to specific gene sequences

Specific mouse gene sequences have been analyzed for the presence of RY and YR dimers at intervals of 10.5 bases in

all possible frames of 148 bases along the sequence (Section II.2.14) and this data has been plotted for the mouse  $\beta$ -globin gene (Figure 9).

As was mentioned in Section III.1.7 the presence of these dimers at regular intervals of 10.5 bases will enhance the flexibility of DNA. Therefore the frame of 148 bases which has the largest number of dimers at these positions will be the most flexible region of the particular sequence. Figures 9a and 9b show two regions of the mouse  $\beta$ -globin gene which are typical of the sequences studied. The individual maxima are shown by arrows and the 10.5 base repeat can be clearly seen.

The most flexible region predicted for this gene starts at base number 1437 (Figure 9a). Once the first nucleosome has been positioned the other nucleosomes should be spaced at regular intervals of about 200 bases from this starting point, and Figure 9c shows that there is a definite variation of maxima of about 200 bases.

#### III.1.7.2 The variation in the C + G content along mouse genes

The same genes used in Section III.1.7.1 have been analyzed for the number of C + G residues which are present in 100-base frames along the sequences (Section II.2.14) and this data is plotted in Figure 10. The most obvious feature of these plots is the wide variation in C + G content along the DNA, varying from 20% to 80% C + G (see also Adams & Eason, 1984). On these figures are depicted the starting positions of nucleosomes predicted using the model of Zhurkin (1983). It is clear from these figures that in many cases the predicted nucleosome positions coincide with regions of high C + G content, and more importantly, very few predicted positions coincide with regions of very low C + G content.

The actual C + G content of predicted nucleosome positions can be calculated using the C + G contents of 148-base frames (not 100-base frames) which start at the same base as the predicted nucleosome (Table 9). In all cases the C + G content of the DNA in the predicted nucleosome position

Figure 9. Predicted flexibility of the  $\beta 1$ -globin gene.

9a: bases 1200 to 1500

9b: bases 121 to 421

9c: bases 50 to 1550

The bendability function ( $F'(M)$ ) is derived by scoring the frequency of RY and YR dimers at positions 1, 5, 11...etc. in 148-base frames along the DNA sequence.

The figures show this data plotted against nucleotide number for two regions of the  $\beta 1$ -globin gene of mouse which are typical of the sequences analyzed.

The 10.5 base variation of maxima is depicted by the arrows. Predicted positions for nucleosomes are shown by boxes.

In Figure 9c the 200 base variation of regions of high predicted flexibility can be clearly seen (continuous line in figure). The open box represents a degree of uncertainty in the nucleosome position.

This method is based upon the observation that YR dimers 'kink' the double helix into the major groove and RY dimers 'kink' the double helix into the minor groove (Dickerson & Drew, 1981). As these minikinks are approximately 5 bp apart only those dimers at positions 1, 5, 11, etc. need to be considered.

$F'(M)$  is simply the number of YR and RY dimers at positions 1, 5, 11 etc. in a given 148 base frame. The maximum value for this function is 29 and the mean value for a random sequence is 12.5 (Zhurkin, 1983).

The starting position of a nucleosome is expected where  $F'(M)$  is a maximum as this 148 base frame has the highest bendability. In figure 9a this is at base number 1438.

Figure 9a.

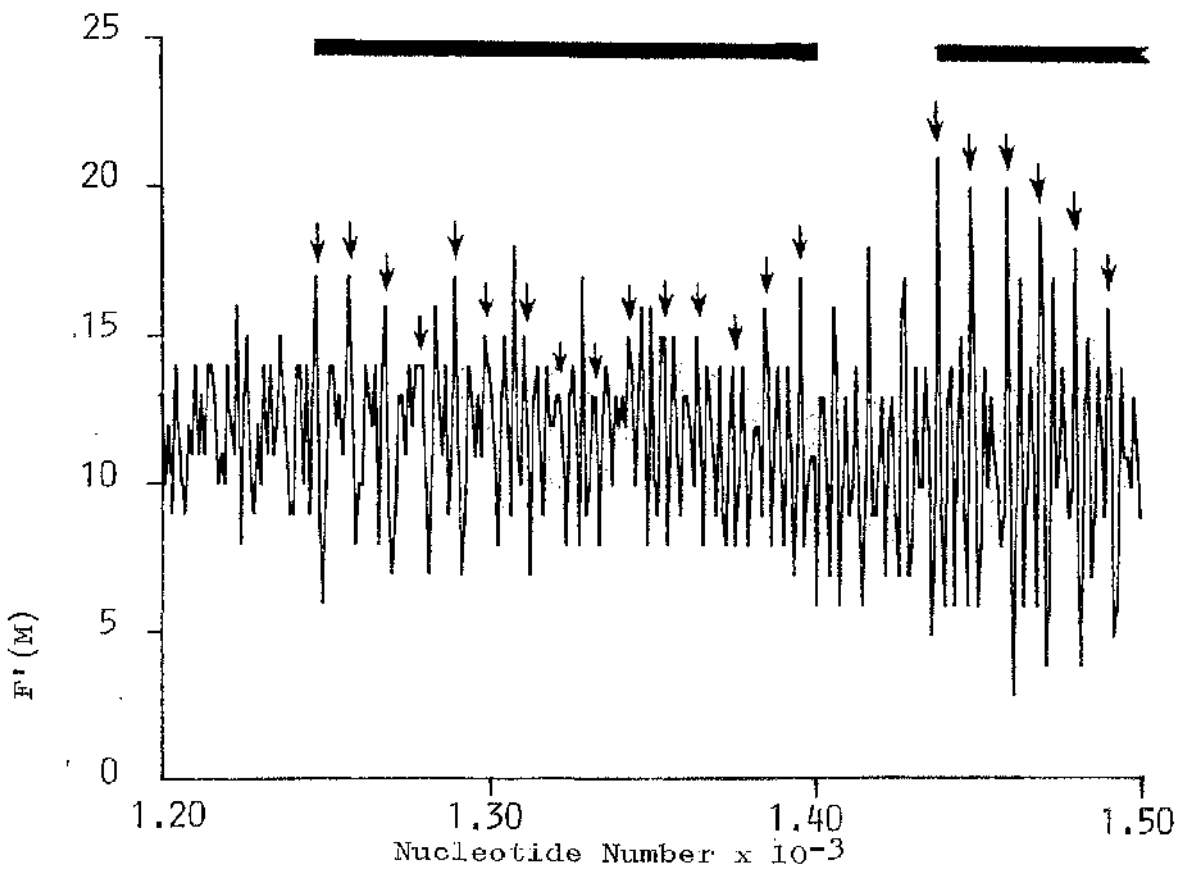


Figure 9b.

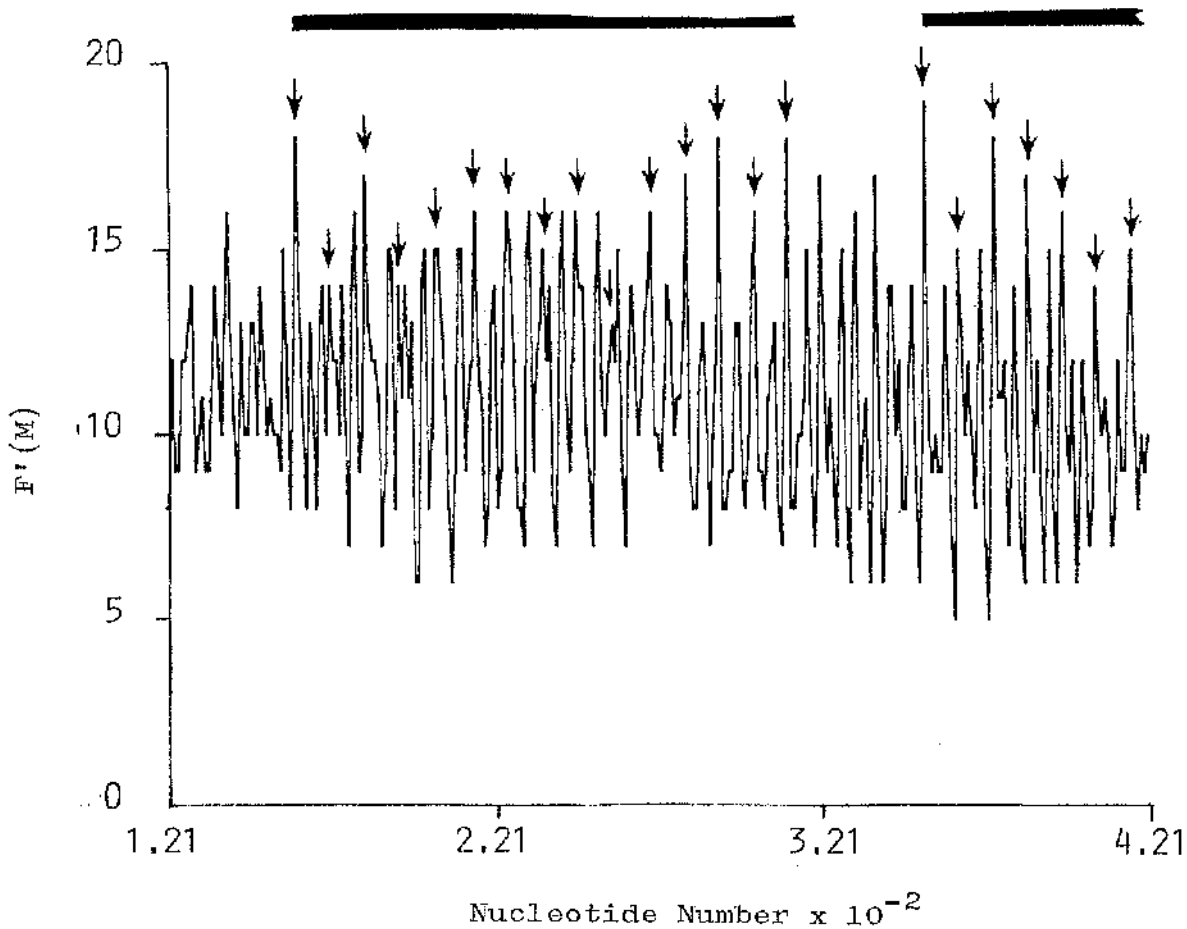


Figure 9c.

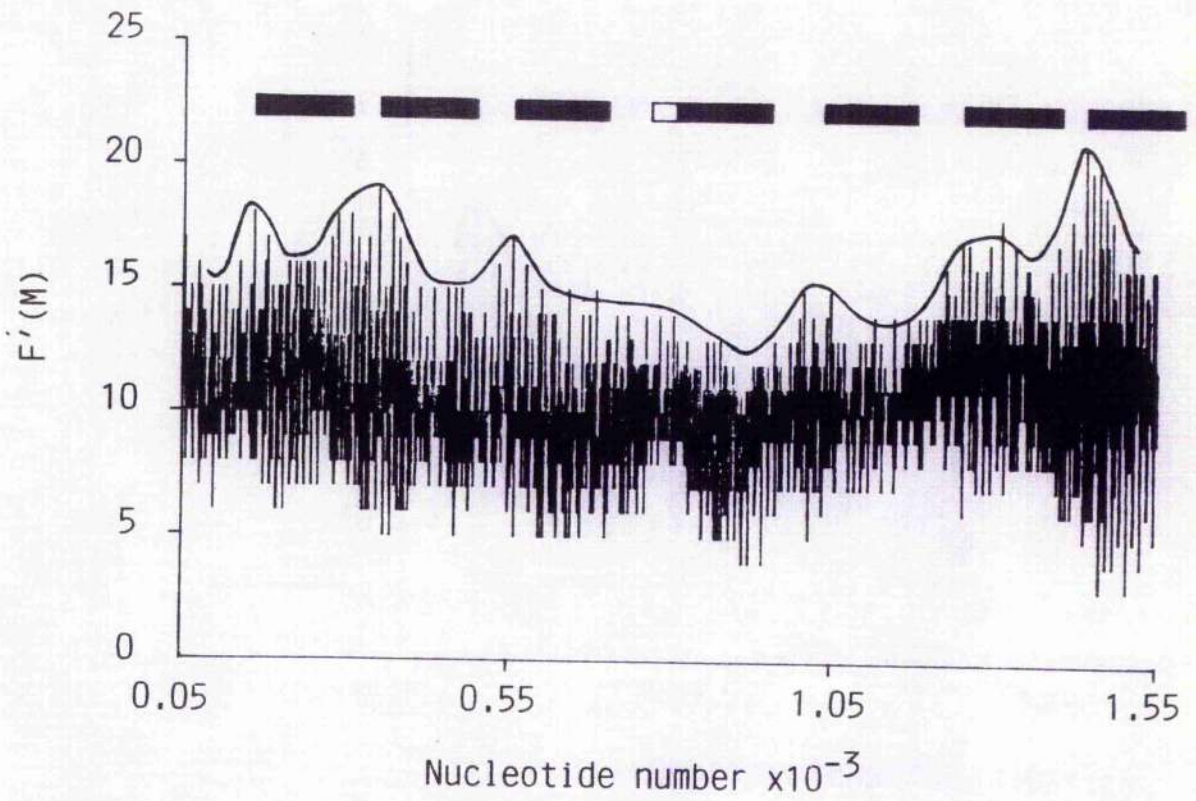
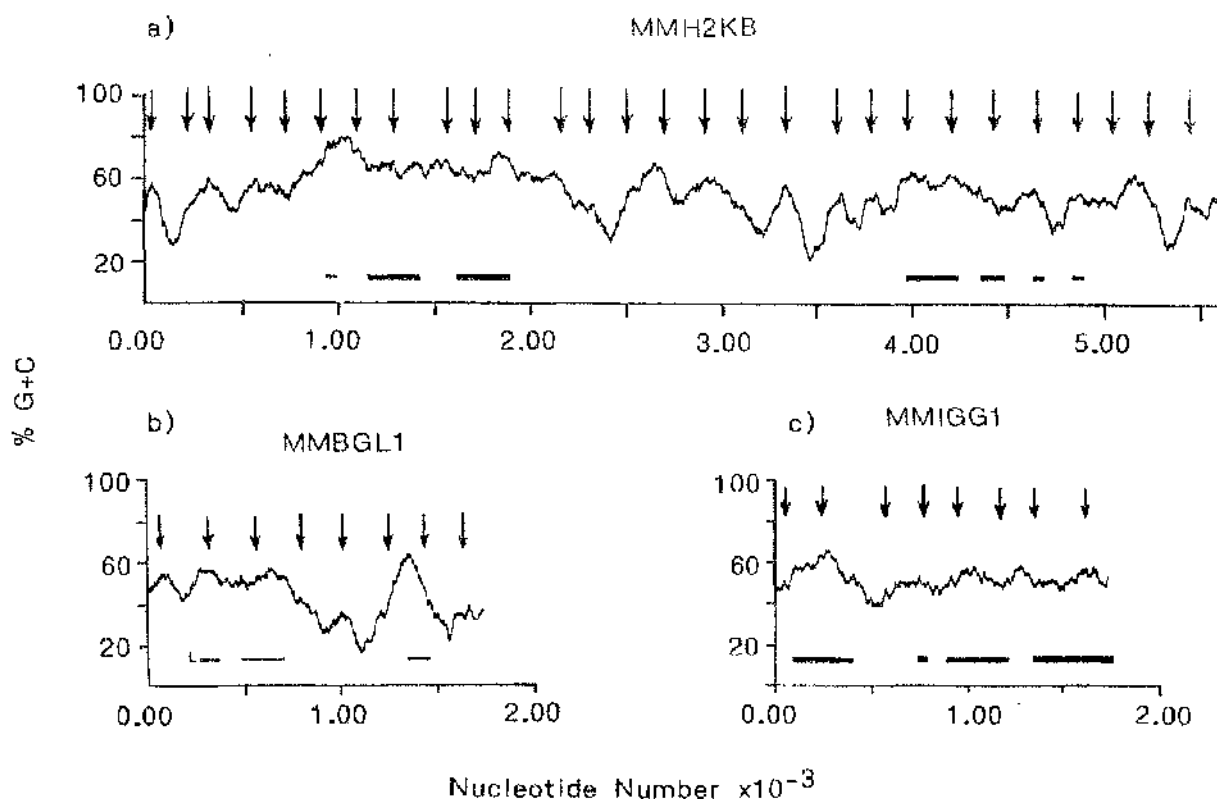


Figure 10. Plot of the variation in C + G content along mouse genes.



10a: MMH2KB, Mouse pseudo gene for the transplantation antigen complex H-2KB.

10b: MMBGL1, Mouse gene for  $\beta 1$ -globin

10c: MMIGG1, Mouse gene for the constant part of the  $\gamma 1$ -immunoglobulin.

The % C + G was determined by scoring the frequency of C and G residues in 100-base frames along the sequence. This data is then plotted against nucleotide number.

In the figures the arrows represent the predicted nucleosome starting positions as determined by the  $F'(M)$  function of Figure 9. The boxes represent exons and the 'L' in Figure 10b signifies the start of the messenger RNA.



Table 9. C + G content of the DNA in predicted nucleosomes positions on mouse DNA sequences.

DNA sequence	Average C + G (%) content of DNA	<sup>+</sup> Predicted C + G (%) content of nucleosomes
MMH2KB	53	56
MMBGL1	44	46
MMIGG1	51	54
MMIG10	51	54
MMRNXX	53	55.5

<sup>+</sup> Data determined from the computer printout of the C + G content of 148-base frames along the DNA sequence.

Key:

MMH2KB, Mouse pseudogene for the transplantation antigen complex H-2KB.

MMBGL1, Mouse gene for the  $\beta$ 1-globin.

MMIGG1, Mouse gene for the constant part of  $\gamma$ 1-immunoglobulin.

MMIG10, Mouse gene for the constant part of  $\gamma$ 2B-immunoglobulin heavy chain.

MMRNXX, Region surrounding the transcription initiation site for mouse 45S pre-rRNA, including parts of non-transcribed spacer and external transcribed spacers 5' to 18 S rRNA.

is higher than the C + G content of the whole sequence. This increase is small, though it is sufficient to account for the increase shown in Table 5 (page 58).

### III.2. In vitro methylation of chromatin

In Section III.1 it was shown that the nucleosomal core DNA is enriched in the bases C + G and also in  $m^5C$ . This  $m^5C$  enrichment may simply reflect the increase in the C + G content (and the CpG content) or it may be that the DNA methylase has a preference for core DNA. This section investigates the effects of chromatin structure on DNA methylation in vitro.

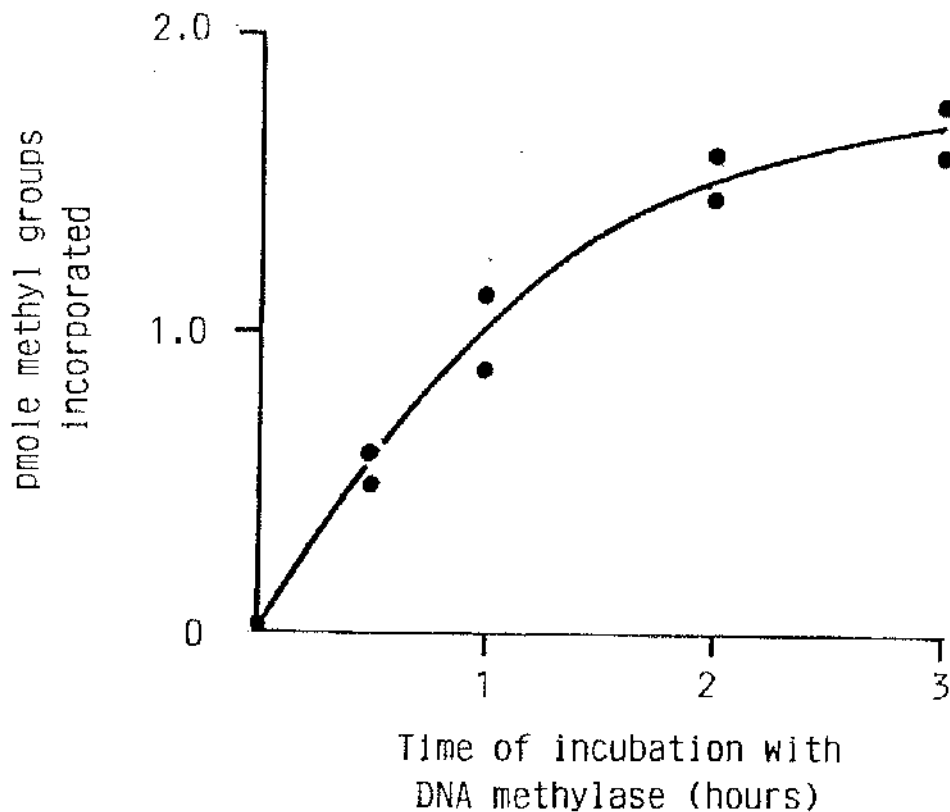
#### III.2.1. Optimization of conditions for methylation

In studying the effects of chromatin structure on DNA methylation in vitro it is important to use a system in which there are no competing effects occurring, i.e. there should be no endogenous methylating activity and no methyl groups present on the DNA at the beginning of the experiment. Such a system is provided by the insect Aedes albopictus (mosquito), the DNA of which has less than 0.02 mol %  $m^5C$  (Adams et al., 1979).

Also the conditions for methylation need to be optimized. Nuclei from mosquito cells were methylated as described in Section II.2.2.1 for various times at 37°C. The number of pmoles of methyl groups incorporated into the DNA was determined at each time (Fig. 11). The amount of methylation achieved increases up to 3 hours, but tails off rapidly between 2 and 3 hours. Thus 3 hours was considered a suitable time for methylation in vitro.

As it is the effects of chromatin structure on DNA methylation that are being studied, assay conditions need to be arranged so that the enzyme is limiting, therefore the substrates need to be in excess if possible. Figure 12 shows that under the conditions used the cofactor AdoMet is in excess at a concentration above 10  $\mu M$ . For the experiments involving methylation in isolated nuclei the lower concentration of 1.5  $\mu M$  (specific activity 15 Ci/mmoles) was used as low

Figure 11. Time course of methylation of mosquito nuclei using ascites DNA methylase.



Nuclei from mosquito cells were methylated in the presence of S-adenosyl-L-[methyl-<sup>3</sup>H]-methionine for various times at 37°C using a partially purified ascites DNA methylase (Section II.2.2.1). DNA was purified and the amount of pmole of methyl group incorporated was determined (Section II.2.4).

Figure 12. Dependence of methylation on the concentration of AdoMet.

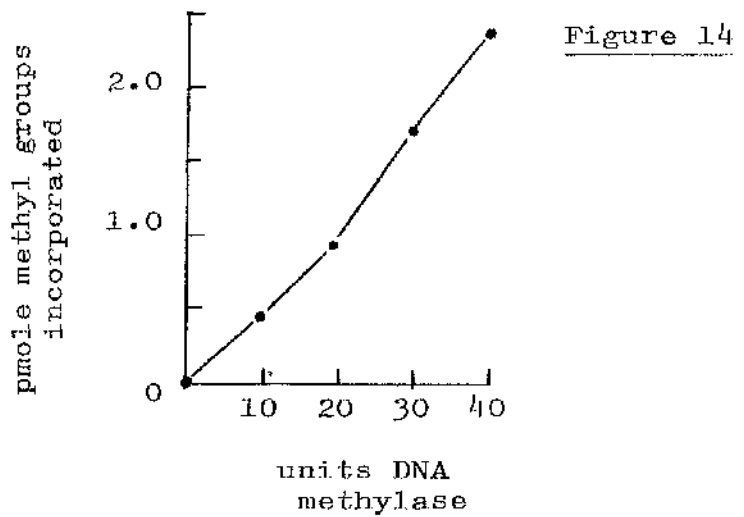
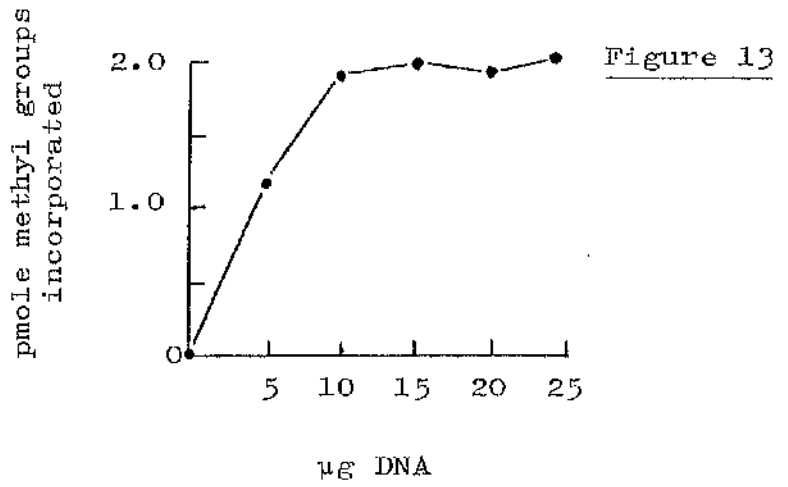
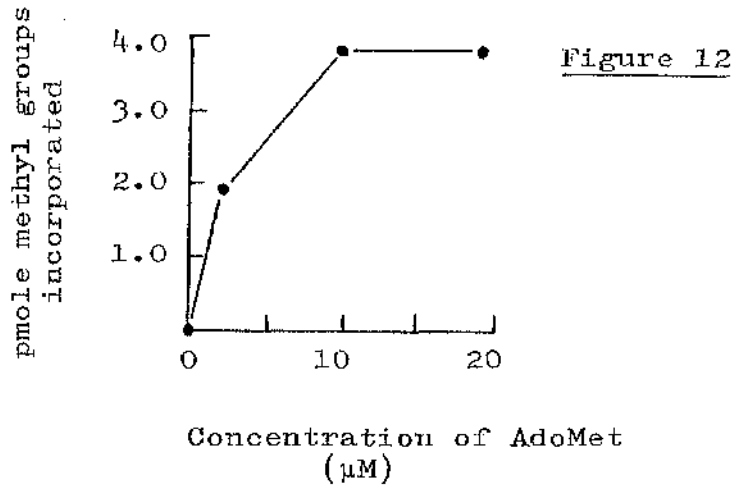
Nuclei from mosquito cells (10  $\mu$ g DNA content) were methylated using 40 units of ascites DNA methylase in the presence of increasing concentrations of [methyl- $^3$ H]-AdoMet for 3 hours at 37 $^{\circ}$ C. The DNA was purified and the amount of methyl groups (pmoles) incorporated was determined (Section II.2.4). The AdoMet was made up by adding increasing amounts of cold AdoMet to 5 $\mu$ Ci (0.33 nmole) of [methyl- $^3$ H]-AdoMet.

Figure 13. Dependence of methylation on the concentration of DNA.

Increasing amounts of nuclei from mosquito cells (0-25  $\mu$ g DNA content) were methylated using 40 units of ascites DNA methylase in the presence of [methyl- $^3$ H]-AdoMet (15 Ci/mmole; 1.5  $\mu$ M) for 3 hours at 37 $^{\circ}$ C. The incorporation of methyl groups was determined as before (Section II.2.4).

Figure 14. Dependence of methylation on the amount of DNA methylase.

Nuclei from mosquito cells (10  $\mu$ g DNA content) were methylated in the presence of [methyl- $^3$ H]-AdoMet (15 Ci/mmole; 1.5  $\mu$ M) for 3 hours at 37 $^{\circ}$ C. with increasing amounts of ascites DNA methylase. The incorporation of methyl groups was determined as before (Section II.2.4).



incorporation of counts was sometimes obtained, especially when using nuclei from log phase L929 cells and endogenous methylase. But for the experiments described in Section III.2.6 the higher concentration of 16  $\mu\text{M}$  (20i/mmole) was used.

Figure 13 shows that using 40 units of enzyme, a DNA concentration (in chromatin) of greater than 10  $\mu\text{g}/200 \mu\text{l}$  assay is saturating. Figure 14 shows that a good incorporation of methyl groups is obtained using 40 units of DNA methylase. more than 40 units of methylase was not used due to having a low supply of the enzyme.

### III.2.2. The sole product of methylation of DNA in vitro is $\text{m}^5\text{C}$

Mosquito nuclei were methylated as described in Section II.2.2.1 and the DNA purified (Section II.2.9). This DNA was pyrolysed to the individual bases which were separated on an Aminex A6 column (Section II.2.11.2). Fractions were collected and the amount of radioactivity in each fraction was determined (Fig. 15).

The figure shows that at least 95% of the radioactivity incorporated into DNA is in the  $\text{m}^5\text{C}$  peak.

### III.2.3. Methylation of mosquito nuclei

When nuclei from mosquito cells are methylated using a partially purified ascites DNA methylase (Section II.2.2.1) and subsequently digested using micrococcal nuclease (Section II.2.3.1), over 80% of the methyl groups are found to have been added to nuclease sensitive DNA, which is assumed to be linker DNA (Figs. 16 & 17). As there is little or no endogenous methylase activity, the methylation that is occurring is at previously unmethylated sites (i.e. it is de novo methylation). This methylation is affected by chromatin structure.

### III.2.4. Methylation of nuclei from stationary phase L929 cells

Another system which has low levels of endogenous DNA methylase activity is that of stationary phase L929 cells (Adams & Hogarth, 1973; Adams, 1974). The DNA of these cells

Figure 15. The sole product of in vitro methylation using ascites DNA methylase is m<sup>5</sup>C.

Nuclei from mosquito cells were methylated using ascites DNA methylase (Section II.2.2.1). The DNA was purified, pyrolysed using 98% formic acid and the bases separated on a column of Aminex A6 resin (Section II.2.11.2). The amount of radioactivity in each fraction was determined.

Figure 15.

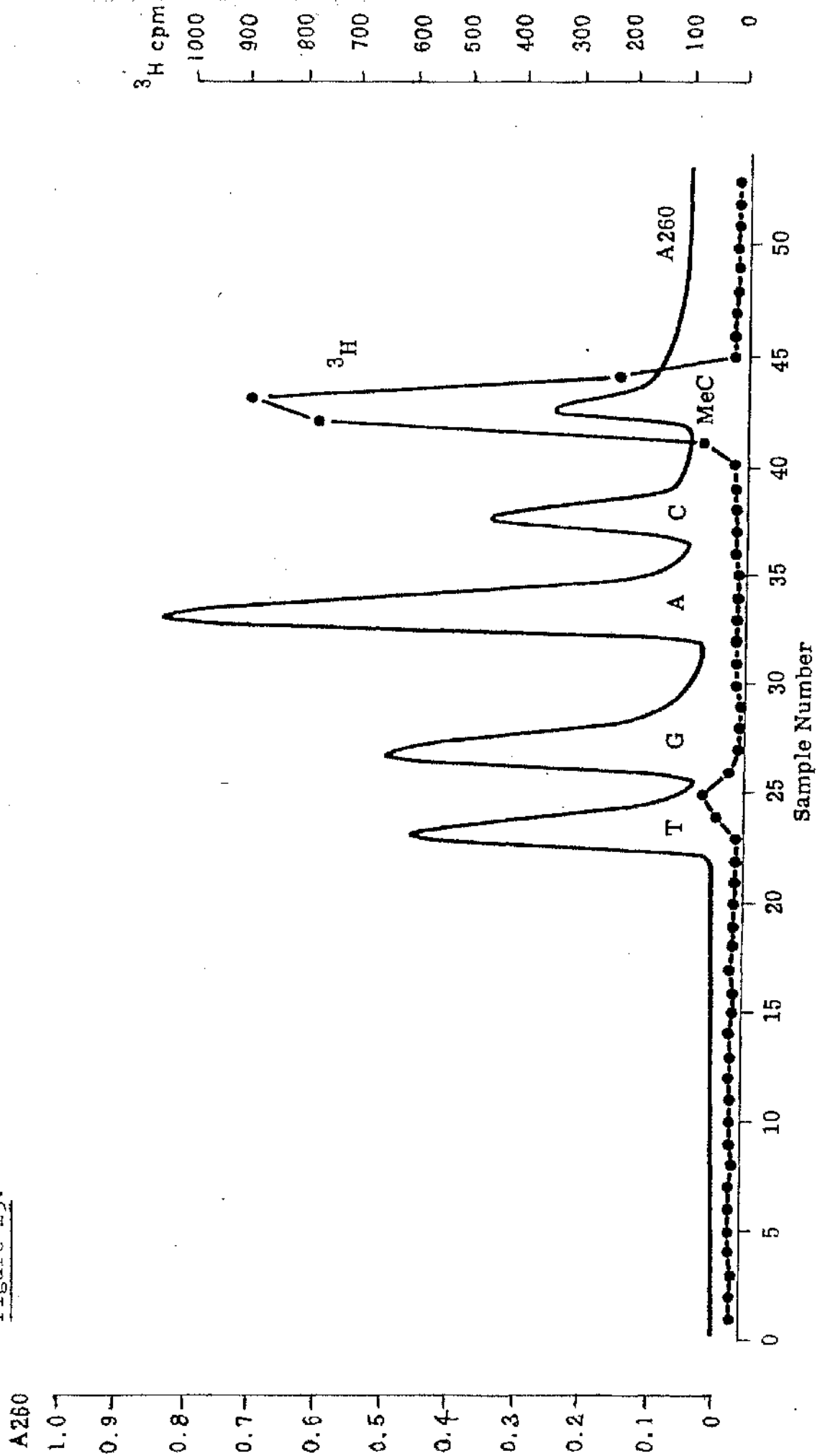




Figure 16. Digestion of in vitro methylated nuclei of mosquito cells using micrococcal nuclease.

- A:-  $^{14}\text{C}$
- B:-  $^3\text{H}$
- C:-  $^3\text{H}/^{14}\text{C}$

Nuclei from deoxy-  $[\text{U-}^{14}\text{C}]$ -cytidine prelabelled mosquito cells were methylated in the presence of S-adenosyl-L-  $[\text{methyl-}^3\text{H}]$ -methionine at  $37^\circ\text{C}$  for 3 hours using ascites DNA methylase (Section II.2.2.1). The nuclei were then digested using micrococcal nuclease (Section II.2.3.1), the DNA purified and the  $^3\text{H}/^{14}\text{C}$  ratio determined (Section II.2.4).

The amounts of  $^3\text{H}$  and  $^{14}\text{C}$  are expressed as a percentage of zero time.

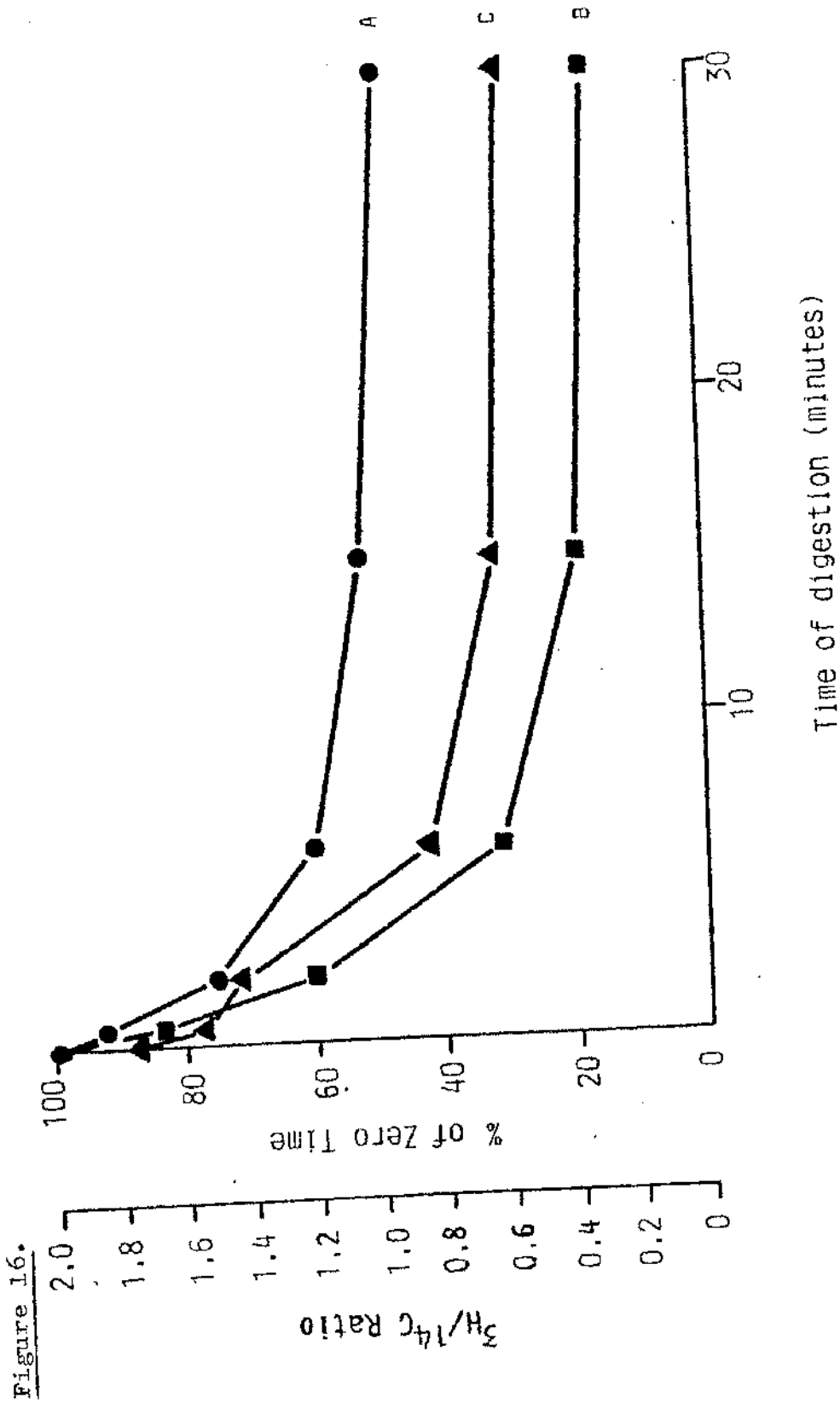
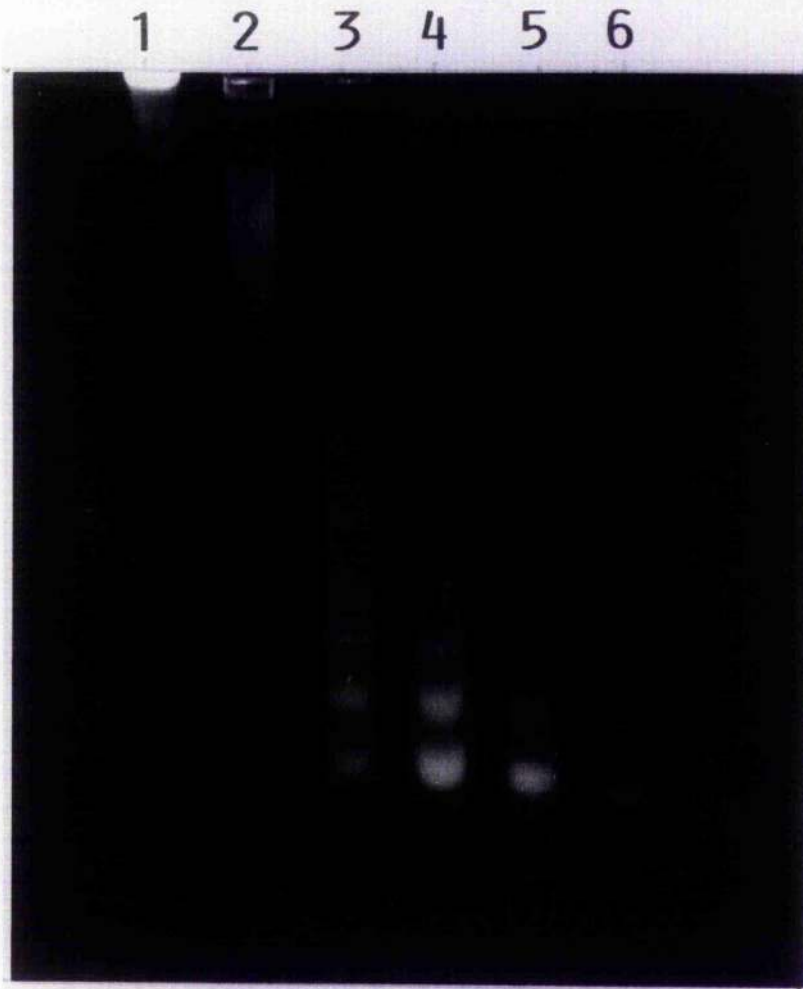


Figure 17. Agarose gel of DNA from mosquito chromatin digested using micrococcal nuclease.



Lanes 1-6: times of digestion; 0, 1/2, 2, 6, 15 and 30 mins.

Nuclei from mosquito cells were methylated and digested as described in Methods (II.2.2.1 and II.2.3.1). Samples were taken at the above times and electrophoresed on a 1.5% agarose gel. DNA was visualized using ethidium bromide.

has about 4% of the cytosines as  $m^5C$  (Table 19, page 114), and about 50% of the CpG dinucleotides are methylated (Navch-Mary & Cedar, 1981), so the DNA differs from the insect system in these respects.

Nuclei from stationary phase L929 cells were methylated as described in Section II.2.2.1, in the presence or absence of ascites DNA methylase, and digested using micrococcal nuclease (Section II.2.3.1).

Figure 18 shows that in the absence of added methylase there is very little endogenous activity and no detectable difference between methylation in core or linker DNA. When ascites methylase is used, there is a seven fold increase in the amount of methylation achieved. This methylation is again de novo methylation and occurs preferentially in linker DNA, though many methyl groups are added to core DNA.

### III.2.5. Methylation of nuclei from log phase L929 cells

Nuclei from log phase cells have an endogenous methylating activity and this activity will be in competition with the added ascites DNA methylase. When these nuclei are methylated in the absence of ascites methylase the methyl groups are added preferentially to core DNA as analyzed by micrococcal nuclease digestion (Fig. 19). This methylation is exclusively at hemimethylated sites (Adams & Hogarth, 1973) and this result reflects the situation seen in vivo (Section III.1.2).

When ascites methylase is added to these nuclei two effects will occur, 1) hemimethylation due to the endogenous activity and 2) de novo methylation due to the added activity. Figure 20 shows that under these conditions the methyl groups are added preferentially to linker DNA (the  $^3H/^{14}C$  ratio decreases). Thus the effects of the added methylase are dominating under the conditions used.

### III.2.6. Comparison of the methylation of chromatin and DNA

#### III.2.6.1. Core particles are poor acceptors of methyl groups

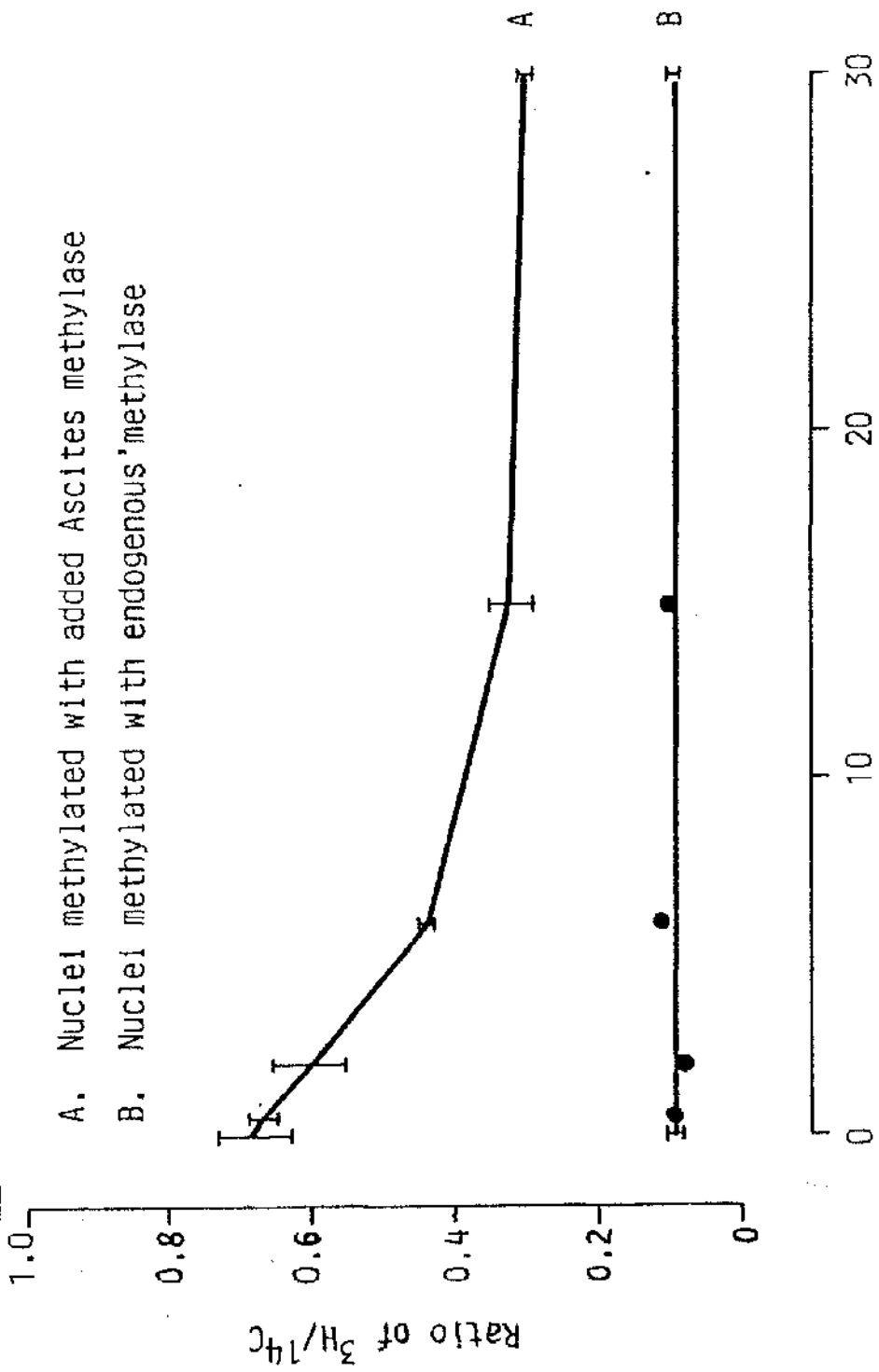
It was mentioned in Section III.2 that nucleosomal core DNA might be a preferred site of methylation for the endogenous methylase and that this might explain the enrichment of  $m^5C$  found in the core DNA in vivo (Section III.1.2).

Figure 18. Digestion of in vitro methylated nuclei from stationary phase L929 cells.

Nuclei from deoxy-  $[U-^{14}C]$ -cytidine prelabelled stationary phase L929 cells were methylated in the presence of S-adenosyl-L-  $[methyl-^3H]$ -methionine at  $37^{\circ}C$  for 3 hours in the presence or absence of ascites methylase (Section II.2.2.1). The nuclei were then digested using micrococcal nuclease (Section II.2.3.1); the DNA purified and the  $^3H/^{14}C$  ratio determined (Section II.2.4).

Only the ratio is shown on the figure.

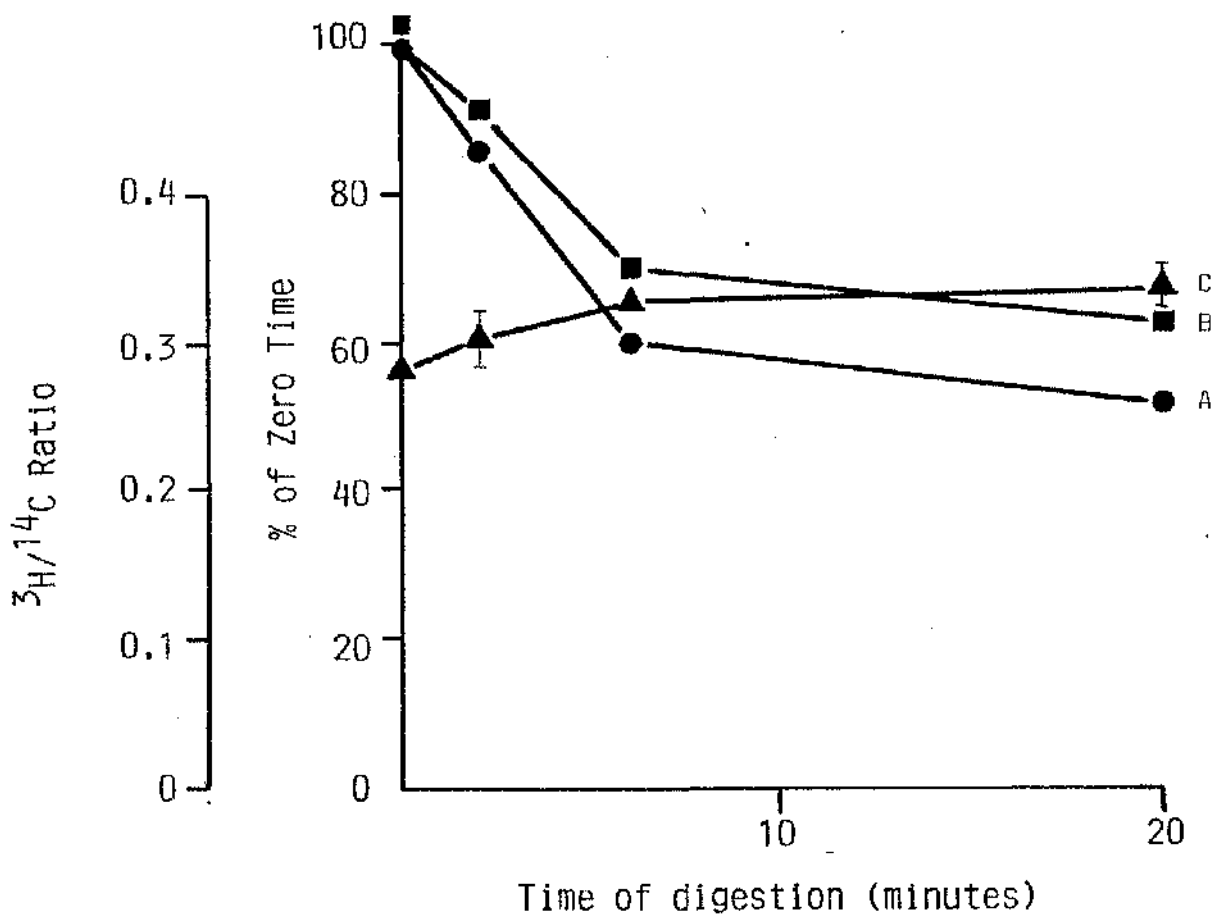
Figure 18.



A. Nuclei methylated with added Ascites methylase  
B. Nuclei methylated with endogenous methylase

Time in minutes of Digestion by Micrococcal Nuclease of  
in vitro Methylated L929 Chromatin

Figure 19. Digestion of nuclei from log phase L929 cells after methylation using the endogenous methylase.



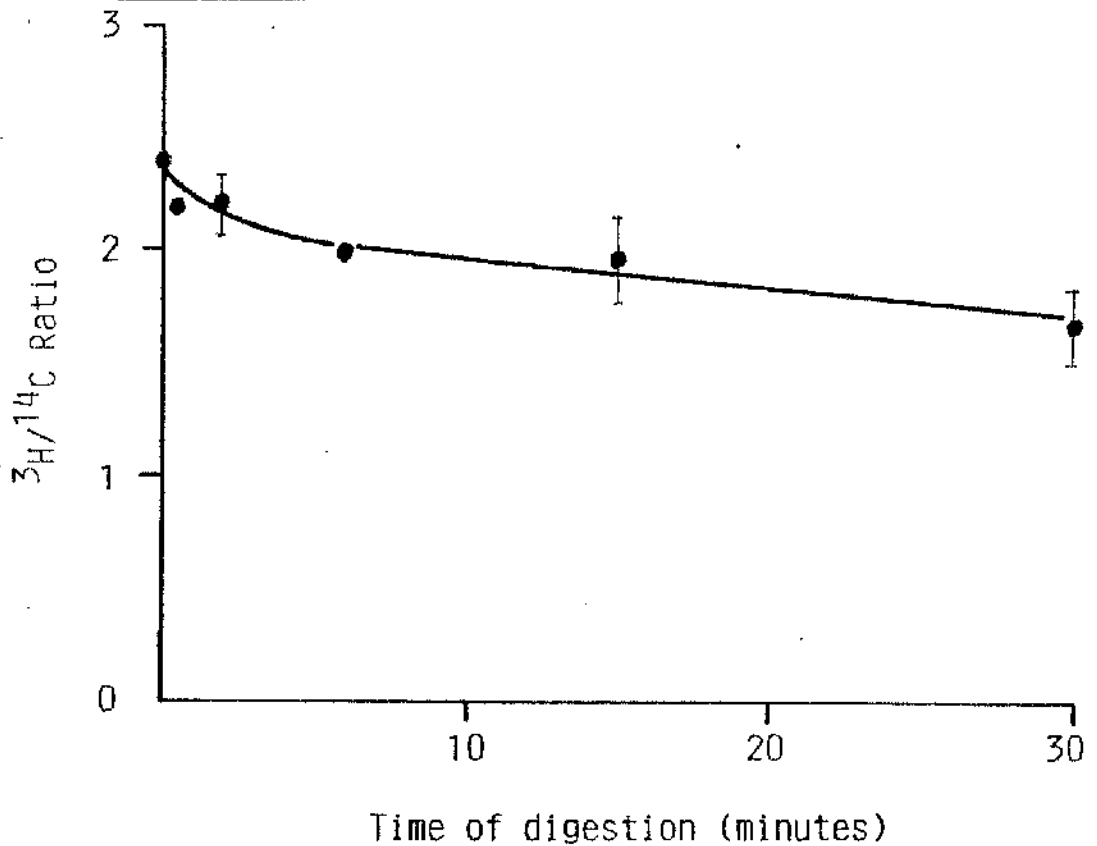
A:-  $^{14}\text{C}$   
B:-  $^3\text{H}$   
C:-  $^3\text{H}/^{14}\text{C}$

Nuclei from deoxy-  $[\text{U}-^{14}\text{C}]$ -cytidine prelabelled log phase L929 cells were methylated in the presence of S-adenosyl-L-  $[\text{methyl}-^3\text{H}]$ -methionine at  $37^\circ\text{C}$  for 3 hours using the endogenous methylase (Section II.2.2.1). The nuclei were then digested using micrococcal nuclease (Section II.2.3.1), the DNA purified and the  $^3\text{H}/^{14}\text{C}$  ratio determined (Section II.2.4).

The amounts of  $^3\text{H}$  and  $^{14}\text{C}$  are expressed as a percentage of zero time.

The range of values is given for the  $^3\text{H}/^{14}\text{C}$  ratio.

Figure 20. Digestion of in vitro methylated nuclei from log phase L929 cells (methylated using ascites methylase).



Nuclei were treated as described in Figure 19 except that ascites methylase was used (Section II.2.2.1). Only the  $^3\text{H}/^{14}\text{C}$  ratio is shown on the figure. (see Fig 19).



To investigate this, total chromatin and core chromatin were methylated using ascites DNA methylase (Section II.2.2.2). Core chromatin was prepared by treatment of nuclei with micrococcal nuclease until 50% of the DNA was rendered acid soluble (Section II.2.3.1). The nuclease action was stopped using EGTA to chelate the  $\text{Ca}^{2+}$  ions.

Figure 21 shows a plot of pmoles of methyl group incorporated against  $\mu\text{g}$  of DNA in the chromatin. Total chromatin is a better substrate for DNA methylase than core chromatin from both stationary and log phase L929 cells. Chromatin from log phase cells is a better substrate than chromatin from stationary phase cells presumably because DNA from log phase cells has some hemimethylated sites which are methylated more rapidly than unmethylated sites (Section I.2.2.1).

Figure 22 is a double reciprocal plot of the data in Figure 21. From this figure the  $V_{\text{max}}$  and  $k_m$  of the enzyme for each substrate has been calculated (Table 10).

Table 10. Kinetic parameters for ascites DNA methylase using chromatin from L929 cells as substrate.

Substrate	$V_{\text{max}}$ pmole/3 hour	$k_m$ $\mu\text{gDNA/ml}$
Total chromatin log phase cells	2.0	17.0
Core chromatin log phase cells	1.6	83.5
Total chromatin stat. phase cells	1.8	38.5
Core chromatin stat. phase cells	1.6	83.5

The calculated  $V_{\text{max}}$  for each substrate varies only slightly, and as the lines drawn in Figure 22 represent the best fit based upon linear regression of the data, these lines could very well all cross the Y axis at the same point. Therefore it is assumed that the  $V_{\text{max}}$  is the same for all

Figure 21. Comparison of the methyl group acceptor ability of various chromatin from L929 cells.

- (●—●) Total chromatin log phase cells
- (○—○) Core " " " "
- (■—■) Total chromatin stationary cells
- (□—□) Core " " "

Samples of nuclei and core preparations (Section II.2.3.1) were methylated for 3 hours (Section II.2.2.2) and the amount of pmoles of methyl groups incorporated into DNA was determined. DNA was assayed by the method of Burton (1956).

Figure 22. LinWeaver-Burke plot of Figure 21.

The same symbols have been used for each chromatin sample.

Figure 21.

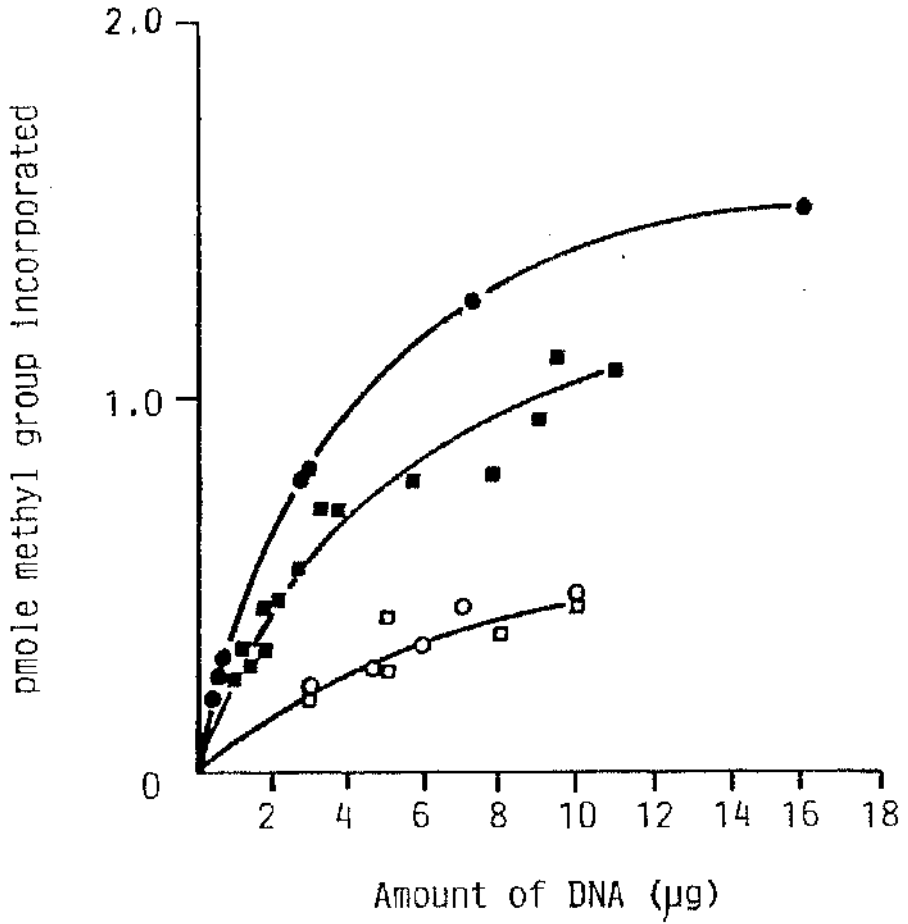
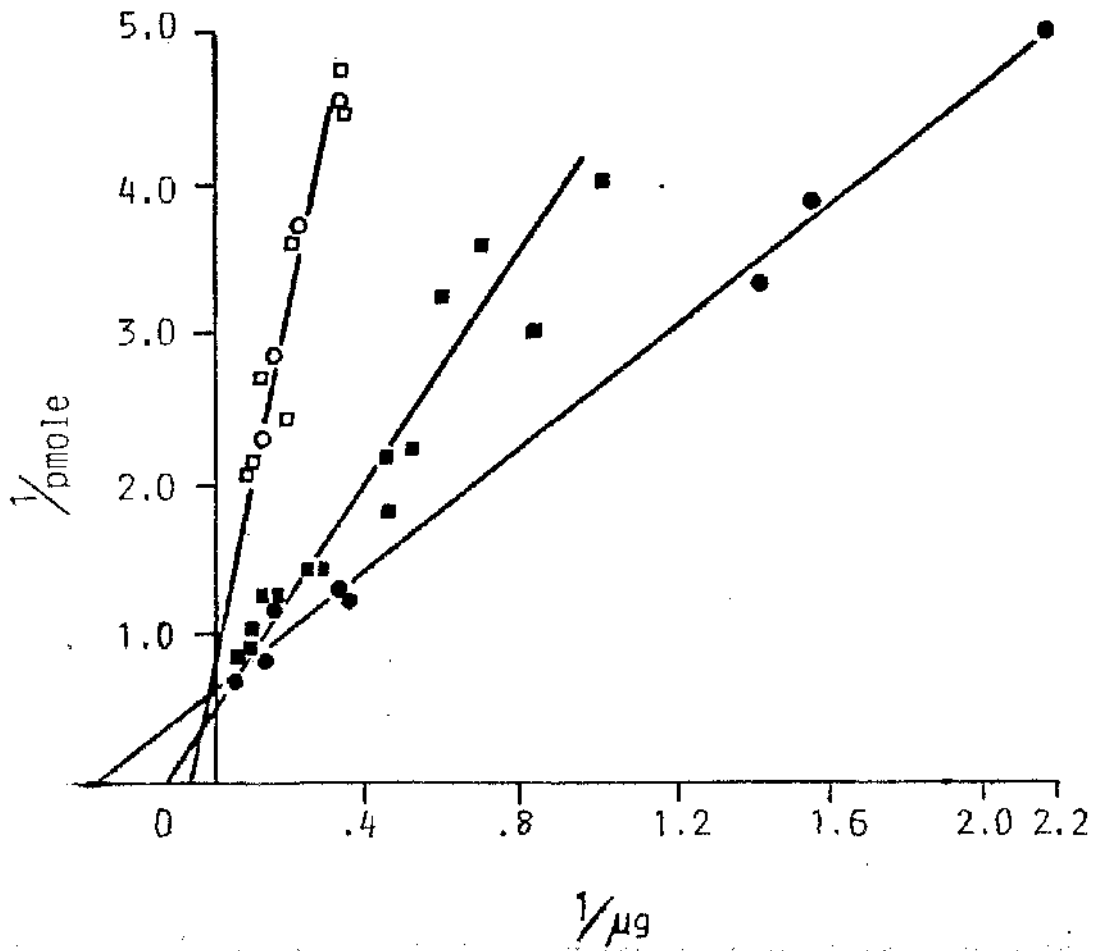


Figure 22.



the substrates. The ascites enzyme has a much lower affinity for core chromatin (higher  $k_m$ ) than for total chromatin.

The same result is observed using chromatin from mosquito cells (Fig. 23). In this case the double reciprocal plots cannot be drawn as the plots are linear in Figure 23. The DNA in chromatin from mosquito cells accepts three times as many methyl groups as the DNA in chromatin from L929 cells.

### III.2.6.2. DNA is a better substrate than chromatin

DNA was purified from total and core chromatin of mosquito and L929 cells (Section II.2.9), methylated using ascites DNA methylase (Section II.2.2.2) and the amount of  $m^5C$  incorporated was determined (Figs. 24 & 25).

The results show that DNA accepts ten times as many methyl groups as chromatin in all these cases. The methylation of DNA from core chromatin is only slightly different from the methylation of total DNA, thus showing that the difference in methylation between total and core chromatin is due to chromatin structure and not due to major differences in the DNA.

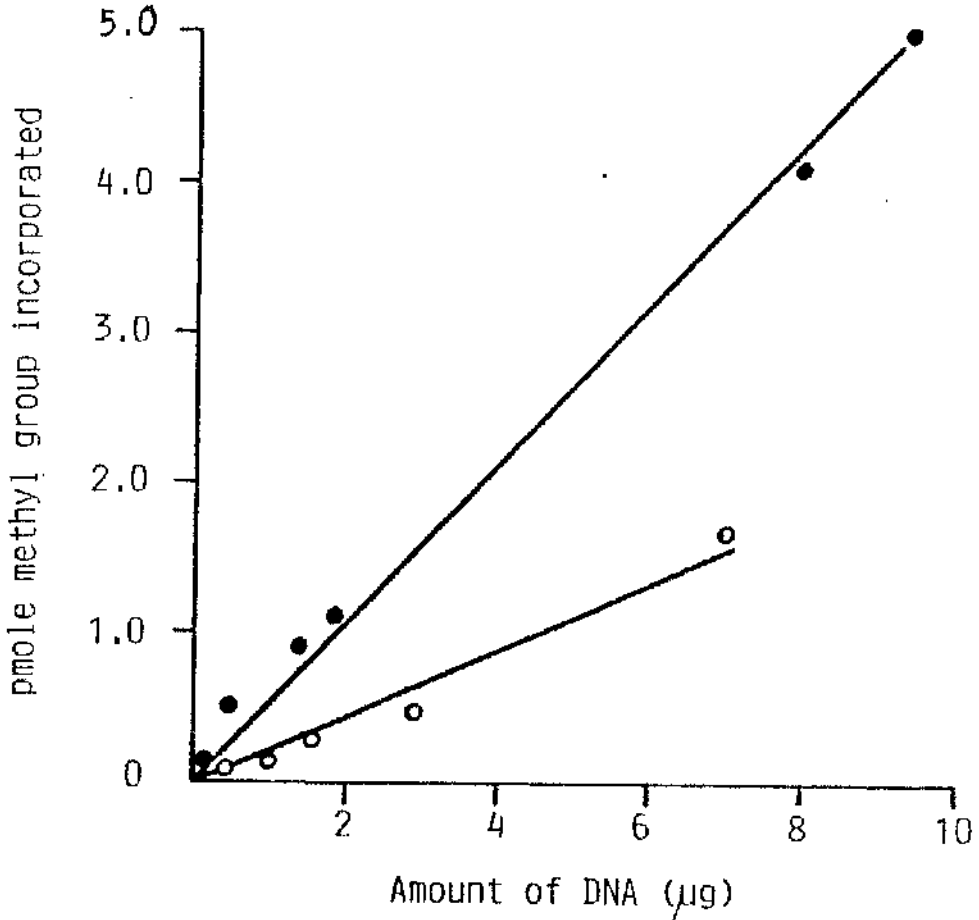
### III.2.7. Histones inhibit DNA methylation in vitro

Total calf thymus histones and Micrococcus luteus DNA were mixed at a protein/DNA ratio ranging from 0 to 2.0 in buffer  $M^+$ . The DNA was then methylated using 40 units of ascites DNA methylase in the presence of 5  $\mu Ci$  of S-adenosyl-L- [methyl- $^3H$ ] -methionine (1.5 Ci/mole) for 1 hour at 37°C. The amount of pmoles of methyl groups incorporated into the DNA was then determined (Fig. 26).

The figure shows that histones inhibit DNA methylation up to a protein/DNA ratio of about 1.0, thereafter adding more histone has little effect.

The histone/DNA ratio of native chromatin is also about 1.0 and the amount of inhibition observed in this experiment is similar to the difference in the methyl group acceptance ability of DNA and chromatin (Section III.2.6).

Figure 23. Methylation of DNA in core and total chromatin from mosquito cells.

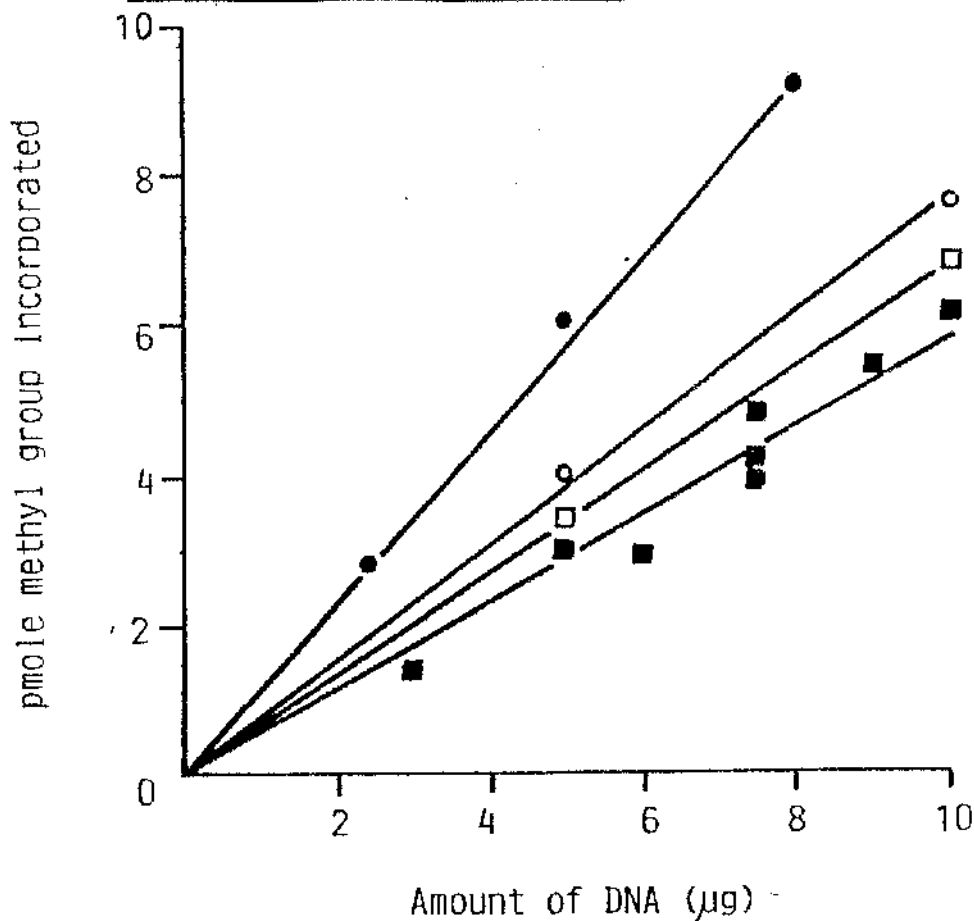


Total and core chromatin from mosquito cells were methylated as described in Section II.2.2.2 and the number of pmoles of methyl group incorporated into DNA was determined. DNA was assayed by the method of Burton (1956).

(●—●): total chromatin

(○—○): core chromatin

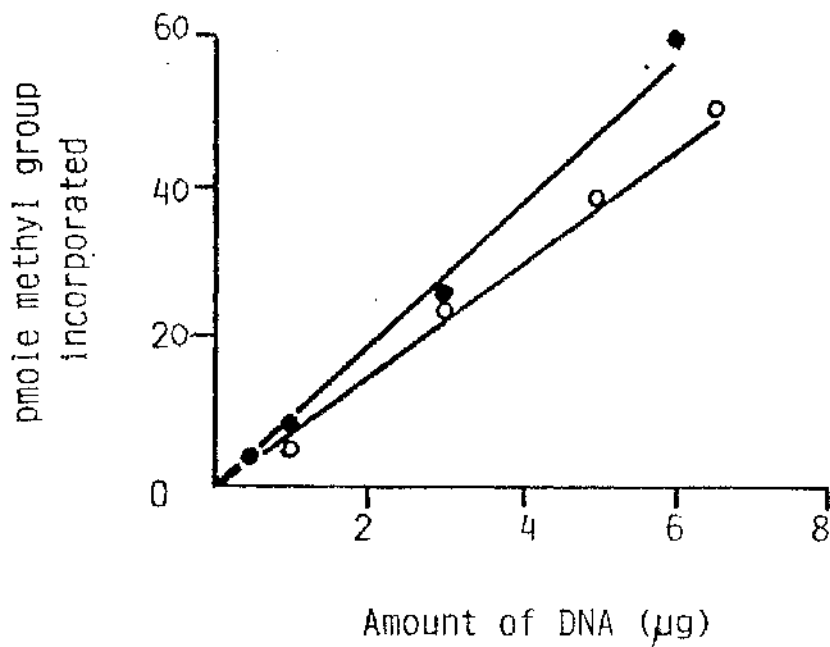
Figure 24. Methylation of DNA purified from total and core chromatins from L929 cells.



DNA from total and core chromatins of L929 cells was methylated as described in Section II.2.2.2 and the amount of pmoles of methyl group incorporated was determined. DNA was assayed by the method of Burton (1956).

The same symbols are used as in Figure 21 (page 88).

Figure 25. Methylation of DNA purified from total and core chromatin of mosquito cells.

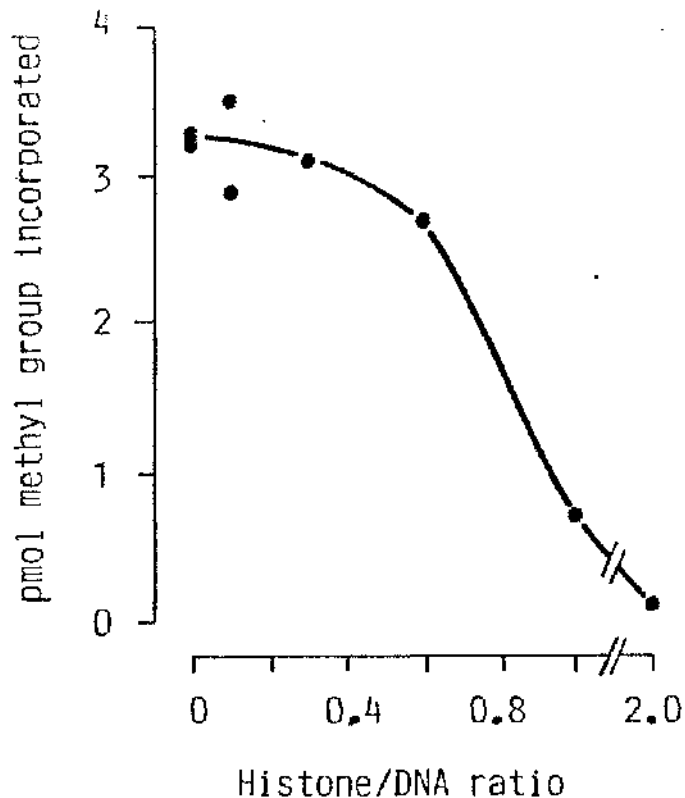


DNA was methylated as described in Figure 24 (page 91).

(●—●): total DNA

(○—○): core DNA

Figure 26. Inhibition of DNA methylation by histones.



Micrococcus luteus DNA (2  $\mu\text{g}$ ) was mixed with total calf thymus histones at a protein/DNA ratio ranging from 0 to 2.0 in buffer  $\text{M}^+$ . The DNA was then methylated using 40 units of ascites DNA methylase in the presence of 5  $\mu\text{Ci}$  S-adenosyl-L- [methyl- $^3\text{H}$ ] -methionine (1.5 Ci/mmol) for 1 hour at  $37^\circ\text{C}$  in a total volume of 70  $\mu\text{l}$ . The DNA was then purified and the amount of pmoles of methyl groups incorporated was determined (Section II.2.4).



### III.2.8. Nuclear non-histone proteins do not inhibit DNA methylation in vitro

Nuclear non-histone proteins were prepared from ascites nuclei by extraction using 0.35M NaCl (Goodwin et al. 1975) followed by dialysis to low salt. The high mobility group proteins were obtained by precipitation of the bulk of the proteins in the 0.35 M NaCl extract by addition of TCA (50% w/v) to a final concentration of 2.5% (w/v); those proteins remaining soluble are the HMG proteins (Goodwin et al. 1975). The 2.5% TCA supernatant was neutralized and dialysed to remove the TCA and salt.

Micrococcus luteus DNA was methylated as described in Section III.2.7 except that the non-histone and HMG proteins were used instead of the histones (Figure 27) (protein was determined by the method of Bradford, 1976).

As can be seen no inhibition of methylation by any of these proteins is observed. Thus the difference in methylation between DNA and chromatin is due mainly to the presence of histones.

### III.3. In vivo methylation of DNA in transcriptionally active chromatin of L929 cells

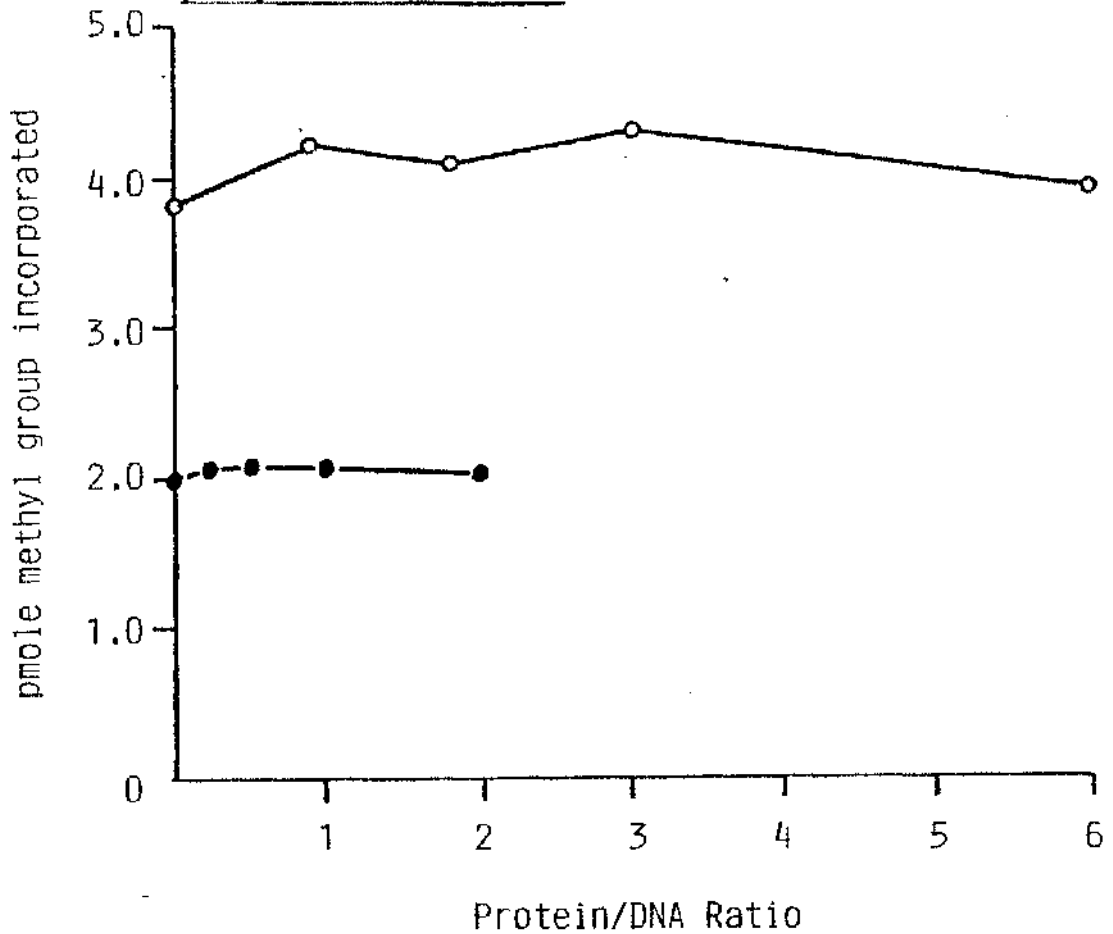
In Section I.1.8 it was discussed that the DNA in transcriptionally active chromatin is undermethylated with respect to bulk DNA (Davie & Saunders, 1981; Naveh-Manly & Cedar, 1981 ).

This section discusses some of the methods which have been used to prepare chromatin from transcriptionally active chromatin and presents data on the methylation of DNA in this chromatin from mouse L929 cells.

#### III.3.1. Use of DNase I to study transcribing chromatin

One of the problems with studying transcriptionally active chromatin in a general sense (i.e. not specific genes) is isolating the transcribing regions from bulk chromatin. Many methods have been used and particular stress has been put on nuclease studies. DNase I has a preference for cutting at transcribing regions initially (Elgin, 1981) so

Figure 27. Effect of nuclear non-histone proteins on DNA methylation in vitro.



Micrococcus luteus DNA (2  $\mu$ g) was mixed with either nuclear non-histone proteins or HMG proteins at a protein/DNA ratio ranging from 0-6.0. The DNA was then methylated as described in the legend to Figure 26.

(○ ——— ○): non-histone proteins

(● ——— ●): HMG proteins

can be used to study these regions. DNase II treatment of chromatin produces nucleosomes, which can be separated into transcribing and non-transcribing regions as the former are soluble in the presence of  $MgCl_2$  at 2 mM (Gottesfeld & Butler, 1977).

These nuclease methods have been used without a great deal of success. When nuclei from deoxy- [ $U-^{14}C$ ] -cytidine prelabelled mosquito cells were methylated (Section II.2.2.1) followed by digestion using DNase I (Section II.2.3.2), the methylated DNA is digested very rapidly (Fig. 28). Whether this methylated DNA is derived from transcriptionally active chromatin cannot be determined from this experiment due to the very rapid digestion. When the rate of digestion is decreased, no definite change in the  $^3H/^{14}C$  ratio can be seen over the first few percent of DNA that is made acid soluble.

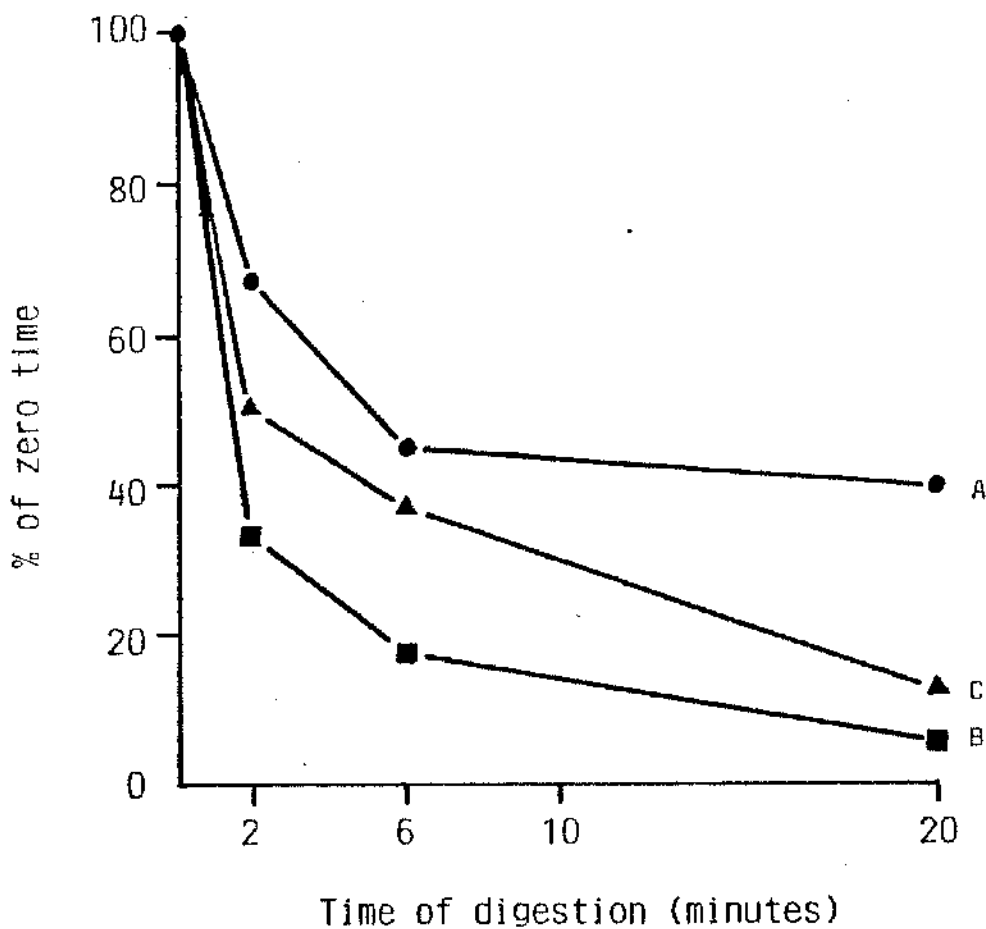
There are three problems with this type of experiment: 1) DNase I only shows a preference for transcribing chromatin, 2) the accuracy is not sufficient to notice a difference in the  $^3H/^{14}C$  ratio when only a few percent of the DNA is made acid soluble and 3) the transcribing regions are lost.

Similar problems have been found using DNase II digestion.

### III.3.2. Sanders' extraction procedure for isolation of transcribing chromatin

When nuclei are briefly digested using micrococcal nuclease this enzyme preferentially cleaves DNA in transcriptionally active regions of chromatin (Gottesfeld & Butler, 1977; Bloom & Anderson, 1978). The liberated nucleosomes from these regions can then be extracted using low concentrations of NaCl (50 mM) and  $Mg^{2+}$  ions (2 mM) (Sanders, 1979). Using a modified procedure Davie & Saunders (1981) isolated transcribing nucleosomes from calf thymus nuclei. These nucleosomes were found to be depleted in histone H1 and  $m^5C$  and enriched in high mobility proteins 14 and 17. This procedure has subsequently been used in this work.

Figure 28. Digestion of in vitro methylated nuclei from mosquito cells using DNase I.



- A: amount of  $^{14}\text{C}$
- B: amount of  $^3\text{H}$
- C: the  $^3\text{H}/^{14}\text{C}$  ratio.

Nuclei from mosquito cells were methylated using added ascites DNA methylase (Section II.2.2.1) and subsequently digested using DNase I (Section II.2.3.2) for several lengths of time. DNA was purified at each time and the amounts of  $^3\text{H}$ ,  $^{14}\text{C}$  and the  $^3\text{H}/^{14}\text{C}$  ratio were determined.

The figures are expressed as a percentage of zero time.

### III.3.3. In vitro transcription of nuclei from L929 cells

Before making extensive use of the Sanders' procedure I felt it necessary to confirm that the nucleosomes isolated by this procedure did actually contain actively transcribed DNA.

Nuclei from log phase L929 cells were incubated in the presence of [ $5\text{-}^3\text{H}$ ]-UTP in order that DNA transcription should continue (Hay *et al.*, 1982) (see legend to Table 11). These nuclei were then extracted as described in Section II.2.6 to isolate the presumptive transcribing regions. The amount of radioactivity in each fraction was then determined. As RNA synthesis in the presence of only one ribonucleotide will not be completed, the nascent RNA should still be associated with its DNA template and should be extracted in the same fraction in the Sanders' procedure. As can be seen in Table 11 the DNA in fractions S0 and S50 has a much higher  $^3\text{H}/^{14}\text{C}$  ratio than the DNA in the other fractions. These fractions are also enriched in the HMG proteins and are deficient in histone H1 (Fig. 29). These fractions were thus taken to be the transcriptionally active chromatin (see also Fig. 30). An enrichment of  $^3\text{H}$  is also seen in fractions S0 and S50 when three or four ribonucleotides are used, though this enrichment is smaller (Table 11. B and C).

### III.3.4. Active regions of chromatin are deficient in $m^5\text{C}$

Mouse L929 cells were labelled *in vivo* with 50  $\mu\text{Ci}$  of [ $6\text{-}^3\text{H}$ ]-uridine for two days and harvested during log phase. Nuclei from these cells were prepared and treated using the Sanders' procedure (Section II.2.6). DNA was purified from each of the fractions and the amount of  $m^5\text{C}$  in each fraction was determined (Section II.2.11.2). The result is given in Table 12.

Table 11. In vitro transcription in nuclei from log phase L929 cells.

	Sample	*% DNA in fraction	$^3\text{H}/^{14}\text{C}$
A	S 0	1.0	0.68
	S 50	7.0	0.67
	S 100	-	-
	S 200	1.0	0.06
	S 400	0.3	0.12
	pellet	91.0	0.03
B	S 0	1.0	0.70
	S 50	6.5	0.68
	S 100	-	-
	S 200	1.0	0.13
	S 400	-	-
	pellet	91.5	0.13
C	S 0	1.0	0.63
	S 50	6.0	0.69
	S 100	-	-
	S 200	1.4	0.22
	S 400	0.3	0.44
	pellet	91.5	0.48

Nuclei from  $[\text{U-}^{14}\text{C}]$ -deoxycytidine prelabelled L929 cells were resuspended in TC buffer at a concentration of about  $5 \times 10^7$  per ml and divided into three fractions.

To each fraction <sup>A-C</sup> was added 50  $\mu\text{Ci}$  of  $[\text{U-}^3\text{H}]$ -uridine-5'-triphosphate.

To fraction B the ribonucleotides ATP and CTP were added to final concentrations of 0.4 mM.

To fraction C the ribonucleotides ATP, CTP and GTP were added to final concentrations of 0.4 mM.

Each fraction was incubated at  $27^\circ\text{C}$  for 5 minutes. The reaction was stopped by cooling to  $0^\circ\text{C}$  on ice and the nuclei were treated according to the Sanders' procedure (Section II.2.6).

\*cpm expressed as % total cpm recovered.

Figure 29. Acrylamide gels of the proteins present in the fractions of a Sanders' extraction.

Figure 29a. 7% acrylamide gel.

Lane	
1	Markers
2	Pellet fraction
3	S 400 "
4	S 200 "
5	S 100 "
6	S 50 "
7	S 0 "
8	2.5% TCA sup of pellet fraction
9	2.5% TCA sup of S 400 "
10	2.5% TCA sup of S 200 "
11	2.5% TCA sup of S 100 "
12	2.5% TCA sup of S 50 "
13	2.5% TCA sup of S 0 "

Figure 29b. 15% acrylamide gel.

Lane	
1	pellet fraction
2	S 400 "
3	S 200 "
4	S 100 "
5	S 50 "
6	S 0 "

Nuclei from L929 cells were treated as described in Section II.2.6. Half of each sample was made 2.5% TCA (w/v) and centrifuged at 850 x g for 10 mins. The supernatants were neutralized and dialysed. Samples of these supernatants and the untreated fractions were electrophoresed on a 7% acrylamide gel. The untreated fractions were electrophoresed on a 15% acrylamide gel also. Gels were stained using the silver stain technique.

Figure 29.

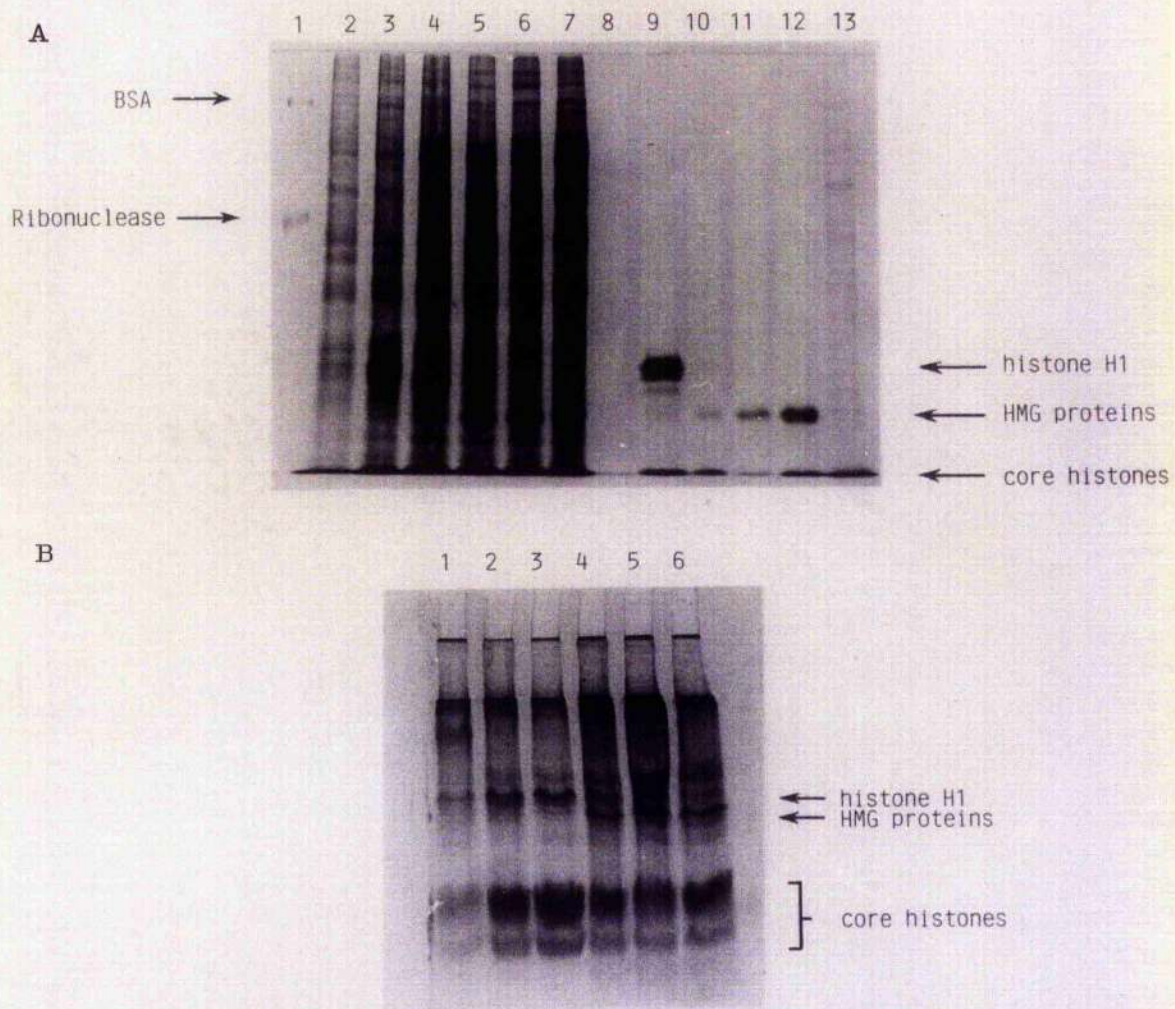
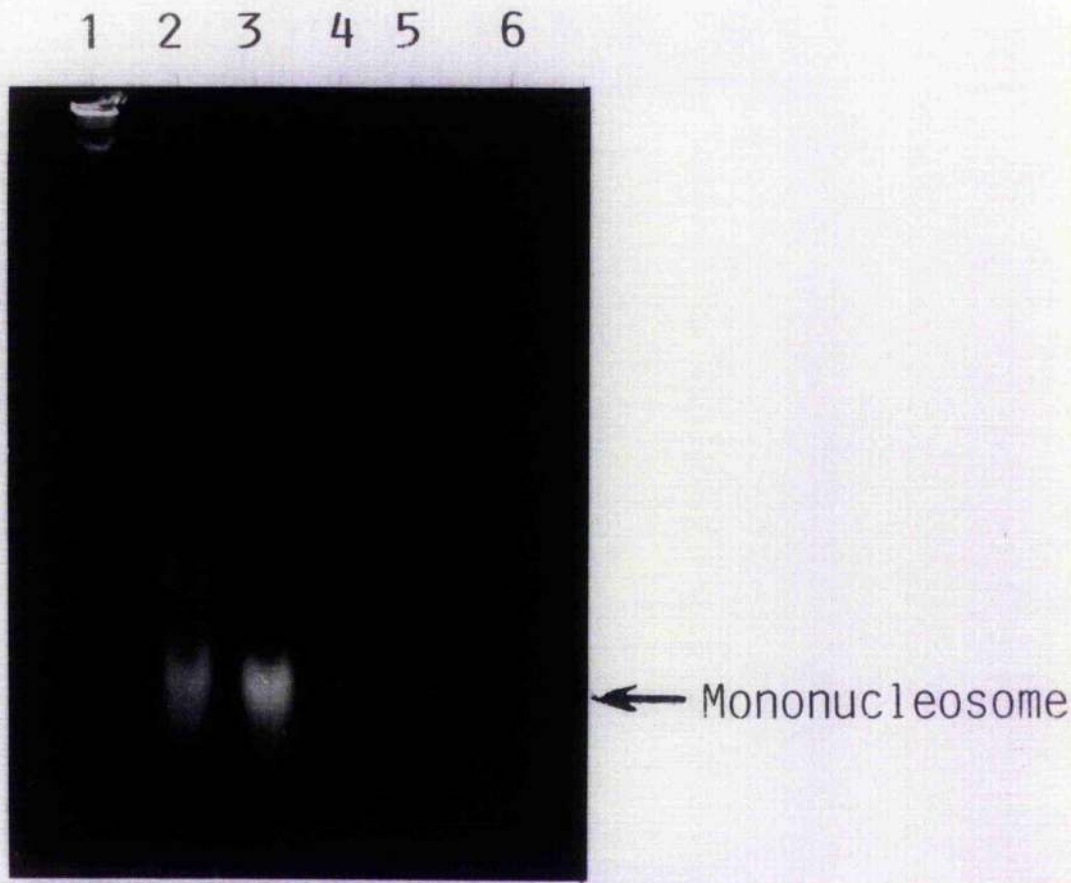




Figure 30. Agarose gel of the DNA present in the fractions of a Sanders' extraction.



Lanes: 1, pellet; 2, S 400; 3, S 200;  
4, S 100; 5, S 50; 6, S 0.

Nuclei from L929 cells were treated as described in Section II.2.6. Samples of each supernatant and the pellet fraction were electrophoresed on a 1.5% agarose gel (Section II.2.12). DNA was visualized using ethidium bromide.

Table 12. Transcribing chromatin is deficient in  $m^5C$ .

	Fraction	<sup>a</sup> % DNA in fraction	* $m^5C$ (%)
<sup>+A</sup> (undigested)	S 0	1.6	0.69
	S 50		
	S 100	5.5	0.63
	S 200		
	S 400		
pellet	93.0	2.95	
<sup>+B</sup> (digested)	S 0	7.0	0.98
	S 50		
	S 100	27.0	2.07
	S 200		
	S 400		
pellet	66.0	3.15	

\*See Section III.1.3.

<sup>+A</sup> is not digested with micrococcal nuclease whereas B is digested (2 minutes).

Fractions refer to the supernatants at different salt concentrations (Section II.2.6).

<sup>a</sup> See table 11, page 99.

As can be seen from Table 12, the DNA in fractions S 0 and S 50 in Section B is deficient in  $m^5C$  compared to the DNA in the other fractions. This result agrees with the result using calf thymus (Davie & Saunders, 1981).

In part A of the table some DNA is present in the low salt extracts even though no nuclease action should have occurred. This is probably due to some endogenous nucleolytic activity. This DNA is also deficient in  $m^5C$  and demonstrates that the procedure is independent of the type of endonuclease used.

It is interesting that the DNA in the fractions S 100 - S 400 is also deficient in  $m^5C$  compared to the DNA in the pellet, and perhaps this DNA contains another class of sequences with an intermediate level of methylation. It is also possible that this DNA is a mixture of transcribing and non-transcribing DNA.

III.3.5. Location of DNA less than 1 hour old

As DNA less than 1 hour after DNA synthesis is under-methylated relative to fully mature DNA (Section I.2.2.1) (see also Table 19, page 114) it is important to know where this DNA is extracted in the Sanders' procedure (nascent DNA is unimportant as this DNA makes up less than 0.5% of the total DNA).

Thus deoxy  $[U-^{14}C]$ -cytidine prelabelled L929 cells in log phase were pulse labelled for 50 minutes with  $[6-^3H]$ -thymidine. Nuclei were prepared and were extracted using the Sanders' procedure (Section II.2.6). The DNA in each fraction was purified and the  $^3H/^{14}C$  ratio was determined (Table 13).

Table 13. Location of DNA less than 1 hour old in a Sanders' extraction.

	Fraction	% DNA in fraction	$^3H/^{14}C$ ratio
+A (undigested)	(a) Total	-	12.50
	S 0	0.20	11.35
	S 50	2.00	10.10
	S 100	1.65	10.50
	S 200	2.55	9.50
	S 400	4.80	9.75
	pellet	88.50	10.95
+B (digested)	(a) Total	-	11.63
	S 0	2.15	9.40
	S 50	5.10	10.05
	S 100	2.25	10.50
	S 200	4.65	10.05
	S 400	9.15	10.25
	pellet	76.50	11.85

(a) Total is a fraction of the nuclei immediately after treatment with micrococcal nuclease.

+ As in Table 12 (page 103).

The result shows that the DNA which is less than 1 hour old is not extracted specifically into any of the fractions. As these cells were all at different stages of the cell cycle (they are asynchronous) the DNA which is synthesized at any given time will represent a random sample of the total DNA, and thus the different amounts of  $m^5C$  in different classes of DNA (Section I.1.8) will not affect the result.

The conclusion is that the DNA extracted into the S 0 and S 50 fractions is from transcribing chromatin and this DNA is deficient in  $m^5C$ .

III.4. Transcribing chromatin is preferentially methylated in vitro

In Section III.3 it has been shown that DNA in transcriptionally active chromatin is undermethylated when compared to the DNA from non-transcribing chromatin. It is possible that this may be due to aspects of the structure of actively transcribed chromatin and this section investigates the methylation of transcribing chromatin in vitro.

III.4.1. Log phase L929 cells

Nuclei from deoxy-  $[U-^{14}C]$  -cytidine prelabelled L929 cells were methylated using the endogenous DNA methylase (Section II.2.2.1) and then extracted using the Sanders' extraction procedure (Section II.2.6). The DNA was purified and the  $^3H/^{14}C$  ratio was determined (Table 14).

Table 14. Methylation of nuclei from log phase L929 cells using endogenous DNA methylase followed by Sanders' extraction.

	Fraction	% DNA in fraction	$^3H/^{14}C$ ratio
*A	(undigested) pellet	100	0.075
B	(digested) S 0	10.7	0.10
	S 50		
	S 100		
	S 200		
	S 400		
	pellet	84.0	0.07

\*Only the pellet fraction included as there was no DNA in the other fractions.

The table shows that the DNA in fractions S 0 and S 50 accepts more methyl groups per unit amount of DNA than the DNA in the other fractions, though this DNA does accept methyl groups.

This result seems to be in apparent contradiction to the in vivo situation (Section III.3.4) and it is possible that some perturbation has been introduced during the preparation of the nuclei.

To try to avoid problems due to perturbations in nuclear preparation a cell lysate was prepared. Mouse L929 cells were suspended in buffer M without glycerol and allowed to swell for 10 minutes. The cells were then gently lysed by pipetting up and down with a Gilson-G200 pipette (Section II.2.1.6).

The lysate was then methylated and extracted as for nuclei and the  $^3\text{H}/^{14}\text{C}$  ratio for each fraction was determined (Table 15).

Table 15. Methylation of a cell lysate of log phase L929 cells using endogenous methylase followed by Sanders' extraction.

Fraction	% DNA in fraction	$^3\text{H}/^{14}\text{C}$ ratio
*A (undigested) pellet	100	0.06
B (digested) S 0 )	6.0	0.12
S 50 )		
S 100 )		
S 200 )		
S 400 )	12.4	0.05
pellet	81.6	0.05

\* As in Table 14 (page 105).

The result obtained is the same as that for nuclei (Table 14). This suggests that the DNA extracted into the S 0 and S 50 fractions is more susceptible to methylation by the endogenous DNA methylase than the DNA in the other fractions.

III.4.2. Stationary phase L929 cells

Nuclei and a cell lysate from stationary phase L929 cells were methylated in the presence or absence of ascites DNA methylase and then extracted using the Sanders' extraction procedure. When no added methylase was used no detectable methylation was achieved. The methylation achieved using added ascites methylase is given in Tables 16 and 17 (page 107).

The result shows that there is no preferential methylation in any of the fractions using the added methylase. This result suggests that in non-dividing cells the transcriptionally active chromatin is not more accessible to added enzymes than is bulk chromatin. (This may not be true for endogenous enzymes).

Table 16. Methylation of nuclei from stationary phase L929 cells using added ascites methylase followed by Sanders' extraction.

	Fraction	% DNA in fraction	$^3\text{H}/^{14}\text{C}$ ratio
*A (undigested)	pellet	100	0.28
B (digested)	S 0 )	12.4	0.23
	S 50 )		
	S 100 )	22.6	0.10
	S 200 )		
	S 400 )		
	pellet	64.7	0.20

Table 17. Methylation of a cell lysate from stationary phase L929 cells using added ascites methylase followed by Sanders' extraction.

	Fraction	% DNA in fraction	$^3\text{H}/^{14}\text{C}$ ratio
*A (undigested)	pellet	100	0.03
B (digested)	S 0 )	12.0	0.03
	S 50 )		
	S 100 )	24.5	0.015
	S 200 )		
	S 400 )		
	pellet	63.5	0.03

\*As in Table 14 (page 105).

III.4.3. Mosquito cells

Nuclei from mosquito cells were methylated using the ascites DNA methylase (Section II.2.2.1) and then extracted using the Sanders' extraction procedure (Section II.2.6). The  $^3\text{H}/^{14}\text{C}$  ratio of the DNA in each fraction was determined (Table 18). In this experiment the extraction with buffer B + 400 mM NaCl has been omitted as at this salt concentration the nuclei formed a gel which could not be resuspended and so no samples could be taken for the agarose gel (Fig. 31).

Only 2% of the total DNA has been extracted into the S 0 and S 50 fractions but this DNA has four times as many methyl groups per unit length as that DNA which remains in the pellet.

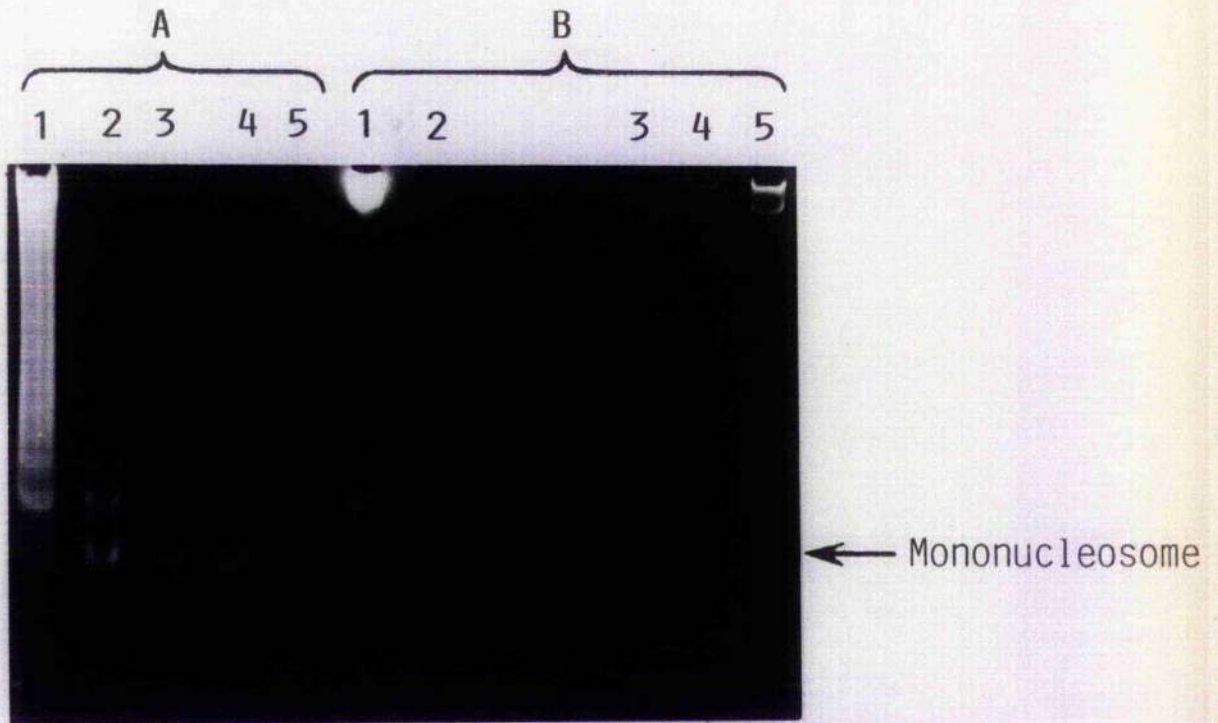
This shows that the transcribing chromatin is more susceptible to a 'soluble' exogenous DNA methylase presumably due to its more open structure. This was not seen with the nuclei from stationary phase L929 cells (Section III.4.2).

Table 18. Methylation of nuclei from mosquito cells followed by Sanders' extraction.

	Fraction	% DNA in fraction	$^3\text{H}/^{14}\text{C}$ ratio
*A (undigested)	pellet	100	0.33
B (digested)	S 0	2.0	1.60
	S 50		
	S100	8.0	0.66
	S200		
	pellet		

\*See Table 14 (page 105).

Figure 31. Agarose gel of the DNA present in the fractions  
of a Sanders' extraction using mosquito nuclei.



Lanes: A, digested nuclei; B, undigested nuclei

1, pellet; 2, S 200; 3, S 100; 4, S 50; 5, S 0.

Nuclei from mosquito cells were treated as described in Section II.2.6. Samples of each supernatant and the pellet fraction were electrophoresed on a 1.5% agarose gel (Section II.2.12). DNA was visualized using ethidium bromide.



III.4.4. DNA methylase is associated preferentially with S 0.

Nuclei were prepared from deoxy-  $[U-^{14}C]$  -cytidine prelabelled L929 cells in log phase. These nuclei were extracted using the Sanders' extraction procedure. Fractions S 50 and S 100 and fractions S 200, S 400 and pellet fraction were pooled. Fraction S 0 and the pooled fractions were centrifuged on 15-30% glycerol gradients at 248,000 x g for 48 hours at 4°C.

Fractions from each gradient were collected and assayed for the presence of DNA ( $^{14}C$  cpm) and DNA methylase (assayed by the method of Turnbull & Adams, 1976). Figure 32 shows the amount of  $^{14}C$  and  $^3H$  (methyl groups) in each sample of the gradients.

With the S 0 fraction five distinct peaks of DNA methylase can be seen, of these four correspond to peaks of DNA. The first methylase peak is presumably free methylase not associated with DNA, whereas the other peaks are due to methylase bound to the DNA in nucleosomes.

Similar patterns can be seen with the other fractions though the proportion of low molecular weight DNA is less than in the S 0 fraction.

The relative amount of methylase in the higher salt fractions is less than in the S 0, i.e. the methylase/DNA ratio is highest in the S 0 fraction. This result shows that DNA methylase is preferentially associated with DNA that is extracted into the low salt fractions of a Sanders' extraction, i.e. with transcriptionally active chromatin.

Figure 32. Glycerol gradients of the fractions of  
a Sanders' extraction of nuclei from  
L929 cells.

A: Fraction S 0.

B: Fractions S 50 and S 100 (pooled).

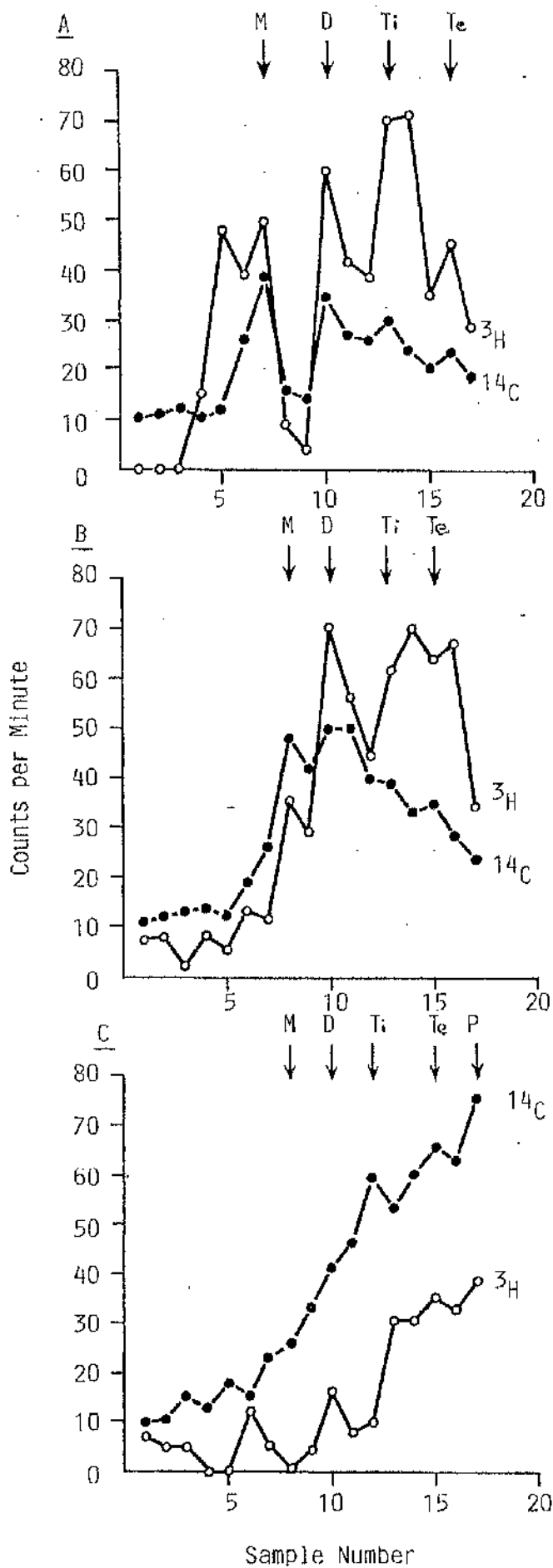
C: Fractions S 200, S 400 and pellet  
(pooled).

M, D, T, T<sub>2</sub> and P designate mono, di, tri,  
tetra and pentanucleosomes respectively.

Nuclei from deoxy- [U-<sup>14</sup>C] -cytidine prelabelled  
L929 cells were extracted using the Sanders' extraction  
procedure (Section II.2.6). Pooled fractions were  
then centrifuged on 15-30% glycerol gradients at  
248,000 x g for 48 hours at 4°C.

Fractions from each gradient were collected and  
assayed for the presence of DNA (<sup>14</sup>C cpm) and DNA  
methylase activity (assayed by the method of Turnbull &  
Adams, 1976).

Figure 32.



### III.5. Delayed methylation

#### III.5.1. DNA methylation is not completed for several hours after DNA synthesis

In Section I.2.2.1 it was mentioned that DNA methylation occurs rapidly after DNA synthesis, but that at least 20% of this methylation is still not complete after several hours.

Mouse L929 cells in log phase were labelled with 50  $\mu$ Ci of [ $6 - ^3\text{H}$ ]-uridine (Section II.2.1.2) for 50 minutes or 2 days at 37°C. The cells were harvested at the same time and nuclei were prepared. Another batch of L929 cells was labelled for 2 days and harvested when the cells had attained stationary phase.

Half the nuclei in each batch were treated using micrococcal nuclease until 50% of the DNA was rendered acid soluble. The remaining DNA was purified and pyrolysed to give the individual bases which were separated on an Aminex A6 column (Section II.2.11.2). The amount of  $m^3\text{C}$  was then determined (Table 19).

The DNA which is synthesized during the 50 minute label is undermethylated by about 30% compared to DNA in cells labelled for 2 days. The DNA which is labelled with tritium in this experiment is that DNA which was synthesized during the 50 minute label. Thus the methylation observed is the methylation of DNA made during that time. This shows that this DNA is undermethylated relative to bulk DNA.

It has been suggested that the level of methylation of DNA from stationary phase cells is higher than in DNA from log phase cells (Rubery & Newton, 1973; Kunnath & Locker, 1982) and in view of the observation that methylation can take many hours to be completed, a small difference might be expected. Table 19 shows that there is no significant difference between the methylation of DNA from log phase or stationary phase cells. Therefore any difference is too small to be detected by this method.

Table 19. New DNA is less methylated than old DNA.

Incubation conditions	$^*m^5C$	
	Total DNA	Nucleosomal core DNA
Log phase cells: 50 minute label	2.61; 2.85	<sup>†</sup> 4.10; 3.72
Log phase cells: 2 day label	3.88; 4.12	4.96; 5.15
Stationary phase cells	4.06; 3.58	5.03; 4.87

\*See Section III.1.3 (page 58 ).

<sup>†</sup>Unreliable result due to bad separation on the Aminex column.

III.5.2. Delayed methylation occurs predominantly in core DNA.

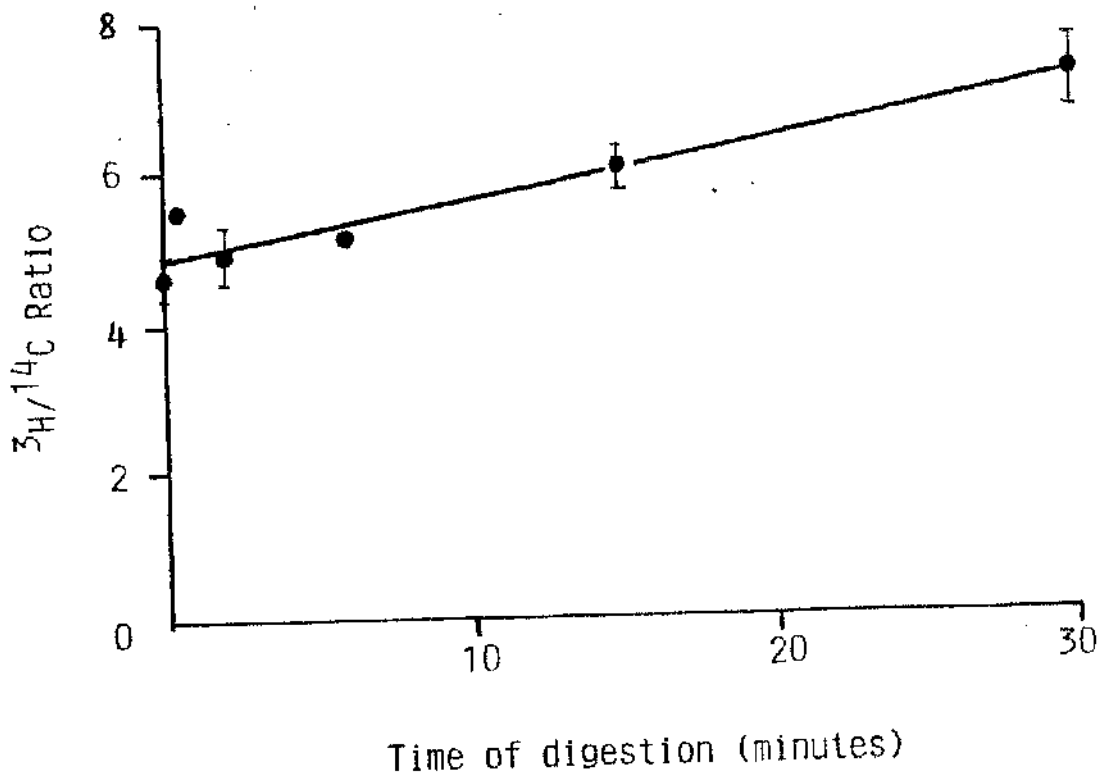
Log phase L929 cells prelabelled with deoxy-  $[U-^{14}C]$  - cytidine were incubated in the presence of hydroxyurea; hypoxanthine; thymidine, aminopterin and sodium formate, and labelled using L-  $[methyl-^3H]$  -methionine as described in Section II.2.1.2. Under these conditions hydroxyurea inhibits DNA synthesis. The aminopterin and sodium formate are used to prevent the terminal methyl group of methionine entering the one carbon pool (Adams, 1971). Hypoxanthine is added because the aminopterin inhibits the synthesis of purines, and the cells still require to synthesize RNA. The end result is that the methyl group of methionine is transferred solely to cytosine in DNA (or to bases in RNA).

The labelled methionine is added 1 hour after DNA synthesis is inhibited so the methyl groups in DNA derived from the labelled methionine is delayed methylation.

Nuclei were prepared from these cells and treated with micrococcal nuclease for various times (Section II.2.3.1). The DNA remaining acid insoluble at each time was purified and the  $^3H/^{14}C$  ratio was determined (Figure 33).

As can be seen the  $^3H/^{14}C$  ratio increases with time showing that the delayed methylation occurring in vivo has a preference for core DNA.

Figure 33. Delayed methylation occurs preferentially in nucleosomal core DNA.



Nuclei were prepared from deoxy-  $[\text{U}-^{14}\text{C}]$ -cytidine prelabelled cells which had been cultured in the presence of L-  $[\text{methyl-}^3\text{H}]$ -methionine as described in Section II.2.1.2. These nuclei were digested using micrococcal nuclease (Section II.2.3.1) the DNA was purified and the  $^3\text{H}/^{14}\text{C}$  ratio was determined.

The bars are the range of values obtained.

### III.5.3. Effect of salt extraction of nuclei on DNA methylation

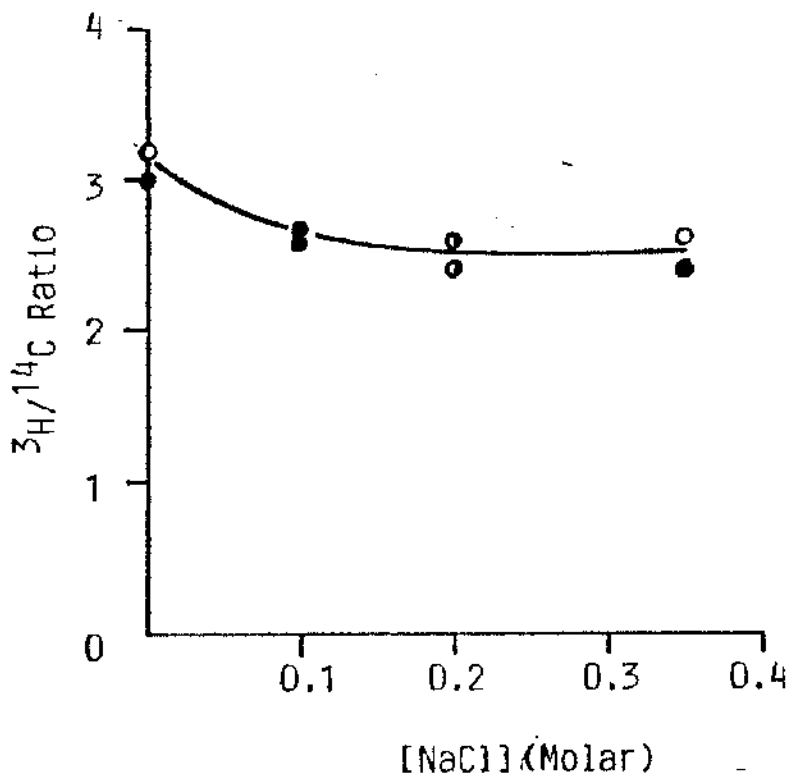
L929 cells in log phase were incubated for 1 hour in the presence of 3 mM hydroxyurea (to inhibit DNA synthesis). Nuclei were prepared from these cells and also from untreated cells, and extracted using buffer M containing NaCl at concentrations from 0 to 0.35 M (this was not a sequential extraction). The nuclei were then washed twice in buffer M and methylated in the presence of 5  $\mu$ Ci of S-adenosyl-L- $[\text{methyl-}^3\text{H}]$ -methionine (1.5 Ci/mmmole) for 2 hours at 37°C. The DNA was purified and the  $^3\text{H}/^{14}\text{C}$  ratio was determined (Figure 34).

In this experiment both the total methylation that is occurring and the delayed methylation are being assayed and Figure 34 shows that the amount of methylation that is occurring is the same in the nuclei from both treated and untreated cells. When nuclei are extracted using 0.35 M NaCl the soluble methylase is extracted (Turnbull & Adams, 1976; Qureshi *et al.*, 1982). Figure 34 shows that the amount of methylation decreases by only 20% thus showing that most of the methylation, both total and delayed methylation, is a product of a tightly bound DNA methylase.

### III.5.4. DNA methylation is associated with the nuclear matrix

L929 cells in log phase were grown in the presence of 3 mM hydroxyurea for 1 hour and then harvested and nuclei were prepared. One half of the nuclei were washed three times with buffer M<sup>+</sup> (5 ml) and resuspended in buffer M<sup>+</sup>. The other half were washed twice with buffer M<sup>+</sup> containing 0.2 M NaCl (5 ml) followed by buffer M<sup>+</sup> (5 ml) and resuspended in buffer M<sup>+</sup>. Both batches were methylated for 2 hours in the presence of 5  $\mu$ Ci S-adenosyl-L- $[\text{methyl-}^3\text{H}]$ -methionine (1.5 Ci/mmmole). The nuclei were then extracted with <sup>buffer</sup> M<sup>+</sup> containing 2 M NaCl and centrifuged at 165,000 x g for 48 hours. The pelleted material was resuspended in buffer M<sup>+</sup> containing 5 mM CaCl<sub>2</sub> and the material digested using micrococcal nuclease at 1000 units per ml for 1 hour

Figure 34. Effect of salt extraction of nuclei on in vitro methylation.



Open circles; Cells cultured in the absence of hydroxyurea.  
Closed circles; Cells cultured in the presence of hydroxyurea.

L929 cells prelabelled with deoxy-  $[\text{U-}^{14}\text{C}]$ -cytidine in log phase were incubated in the presence or absence of 3 mM hydroxyurea for 1 hour at  $37^\circ\text{C}$ . Nuclei were prepared from these cells and extracted using buffer M containing NaCl at concentrations from 0 to 0.35 M (not sequentially). The nuclei were then methylated in the presence of 5  $\mu\text{Ci}$  of S-adenosyl-L-  $[\text{methyl-}^3\text{H}]$ -methionine (1.5 Ci/mmole) for 2 hours at  $37^\circ\text{C}$ . The DNA was purified and the  $^3\text{H}/^{14}\text{C}$  ratio was determined (Section II.2.4).



at 37°C. Samples were taken at each stage of the procedure and the amount of DNA and the  $^3\text{H}/^{14}\text{C}$  ratio were determined (Table 20).

The result shows that there is an enrichment of methylated DNA associated with the matrix. Very little difference is seen when the soluble methylase is extracted prior to the treatment.

Table 20. Methylation occurs preferentially in DNA associated with the nuclear matrix.

Sample	Recovery of DNA (percent)	$^3\text{H}/^{14}\text{C}$ ratio
A Total	100	0.11
B Total	100	0.095
A 2 M NaCl pellet	71	0.16
B 2 M NaCl pellet	39	0.18
A matrix	14	0.22
B matrix	7	0.38

A untreated nuclei

B nuclei extracted with 0.2 M NaCl prior to procedure.

As the previous experiment has shown that all the methylation that occurs in isolated nuclei is delayed methylation (Fig. 34) it is concluded that DNA methylation is a product of a bound methylase activity which is associated with the nuclear matrix.

### III.5.5. Effect of 5-azadeoxycytidine on in vitro methylation

Indirect evidence supporting the idea that methylation is a product of a bound DNA methylase comes from the use of the methylation inhibitor 5-azacytidine. When L929 cells are released from stationary phase and cultured for 10 hours, followed by 10 hours in the presence of 1  $\mu\text{M}$  5-azadeoxycytidine, the soluble methylase activity decreases by five fold whereas the bound activity increases (Qureshi, 1983). When nuclei from these cells and untreated cells are methylated

using the endogenous methylase no difference in methylation is observed (Table 21).

Table 21. Methylation of DNA in nuclei from L929 cells treated with 5-azadeoxycytidine.

Sample	pmoles methyl group incorporated
L929 nuclei	0.18
L929 nuclei + ascites methylase	0.21
*Laza nuclei	0.20
*Laza nuclei + ascites methylase	1.00

\*Laza is an abbreviation for L929 cells treated with 5-azadeoxycytidine.

The amount of DNA in each sample was 1 µg.

This result shows that DNA methylation is not a product of the soluble DNA methylase.

When these nuclei are methylated using ascites methylase five times as many methyl groups are incorporated into Laza DNA than into DNA from control cells. This shows that there are many more hemimethylated sites and serves as an indicator that inhibition of methylation has occurred in vivo.

### III.5.6. The methylation in Laza nuclei is preferentially in core DNA

Laza nuclei from Section III.5.5 were methylated using the endogenous methylase and digested using micrococcal nuclease (Section II.2.3.1). The DNA remaining acid insoluble was purified and the  $^3\text{H}/^{14}\text{C}$  ratio was determined (Fig. 35). As can be seen, the ratio increases with time thus showing that when the soluble methylase is inhibited the methylation which occurs is still similar to the methylation occurring in control cells (Fig. 19, page 84).

Figure 35. Digestion of in vitro methylated nuclei  
of L929 cells prelabelled with  
5-azadeoxycytidine.

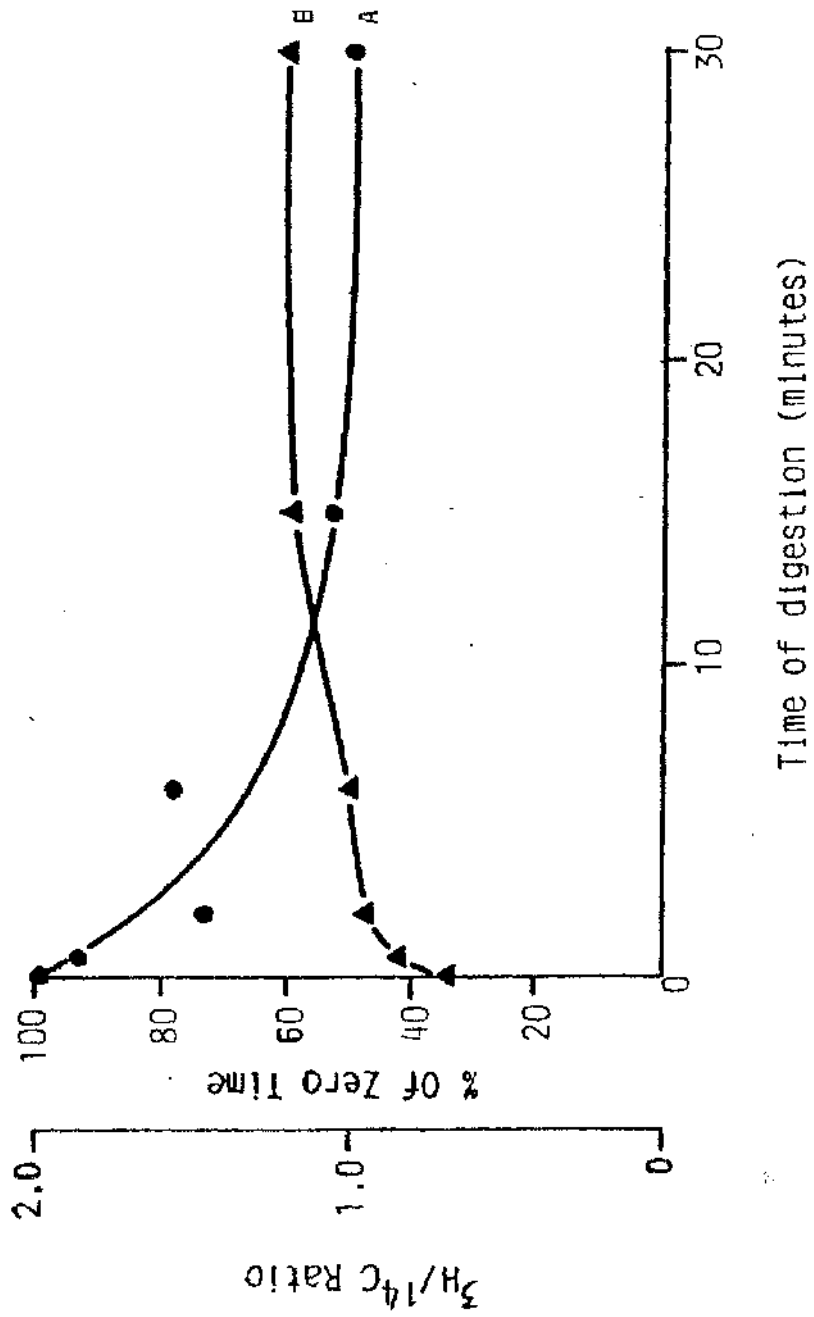
A:  $^{14}\text{C}$

B:  $^3\text{H}/^{14}\text{C}$ .

Nuclei from deoxy-  $[\text{U-}^{14}\text{C}]$ -cytidine prelabelled L929 cells which had been labelled with 5-azadeoxycytidine (Section II.2.1.3) were methylated in the presence of S-adenosyl-L-  $[\text{methyl-}^3\text{H}]$ -methionine at  $37^\circ\text{C}$  for 3 hours using the endogenous methylase (Section II.2.2.1). The nuclei were then digested using micrococcal nuclease (Section II.2.3.1), the DNA purified and the  $^3\text{H}/^{14}\text{C}$  ratio determined (Section II.2.4).

The amount of  $^{14}\text{C}$  is expressed as a percentage of zero time.

Figure 35.



IV. DISCUSSION

#### IV.1. Base composition of nucleosomal core DNA

##### IV.1.1. Core DNA is enriched in C + G residues

It has been shown in Section III.1 that the DNA in nucleosomal core particles prepared from mouse L929 chromatin is enriched in the base  $m^5C$ . This result is in agreement with the result for chicken and calf thymus chromatin (Razin & Cedar, 1977), though these authors found a greater increase. Such an increase however, was not found in the core DNA from the chromatin of Chinese hamster ovary cells (Adams *et al.*, 1977). These authors however, stated that no difference would have been detected even if all of the  $m^5C$  had been in the core DNA of this chromatin, as the methods were not sensitive enough to detect such an increase. The nucleosomal DNA is also enriched in the bases cytosine and guanine and in the dinucleotide CpG. The enrichment in  $m^5C$  and CpG may simply reflect the enrichment in C + G content (Adams & Eason, 1984).

##### IV.1.2. Problems inherent in these studies

Many studies have been carried out to determine whether nucleosomes are located on specific DNA sequences in a non-random manner (Section I.5.5). These studies have usually involved the use of micrococcal nuclease to cleave the DNA between adjacent nucleosomes, and to relate these cleavage sites to nearby restriction nuclease sites, by examination of the bands produced using electrophoresis on agarose gels. These studies have given both positive and negative results (Zachau & Igo-Kemenes, 1981).

There are however, several problems inherent in these studies. First, micrococcal nuclease has specificities in its mode of cleavage of DNA, it cleaves preferentially in A + T rich regions (Nelson *et al.*, 1979a; Dingwall *et al.*, 1981; Horz & Altenberger, 1981). Second, this nuclease will also cleave in the core DNA and does so in A + T rich regions (McGhee & Felsenfeld, 1983), so many of the potential 'phases' will be lost in studies using this enzyme, as

only the 146 base pair DNA fragment isolated from agarose gels is examined.

These problems are inherent in any study when micrococcal nuclease is used. In the present study, three ways of avoiding these problems were used. First, when 'naked' DNA is cleaved using micrococcal nuclease, it has been reported that the enzyme cleaves preferentially at A + T rich regions, as in chromatin (Nelson et al., 1979a; Keene & Elgin, 1981; Udvardy & Schedl, 1983) and the DNA which remains acid insoluble is therefore C + G rich (Nelson et al., 1979a). In the present work, when purified DNA was digested until 50% of the DNA was made acid soluble, no enrichment in m<sup>5</sup>C was seen in the resistant material (Section III.1.4). This suggests that even though the initial endonucleolytic cleavages are in A + T rich regions, the subsequent exonucleolytic activity is not selective. This is supported by the observation that no enrichment of calf thymus DNA in C + G during digestion using micrococcal nuclease is found (R.L.P. Adams, personal communication). Extensive studies by Cockell et al. (1983) provide further support. Second, when chromatin was digested, the digestions did not proceed significantly beyond 50% acid solubility, and intact 146 base pair fragments were not isolated for study. Rather, that DNA precipitable using 70% (v/v) ethanol was used. This DNA contained fragments which were smaller than 146 base pairs (data not shown), thus those core fragments which are cleaved at A + T rich sites are recovered as more than one fragment. In this way the loss of 146 base pair fragments is hopefully minimized, and the artifact reported by McGhee & Felsenfeld (1983) will be avoided. Third, an agent which cleaves DNA in a random manner was used to digest the DNA in chromatin (Section III.1.5, page 60). This reagent is methidiumpropyl-EDTA.iron (II), which in the presence of oxygen will cleave DNA (Hertzberg & Dervan, 1982). Extensive studies using this reagent have shown that it has no base specificities in its cleavage (van Dyke & Dervan, 1983a; 1983b), and it has a strong preference for cleavage in the linker regions of chromatin

(Cartwright *et al.*, 1983). By the use of this reagent, I have shown that a core particle containing about 140 base pairs of DNA can be produced (Fig. 7, page 61). This DNA is again enriched in  $m^5C$  when compared with total DNA (Table 7, page 61). The core particle is not thought to be the product of an endogenous nuclease (e.g. the  $Ca^{2+}/Mg^{2+}$  dependent nuclease) as EDTA was present in the reaction mixture at 2 mM. This has little effect on the methidium-propyl-EDTA.iron (II) reaction (Cartwright *et al.*, 1983).

Thus it appears that nucleosomal core DNA is indeed enriched in the bases C, G and  $m^5C$ .

#### IV.1.3. Reconstitution of chromatin

When mouse L929 cell DNA and histones are reconstituted to give chromatin, and this chromatin digested using micrococcal nuclease, the resistant DNA is again found to be enriched in  $m^5C$  (Table 8, page 63). This result is in contrast to that of Razin & Cedar (1977) who found a randomization of  $m^5C$  after reconstitution of chromatin. This result suggests that nucleosomes are binding preferentially to C + G rich regions of DNA and suggests that there is some property of this DNA which favours nucleosome formation.

#### IV.1.4. Application of DNA flexibility models

Two models for DNA flexibility based upon aspects of the DNA sequence have been described (Trifonov, 1980; Zhurkin, 1983) and these are discussed in Section III.1.7.

The model of Zhurkin has been used for study because of the reasons outlined in Section III.1.7. Also the model of Trifonov does not seem to hold over stretches of DNA longer than 60 to 100 base pairs, whereas that of Zhurkin holds over many kilobases. When this model is applied to a mouse DNA sequence, a distinct periodicity of about 10.5 base pairs is seen (Fig. 9, page 66). If the presence of these RY and YR dimers at these positions does indeed enhance the flexibility of DNA in nucleosomes, then the regions of the DNA which contain this periodicity will be more flexible than DNA not containing this periodicity. This is because these maxima



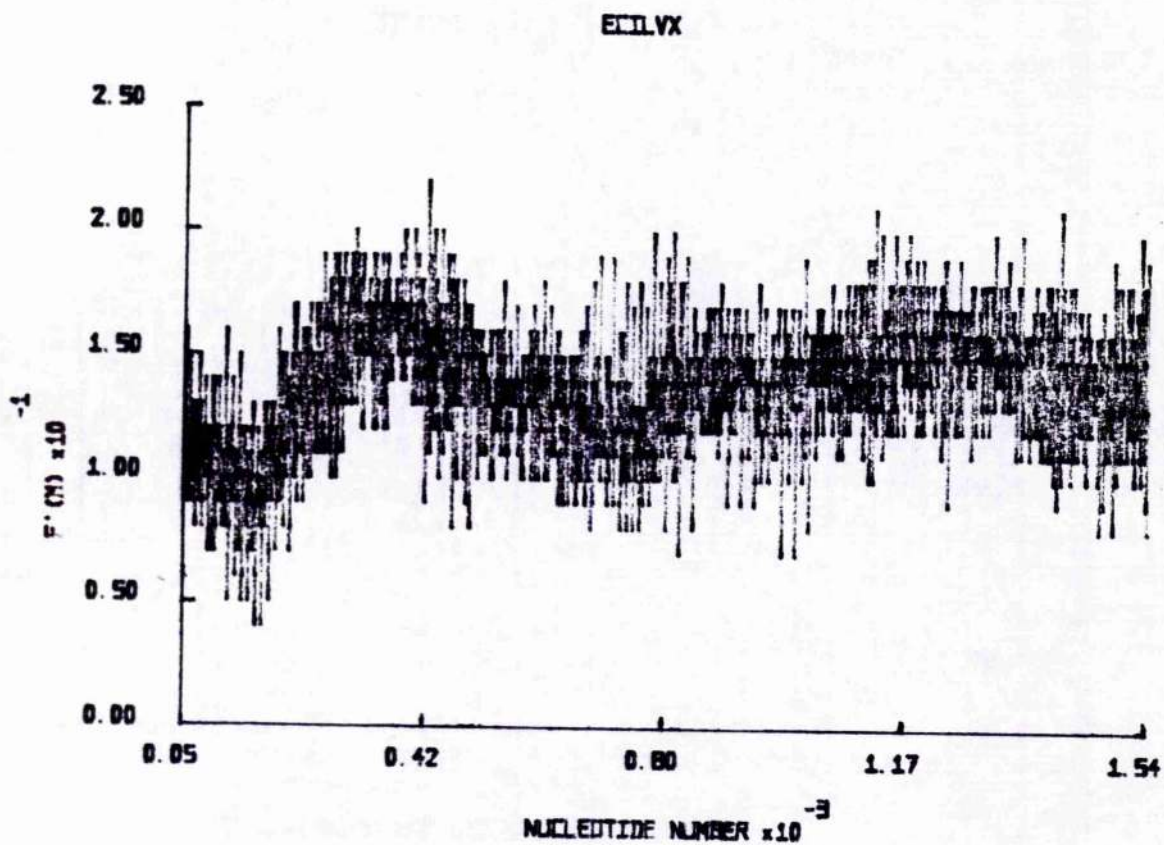
will reinforce each other. If on the other hand maxima are found outwith this periodicity, they will have a tendency to oppose the flexibility. For the mouse sequences this is rarely seen. However when an E.coli sequence is analyzed for this periodicity, maxima are not arranged at 10.5 base pair intervals, but in some regions occur seemingly at random (Fig. 36). This shows that the reinforcement of flexibility is not observed.

The positions of nucleosomes can be predicted on the basis of Zhurkin's model (Fig. 9, page 66). Zhurkin has applied his model to DNA sequences on which nucleosomes have been located using the nuclease methods described in Section I.5.5. On the whole the predictions match the experimental results (Zhurkin, 1983). Thus this seems to be a valid model for prediction of nucleosome binding sites.

The DNA sequences used in the flexibility studies have also been analyzed for the presence of C and G residues in 100 and 148 base frames along the sequences (Fig. 10, page 69). The figures show that there is a marked variation in the percent C + G as one progresses along the sequences (see also Adams & Eason, 1984). This variation is more obvious in introns than in exons. In the former, many regions are very C + G poor, this is also seen for the 5' regions immediately preceding the messenger RNA starting point. When the locations of nucleosomes predicted by the model of Zhurkin are superimposed onto these figures, it is clear that in many cases predicted nucleosome positions coincide with regions of high C + G content, or more particularly, very few predicted positions occur on regions of very high A + T content. When the C + G content of predicted nucleosomes is determined and compared to the whole sequence, an increase is seen. This increase is small (about 9%) but it is the same as the increase seen for mouse chromatin (Table 5, page 59). Thus this evidence supports the experimental observation that the DNA in nucleosomes is C + G rich.

Figure 36. Predicted flexibility of the E.coli ilv locus (ilvL- ilvE).

Figure 36a.

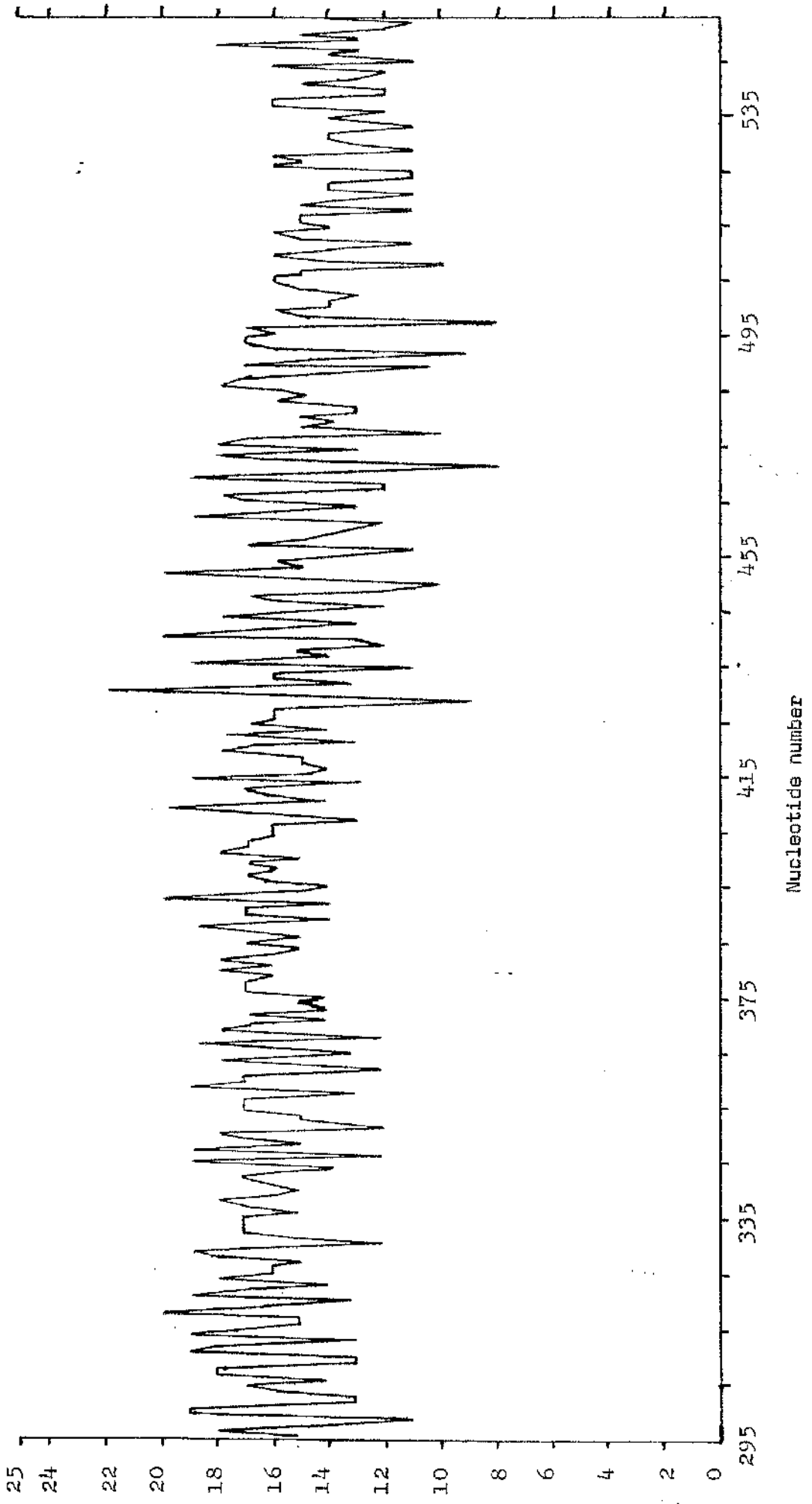


36a: bases 50 to 1540

36b: bases 295 to 555

The bendability function ( $F'(M)$ ) is derived as described for Figure 9.

Figure 36b.



#### IV.1.5. Implications of these studies

From the previous sections the overall conclusion is that nucleosomes have a tendency to form on C + G rich regions of DNA due to the enhanced flexibility of such regions. How does this compare to other studies on DNA folding?

The result seems to be in conflict with the results of Hogan et al. (1983). These authors measured the Young's modulus (measure of stiffness) of the homopolymers poly(dG).poly(dC) and poly(dA).poly(dT) and found that the value for the former is higher. This shows that poly(dG).poly(dC) is a stiffer molecule and suggests that it is less likely to be able to bend around the histone octamer in a nucleosome than the poly(dA).poly(dT) copolymer. It has been shown however, that neither of the homopolymers form a chromatin-like structure when associated with the four core histones (Rhodes, 1979; Simpson & Kunzler, 1979; Kunkel & Martinson, 1981). Therefore the result of Hogan et al. (1983) may not be applicable to DNA which has all four bases. These same studies have shown that the copolymers poly(dA.dT).poly(dA.dT) and poly(dG.dC).poly(dG.dC) will both form well defined nucleosomes.

It is interesting that the flexibility model of Trifonov (1980) predicts that the poly(dA).poly(dT) copolymer would be very flexible and capable of forming a nucleosome, whereas the model of Zhurkin (1983) does not. This gives further support for the use of the latter model in the flexibility studies in Section III.1.7.

An interesting study has been made of the trypanosome kinetoplast DNA (Wu & Crothers, 1984). A segment of this DNA behaves anomalously when electrophoresed on agarose gels. This behaviour has been related to the occurrence of a bending locus in the DNA by these authors (Marini et al., 1982; Wu & Crothers, 1984). This bending locus has tracts of a simple repeat sequence (CA<sub>5-6</sub>T) symmetrically distributed about it, with a repeat interval of 10 base pairs. These sequences are interspersed with C + G rich sequences. Whether

this result has any significance to studies on mouse genes is unclear, but other sequences of this type in kinetoplast DNA are not associated with bending loci.

In support of the data in Section III.1, the 5'-region of the Drosophila Hsp 70 heat shock gene is very A + T rich (Torok & Karch, 1980) and this region has a very low predicted flexibility (Zhurkin, 1983). It is in this region that the DNase I hypersensitive sites are located (Wu, 1980), and it is perhaps reasonable to expect that nucleosomes are not associated with this DNA. Indeed it has been shown that nucleosomes do not form readily on DNA containing runs of A + T residues (Kunkel & Martinson, 1981). Such runs will occur on C + G poor regions of DNA showing low bendability, i.e. they are poor in ApT and TpA dinucleotides at the correct positions. Many such A or T runs are seen in the 5'-region of the Hsp 70 gene (Torok & Karch, 1980). These runs are short, but occur many times, and perhaps this is sufficient to inhibit nucleosome formation.

The estrogen-regulated chicken vitellogenin gene also has a very A + T rich region in the 5' flanking sequence (Geiser et al., 1983). In this region the estrogen receptor binds, so it would perhaps be necessary for this region to be nucleosome free. It is interesting that all of the A + T rich regions in this 5' flanking sequence can form hairpins (Geiser et al., 1983). This is also seen in the Hsp 70 gene (Torok & Karch, 1980). These observations tie in with the observation that hairpins do not associate with nucleosomes (Nickol & Martin, 1983). Structures sensitive to single strand specific nucleases have recently been shown to occur in the 5' region of the  $\beta$ A-globin gene of chicken in native chromatin (Larsen & Weintraub, 1982; Weintraub, 1983), and may be due to hairpin-like elements. These sites correlate with the DNase I hypersensitive sites in this gene.

Furthermore, when micrococcal nuclease is used as a probe for DNA structure, preferred sites are found at intervals of about 200 base pairs, and are located at A + T rich segments (Kecne & Elgin, 1981; 1984; Udvardy & Schedl, 1983). It has been suggested that these regions occur in linker DNA.

Thus it seems that nucleosomes would not bind to A + T rich regions of DNA. This model would allow 5' control regions to be nucleosome free if they had a high content of A and T residues. This is clearly seen for Hsp 70, vitellogenin and the  $\beta 1$ -globin gene (of mouse) (Fig. 10, page 69). This is also predicted for the H2KB complex of mouse (Fig. 10, page 69).

#### IV.1.6. Alternative explanations

Alternative explanations for the locations of nucleosomes on C + G rich DNA may exist. For instance, the increased stability of the CG base pair and the increased hydrophobicity of methylcytosine relative to cytosine may lead to a greater stability of nucleosomes in C + G rich DNA, as may the different helical twist angles existing between adjacent bases.

It is even possible that nucleosomes located on specific sequences decrease the propensity of methylcytosines in these (stabilized) regions to undergo deamination. This would lead to linker regions becoming A + T rich as  $m^5CpG$  dinucleotides are converted to TpG and CpA (Bird, 1980). Such an explanation would require nucleosomes to have adopted positions on the DNA which would need to have been conserved for millions of years for a regular pattern to be established.

#### IV.2. Methylation of chromatin and DNA in vitro

##### IV.2.1. Effect of chromatin structure on DNA methylation

In Section III.1 (page 55) it has been shown that nucleosomal core DNA is enriched in  $m^5C$ , C and G. The possible reasons for this enrichment are discussed in Section IV.1.4.

An alternative explanation based upon preference of the methylase for core regions of chromatin is also possible. Therefore experiments were performed to test this idea.

When nuclei from mosquito cells are methylated using ascites DNA methylase, most of the methyl groups are added to nuclease sensitive (linker) DNA. This result is the same as that obtained by Bloch & Cedar (1976) who used E. coli methylase and nuclei from chicken erythrocytes. This suggests

that the linker DNA is more accessible than the core DNA to the methylase and is in apparent contradiction to the observed distribution of  $m^5C$  in vivo. The same conclusion as to the accessibility of linker DNA to enzymes has come from studies using micrococcal nuclease (Noll, 1974a; Rill & van Holde, 1974). Thus linker has a more open structure. When ascites methylase is used to methylate chromatin in nuclei from stationary phase L929 cells, the same result is observed, though with these nuclei some methylation occurs in core DNA. On the other hand, when nuclei from log phase cells are methylated using the endogenous methylase the opposite result is obtained, i.e. the core DNA is preferentially methylated. This latter result reflects the in vivo distribution of  $m^5C$  in chromatin. However, when ascites DNA methylase is used to methylate nuclei from log phase cells, linker DNA is again preferentially methylated (the ratio of  $^3H/^{14}C$  now decreases).

These results tell us several things about the nature of DNA methylation in chromatin. First, methylation using the endogenous methylase is hemimethylation (Section I.2.2.1) and this occurs preferentially in core DNA. Second, methylation using the added ascites DNA methylase can be both de novo and hemimethylation. With mosquito nuclei this action is de novo, as the amount of methyl groups added by the enzyme is greater than the amount already present in the DNA (0.5 pmoles/ $\mu g$  DNA compared with less than 0.1 pmole/ $\mu g$  DNA). In this case the enzyme is methylating the more accessible linker DNA. Using stationary phase cells the same conclusion can be made, except that very few methyl groups are incorporated compared to the amount present in vivo (0.2 pmoles/ $\mu g$  DNA compared with 20 pmoles/ $\mu g$  DNA). Third, the conclusion is that for some reason the hemimethylated sites in the DNA of log phase L929 cells are preferentially in core DNA (though many sites are in linker DNA).

One possible reason for this is the lag of methylation at DNA replication (Section II.2.2.1). Nucleosomes are positioned on DNA very soon after DNA replication (De Pamphilis &

Wassarman, 1980; Herman *et al.*, 1981) and perhaps this process precedes DNA methylation. If so, then the hemimethylated sites in the linker DNA will be more rapidly methylated than those in the core DNA. Therefore when nuclei are isolated, there will be more hemimethylated sites remaining in the core DNA than in the linker DNA, and it is these sites that are methylated in vitro.

An alternative explanation is that the methylase has a higher affinity for core chromatin than for linker chromatin. This would tend to be excluded based on the results using exogenous methylase. Nevertheless a more detailed study was undertaken.

Chromatin, core chromatin and DNA were methylated using the ascites methylase under conditions where the methylase was in excess. The results are given in Figures 21-25 (pages 87, 90-92). From the results several statements can be made. First, chromatin from log phase cells is a better substrate than chromatin from stationary phase cells. Second, core chromatins are a poor substrate for the enzyme. Third, the DNA from all sources is a good substrate, and there is little difference between total and core DNA. Fourth, in all cases the insect substrates are better acceptors of methyl groups than the L929 substrates. Fifth, DNA is a ten fold better acceptor than chromatin.

From these results it can be concluded that: one, the difference between the methylation of total and core chromatin is not due to major differences in the DNA, two, the methylase does not have a higher affinity for core chromatin, and three, the DNA in linker has a lower accessibility to methylase than the purified DNA.

One of the problems with these comparisons is that the endogenous methylase does not efficiently methylate de novo sites in vivo (Adams & Hogarth, 1973) whereas the added ascites methylase does so. This may be due to the endogenous methylase being in a different environment or in a different form, and studies on 'bound' and 'soluble' forms of methylase would tend to support this (Section III.5). Alternatively



the enzyme may have been activated in some way. Stimulation of methylase activity by limited proteolysis using trypsin has been reported (Adams et al., 1983b) though this should not happen during purification of the methylase as the trypsin inhibitor PMSF is present.

#### IV.2.2. Effects of proteins on DNA methylation in vitro

The previous section has shown that chromatin structure affects methylation in vitro as DNA accepts ten times as many methyl groups as chromatin. Is this difference due to the chromatin structure or simply due to the presence of chromatin proteins?

When DNA is methylated in vitro in the presence of increasing concentrations of histones, inhibition of methylation occurs (Fig. 26, page 93). This inhibition increases up to a protein/DNA ratio (w/w) of 1.0, thereafter adding more histone has little effect. The protein/DNA ratio (w/w) of native chromatin is also about 1.0. The difference between the methyl acceptor ability of DNA and chromatin is about ten fold, this is the same difference as is obtained when histones are simply added to a DNA solution at the same protein/DNA ratio. Thus it seems that it is mainly the presence of histones which protects the DNA from methylation. This cannot be completely true as linker DNA is also protected to some extent. The effect of nuclear non-histone proteins was also studied. No inhibition of methylation by the HMG proteins or the nuclear non-histones in an 0.35 M NaCl extract was observed.

This result is in conflict with the results of Huang et al. (1984) who showed that HMG 14 inhibited the methylation of native Micrococcus luteus DNA, denatured M.luteus DNA and hemimethylated phage XP-12 DNA by more than 90% at a protein/DNA ratio (w/w) of 2.0. The HMG proteins 1 and 2 had limited effects on the methylation of M.luteus DNA but not XP-12 DNA.

As these proteins are purified using TCA, it is possible that the result observed is due to insufficient dialysis to remove the TCA. The same result was observed in the present

work, but when the HMG preparation was neutralized prior to dialysis no inhibition was observed (R.L.P. Adams, personal communication).

The HMG proteins bind preferentially to nucleosomes (not DNA) (Weisbrod et al., 1980; Weisbrod & Weintraub, 1981) so it is possible that no inhibitory effects of these proteins will be seen in the absence of nucleosomes. Therefore the effects of these proteins on the methylation of chromatin should be determined. Preliminary experiments from such experiments show that the removal of non-histone proteins from mosquito nuclei results in an increased methylation of DNA in chromatin compared with control nuclei (data not shown). Also when an 0.35 M NaCl extract of mosquito nuclei is added to mosquito nuclei which are then methylated using ascites methylase, up to 50% inhibition of methylation is observed (data not shown), though this did not occur when L929 nuclei were used (R.L.P. Adams, personal communication). These experiments using insect cells have problems due to the lack of HMG 14/17-like proteins in insects (Franco et al., 1977; Bassuk & Mayfield, 1982).

#### IV.2.3. Conclusions

Several conclusions concerning in vivo methylation can be made based upon these in vitro studies. First, the preferential methylation of DNA in the nucleosomal core particle is not due to major differences in this DNA. Second, DNA methylase does not have a greater affinity for core particles than for total chromatin. Third, histones inhibit DNA methylation in vitro. Fourth, non-histones may need the presence of chromatin structure (i.e. nucleosomes) before they affect methylation.

The overall conclusion is that the increased methylation in core DNA is not due to preference of the methylase for core regions of chromatin.

### IV.3. Methylation of DNA in transcriptionally active chromatin

#### IV.3.1. Isolation of transcribing chromatin

As was discussed in Section III.3, one of the problems with studying transcriptionally active chromatin is the isolation of these regions from bulk chromatin. Deoxyribonuclease I studies were discarded due to the problems outlined in Section III.3.1, whereas DNase II studies had problems due to pH changes, though this method worked fairly well the results obtained from its use were prone to experimental errors.

The Sanders' procedure (outlined in Section II.2.6) was used because of its simplicity and reproducibility. The method works something like the DNase II/Mg<sup>2+</sup> solubility method (Gottesfeld & Butler, 1977). The enzyme micrococcal nuclease is used to cleave the chromatin into nucleosomes. The transcriptionally active nucleosomes are then extracted from the nuclei using NaCl and Mg<sup>2+</sup> ions. The latter are essential as the extraction procedure does not work in the presence of EDTA. One of the advantages of this procedure is that the digestion time is relatively unimportant and does not affect the distribution of nucleosomes between the supernatant fractions (Davie & Saunders, 1981).

The amount of material recovered in the low salt fractions increases with digestion at early digestion times, but reaches a plateau at about 10% of the total DNA (data not shown). The reason why the nucleosomes are present mainly as mononucleosomes in the S0 and S50 fractions (Fig. 30, page 102) is due to the preferential cleavage sites for this nuclease in the transcriptionally active chromatin (Gottesfeld & Butler, 1977; Bloom & Anderson, 1978).

#### IV.3.2. Evidence that the nucleosomes in S0 and S50 are derived from transcribing chromatin

Two lines of evidence are provided. First, when nuclei are allowed to undergo transcription in vitro in the presence of only one ribonucleotide and then extracted using the Sanders' extraction procedure, most of the radioactivity

(derived from [ $^3\text{H}$ ]-UTP) is found in the fractions S0 and S50 (Table 11, page 99), thus showing that nascent RNA is present in these fractions. Second, when samples of each extract are electrophoresed on acrylamide gels, fractions S0 and S50 are enriched in HMG proteins and depleted in histone H1 (Fig. 29, page 100), this is characteristic of transcribing chromatin (Section I.5.6).

The problem with these gels is that the HMG 14 and 17 bands run among the core histone bands and only HMG 1 and 2 are visible.

These observations are in agreement with those of Davie & Saunders (1981). These authors also showed that transcribing nucleosomes are enriched in transcriptionally active sequences.

#### IV.3.3. Characterization of the DNA in fractions S0 and S50

Nuclei from mouse L929 cells were digested briefly with micrococcal nuclease and treated using the Sanders' procedure. DNA was purified and the amount of  $\text{m}^5\text{C}$  determined (Table 12, page 103).

The results show that the DNA in fractions S0 and S50 has an  $\text{m}^5\text{C}$  content of less than 1.0%, i.e. less than 1.0% of the cytosines are  $\text{m}^5\text{C}$ . This compares with 3.5% in total DNA (Section III.1.2). This result is similar to the result for calf thymus (Davie & Saunders, 1981) except that the deficiency is much greater for mouse cells. The DNA in the intermediate fractions is also deficient in  $\text{m}^5\text{C}$  and may be due to another class of DNA sequences or may be a mixture of transcribing and non-transcribing DNA. This latter would only occur if the chromatin has been insufficiently digested, and the lack of DNA greater than 400 base pairs in length in the S100-S400 fractions would tend to exclude this possibility.

The obvious problem in this experiment is that the level of methylation of total DNA (that DNA present in the pellet fraction when no digestion has taken place) is lower than that obtained for total DNA in other experiments. This latter figure has been derived using data from many different

determinations and should therefore be accurate. Figures as low as 2.9% have occasionally been obtained, and thus the figure obtained here may simply be from the extreme end of the normal distribution curve.

The conclusion is that DNA from transcriptionally active chromatin is deficient in  $m^5C$ .

As DNA methylation takes more than 1 hour to be completed, there will be DNA which is undermethylated compared with mature DNA and the S0 and S50 fractions may be enriched in this DNA. Therefore cells were labelled for 50 minutes in the presence of  $[6 - ^3H]$ -thymidine and the chromatin was fractionated as before. The DNA which is labelled is DNA less than 1 hour old. This DNA is not preferentially extracted in any of the fractions of the Sanders' fractionation. As the cells were asynchronous, the DNA being synthesized at any given time will be a random sample of the DNA sequences present in mouse cells. Thus the different amounts of  $m^5C$  in different DNAs will not bias the result (Section I.1.8.). This result will not be affected by the extraction of nascent DNA, as this DNA is present in a very small amount. Nascent DNA is that DNA less than two minutes old, and is not yet methylated.

#### IV.3.4. Methylation of transcribing chromatin in vitro

The  $m^5C$  content of the DNA in transcribing chromatin is less than in the DNA in non-transcribing chromatin. This is not due to a deficiency of methylatable sites, as only 40% of the CpG dinucleotides are methylated in this DNA in L929 cells (Naveh-Many & Cedar, 1981). The reason that transcribing DNA is undermethylated may be that the DNA methylase simply does not methylate at de novo sites to any great extent in vivo. As de novo methylation can occur in vitro, the question becomes, why does de novo methylation not occur in vivo in transcribing regions? A possible answer may simply be that the methylase has no access to the unmethylated sites in transcribing DNA. The effects of chromatin structure on methylation in vitro were examined in Section III.2 and these experiments can be extended to study transcribing and non-transcribing regions of chromatin.

When nuclei from log phase L929 cells are methylated using the endogenous DNA methylase, and the transcribing regions isolated, this DNA is found to have accepted more methyl groups than DNA in non-transcribing regions (Table 14, page 105). This seems to be a contradiction to the situation seen in vivo. A possible answer is that some perturbation has been introduced during nuclear preparation. As the use of a cell lysate of these cells gives the same result (Table 15, page 106), this shows that no such perturbation exists, or that this perturbation is not a result of nuclear preparation.

When nuclei from stationary phase cells are used, and methylated using the added ascites methylase, no such difference is observed (Tables 16 & 17, page 107). The conclusion is that transcribing chromatin is not more accessible to enzymes in stationary phase cells, whereas this chromatin is more accessible to endogenous methylase in log phase cells. The problem with these experiments is that the control experiment, that of log cells and added ascites methylase was not done. An analogous situation to log phase cells are mosquito cells, these cells are also in log phase. When ascites methylase is used to methylate the DNA in these nuclei, most of the methyl groups are added to transcribing DNA.

Thus it can be concluded that in nuclei from log phase cells the DNA in transcribing chromatin is more accessible to methylase in vitro than is non-transcribing chromatin.

Using L929 cells, this methylation is hemimethylation, and for some reason there are more of these sites in transcribing chromatin. Using mosquito cells on the other hand, this methylation is de novo and suggests that transcribing chromatin contains no inhibitors of DNA methylation which are not present in non-transcribing chromatin.

A possible reason for the presence of more hemimethylated sites in transcribing regions could be due to the on-going transcription. Maybe the transcriptional process competes with the methylase, so methylation remains incomplete longer in the transcribing regions. This would explain why no difference was observed using stationary phase cells as

DNA synthesis has stopped and DNA methylation will continue, so allowing methylation to be completed.

An experiment to test the idea that transcription competes with methylation was done using  $\alpha$ -amanitin to inhibit transcription in vivo. Nuclei were isolated and the chromatin was fractionated into transcribing and non-transcribing regions. The level of methylation was measured in each chromatin. No results are given, as this experiment unfortunately failed to work. Another experiment which needs to be done is to inhibit transcription in vivo, and methylate the nuclei in vitro using the endogenous methylase. This would show whether or not the methylation had been completed, i.e. the hemimethylated sites had now been filled.

A trivial reason why the transcribing DNA becomes more methylated in vitro using endogenous methylase, may be due to the nascent DNA being extracted into the S0 fraction in the Sanders' procedure. This DNA would be extensively hemimethylated. This latter has been shown to be the case (R.L.P. Adams, personal communication). This would explain the observation that the methylase is preferentially extracted in the S0 fraction (Fig. 32, page 111). This is in contrast to Creusot & Christman (1981) who found the methylase preferentially in non-transcribing chromatin. This again may be related to the newly synthesized regions being extracted into S0.

#### IV.4. Studies on delayed methylation

##### IV.4.1. DNA methylation can be delayed

Many studies have shown that DNA methylation takes many hours to be completed (Section I.2.2.1). In this work it has been confirmed that the level of methylation in DNA less than one hour after its synthesis is only 70% of the level in the DNA from stationary phase cells (Section III.5.1). The methylation which has occurred is preferentially in the core DNA of chromatin, thus it is similar to total methylation in this respect. In contrast to observations by Rubery & Newton (1973) and Kunnath & Locker (1982), there is no significant difference in the level of methylation between

log phase and stationary phase L929 cells (Table 19, page 114), and considering that at least 70% of the methylation occurs within a few minutes of DNA replication, it is not surprising that no difference can be detected. As these cells were asynchronous, there will not be any gross difference in the level of methylation in these cells due to the different classes of DNA which have different levels of methylation. These DNAs are known to be replicated at different times during S-phase (Adams, 1971; Comings, 1972; Schneiderman & Billen, 1973).

Due to an inaccuracy in one of the figures in Table 19, it is difficult to state that delayed methylation occurs preferentially in core or in linker DNA. Therefore an experiment to test this was carried out. L929 cells in log phase were incubated in the presence of hydroxyurea for one hour to inhibit DNA synthesis. Then L- [methyl- $^3\text{H}$ ] -methionine was added and the cells incubated for a further 3 hours. Nuclei were prepared from these cells and digested using micrococcal nuclease. The ratio of  $^3\text{H}/^{14}\text{C}$  was determined at each digestion time (Fig. 33, page 115). This ratio increases, showing that delayed methylation occurs preferentially in core DNA.

#### IV.4.2. Effect of salt extraction

When nuclei from L929 cells are extracted with low salt concentrations (0.35 M) most of the DNA methylase activity is solubilized (Section I.2.2.1). A fraction of the methylase cannot be solubilized at a salt concentration of 2.0 M (Qureshi *et al.*, 1982; Qureshi, 1983). In Section I.2.2.1 it is mentioned that this latter activity may be responsible for the methylation which occurs immediately after DNA synthesis, and the soluble methylase may be responsible for the delayed methylation. This experiment is designed to test this idea.

L929 cells were grown in the presence of hydroxyurea for one hour prior to harvesting. Another batch of cells were not so treated. In this way both delayed methylation and total methylation can be examined in vitro. The result is



given in Figure 34 (page 117). As can be seen the amount of methylation occurring in both batches of nuclei is the same. This shows that all the methylation that occurs in isolated nuclei is delayed methylation. The nascent DNA (less than 2 minutes old) is presumably methylated prior to nuclear preparation. Extraction of these nuclei with increasing salt concentrations decreases the methylation achieved by only 20%, the change is the same in both batches of nuclei. This shows that at least 80% of the delayed methylation is the product of the tightly bound DNA methylase activity described by Qureshi (1983). As this activity can only be solubilized by extensive treatment with micrococcal nuclease, it is assumed that this activity is associated with the nuclear matrix (Qureshi, 1983). This experiment does not study the methylation of nascent DNA, though as this DNA has been shown to be associated with the nuclear matrix (Berezney & Coffey, 1975; Keller & Riley, 1976; Miller *et al.*, 1978) it is assumed that this methylation is also a product of the bound activity.

#### IV.4.3. Methylation occurs in the nuclear matrix

The nuclear matrix from L929 cells was isolated by treatment of nuclei with 2.0 M NaCl followed by digestion using micrococcal nuclease. The resistant DNA is matrix DNA (Berezney & Coffey, 1974; 1977; Mitchelson *et al.*, 1979).

About 10% of the total DNA remains in this fraction. When the nuclei are methylated prior to the isolation of matrix DNA, the methyl groups are found preferentially associated with this DNA (Table 20, page 118). If the nuclei are extracted with 0.2 M NaCl to extract the soluble methylase prior to methylation, very little difference is observed. Therefore the methylation occurring in isolated nuclei is in the DNA associated with the nuclear matrix and is a product of the bound DNA methylase (Table 20, page 118). Some methylation does occur in non-matrix DNA and this may be a product of the soluble DNA methylase.

#### IV.4.4. Studies using 5-azadeoxycytidine

When cells are treated with 5-azadeoxycytidine, inhibition of DNA methylation in vivo occurs (Jones & Taylor, 1980). This inhibition is thought to involve irreversible binding of the methylase to the 5-azacytosine residues in DNA (Adams et al., 1982; Creusot et al., 1982; Taylor & Jones, 1982; Santi et al., 1983). Thus it was thought that the use of this methylation inhibitor might be useful in the study of bound DNA methylase, and Qureshi (1983) has shown that when cells are grown in the presence of 5-azadeoxycytidine, the amount of soluble methylase activity decreases, whereas the amount of bound activity increases. The conclusion from this experiment is that the bound enzyme is inhibited due to binding to the 5-azacytosine residues in DNA, and so more soluble methylase becomes bound in the matrix to compensate.

Therefore this system has been used to study the soluble and bound activities.

When nuclei from cells cultured in the presence of 5-azadeoxycytidine (Section II.2.1.3) are methylated in vitro using the endogenous DNA methylase, a slight increase in methylation is observed when compared to nuclei from control cells (Table 21, page 119). This increase is not five-fold as might be expected from the observation that five times as much methylase activity is now bound in the matrix (Qureshi, 1983). This may be due to the decrease in soluble activity. Moreover, no decrease is observed, therefore the soluble enzyme is not responsible for the methylation that is occurring, as the soluble activity is decreased more than five-fold under these conditions (Qureshi, 1983).

The control experiment is the methylation of these nuclei using ascites DNA methylase. No increase in methylation is seen in nuclei from control cells, whereas a five-fold increase is seen with nuclei from treated cells. This shows that DNA methylation has indeed been inhibited by culturing the cells in the presence of 5-azadeoxycytidine. These extra methylation sites should be hemimethylation sites due to there being less than one generation since the cells were released from stationary phase.

The methylation which does occur in nuclei from treated cells using the endogenous methylase is also preferentially in the core DNA (Fig. 35, page 120) and shows that the methylation is the same as the total methylation in control cells.

#### IV.5. Conclusions and further work

The conclusions from Section IV.3.4 are as follows:

One, when nuclei from stationary phase L929 cells are methylated in vitro using the added ascites enzyme there is no preference for either transcribing or non-transcribing chromatin. This methylation is de novo due to there being no hemimethylation sites in this chromatin. This shows that there is probably no specific inhibitors of methylation restricted to specific regions in chromatin. This could be tested by examining the effects of HMG proteins on the methylation in this system.

Two, using the endogenous methylase in nuclei from log phase cells, the methylation is preferentially in DNA from transcribing regions of chromatin. This may be due to the presence of the transcriptional apparatus inhibiting methylation. <sup>in vivo,</sup> therefore there will be more hemimethylated sites in this chromatin compared to non-transcribing chromatin. This could be tested by inhibiting transcription in vivo for several hours and then methylating in vitro. The methylation should no longer be preferential. Alternatively the DNA extracted into the transcribing fraction may contain undermethylated nascent DNA, and this should be examined in more detail.

Three, an exogenous (ascites) DNA methylase preferentially methylates transcribing regions of chromatin in mosquito nuclei. As this methylation is de novo due to the low endogenous methylation, the result suggests that this chromatin is in a more open structure and that no specific inhibitors of methylation are present, at least in insects. Again the effects of non-histone proteins should be examined. Unfortunately this same experiment was not done using nuclei

from log phase L929 cells.

Other experiments relevant to methylation of nuclei are described in Section IV.2.2.

The methylation which occurs in isolated nuclei is delayed methylation and occurs in the nuclear matrix (Section IV.4). Both nascent DNA and transcriptionally active DNA are in the nuclear matrix. As methylation follows soon after DNA synthesis (except delayed methylation) the methylase is also thought to be in the matrix. Experiments have shown that some methylase is present in the matrix and this is responsible for the delayed methylation and presumably the rapid methylation. The function of the soluble methylase is obscure, but may be responsible for the limited amount of de novo methylation which is thought to occur in vivo. No evidence for this is presented, but de novo methylation can be obtained using a purified methylase. This methylase may have the restrictions inhibiting de novo methylation removed due to being in a different environment than the soluble methylase in the nuclei, or this may simply be due to using an excess (40-fold) of purified methylase. This excess may be sufficient to overcome any inhibitions barring de novo methylation of chromatin in nuclei.

Experiments to investigate the functions of the soluble methylase should be considered.

REFERENCES

- Aaronson, P.R. & Blobel, G. (1975) Proc. Natl. Acad. Sci. USA. 72, 1007-1011.
- Adams, R.L.P. (1971) Biochim. Biophys. Acta. 254, 205-212.
- Adams, R.L.P. (1973) Nature New Biol. 244, 27-29.
- Adams, R.L.P. (1974) Biochim. Biophys. Acta. 335, 365-373.
- Adams, R.L.P., Burdon, R.H. & Fulton, J. (1983a) Biochem. Biophys. Res. Comm. 113, 695-702.
- Adams, R.L.P., Burdon, R.H., McKinnon, K. & Rinaldi, A. (1983b) FEBS.Lett. 163, 194-198.
- Adams, R.L.P., Davis, T., Fulton, J., Kirk, D., Qureshi, M.A. & Burdon, R.H. (1984) Curr. Top. Microbiol. Immunol. 108, 143-156.
- Adams, R.L.P. & Eason, R. (1984). Nucleic Acid.Res. (in Press).
- Adams, R.L.P., Fulton, J. & Kirk, D. (1982) Biochim. Biophys. Acta. 697, 286-294.
- Adams, R.L.P. & Hogarth, C. (1973) Biochim. Biophys. Acta. 331, 214-220.
- Adams, R.L.P., McKay, E.L., Craig, L.M. & Burdon, R.H. (1979) Biochim. Biophys. Acta. 563, 72-81.
- Adams, R.L.P., McKay, E.L., Douglas, J.T. & Burdon, R.H. (1977) Nucleic Acid.Res. 4, 3097-3108.
- Adolph, K.W. (1980) J. Cell Sci. 42, 291-304.
- Adolph, K.W., Cheng, S.M. & Laemmli, U.K. (1977a) Cell 12, 805-816.
- Adolph, K.W., Cheng, S.M., Paulson, J.R. & Laemmli, U.K. (1977b) Proc. Natl. Acad. Sci. USA. 74, 4937-4941.
- Aelen, M.A., Opstellen, J.G. & Wanka, F. (1983) Nucleic Acid. Res. 11, 1181-1195.
- Albright, S.C., Nelson, P.P. & Garrard, W.T. (1979) J. Biol. Chem. 254, 1065-1073.
- Andrews, G.K., Dziadek, M. & Tamaoki, T. (1982) J. Biol. Chem. 257, 5148-5153.
- Arber, W. & Linn, S. (1969) Ann. Rev. Biochem. 38, 467-500.
- Axel, R. (1976) Prog. Nucleic Acid.Res. Mol. Biol. 19, 335-371. Academic Press. NY.

- Axel, R., Cedar, H. & Felsenfeld, G. (1975) *Biochem.* 14, 2489-2495.
- Bachi, B., Reiser, J. & Pirrotta, V. (1979) *J. Mol. Biol.* 128, 143-163.
- Bassuk, J.A. & Mayfield, J.E. (1982) *Biochem.* 21, 1024-1027.
- Behe, M. & Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. USA.* 78, 1619-1623.
- Bently, G.A., Finch, J.T. & Lewit-Bentley, A. (1981) *J. Mol. Biol.* 145, 771-784.
- Berezney, R. (1980) *J. Cell Biol.* 85, 641-650.
- Berezney, R. & Coffey, D.S. (1974) *Biochem. Biophys. Res. Comm.* 60, 1410-1417.
- Berezney, R. & Coffey, D.S. (1975) *Science* 189, 291-293.
- Berezney, R. & Coffey, D.S. (1977) *J. Cell Biol.* 73, 616-637.
- Bird, A.P. (1978) *J. Mol. Biol.* 118, 49-60.
- Bird, A.P. (1980) *Nucleic Acid. Res.* 8, 1499-1504.
- Bird, A.P. (1984) *Curr. Top. Microbiol. Immunol.* 108, 129-142.
- Bird, A.P. & Taggart, M.H. (1980) *Nucleic Acid. Res.* 8, 1485-1497.
- Bird, A.P., Taggart, M.H. & MacLeod, D. (1981) *Cell* 26, 381-390.
- Bird, A.P., Taggart, M.H. & Smith, B.A. (1979) *Cell* 17, 889-901.
- Bloch, S. & Cedar, H. (1976) *Nucleic Acid. Res.* 6, 1507-1519.
- Bloom, K.S. & Anderson, J.N. (1978) *Cell* 15, 141-150.
- Boehm, T.L.J. & Drahovsky, D. (1980) *Z. Naturforsch.* 34c, 611-620.
- Borchsenius, S., Bonven, B., Leer, J.C. & Westergaard, O. (1981) *Eur. J. Biochem.* 117, 245-250.
- Borkhardt, B. & Nielsen, O.F. (1981) *Chromosoma* 84, 131-143.
- Bower, D.J., Errington, L.H., Cooper, D.N., Morris, S. & Clayton, R.M. (1983) *Nucleic Acid. Res.* 11, 2513-2527.
- Bradford, M.M. (1976) *Analyt. Biochem.* 72, 248-254.
- Brahic, M. & Fraser, M.J. (1971) *Biochim. Biophys. Acta.* 240, 23-36.
- Browne, M.J. & Burdon, R.H. (1977) *Nucleic Acid. Res.* 4, 1025-1037.
- Browne, M.J., Cato, A.C.B. & Burdon, R.H. (1979) *FEBS. Lett.* 91, 69-73.

- Browne, M.J., Turnbull, J.F., McKay, E.L., Adams, R.L.P. & Burdon, R.H. (1977) *Nucleic Acid. Res.* 4, 1039-1045.
- Bryan, P.N., Hofstetter, H. & Birnstiel, M.L. (1981) *Cell* 27, 459-466.
- Bryan, P.N., Wright, E.B. & Olins, D.E. (1979) *Nucleic Acid. Res.* 6, 1449-1465.
- Burdon, R.H. & Adams, R.L.P. (1969) *Biochim. Biophys. Acta.* 174, 322-329.
- Burton, K. (1956) *Biochem. J.* 62, 315-323.
- Burton, W.G., Grabowy, C.T. & Sager, R. (1979) *Proc. Natl. Acad. Sci. USA.* 76, 1390-1394.
- Burton, W.G., Roberts, R.J., Myers, P.A. & Sager, R. (1977) *Proc. Natl. Acad. Sci. USA.* 74, 2687-2691.
- Busby, D.W.G., House, W. & McDonald, J.R. (1964) *Virological Techniques*. p.218. Churchill, London.
- Camerini-Otero, R.D., Sollner-Webb, B. & Felsenfeld, G. (1976) *Cell* 8, 333-347.
- Cartwright, I.L., Hertzberg, R.P., Dervan, P.B. & Elgin, S.C.R. (1983) *Proc. Natl. Acad. Sci. USA.* 80, 3213-3217.
- Chalkley, R. (1975) *Biochem. Biophys. Res. Comm.* 64, 587-594.
- Chapman, G.E., Aviles, F.J., Crane-Robinson, C. & Bradbury, E.M. (1978) *Eur. J. Biochem.* 90, 287-296.
- Christman, J.K., Price, P., Pedrinan, L. & Acs, G. (1977) *Eur. J. Biochem.* 81, 53-61.
- Christy, B. & Scangos, G. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 6299-6303.
- Cockell, M., Rhodes, D. & Klug, A. (1983) *J. Mol. Biol.* 170, 423-446.
- Cohen, J.C. (1980) *Cell* 19, 653-662.
- Cole, R.D., Lawson, G.M. & Hsiang, M.W. (1977) *Cold Spring Harb. Symp. Quant. Biol.* 42, 253-263.
- Colgan, V., Elbrecht, A., Goldman, P., Lazier, C.B. & Deeley, R. (1982) *J. Biol. Chem.* 257, 14453-14460.
- Comings, D.E. (1972) *Exp. Cell Res.* 74 383-390.
- Compere, S. & Palmiter, R.D. (1981) *Cell* 25, 233-240.

- Compton, J.L., Bellard, M. & Chambon, P. (1976) Proc. Natl. Acad. Sci. USA. 73, 4382-4386.
- Coulondre, C., Millar, J.H., Farabough, J.H. & Gilbert, W. (1978) Nature 274, 775-780.
- Creusot, F., Acs, G. & Christman, J.K. (1982) J. Biol. Chem. 257, 2041-2048.
- Creusot, F. & Christman, K. (1981) Nucleic Acid. Res. 9, 5359-5381.
- Crick, F.H.C. & Klug, A. (1975) Nature 255, 530-533.
- Cummings, D.J., Tait, A. & Goddard, J.M. (1974) Biochim. Biophys. Acta. 374, 1-11.
- Davie, J.R. & Saunders, C.A. (1981) J. Biol. Chem. 256, 12574-12580.
- Dawid, I.B. (1974) Science 184, 80-81.
- Dawid, I.B., Brown, D.D. & Reeder, R.H. (1970) J. Mol. Biol. 51, 341-360.
- Deobagkar, D.N., Muralidharan, K., Devare, S.G., Kalghatgi, K.K. & Chandra, H.S. (1982) J. Biosci. 4, 513-526.
- DePamphilis, M.L. & Wassarman, P.M. (1980) Ann. Rev. Biochem. 49, 627-666.
- Desrosiers, R.C., Mulder, C. & Fleckenstein, B. (1979) Proc. Natl. Acad. Sci. USA. 76, 3839-3843.
- Dingwall, C., Lomonosoff, G.P. & Laskey, R.A. (1981) Nucleic Acid. Res. 9, 2659-2673.
- Doskocil, J. & Šormova, Z. (1965) Biochim. Biophys. Acta. 95, 513-515.
- Drahovsky, D., Boehm, T.L.G. & Kreis, W. (1979) Biochim. Biophys. Acta. 563, 28-35.
- Drahovsky, D. & Morris, N.R. (1971) J. Mol. Biol. 57, 475-489.
- Dunn, D.B. & Smith, J.D. (1955) Nature 175, 336-337.
- Earle, W.R. (1943) J. Natl. Cancer Inst. 4, 165-212.
- Earnshaw, W.C. & Laemmli, U.K. (1983) J. Cell Biol. 96, 84-93.
- Egan, P.A. & Levy-Wilson, B. (1981) Biochem. 20, 3695-3702.
- Ehrlich, M., Ehrlich, K. & Mayo, J.A. (1975) Biochim. Biophys. Acta. 395, 109-119.
- Dickerson, R.E. & Drew, H.R. (1981) J. Mol. Biol. 149, 761-786.



- Ehrlich, M., Gama-Sosa, M.A., Huang, L.-H, Midgett, R.M.,  
Kuo, K.C., McCune, R.A. & Gehrke, C. (1982) *Nucleic  
Acid. Res.* 10, 2709-2721.
- Elgin, S.C.R. (1981) *Cell* 27, 413-415.
- Engel, J.D. & von Hippel, P.H. (1978) *J. Biol. Chem.* 253,  
927-934.
- Eskin, B. & Linn, S. (1972a) *J. Biol. Chem.* 247, 6183-6191.
- Eskin, B. & Linn, S. (1972b) *J. Biol. Chem.* 247, 6192-6196.
- Evans, H.H., Evans, T.E. & Littman, S. (1973) *J. Mol. Biol.*  
74, 563-572.
- Faber, A.J., Cook, H. & Hancock, R. (1981) *Eur. J. Biochem.*  
120, 357-361.
- Feinberg, A.P. & Vogelstein, B. (1983a) *Nature* 301, 89-91.
- Feinberg, A.P. & Vogelstein, B. (1983b) *Biochem. Biophys.*  
*Res. Comm.* 111, 47-54.
- Finch, J.T., Brown, R.S., Rhodes, D., Richmond, T., Rushton,  
B., Lutter, L.G. & Klug, A. (1981) *J. Mol. Biol.*  
145, 757-769.
- Finch, J.T. & Klug, A. (1976) *Proc. Natl. Acad. Sci. USA.*  
73, 1897-1901.
- Fisher, P.A., Berrios, M. & Blobel, G. (1982) *J. Cell Biol.*  
92, 674-686.
- Fittler, F. & Zachau, H.G. (1979) *Nucleic Acid. Res.* 7, 1-13.
- Foe, V.E., Wilkinson, L.E. & Laird, C.D. (1976) *Cell* 2,  
131-146.
- Fradin, A., Manley, J.L. & Prives, C.L. (1982) *Proc. Natl.*  
*Acad. Sci. USA.* 79, 5142-5146.
- Frado, L.-L., Mura, C.V., Stollar, B.D. & Woodcock, C.L.F.  
(1983) *J. Biol. Chem.* 258, 11984-11990.
- Franco, L., Mantero, F. & Rodriguez-Molina, J.S. (1977)  
*FEBS. Lett.* 78, 317-320.
- Friedman, J. & Razin, A. (1976) *Nucleic Acid. Res.* 2, 2665-  
2675.
- Gabrielli, F., Hancock, R. & Faber, A.J. (1981) *Eur. J.*  
*Biochem.* 120, 363-369.
- Garel, A. & Axel, R. (1976) *Proc. Natl. Acad. Sci. USA.* 73,  
3966-3970.

- Gattoni, S., Kirschmeier, P., Weinstein, B., Escobedo, J. & Dina, D. (1982) *Mol. Cell. Biol.* 2, 42-51.
- Geier, G.E. & Modrich, P. (1979) *J. Biol. Chem.* 254, 1408-1413.
- Geiser, M., Mattaj, I.W., Wilks, A.F., Seldran, M. & Jost, J.-P. (1983) *J. Biol. Chem.* 258, 9024-9030.
- Gerace, L. & Blobel, G. (1980) *Cell* 19, 277-287.
- Geraci, D., Eremenko, T., Cocchiara, A., Scarano, E. & Volpe, P. (1974) *Biochem. Biophys. Res. Comm.* 57, 353-358.
- Gerber-Huber, S., May, F.E.B., Westley, B.R., Felber, B.K., Hosbach, H.A., Andres, A.-C. & Ryffel, G.U. (1983) *Cell* 33, 43-51.
- Gill, J.E., Mazrimas, J.A. & Bishop, C.C. (1974) *Biochim. Biophys. Acta.* 335, 330-348.
- Gjerset, R.A. & Martin, Jr., D.W. (1982) *J. Biol. Chem.* 257, 8581-8583.
- Glickman, B.W. & Radman, M. (1980) *Proc. Natl. Acad. Sci. USA.* 77, 1063-1067.
- Goldsbrough, P.B., Ellis, T.H.N. & Lomonossoff, G.P. (1982) *Nucleic Acid. Res.* 10, 4501-4513.
- Goodwin, G.H., Nicolas, R.H. & Johns, E.W. (1975) *Biochim. Biophys. Acta.* 405, 280-291.
- Gorovsky, M.A., Hattman, S. & Pleger, G.L. (1973) *J. Cell Biol.* 56, 697-701.
- Gottesfeld, J.M. & Bloomer, L.S. (1980) *Cell* 21, 751-760.
- Gottesfeld, J.M. & Butler, P.J.G. (1977) *Nucleic Acid. Res.* 9, 3155-3173.
- Gottesfeld, J.M. & Partington, G.A. (1977) *Cell* 12, 953-962.
- Grainger, R.M. & Ogle, R.C. (1978) *Chromosoma* 65, 115-126.
- Graves, J.A.M. (1982) *Exp. Cell Res.* 141, 99-105.
- Groot, G.S.P. & Kroon, A.M. (1979) *Biochim. Biophys. Acta.* 564, 355-357.
- Groudine, M., Eisenman, R. & Weintraub, H. (1981) *Nature* 292, 311-317.
- Gruenbaum, Y., Cedar, H. & Razin, A. (1982) *Nature* 295, 620-622.
- Gruenbaum, Y., Naveh-Many, T., Cedar, H. & Razin, A. (1981) *Nature* 292, 860-862.

- Guntaka, R.V., Rao, P.Y., Mitsians, S.A. & Katz, R. (1980)  
J. Virol. 34, 569-572.
- Gunthert, U., Schweiger, M., Stupp, M. & Doerfler, W. (1976)  
Proc. Natl. Acad. Sci. USA. 73, 3923-3927.
- Guseinov, V.A., Mazin, A.C., Vanyushin, B.F. & Belozerskii,  
A.N. (1972) Biochemistry (Biokimiya) 37, 312-318.
- Harrison, J.J., Anisowicz, A., Gadi, I.K., Raffeld, M. &  
Sager, R. (1983) Proc. Natl. Acad. Sci. USA. 80,  
6606-6610.
- Hattman, S., Brooks, J.E. & Masurekar, M. (1978a) J. Mol.  
Biol. 126, 367-380.
- Hattman, S., Kenny, C., Berger, L. & Pratt, P. (1978b)  
J. Bacteriol. 135, 1156-1157.
- Hay, N., Skolnik-David, H. & Aloni, Y. (1982) Cell 29, 183-193.
- Hayward, G. (1982) Virology 49, 342-344.
- Hentzen, P.C., Rho, J.H. & Bekhov, I. (1984) Proc. Natl. Acad.  
Sci. USA. 81, 304-307.
- Herman, T.M., DePamphilis, M.L. & Wassarman, P.M. (1981)  
Biochem. 20, 621-630.
- Hertzberg, R.P. & Dervañ, P.B. (1982) J. Amer. Chem. Soc.  
104, 313-315.
- Hewish, D.R. & Burgoyne, L.A. (1973) Biochem. Biophys. Res.  
Comm. 52, 504-510.
- Hill, R.J. & Stollar, B.D. (1983) Nature 305, 338-340.
- Hjelle, B.L., Phillips, III, S.A. & Seeburg, P.M. (1982)  
Nucleic Acid. Res. 10, 3459-3473.
- Hogan, M., LeGrange, J. & Austin, B. (1983) Nature 304,  
752-754.
- Holliday, R. & Pugh, J.E. (1975) Science 187, 226-232.
- Hörz, W. & Altenberger, W. (1981) Nucleic Acid. Res. 9,  
2643-2658.
- Hörz, W., Fittler, F. & Zachau, H.G. (1983) Nucleic Acid. Res.  
11, 4275-4285.
- Hotchkiss, R.D. (1948). J. Biol. Chem. 175, 315-332.
- Hotta, Y. & Hecht, N. (1971) Biochim. Biophys. Acta. 238,  
50-59.

- Howell, W.M. & Hsu, T.C. (1979) *Chromosoma* 73, 61-66.
- Huang, L.-H., Wang, R., Gama-Sosa, M.A., Shenoy, S. & Ehrlich, M. (1984) *Nature* 308, 293-295.
- Hunt, B.F. & Vogelstein, B. (1981) *Nucleic Acid.Res.* 2, 349-363.
- Hynes, N.E., Rahmsdorf, U., Kennedy, N., Fabiani, L., Michalides, R., Musse, R. & Groner, B. (1981) *Gene* 15, 307-317.
- Ikeda, K., Frazier, J. & Miles, H.T. (1970) *J. Mol. Biol.* 54, 59-84.
- Jackson, D.A., McCready, S.J. & Cook, P.R. (1981) *Nature* 292, 552-555.
- Janulaitis, A., Klimasaukas, S., Petrušyte, M. & Butkus, V. (1983) *FEBS.Lett.* 161, 131-134.
- Jeppeson, P.G.N. & Bankier, A.T. (1979) *Nucleic Acid.Res.* 7, 49-67.
- Johnson, E.M., Allfrey, V.G., Bradbury, E.M. & Matthews, H.R. (1978) *Proc. Natl. Acad. Sci. USA.* 75, 1116-1120.
- Johnson, T.B. & Coghill, R.D. (1925) *J. Amer. Chem. Soc.* 47, 2838-2844.
- Jones, P.A. & Taylor, S.M. (1980) *Cell* 20, 85-93.
- Jones, P.A., Taylor, S.M., Mohandas, T. & Shapiro, L.J. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 1215-1219.
- Jones, A.S. & Walker, R.T. (1963) *J. Gen. Microbiol.* 31, 187-194.
- Kahan, B. & De Mars, R. (1980) *Somatic Cell Genet.* 6, 309-323.
- Kahman, R. (1983) *Cold Spring Harb. Symp. Quant. Biol.* 47, 647-653.
- Kahman, R. (1984) *Curr. Top. Microbiol. Immunol.* 108, 29-48.
- Kalousek, F. & Morris, N.R. (1968) *J. Biol. Chem.* 243, 2440-2442.
- Kappler, J. (1970) *J. Cell.Physiol.* 75, 21-32.
- Kappler, J. (1971) *J. Cell.Physiol.* 78, 33-36.
- Kaput, J. & Sneider, T.W. (1979) *Nucleic Acid.Res.* 7, 2303-2322.
- Kaye, A.M. & Winocour, E. (1967) *J. Mol. Biol.* 24, 475-478.
- Keene, M.A. & Elgin, S.C.R. (1981) *Cell* 27, 57-64.

- Kecne, M.A. & Elgin, S.C.R. (1984) *Cell* 36, 121-129.
- Keller, J.M. & Riley, D.E. (1976) *Science* 193, 399-401.
- Kiryanov, G.I., Kirnos, M.D., Demidkina, N.P., Alexandarushkina, N.I. & Vanyushin, B.F. (1980) *FEBS.Lett.* 112, 225-228.
- Korba, B.E. & Hays, J.D. (1982a) *J. Mol. Biol.* 157, 213-235.
- Korba, B.E. & Hays, J.B. (1982b) *Cell* 28, 531-541.
- Kornberg, R.D. (1974) *Science* 184, 868-871.
- Kornberg, R.D. (1977) *Cold Spring Harb. Symp. Quant. Biol.* 42, 103-108.
- Kornberg, R.D. & Thomas, J.O. (1974) *Science* 184, 865-868.
- Kuchl, L., Lyness, T., Dixon, G.H. & Levy-Wilson, B. (1980) *J. Biol. Chem.* 255, 1090-1095.
- Kunkel, G.R. & Martinson, H.G. (1981) *Nucleic Acid.Res.* 9, 6869-6888.
- Kunnath, L. & Locker, J. (1982) *Biochim. Biophys. Acta.* 699, 264-271.
- Kuo, M.T. (1982) *J. Cell Biol.* 93, 278-284.
- Kuo, M.T., Sahasrabudde, C.G. & Saunders, M. (1976) *Proc. Natl. Acad. Sci. USA.* 73, 1572-1575.
- Kuo, T.-T., Huang, T.-C. & Teng, M.-H. (1968) *J. Mol. Biol.* 34, 373-375.
- Labhardt, P. & Koller, T. (1982) *Cell* 28, 279-292.
- Lacy, E. & Axel, R. (1975) *Proc. Natl. Acad. Sci. USA.* 72, 3978-3982.
- LaPeyre, J.N. & Becker, F.F. (1979) *Biochem. Biophys. Res. Comm.* 87, 698-705.
- Larsen, A. & Weintraub, H. (1982) *Cell* 29, 609-622.
- Lautenberger, J.A. & Linn, S. (1972) *J. Biol. Chem.* 247, 6167-6182.
- Lawrence, J. & Goeltz, P. (1981) *Nucleic Acid.Res.* 9, 858-866.
- Lebkowski, J.S. & Laemmli, U.K. (1982) *J. Mol. Biol.* 156, 309-324.
- Levy-Wilson, B. & Dixon, G.H. (1979) *Proc. Natl. Acad. Sci. USA.* 76, 1682-1686.
- Lewis, C.D. & Laemmli, U.K. (1982) *Cell* 29, 171-181.
- Liu, L.F. & Wang, J.C. (1978) *Cell* 15, 979-984.

- Lohr, D.E., Kovacic, R.T. & Van Holde, K.E. (1977) *Biochem.* 16, 463-471.
- Louis, C., Schedl, P., Samal, B. & Worcel, A. (1980) *Cell* 22, 387-392.
- Low, M., Hay, J. & Keir, H.M. (1969) *J. Mol. Biol.* 46, 205-207.
- Lubit, B.W., Pham, T.D., Miller, O.J. & Erlanger, B.F. (1976) *Cell* 9, 503-509.
- Lubit, B.W., Schrek, R.R., Miller, O.J. & Erlanger, B.F. (1974) *Exp. Cell Res.* 89, 426-429.
- Lutter, L.C. (1979) *Nucleic Acid Res.* 6, 41-56.
- Lyon, M.F. (1961) *Nature* 190, 372-373.
- MacLeod, D. & Bird, A.P. (1983) *Nature* 306, 200-203.
- McGhee, J.D. & Felsenfeld, G. (1983) *Cell* 32, 1205-1215.
- McGhee, J.D. & Ginder, G.D. (1979) *Nature* 280, 419-420.
- McKeon, C., Ohkubo, H., Pastan, I. & De Crombrughe, B. (1982) *Cell* 29, 203-210.
- McKnight, S.L., Sultivan, N.L. & Miller, O.L. (1976) *Prog. Nucleic Acid Res. Mol. Biol.* 19, 313-334. Academic Press, NY.
- Mamelak, L. & Boyer, H.W. (1970) *J. Bacteriol.* 104, 57-62.
- Mandel, J.L. & Chambon, P. (1979) *Nucleic Acid Res.* 7, 2081-2103.
- Manes, C. & Menzel, P. (1981) *Nature* 293, 589-590.
- Maniatis, T., Jeffrey, A. & Kleid, D.G. (1975) *Proc. Natl. Acad. Sci. USA.* 72, 1184-1188.
- Mann, M.B. & Smith, H.O. (1978) *Nucleic Acid Res.* 4, 4211-4221.
- Margison, G.P. & Pegg, A.E. (1981) *Proc. Natl. Acad. Sci. USA.* 78, 861-865.
- Marini, J., Levene, S., Crothers, D.M. & Englund, P.T. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 7664-7668.
- Marinus, M.G. & Konrad, B. (1976) *Mol. Gen. Genet.* 149, 273-277.
- Marinus, M.G. & Morris, N.R. (1974) *J. Mol. Biol.* 85, 309-322.
- Marinus, M.G. & Morris, N.R. (1975) *Mutat. Res.* 28, 15-26.

- Marsden, M.P.F. & Laemmli, U.K. (1979) *Cell* 17, 849-858.
- Mathis, D.J. & Gorovsky, M.A. (1976) *Biochem.* 15, 750-755.
- Meijlink, F.C.P.W., Philipson, J.N.J., Gruber, M. & Ab, G.  
(1983) *Nucleic Acid.Res.* 11, 1361-1373.
- Mengeritsky, G. & Trifonov, E.N. (1983) *Nucleic Acid.Res.*  
11, 3833-3851.
- Mermod, J.-J., Bourgeois, S., Defer, N. & Crepin, M. (1983)  
*Proc. Natl. Acad. Sci. USA.* 80, 110-114.
- Miki, B.L.A., Gurd, J.W. & Brown, I.R. (1980) *Can. J. Biochem.*  
58, 1261-1269.
- Miller, T.E., Huang, C. & Pogo, A.O. (1978) *J. Cell Biol.* 76,  
692-704.
- Mitchelson, K.R., Bekers, A.G.M. & Wanka, F. (1979) *J. Cell  
Sci.* 39, 247-256.
- Mitsuhashi, J. & Maramorosch, K. (1964) *Contrib. Boyce  
Thompson Inst.* 22, 435-460.
- Modrich, P. (1982) *CRC. Crit. Rev. Biochem.* 13, 287-323.
- Mohandas, T., Sparkes, R.S., & Shapiro, L.J. (1981) *Science*  
211, 393-396.
- Molitor, H., Drahovsky, D. & Walker, A. (1976) *Biochim.  
Biophys. Acta.* 432, 28-36.
- Morris, N.R. (1976a) *Cell* 8, 357-363.
- Morris, N.R. (1976b) *Cell* 9, 627-632.
- Musich, P.R., Maio, J.J. & Brown, F.L. (1977) *J. Mol. Biol.*  
117, 657-677.
- Nakhasi, H.L., Lynch, K.R., Dolan, K.P., Unterman, R.,  
Antakly, T. & Feigelson, P. (1982) *J. Biol. Chem.*  
257, 2726-2729.
- Nass, M.M.K. (1973) *J. Mol. Biol.* 80, 155-175.
- Naveh-Many, T. & Cedar, H. (1981) *Proc. Natl. Acad. Sci. USA.*  
78, 4246-4250.
- Nelson, P.P., Albright, S.C. & Garrard, W.T. (1979a) *J. Biol.  
Chem.* 254, 9194-9199.
- Nelson, D., Perry, M.E. & Chalkley, R. (1979b) *Nucleic Acid.  
Res.* 6, 561-574.
- Nickol, J. & Martin, R.G. (1983) *Proc. Natl. Acad. Sci. USA.*  
80, 4669-4673.

- Nikolskaya, I.I., Lapatina, N.G. & Debov, S.S. (1976)  
Biochim. Biophys. Acta. 435, 206-210.
- Niwa, O. & Sugahara, T. (1981) Proc. Natl. Acad. Sci. USA.  
78, 6290-6294.
- Noll, M. (1974a) Nature 251, 249-251.
- Noll, M. (1974b) Nucleic Acid. Res. 1, 1573-1578.
- Noll, M. (1976) Cell 8, 349-355.
- Noll, M. & Kornberg, R.D. (1977) J. Mol. Biol. 109, 393-404.
- Nordheim, A., Pardue, M.L., Lafer, E.M., Müller, A., Stollar, B.D. & Rich, A. (1981) Nature 294, 417-422.
- Oda, K. & Marmur, J. (1966) Biochem. 5, 761-773.
- Ohno, S., Kaplan, W.D. & Kinoshita, R. (1959) Exp. Cell Res. 18, 415-418.
- Olins, A.L. & Olins, D.E. (1974) Science 183, 330-332.
- Olins, D.E. & Wright, E.B. (1973) J. Cell Biol. 59, 304-309.
- Oudet, P., Gross-Bellard, M. & Chambon, P. (1975) Cell 4, 281-300.
- Palen, T.E. & Cech, T.R. (1984) Cell 36, 933-942.
- Panet, A. & Cedar, H. (1977) Cell 11, 933-940.
- Pardon, J.F., Worcester, D.L., Wooley, J.C., Cotter, R.I., Lilley, D.M.J. & Richards, B.M. (1977) Nucleic Acid. Res. 4, 3199-3214.
- Paulson, J.R. & Laemmli, U.K. (1977) Cell 12, 817-828.
- Peacock, W.J., Miklos, G.L.G. & Goodchild, D.J. (1975) Genetics 79, 613-634.
- Pegg, A.E. & Hui, G. (1978) Biochem. J. 173, 739-748.
- Pfeiffer, G.P., Grünwald, S., Boehm, T.L.J. & Drahovsky, D. (1983) Biochim. Biophys. Acta. 740, 323-330.
- Piper, P.W., Celis, J., Kattoff, K., Leer, J.C., Nielson, O.F. & Westergaard, O. (1976) Nucleic Acid. Res. 3, 493-505.
- Plasterk, R.H.A., Vrieling, H., Van de Putte, P. (1983) Nature 301, 344-347.
- Pollock, Jr., J.M., Swihart, M. & Taylor, J.H. (1978) Nucleic Acid. Res. 5, 4855-4863.
- Pruitt, S.C. & Grainger, R.M. (1981) Cell 23, 711-720.
- Pukkila, P.J., Peterson, J., Herman, G., Modrich, P. & Meselson, M. (1983) Genetics 104, 571-582.



- Qureshi, M.A. (1983) Ph.D. Thesis, University of Glasgow.
- Qureshi, M.A.; Adams, R.L.P. & Burdon, R.H. (1982) *Biochem. Soc. Trans.* 10, 455-456.
- Radman, M. (1981) *Biochem. Soc. Trans.* 9, 78p.
- Radman, M., Van den Elsen, P. & Glickman, B. (1978) *Mol. Gen. Genet.* 163, 307-312.
- Rac, P.M.M. & Steele, R.E. (1979) *Nucleic Acid. Res.* 6, 2987-2995.
- Rahe, B., Ericksson, R.P. & Quinto, M. (1983) *Nucleic Acid. Res.* 11, 7947-7960.
- Rall, S.C., Okinaka, R.T. & Striniste, G.F. (1977) *Biochem.* 16, 4940-4944.
- Ravetch, J.V., Horiuchi, K. & Zinder, N.D. (1978) *Proc. Natl. Acad. Sci. USA.* 75, 2266-2270.
- Razin, A. (1973) *Proc. Natl. Acad. Sci. USA.* 70, 3773-3775.
- Razin, A. & Cedar, H. (1977) *Proc. Natl. Acad. Sci. USA.* 74, 2725-2728.
- Razin, A. & Friedman, J. (1981) *Prog. Nucleic Acid. Res. Mol. Biol.* 25, 33-52. Academic Press. NY.
- Razin, A., Goren, D. & Friedman, J. (1975) *Nucleic Acid. Res.* 2, 1967-1974.
- Razin, A., Matieva, V.L. & Georgiev, G.P. (1979) *Nucleic Acid. Res.* 7, 1713-1735.
- Razin, A. & Razin, S. (1980) *Nucleic Acid. Res.* 8, 1383-1390.
- Razin, A., Sedat, J.W. & Sinsheimer, R.L. (1970) *J. Mol. Biol.* 53 251-259.
- Razin, A., Urieli, S., Pollack, Y., Gruenbaum, Y. & Glaser, G. (1980) *Nucleic Acid. Res.* 8, 1783-1792.
- Reiser, J. & Yuan, R. (1977) *J. Biol. Chem.* 252, 451-456.
- Rhodes, D. (1979) *Nucleic Acid. Res.* 6, 1805-1816.
- Richards, B.M., Pardon, J.F., Lilley, D.M.J., Cotter, R.I., Wooley, J.C. & Worcester, D.L. (1977) *Cell Biol. Int. Rep.* 1, 107-116.
- Riggs, A.D. (1975) *Cytogenet. Cell Genet.* 14, 9-25.
- Rill, R. & Van Holde, K.E. (1973) *J. Biol. Chem.* 248, 1080-1083.

- Roberts, R.J. (1976) CRC. Crit. Rev. Biochem. 4, 123-164.
- Roberts, R.J. (1980) Nucleic Acid. Res. 8, r63-r80.
- Robinson, S.I., Nelkin, B.D. & Vogelstein, B. (1982) Cell 28, 99-106.
- Robinson, S.I., Small, D., Idzerda, R., McKnight, G.S. & Vogelstein, B. (1983) Nucleic Acid. Res. 11, 5113-5130.
- Rogers, J. & Wall, W. (1981) Proc. Natl. Acad. Sci. USA. 78, 7497-7501.
- Romanov, G. & Vanyushin, B.F. (1981) Biochim. Biophys. Acta. 653, 204-218.
- Roy, P.H. & Weissbach, A. (1975) Nucleic Acid. Res. 2, 1669-1684.
- Rubery, E.D. & Newton, A.A. (1973) Biochim. Biophys. Acta. 324, 24-36.
- Ruiz-Carillo, A., Wamgh, L.J. & Allfrey, V.G. (1975) Science 190, 117-128.
- Sager, R. (1954) Proc. Natl. Acad. Sci. USA. 40, 356-363.
- Sager, R., Grabowy, C. & Sano, H. (1981) Cell 24, 41-47.
- Sager, R. & Kitchin, R. (1975) Science 189, 426-433.
- Sager, R. & Lane, D. (1972) Proc. Natl. Acad. Sci. USA. 69 2410-2413.
- Sager, R. & Ramanis, Z. (1973) Theor. Appl. Genet. 43, 101-108.
- Sager, R., Sano, H. & Grabowy, C.T. (1984) Curr. Top. Microbiol. Immunol. 108, 157-172.
- Samal, B., Worcel, A., Louis, C. & Schedl, P. (1981) Cell 23. 401-409.
- Sanders, M.M. (1978) J. Cell Biol. 79, 97-109.
- Sanford, K.K., Earle, W.R. & Likely, G.D. (1948) J. Natl. Cancer Inst. 2, 229-246.
- Sanford, J., Forrester, L. & Chapman, V. (1984) Nucleic Acid. Res. 12, 2823-2836.
- Sano, H., Grabowy, C.T. & Sager, R. (1981) Proc. Natl. Acad. Sci. USA. 78, 3118-3122.
- Sano, H., Noguchi, H. & Sager, R. (1983) Eur. J. Biochem. 135, 181-185.
- Santi, D.V., Garret, C.E. & Barr, P.J. (1983) Cell 33, 9-10.
- Scarano, E. (1969) Ann. Embryol. Morphogen. 1 (suppl.), 51-61.

- Scheer, U. (1978) *Cell* 13, 535-549.
- Schlagman, S., Hattman, S., May, M.S. & Berger, L. (1976) *J. Bacteriol.* 126, 990-996.
- Schneiderman, M.H. & Billen, D. (1973) *Biochim. Biophys. Acta.* 308, 352-360.
- Schnedl, W., Dev, V.G., Tantravahi, R., Miller, D.A., Erlanger, B.F. & Miller, O.J. (1975) *Chromosoma* 52, 59-66.
- Schreck, R.R., Erlanger, B.F. & Miller, O.J. (1977) *Exp. Cell Res.* 108, 403-411.
- Seyedin, S.M. & Kistler, W.S. (1979) *Biochem.* 18, 1371-1375.
- Shapiro, L.J. & Mohandas, T. (1982) *Cold Spring Harb. Symp. Quant. Biol.* 47, 631-638.
- Shatkin, A.J. (1969) *Fundamental Techniques in Virology* (Ed. Habel, K. & Salzman, N.P.) Academic Press. NY. p.496-501.
- Shen, C.K.J. & Maniatis, T. (1980) *Proc. Natl. Acad. Sci. USA.* 77, 6634-6638.
- Shmookler Reis, R., Timmis, J.N. & Ingle, J. (1981) *Biochem. J.* 195, 723-734.
- Simon, D., Grunert, F., Acken, U.V., Doring, H.P. & Kroger, H. (1978) *Nucleic Acid Res.* 5, 2153-2168.
- Simon, D., Stuhlman, H., Jahner, D., Wagner, H., Warner, E. & Jaenisch, R. (1983) *Nature* 304, 275-277.
- Simpson, R.T. (1978) *Biochem.* 17, 5524-5531.
- Simpson, R.T. & Kunzler, P. (1979) *Nucleic Acid Res.* 6, 1387-1415.
- Simpson, R.T. & Whitlock, Jr., J.P. (1976) *Cell* 9, 347-353.
- Singer, D.S. (1979) *J. Biol. Chem.* 254, 5506-5514.
- Singer, T., Robert-Ems, J., Luthardt, F.W. & Riggs, A.D. (1979) *Nucleic Acid Res.* 7, 2369-2385.
- Singh, K.R.P. (1967) *Curr. Sci.* 36, 506-508.
- Sinsheimer, R. (1955) *J. Biol. Chem.* 215, 579-583.
- Smerdon, M.J. & Isenberg, I. (1976) *Biochem.* 15, 4233-4241.
- Sneider, T.W. (1972) *J. Biol. Chem.* 247, 2872-2875.
- Sneider, T.W. (1980) *Nucleic Acid Res.* 8, 3829-3840.
- Sneider, T.W. & Potter, V.R. (1969) *J. Mol. Biol.* 42, 271-284.
- Sobell, H.M., Tsai, C.C., Gilbert, S.G., Jain, S.C. & Sakore, T.D. (1976) *Proc. Natl. Acad. Sci. USA.* 73, 3068-3072.

- Sollner-Webb, B., Camerini-Otero, R.D. & Felsenfeld, G.  
(1976) Cell 9, 179-193.
- Sollner-Webb, B. & Felsenfeld, G. (1975) Biochem. 14,  
2915-2920.
- Spadafora, C., Bellard, M., Compton, J.L. & Chambon, P.  
(1976) FEBS.Lett. 69, 281-285.
- Spadafora, C. & Crippa, M. (1984) Nucleic Acid.Res. 12,  
2691-2704.
- Spector, D., Varmus, H.E. & Bishop, J.M. (1978) Proc. Natl.  
Acad. Sci. USA. 75, 4102-4106.
- Spinelli, G., Ailanes, I., Anello, L., Ciaccio, M. & DiLiegro,  
I. (1982) Nucleic Acid.Res. 10, 7971-7991.
- Stehelin, D., Varmus, H.E., Bishop, J.M. & Vogt, P.K. (1976)  
Nature 260, 170-173.
- Stein, R., Sciaky-Gallili, N., Razin, A. & Cedar, H. (1983)  
Proc. Natl. Acad. Sci. USA. 80, 2422-2426.
- Stein, A. & Townsend, T. (1983) Nucleic Acid.Res. 11, 6803-6819.
- Strauss, F. & Prunell, A. (1980) Nucleic Acid.Res. 10,  
2275-2293.
- Sturm, K.S. & Taylor, J.H. (1981) Nucleic Acid.Res. 9, 4537-  
4546.
- Subramanian, K.N. (1982) Nucleic Acid.Res. 10, 3475-3485.
- Sutter, D. & Doerfler, W. (1980) Proc. Natl. Acad. Sci. USA.  
77, 253-256.
- Tantravahi, U., Guntaka, R.V., Erlanger, B.F. & Miller, O.J.  
(1981) Proc. Natl. Acad. Sci. USA. 78, 489-493.
- Taylor, S.M. & Jones, P.A. (1982) J. Mol. Biol. 162, 679-692.
- Teng, C.S., Andrews, G.K. & Teng, C.T. (1979) Biochem. J.  
181, 585-591.
- Tewari, K.K. & Wildman, S.G. (1966) Science 153, 1269-1271.
- Thoma, F. & Koller, T. (1977) Cell 12, 101-107.
- Thoma, F., Koller, T. & Klug, A. (1979) J. Cell Biol. 83,  
403-427.
- Thomas, J.O. & Thompson, R.J. (1977) Cell 10, 633-640.
- Torok, I. & Karch, F. (1980) Nucleic Acid.Res. 8, 3105-3123.
- Tosi, L., Granieri, A. & Scarano, E. (1972) Exp. Cell Res.  
72, 257-264.

- Toussaint, A. (1976) *Virology* 70, 17-27.
- Toussaint, A. (1977) *J. Virology* 23, 825-826.
- Trifonov, E.N. (1980) *Nucleic Acid.Res.* 8, 4041-4053.
- Turnbull, J.F. & Adams, R.L.P. (1976) *Nucleic Acid.Res.* 3, 677-695.
- Udvardy, A. & Schedl, P. (1983) *J. Mol. Biol.* 166, 159-181.
- Van der Ploeg, L.H.T. & Flavell, R.A. (1980) *Cell* 19, 947-958.
- Van Dyke, M.W. & Dervan, P.B. (1983a) *Biochem.* 22, 2373-2377.
- Van Dyke, M.W. & Dervan, P.B. (1983b) *Nucleic Acid.Res.* 11, 5555-5566.
- Van Lier, J.J.C., Smits, M.T. & Buck, H.M. (1983) *Eur. J. Biochem.* 55-62.
- Vanyushin, B.F. (1968) *Uspekhi Sovremenoi Biologii* 65, 163-185.
- Vanyushin, B.F., Belozersky, A.N., Kokurina, N.A. & Kadirova, D.X. (1968) *Nature* 218, 1066-1068.
- Vanyushin, B.F., Kokurina, N.A. & Belozersky, A.N. (1965) *Dokl. Akad. Nauk. S.S.R.* 161, 1451-1454.
- Vanyushin, B.F., Mazin, A.L., Vasilyev, V.K. & Belozersky, A.N. (1973a) *Biochim. Biophys. Acta.* 299, 397-403.
- Vanyushin, B.F., Nemirovsky, L.E., Klimenko, V.V., Vasilyev, V.K. & Belozersky, A.N. (1973b) *Gerontologia* 19, 138-152.
- Vanyushin, B.F., Tkachera, S.G. & Belozersky, A.N. (1970) *Nature* 225, 948-949.
- Vardimon, L., Kressman, A., Cedar, H., Maechler, M. & Doerfler, W. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 1073-1077.
- Varshavsky, A.J., Bakayev, V.V. & Georgiev, G.P. (1976) *Nucleic Acid.Res.* 3, 477-492.
- Venolia, L., Gartler, S.M., Wassman, E.R., Yen, P., Mohandas, T. & Shapiro, L.J. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 2352-2354.
- Viegas-Pequignot, E., Derbin, C., Malfroy, B., Taillandier, E., Leng, M. & Dutrillaux, B. (1983) *Proc. Natl. Acad. Sci. USA.* 80, 5890-5894.
- Vogelstein, B., Pardoll, D.M. & Coffey, D.S. (1980) *Cell* 22, 79-85.

- Von Acken, U., Simon, D., Gruncert, F., Doring, H.P. & Kroger, H. (1979) *Virology* 99, 152-157.
- Von Holt, G., Strickland, W.N., Brandt, W.F. & Strickland, M.S. (1979) *FEBS.Lett.* 100, 201-218.
- Waalwijk, C. & Flavell, R.A. (1978) *Nucleic Acid.Res.* 5, 4631-4641.
- Wachsmuth, E.D. & Jost, J.-P. (1976) *Biochim. Biophys. Acta.* 437, 454-461.
- Waechter, D.E. & Baserga, R. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 1106-1110.
- Wagner, I. & Capesius, I. (1981) *Biochim. Biophys. Acta.* 654, 52-56.
- Wang, R.Y.-H., Huang, L.-H. & Ehrlich, M. (1982) *Anal. Biochem.* 122, 89-93.
- Warron, W., Crathorn, A.R. & Shooter, K.V. (1979) *Biochim. Biophys. Acta.* 513, 82-88.
- Weintraub, H. (1983) *Cell* 32, 1191-1203.
- Weintraub, H., Worcel, A. & Alberts, B. (1976) *Cell* 9, 409-417.
- Weisbrod, S., Groudine, M. & Weintraub, H. (1980) *Cell* 19, 289-301.
- Weisbrod, S. & Weintraub, H. (1981) *Cell* 23, 391-400.
- Weischet, W.O., Glotov, B.O. & Zachau, H.G. (1983) *Nucleic Acid.Res.* 11, 3613-3630.
- Westmaas, G.C., Van der Maas, W.L. & Van de Putte, P. (1976) *Mol. Gen. Genet.* 145, 81-87.
- Whitlock, J.P. & Simpson, R.T. (1976) *Biochem.* 15, 3307-3313.
- Whittaker, P.A. & Hardman, N. (1980) *Biochem. J.* 191, 859-862.
- Whittaker, P.A., McLachlan, A. & Hardman, N. (1981) *Nucleic Acid.Res.* 9, 801-814.
- Wigler, M.H. (1981) *Cell* 24, 285-286.
- Wilks, A.F., Cozens, P.J., Mattaj, I.W. & Jost, J.-P. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 4252-4255.
- Wilks, A.F., Seldran, M. & Jost, J.-P. (1984) *Nucleic Acid. Res.* 12, 1163-1177.
- Willis, D.B. & Granoff, A. (1980) *Virology* 107, 250-257.
- Wilson, V.L. & Jones, P.A. (1983) *Cell* 32, 239-246.

- Wittig, B. & Wittig, S. (1979) *Cell* 18, 1173-1183.
- Woodcock, C.L.F., Frado, L.-L.Y., Hatch, C.L. & Ricciardiello, L. (1976) *Chromosoma* 58, 33-39.
- Wu, C. (1980) *Nature* 286, 854-860.
- Wu, H.-M. & Crothers, D.M. (1984) *Nature* 308, 509-513.
- Wunderlich, F. & Herlan, G. (1977) *J. Cell Biol.* 73, 271-278.
- Wyatt, G.R. (1951) *Biochem. J.* 48, 581-584.
- Wyatt, G.R. & Cohen, S.S. (1953) *Biochem. J.* 55, 774-782.
- Yagi, M. & Koshland, M.E. (1981) *Proc. Natl. Acad. Sci. USA.* 78, 4907-4911.
- Youssoufian, H., Hammer, S.H., Hirsch, M.S. & Mulder, C. (1982) *Proc. Natl. Acad. Sci. USA.* 79, 2207-2210.
- Yuan, R. (1981) *Ann. Rev. Biochem.* 50, 285-315.
- Zachau, H.G. & Igo-Kemenes, T. (1981) *Cell* 24, 597-598.
- Zhang, X.-Y., Fittler, F. & Hürz, W. (1983) *Nucleic Acid.Res.* 11, 4287-4306.
- Zhurkin, V.B. (1981) *Nucleic Acid.Res.* 9, 1963-1971.
- Zhurkin, V.B. (1983) *FEBS.Lett.* 158, 293-297.

INSTITUTE OF BIOCHEMISTRY  
GLASGOW UNIVERSITY