

THE STRUCTURE OF *SCHIZOSACCHAROMYCES POMBE* DNA IN
MOUSE CELLS

JOHN M^c MANUS

Ph.D.

UNIVERSITY OF EDINBURGH

1993



ABSTRACT.

Integrated *Schizosaccharomyces pombe* (*S. pombe*) transgenomes in murine cells display a cytological constriction at metaphase. That the metaphase constriction is common to all integrated *S. pombe* transgenomes suggests that interactions between *S. pombe* sequences and murine proteins are the cause of these apparently aberrant structures.

To investigate which aspects of chromatin packaging within the *S. pombe* transgenomes are related to this constriction I have studied one of the hybrid cell lines F1.1. I have analysed the amount of *S. pombe* DNA forming this transgenome and selected regions of it for further investigation. DNA methylation, nucleosome packaging and nucleoskeleton attachment were investigated as potential causes of the constriction.

It was found that DNA methylation accumulated within the transgenome during culture, although no correlating alteration in metaphase structure was detected. Methylation was therefore felt not to be related to the observed structure of the transgenome. The transgenome was found to adopt a murine nucleosomal repeat of 185bp, rather than retain its original 160bp *S. pombe* repeat. Adoption of the host nucleosomal repeat by introduced sequences was found to be a general phenomenon, independent of the structure of introduced DNA. This showed that nucleosome structure was dependent upon the host repertoire of histone and accessory factors and not on the underlying DNA sequence. It was possible to demonstrate differential attachment of both murine and *S. pombe* sequences to the mouse nucleoskeleton. Analysis of an extended region of *S. pombe* DNA suggests that attachment may be closer, <20kb apart, in the *S. pombe* transgenome, than the 80kb separation expected for mouse DNA. Packaging of these smaller loops at metaphase may lead to the observed constricted structure of the *S. pombe* transgenomes.

ACKNOWLEDGEMENTS.

I would like to thank everyone from the MRC who helped me throughout my Ph.D.. Special thanks go to Wendy Bickmore for her expert lab knowledge and tireless reviewing of various drafts. I would also like to thank Robin Allshire for starting me off on this arduous trail. To everyone else who helped me, many thanks, especially Sheila Mould and Helen Moffat from the MRC library for one or two heavy photocopying sessions and Harry Edwards in the Dunn school for the photography.

To everybody else, too numerous to mention, who made Edinburgh such a wonderful and memorable few years thank you. You will always be remembered.

Particular praise must go to Fiona who picked me up when times seemed hard and pushed me on with a warm heart and cold feet.

I would like to dedicate this thesis to my parents who have always supported me and allowed me to find my own direction, without them none of this would have been possible.

TABLE OF CONTENTS

TITLE PAGE.....	1
ABSTRACT	2
DECLARATION	3
ACKNOWLEDGEMENTS	4
TABLE OF CONTENTS.....	5
LIST OF FIGURES	10
LIST OF TABLES.....	12
ABBREVIATIONS.....	13
CHAPTER. 1. INTRODUCTION.....	14
1.1 Background.....	14
1.1.2 Hybrid cell line establishment.....	15
1.2. Methods of DNA introduction for transgenome formation.....	16
1.3. DNA modifications	22
1.3.1 Distribution of methylation.....	22
1.3.2 De novo methylation	23
1.3.3 Gene expression.....	24
1.3.4 Methylation and chromatin structure.	25
1.4. Chromatin.....	26
1.4.1 The nucleosome.....	26
1.4.2 Histone structure.....	27
1.4.3 Histone modification.....	28
1.4.4 Histones and gene expression.....	29
1.4.5. Nucleosome positioning	30
1.4.6 Nucleosome repeat length	31
1.4.7 Chromatin replication.	32
1.5 Higher-order structure.....	33
1.5.1 Nucleosome organisation.....	33
1.5.2 Chromosomal loops	36
1.5.3 Loop size.....	36
1.5.4 Transcription and replication occurs on nuclear cages.....	38
1.5.5 The nucleoskeleton, transcription and replication.....	39
1.5.6 Scaffold attachment.....	40
1.5.7 Scaffold binding conservation.	41
1.5.8 Topoisomerase II and scaffold attachment.....	41
1.5.9 Scaffold loop boundaries.....	42
1.5.10 Scaffold proteins.....	43

1.5.11 Packaging loops into chromosomes	44
1.6 Linear differentiation of chromosomes.....	44
1.6.1 Secondary constrictions	44
1.6.2 Chromomeres.....	45
1.6.3 Fragile sites	45
1.6.4 Chromosome banding	46
1.7 Centromeres.....	47
1.7.1 Primary constrictions.....	47
1.7.2 Centromere proteins.....	47
1.7.3 Centromeric DNA	48
1.7.4 Yeast centromeres.....	49
1.8 Telomeres.....	51
CHAPTER 2. MATERIALS AND METHODS.....	54
2.1 Mammalian cell culture.....	54
2.1.1 Thawing of cells stored in liquid nitrogen.....	54
2.1.2 Maintenance of cell lines in culture.....	54
2.1.3 Transfer and harvesting of cells.	54
2.1.4 Chromosome preparation.....	55
2.2 Bacterial cell culture	55
2.2.1 Media and additives	55
2.2.2 Bacterial strain used	55
2.2.3 Bacterial cell culture.....	55
2.3 Yeast cell culture	56
2.3.1 <i>S. pombe</i> strain used.....	56
2.3.2 Solutions.....	56
2.3.3 Culture of <i>S. pombe Int-5</i> cells.....	56
2.4 Manipulation of plasmid DNA	57
2.4.1 Small scale plasmid preparations	57
2.4.2 Large scale plasmid preparation.....	57
2.5 Screening of the cosmid library.....	57
2.5.1 Plating of the library.....	57
2.5.2 Replicating the cosmid library.	58
2.6 Preparation of DNA from cultured cells.....	58
2.6.1 Preparation of DNA from mammalian cells.....	58
2.7 Preparation of DNA from <i>S. pombe Int-5</i> cells.....	58
2.8 Manipulation of DNA by enzymes.....	59
2.9 Separation of DNA fragments by gel electrophoresis.....	59

2.10	Transfer of DNA to Nylon and Nitrocellulose Membranes.....	60
2.10.1	Sothern transfer of DNA from gel to membranes	60
2.10.2	Immobilisation of DNA from bacterial colonies onto Hybond-N.....	61
2.11	DNA Probes.....	61
2.12	Radiolabelling of DNA.....	62
2.12.1	Random prime reaction.....	62
2.12.2	End labelling of DNA.....	62
2.13	Nucleic Acid Hybridisation.....	63
2.14	Photography.....	63
2.15	Autoradiography.....	63
2.15.1	Fixation of agarose gels.....	64
2.15.2	Autoradiograph development.....	64
2.15.3	Densitometry of autoradiographs.....	64
2.16	Mammalian nuclei preparation and micrococcal nuclease digestion.....	64
2.16.1	Mammalian nuclei preparation.	64
2.16.2	Mammalian micrococcal nuclease digestion.	64
2.17	Yeast nuclei preparation and nuclease digestion.	65
2.18	Manipulation of cells in agarose beads.....	66
2.18.1	Encapsulation of cells in agarose beads	66
2.18.2	Permeabilisation and digestion of cells in agarose beads.	66
2.18.3	Electrophoresis of unattached chromatin.....	67
2.18.4	Recovery from beads.	67
2.18.5	Recovery of loop DNA from gel.....	68
2.18.6	Analysis of DNAs	68
2.19	Manipulation of <i>S. pombe</i> in agarose beads.....	68
2.19.1	Encapsulation of <i>S. pombe</i> cells in agarose beads.	68
2.19.2	Lysis of <i>S. pombe</i> cell wall in agarose beads.	68
2.19.3	<i>S. pombe</i> nuclei preparation in agarose beads.....	69
2.19.4	Digestion of <i>S. pombe</i> chromatin in agarose beads.....	69
2.19.5	Recovery of DNAs from beads and gel.....	69

CHAPTER 3. ANALYSIS OF CONTENT AND STABILITY OF *S. pombe*

SEQUENCES IN MOUSE / <i>S. pombe</i> HYBRID CELL LINES.....	70
3.1 Introduction	70
3.2 Analysis of the <i>S. pombe</i> content of hybrid cell lines	71
3.2.1 Probes used from <i>S. pombe</i> <i>Int 5</i> genome.....	71
3.2.2 <i>S. pombe</i> chromosome III marker content of hybrid cell lines F1.1 and F7.2.....	71
3.2.3 <i>S. pombe</i> Chromosome I and II marker content of hybrid cell lines F1.1 and F7.2.....	76

3.2.4 Analysis of <i>S. pombe</i> centromeres and telomeres in hybrid cell lines F1.1 and F7.2.....	78
3.3 Stability of hybrid cell line transgenomes.....	82
3.4 Discussion.....	82
CHAPTER 4. SCREENING OF AN <i>S. pombe</i> COSMID LIBRARY.....	87
4.1 Introduction.....	87
4.2 Screening of the <i>S. pombe</i> cosmid library.....	87
4.3 Secondary screening of cosmid clones.....	88
4.3.1 rDNA containing cosmids.....	88
4.3.2 Cosmids selected for the <i>ade6</i> locus.....	88
4.3.2 Cosmids from the <i>ura4</i> locus.....	90
4.4 Cosmids tested on cell lines.....	90
CHAPTER 5 METHYLATION OF <i>S. pombe</i> SEQUENCES IN THE F1.1 HYBRID CELL LINE.....	98
5.1 Introduction.....	98
5.2 Quantitative analysis of methylation of cell lines.....	100
5.3 Methylation levels in the <i>S. pombe</i> insert of F1.1.....	102
5.4 Time course study of methylation in the F1.1 insert.....	107
5.5 Cytology of the <i>S. pombe</i> insert in F1.1.....	116
5.6 Conclusions.....	116
CHAPTER 6. NUCLEOSOME STRUCTURE OF <i>S. pombe</i> SEQUENCES IN HYBRID CELL LINES.....	120
6.1 Introduction.....	120
6.2 Nuclease digestions.....	121
6.3 Analysis of <i>S. pombe</i> nucleosome ladders.....	122
6.4 Nucleosome structure of <i>S. pombe</i> sequences in hybrid cell lines.....	130
6.4 Conclusions.....	142
CHAPTER 7. ATTACHMENT OF <i>S. POMBE</i> SEQUENCES TO THE NUCLEOSKELETON OF THE F1.1 HYBRID CELL LINE.....	144
7.1 Introduction.....	144
7.2 Nucleoskeleton procedures.....	147
7.3 Digestion of chromatin in permeabilised cells.....	149
7.4 Attachment of mouse sequences to the nucleoskeleton.....	151
7.5 Attachment of <i>S. pombe</i> sequences to the nucleoskeleton of F1.1.....	155
7.6 Attachment over an extended region of <i>S. pombe</i> sequences to the F1.1 nucleoskeleton.....	157

7.7 Conclusions.....	158
CHAPTER 8 . DISCUSSION.....	163
8.1 Aims of project.	163
8.2 Hybrid cell line content.....	163
8.3 Methylation.....	166
8.4 Nucleosome repeat length.....	168
8.5 Nucleoskeleton attachment.....	171
8.6 Transcription.....	175
8.7 Replication.....	176
8.8 Metaphase chromosome structure.....	177
8.9 Final conclusions.....	178
REFERENCES	180

LIST OF FIGURES.

1.1 CYTOLOGY OF <i>S. POMBE</i> SEQUENCES IN THE F1.1 HYBRID CELL LINE.....	17
1.2 CYTOLOGY OF <i>S. POMBE</i> SEQUENCES IN HYBRID CELL LINES	18
1.3 E.M. ANALYSIS OF THE F1.1 <i>S. POMBE</i> - MOUSE HYBRID CHROMOSOME	19
1.4 MODELS OF HIGHER ORDER NUCLEOSOME PACKAGING	35
1.5 THE STRUCTURE OF <i>S.POMBE</i> CENTROMERES.....	50
3.1 PHYSICAL MAP OF <i>S. POMBE INT-5</i> GENOME.....	72
3.2 <i>S. POMBE</i> CHROMOSOME III MARKER CONTENT OF HYBRID CELL LINES.....	73
3.3 <i>S. POMBE</i> CHROMOSOMES I AND II MARKER CONTENT OF HYBRID CELL LINES	77
3.4 <i>S. POMBE</i> CENTROMERIC, TELOMERIC CONTENT OF HYBRID CELL LINES.....	79
3.5 STABILITY OF <i>S. POMBE</i> SEQUENCES IN THE HYBRID CELL LINES F1.1 AND F7.2.	83
4.1 SECONDARY SCREENING OF <i>S. POMBE</i> rDNA COSMIDS	89
4.2 SECONDARY SCREENING OF Cade6-I.....	91
4.3 SECONDARY SCREENING OF <i>S. POMBE ura4</i> COSMIDS.....	92
4.4 Cade6-I CONTENT OF FUSION AND TRANSFECTANT CELL LINES.....	95
4.5 Cura4-II CONTENT OF FUSION AND TRANSFECTANT CELL LINES.....	96
5.1 COMPARISON OF OVERALL CELLULAR METHYLATION LEVELS	101
5.2 METHYLATION STUDY OF THE <i>nda3</i> REGION OF THE F1.1 <i>S. POMBE</i> INSERT AT PASSAGE 16.....	103
5.3 METHYLATION STUDY OF THE Cade6-I REGION OF THE F1.1 <i>S. POMBE</i> INSERT AT PASSAGE 30	105
5.4 METHYLATION STUDY OF THE Cura4-II REGION OF THE F1.1 <i>S. POMBE</i> INSERT AT PASSAGE 50	106
5.5 TIME COURSE METHYLATION STUDY OF MOUSE <i>con</i> SEQUENCES IN F1.1	109
5.6 TIME COURSE METHYLATION STUDY OF <i>S. POMBE 0.7kb rNTS</i> SEQUENCES IN F1.1	110
5.7 TIME COURSE METHYLATION STUDY OF <i>S. POMBE Cade6-I</i> SEQUENCES IN F1.1	111
5.8 TIME COURSE METHYLATION STUDY OF <i>S. POMBE Cura4-II</i> SEQUENCES IN F1.1.	113

5.9 TIME COURSE METHYLATION STUDY OF <i>S. POMBE</i> dh SEQUENCES IN F1.1	114
5.10 EXTENDED TIME COURSE METHYLATION STUDY OF <i>S. POMBE</i> Cura4-II SEQUENCES IN F1.1.....	115
5.11 CYTOLOGICAL ANALYSIS OF THE <i>S. POMBE</i> INSERT IN F1.1 OVER TIME.	117
6.1 SCHEMATIC DIAGRAM OF MICROCOCCAL NUCLEASE DIGESTS.....	123
6.2 SCHEMATIC COMPARISON OF <i>S. POMBE</i> AND MOUSE NUCLEOSOME LADDERS.	124
6.3 MICROCOCCAL NUCLEASE DIGESTION OF <i>S. POMBE</i> NUCLEI.....	126
6.4 MICROCOCCAL NUCLEASE DIGESTION OF <i>S. POMBE</i> NUCLEI PROBED WITH dh SEQUENCE	128
6.5 MICROCOCCAL NUCLEASE DIGESTION OF <i>S. POMBE</i> NUCLEI PROBED WITH SV ₂ neo, Cade6-I and Cura4-II.....	129
6.6 MICROCOCCAL NUCLEASE DIGESTION OF F1.1 NUCLEI.....	131
6.7 MICROCOCCAL NUCLEASE DIGESTIONS OF F1;1, F7.2 AND C127 NUCLEI PROBED WITH SV ₂ neo SEQUENCES	133
6.8 MICROCOCCAL NUCLEASE DIGESTS OF F1.1, F7.2 AND C127 NUCLEI PROBED WITH Cura4-II COSMID.	134
6.9 MICROCOCCAL NUCLEASE DIGESTS OF F1.1 AND C127 NUCLEI PROBED WITH <i>S. POMBE ade6</i> ..	136
6.10 MICROCOCCAL NUCLEASE DIGESTS OF F1.1, DC:1 AND C127 NUCLEI PROBED WITH <i>S. POMBE ura4</i>	137
6.11 MICROCOCCAL NUCLEASE DIGESTS OF F7.2 AND C127 NUCLEI PROBED WITH <i>S. POMBE ura4</i>	138
6.12 MICROCOCCAL NUCLEASE DIGESTS OF F1.1, F7.2 AND C127 NUCLEI PROBED WITH <i>S. POMBE</i> dh SEQUENCE.....	140
6.13 MICROCOCCAL NUCLEASE DIGESTS OF F1.1, DC:1 AND C127 NUCLEI PROBED WITH <i>S. POMBE</i> dh SEQUENCE.....	141
7.1. PROCEDURE FOR NUCLEOSKELETON PREPARATIONS.....	148
7.2 DIGESTION OF NUCLEOSKELETON PREPARATIONS WITH HAE III.....	150
7.3 COMPLETE HAE III DIGESTION OF NUCLEOSKELETON PREPARATIONS AFTER RECOVERY	152
7.4 ATTACHMENT OF MOUSE SEQUENCES TO THE F1.1 NUCLEOSKELETON.....	153
7.5.ATTACHMENT OF <i>S. POMBE</i> SEQUENCES TO THE F1.1 NUCLEOSKELETON.....	156
7.6 ATTACHMENT OF Cura4-II SEQUENCES TO THE F1.1 NUCLEOSKELETON	159
7.7 Cura4-II COSMID CUT WITH VARIOUS ENZYMES AND PROBED WITH THE <i>ura4</i> GENE SEQUENCE.....	160

LIST OF TABLES

1.1 METHODS FOR NUCLEAR ATTACHMENT ANALYSIS	37
3.1 <i>S. POMBE</i> SEQUENCE CONTENT OF HYBRID CELL LINES.....	84
4.1 SIZES OF SELECTED COSMIDS FROM AN <i>S.POMBE</i> COSMID LIBRARY.....	93

ABBREVIATIONS

APRT	: Adenine phosphoribosyltransferase
ARS	: Autonomously Replicating Sequence
5- Aza C	: 5 Azacytidine
BrdU	: 5-Bromo-2'-deoxy uridine
bp	: base pairs of DNA
CENP	: Centromere specific protein
CHO	: Chinese hamster ovary cells
CMGT	: Chromosome Mediated Gene Transfer
dATP	: Deoxyadenosine triphosphate
dCTP	: Deoxycytidine triphosphate
dGTP	: Deoxyguanosine triphosphate
dTTP	: Deoxythymidine triphosphate
dUTP	: Deoxy uridine triphosphate
<i>E. coli</i>	: <i>Escherichia. coli</i>
EDTA	: Ethylenediaminetetra-acetic acid
E.M.	: Electron Microscopy
EtBr	: Ethidium Bromide
FITC	: Fluorescein isothiocyanate
G6PD	: Glucose 6- phosphate dehydrogenase
GART	: Phosphoribosylglycinamide formyltransferase
HPRT	: Hypoxanthine phosphoribosyl transferase
kb	: kilobase pairs of DNA
Krpm	: x 1000 revolutions per minute
LIS	: Lithium-3' 5' diiodosalicylate
Mb	: Megabase pairs of DNA
MECP	: Methylated CpG binding protein
o/n	: over night
PBS	: Phosphate buffered saline
PEG	: Polyethylene glycol
PIPES	: Piperazine-N-N'-bis[2-ethane sulfonic acid]
PMSF	: Phenylmethylsulfonyl fluoride
rt	: room temperature
SAR	: Scaffold attachment region
SDS	: Sodium dodecyl sulphate
<i>S. pombe</i>	: <i>Schizosacharomyces pombe</i>
TCA	: Trichloroacetic acid
TE	: Tris EDTA buffer

1.INTRODUCTION

1.1 BACKGROUND

Eukaryotic chromosomes are the functional units, that allow a cell to accurately pass on its genome to its progeny. This is achieved by a vast number of molecular interactions involving both nucleic acids and proteins which allow for packaging of the DNA so that the mechanisms of cell division can segregate the stored information appropriately. Chromosome packaging is also required, so that the vast amount of DNA in a eukaryotic cell can be transcribed and replicated efficiently.

In higher eukaryotic nuclei roughly 2m, fully extended, of DNA must be compacted into a sphere of 10 μ m diameter. This is achieved by the interaction of DNA with histones and other non-histone proteins. The protein DNA interactions involved in this condensation process lead to a 10000 fold packaging ratio of metaphase chromosomes compared to their fully extended length of naked DNA.

The reasons for studying chromosome structure are many fold. It is becoming more apparent that gene expression is not solely dependent on the availability of the relevant trans-acting regulators, but also on the chromatin structure of the domain in which a gene is located. Therefore the study of chromatin structure, will provide a better understanding of how chromatin interacts with regulatory molecules to produce the high levels of gene regulation observed *in vivo*. An understanding of how sequences interact with chromatin proteins, in heterologous systems, is also important for the purposes of efficient transgenic and gene therapy programmes. To understand these heterologous interactions should allow more efficient study of regulatory sequences *in vivo*, and also more efficient production of transgenics through the removal of the variance in gene expression levels caused by cis-acting affects related to the site of integration. Chromosome structure analysis is also important in answering which of the factors involved in chromosome packaging may affect the apparent morphology and linear differentiation of chromosomes at metaphase.

Various techniques used to band mammalian chromosomes have demonstrated that chromosomes may not be uniformly packaged (Bickmore and Sumner, 1989). Genes appear to be concentrated in (R-bands) interchromomeric regions of chromosomes. Linear differentiation of chromosomes is also apparent with replication analogues, which demonstrate differences in the replication timing of chromosomal regions.

Inhibition of the packaging process can be induced by high levels of transcription in specific regions, such as the nucleolar organisers. Factors that delay replication of DNA or shorten the apparent G2-phase of the cell, as used in fragile site expression, can also result in regions of apparent undercondensation. Since the apparent chromosome morphology appears to reflect both the transcriptional and replicative activity in chromosomal regions, it is of prime importance to study chromatin structure, to understand how these metabolic activities relate to the observed chromosome morphology.

The study of the factors that directly affect chromosome morphology is most easily studied in higher eukaryotes, which have the benefit of direct cytological assays, due to the ease of observation of such large chromosomes. However, higher eukaryotes, due to the complexity of their genomes, do not easily lend themselves to biochemical or genetic analysis of large genomic regions. Yeasts have simple, largely unique sequence genomes, with excellent genetics which favours well for regional analysis of chromosomes. However their chromosomes are relatively small and largely non-analysable by light microscopy. Very recently chromosomal *in situ* hybridisation has been achieved for the fission yeast *Schizosaccharomyces pombe* (*S. pombe*) (Uzawa and Yanagida, 1992). Analysing the relationship between the yeast DNA sequences and the mammalian proteins in hybrid cell lines, derived by fusing yeast and mammalian cells, would allow elucidation of the important factors involved in chromosome condensation. The effects of the relationship between the yeast DNA and mammalian proteins, on chromosome structure can then be evaluated using both molecular biology and cytogenetics.

1.1.2 . Hybrid cell line establishment.

With a view to the production of mammalian artificial chromosomes Allshire *et al.*, (1987) investigated whether the chromosomes of *S. pombe* could replicate as linear autonomous elements in mammalian cells. This could lead to the production of a vector system, allowing the cloning of large mammalian chromosomal elements such as telomeres and centromeres. To select for yeast chromosomes in mammalian cells the *SV2neo* gene was targeted into the *ura4* gene of *S. pombe* chromosome III. Protoplasts were made from this *S. pombe Int5* strain and fused to mouse C127 cells. Hybrid cell lines resistant to G418 were selected. These hybrids fell into two classes, those in which the yeast DNA was retained only under selection as an extrachromosomal autonomous element, and those in which the yeast DNA was stably integrated into the mouse chromosomes. In the integrated hybrid cell line F1.1 it was observed by *in situ* hybridisation on metaphase chromosomes, that the site of integration of the yeast DNA

exhibited a marked constriction in the chromosome concerned (Perry, unpublished observations) (Figure 1.1). This unusual chromosome structure at the site of *S. pombe* DNA integration has been shown to be a characteristic of many such stably transformed lines (Perry *et al.*, unpublished observations) (Figure 1.2). This thinning of the mouse chromosomal structure at the site of *S. pombe* integration is also supported by scanning electron microscopy where a pronounced thinning of the chromosome is detected at the site of presumed *S. pombe* integration, in F1.1 (Sumner, unpublished) (Figure 1.3). The hybrid line F1.1 is the major cell line analysed in these studies. The constricted appearance of these regions resembles to some extent the cytology of fragile sites in mammalian chromosomes. Fragile sites are susceptible to breakage. Mutagen sensitivity assays have been carried out on F1.1 which show that the yeast insert exhibits a high frequency of sister chromatid exchanges after mitomycin C treatment compared to an equivalently sized region of the mouse genome (Perry *et al.*, unpublished). This suggests that this region is indeed prone to more frequent breakage than that of a similarly sized region of the mouse genome.

In order to investigate the molecular basis for the non-typical morphology in these insert regions, I have analysed three factors that may contribute to the observed morphology of the yeast transgenomes.

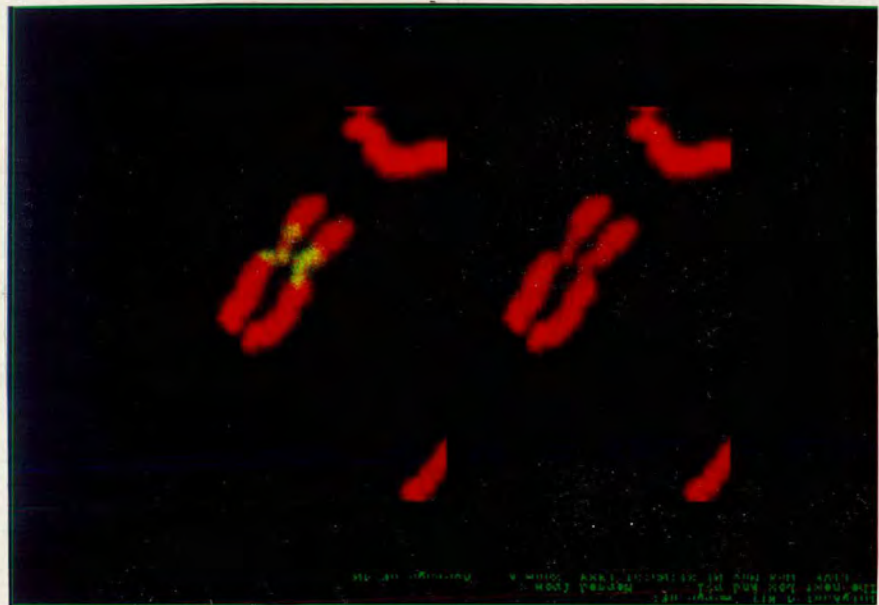
1. Methylation
2. Nucleosome repeats
3. Attachment to the nucleoskeleton.

1.2 METHODS OF DNA INTRODUCTION FOR TRANSGENOME FORMATION.

DNA sequences have been introduced to cells, to produce transgenomes, by a range of procedures. DNA mediated gene transfer techniques have included calcium phosphate precipitation (Kucherlapati and Skoulchi, 1984), electroporation (Boggs *et al.*, 1986), lipofection (Gnirke and Huxley, 1991) and microinjection (Gnirke *et al.*, 91). Sequences have also been introduced as chromatin by chromosome mediated gene transfer (CMGT) (Porteous, 1987) and polyethylene glycol induced cell fusion (Harris, 1970; Ward *et al.*, 1986).

For the purposes of this work I will focus on the techniques that have been used to

FIGURE 1:1. Cytology of *S. pombe* sequences in the F1.1 hybrid cell line.



A

B

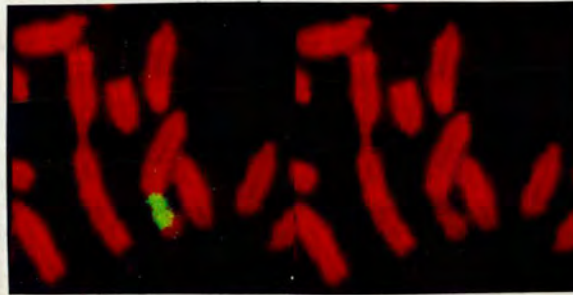
Figure 1:1. F1.1 metpahase chromosomes were prepared and hybridised with biotinylated total *S. pombe* DNA. Biotin was detected by avadin-FITC. DNA was stained with propidium iodide. Images were captured by confocal microscopy.

A= merged FITC+ propidium iodide

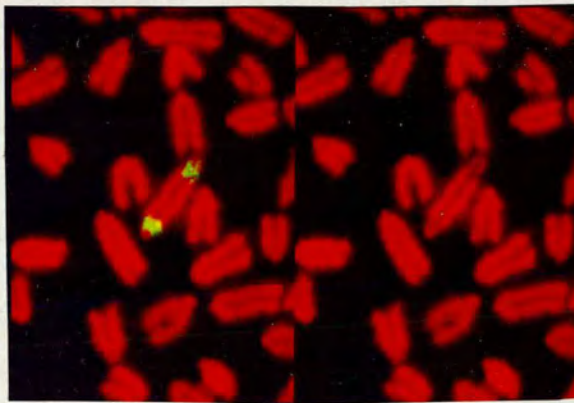
B= propidium iodide

Note the constriction at the site of integration.

FIGURE 1:2. Cytology of *S. pombe* sequences in hybrid cell lines.



A



B

Figure 1:2. Chromosomes prepared from hybrid cell lines F48A1 and F48C2 were hybridised with biotinylated total *S. pombe* DNA. Biotin was detected by avidin-FITC. DNA was stained with propidium iodide. Images were captured by confocal microscopy. F48A1 and F48C2 are *S. pombe* mouse hybrid cell lines from the same experiment as F1.1

A= F48A1

B= F48C2

Note the constrictions at the site of integration, as seen for F1.1 in Fig1.1 .

FIGURE 1:3. Scanning electron microscopy analysis of the F1.1 *S. pombe*- mouse hybrid chromosome.



Figure1:3. F1.1 cells were prepared for analysis by Dr Perry. Scanning E.M. analysis was carried out by Dr Sumner. The centromere is seen as a constriction in the lower right hand corner. The other constriction is believed to be the site of *S. pombe* integration into F1.1 chromosome, as seen in Fig 1.1.

transfer yeast artificial chromosomes (YACs) from the yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) to mammalian cells. The YACs used in these studies are generally derived from human sequence containing libraries. Special interest is taken in the structures of the transgenomes derived from these procedures and also in the degree of *S. cerevisiae* sequence cointegrated with the YAC sequences. Yeast sequences other than those derived from the YAC may either be cointroduced if YAC DNA is not separated from the yeast background during DNA preparation, or introduction during fusion techniques. Unfortunately no extensive cytological or molecular study of the yeast sequences cointegrated has been undertaken, which would have been useful for comparison to our own findings.

DNA mediated gene transfer procedures require that DNA is isolated prior to introduction into the host cell. Therefore shearing of the DNA may occur prior to introduction and this together with passage of DNA through host cytoplasm, could lead to the presence of rearranged transgenomes. Calcium phosphate precipitation has been used for introduction of short 40-50kb YACS (Eliceiri *et al.*, 1991) containing the human glucose 6 phosphate dehydrogenase (G6PD) gene and the YAC sequences were apparently transferred intact. Previous findings (Kucherlapati and Skoulchi, 1984) have demonstrated that calcium phosphate precipitation leads to transfer of ~100kb of concatemered DNA to the recipient cell. Lipofection has been used to transfer a 600kb YAC containing the human phosphoribosylglycinamide formyltransferase (GART) gene (Gnirke and Huxley, 1991) into CHO cells. Hybridisation with an Alu probe suggested that 40% of the transgenomes exhibit a degree of rearrangement. Hybridisation with yeast *Ty* element showed that a variable amount of *S. cerevisiae* DNA had cointegrated in different cell lines. Electroporation has been used to transfer a GART containing YAC (Huxley and Gnirke, 91) to mammalian cells with apparently no cointegration of yeast DNA and with an unrearranged transgenome. Gnirke and Huxley (1991) used microinjection to introduce DNA from a 660kb HPRT-containing YAC into mammalian cells. Due to shearing during DNA preparation the average size of introduced molecules was 100kb. Yeast sequences were also introduced since the YAC was not separated from the rest of the yeast genome. Although high levels of transgenome establishment were achieved rearrangements were apparent in most cell lines. This rearrangement may be due to the breakage of sequences prior to introduction or breakage and religation after entry. In conclusion, with DNA mediated gene transfer procedures, it appears possible to transfer large DNA molecules and obtain an unrearranged transgenome although rearrangements occur in a high percentage of cases.

Transfer of DNA in its more normal state, as chromatin, has been achieved by

chromosome mediated gene transfer (CMGT) and polyethylene glycol (PEG) induced cell fusion. In CMGT chromosomes are fractured and then transferred to the recipient cells by calcium phosphate precipitation (Porteous, 1987). These studies have largely been carried out with host and donor from the same kingdom as opposed to the inter kingdom yeast to mammalian procedure. It is believed (Simmons *et al.*, 1978) that sequences introduced by CMGT are degraded within the cytoplasm and may be religated in the nucleus to form a stable transgenome. The level of degradation and religation of the introduced material would determine the degree of rearrangement observed within these transgenomes. CMGT derived transgenomes have been shown to exhibit wide variation in the degree of rearrangement of their transgenomes (Bickmore *et al.*, 1989). CMGT transgenomes have been found to include microdeletions and extreme levels of rearrangement.

PEG mediated fusion has been used for both intra and inter kingdom cell fusion procedures. After fusion of intra kingdom species (Harris, 1970) heterokaryons are formed. These heterokaryons establish synchrony and enter mitosis simultaneously, the chromosomes mix and align upon a single metaphase plate. Transgenomes established have been found to be linearly intact with no apparent microdeletions (Bickmore *et al.*, 1989). This lack of transgenome rearrangement presumably arises due the coexistence of vertebrate nuclei within a stable heterokaryon, thus protecting the donor DNA from nuclease attack. Ward *et al.*, (1986) demonstrated that heterokaryons could also be formed during the inter kingdom fusion of *S. cerevisiae* and mammalian cells. Whether the heterokaryons are stable and follow a similar route to synchrony and transgenome formation remains to be seen. PEG fusion has been used to transfer YACs from *S. cerevisiae* to mammalian cells. Pachnis *et al.*, (1990) introduced a 450 kb human sequence *neo* resistance gene containing YAC into mouse L cells. Pavan *et al.*, (1990) introduced a 360kb *neo* bearing human sequence YAC into embryonal carcinoma cell lines and Huxley *et al.*, (1991) introduced an HPRT gene bearing YAC into mouse cells. Transgenic mice have been produced by fusion of YAC containing spheroplasts with ES cells which were microinjected into blastocysts (Jakobovitz *et al.*, 1993). In all of the above cases a degree of rearrangement was observed within the transgenomes as assessed by *Alu* hybridisation. In general 50% of hybrid cell lines exhibited a degree of rearrangement of human introduced sequences, although these rearrangements were usually slight, involving only a few bands. Hybridisation with YAC vector sequences showed that in almost all hybrid cell lines that YAC derived telomeric sequences were absent from the derived transgenomes. Varying amounts of yeast DNA had also entered these hybrids, as assessed by Ty hybridisation and the level and regions of the *S. cerevisiae* genome which entered was independent of YAC transgenome structure.

Huxley *et al.* (1991) used a yeast subtelomeric Y' sequence to demonstrate that the yeast telomeres were generally absent from the genome. It would appear therefore that telomeric sequences are generally lost prior to integration of sequences. This may be related to the rearrangements required for integration of introduced sequences or may suggest that the yeast telomeres do not form appropriate telomeric structures in mammalian cells and are therefore degraded. Unfortunately little cytology was carried out in the experiments described and so comparison of our transgenome structures with those derived from *S. cerevisiae* introduction is not possible.

It would appear that some rearrangement of introduced DNA is common to most introduction procedures, apart from fusion, when DNA is transferred between species. Lack of cytological data on the cointroduced yeast DNA is unfortunate. Huxley *et al.* (1991) have carried out in-situ hybridisations to the YAC transgenomes but have not observed any unusual structure and have not investigated the cointegrated yeast sequences. Human sequence bearing transgenomes derived from the YACs may not yield any cytological differences in structure that may be provided by yeast sequences.

1.3 . DNA Modifications

1.3.1 Distribution of methylation

DNA methylation is a modification that shows a striking difference between vertebrates and simple eukaryotes. Vertebrate DNA exhibits methylation of cytosine residues at a majority of CpG dinucleotides, while simple eukaryotes, including *S. pombe*, do not exhibit DNA methylation at detectable levels (Bird, 1987). The CpG dinucleotide is depleted in vertebrate genomes, and can be as low as 25% of the expected frequency (Swartz *et al.*, 1962), presumably due to mutation created by the deamination of 5-methylcytosine, which leads to its transition to thymine. The level of CpG deficiency correlates well with the level of methylation in a genome, and also inversely with excesses in the level of the mutation products: TpG and CpA (Bird, 1980). In vertebrates, 80% of CpG dinucleotides are methylated at the cytosine, while the other 20% show no methylation. This non-methylated fraction appears to be found almost exclusively at the 5' end of genes, in regions known as CpG islands. The CpG islands are so called because they show no CpG deficiency. A detailed study has shown that all known housekeeping genes have CpG islands (Gardiner-Garden and Frommer, 1987). Many tissue specific genes also have CpG islands.

A gene's methylation state appears to be established in the germ line (Frank *et al.*, 1991) where previous methylation is modified by the addition or loss of methylation

from CpG dinucleotides (Ariel *et al.*, 1991). It has been shown (Monk *et al.*, 1987) that during mouse pre-implantation development the genome undergoes global demethylation, as assessed by studying methylation of the highly repeated MIF sequence. *De novo* methylation activity then establishes a methylation pattern on this sequence as somatic development occurs. The methylation pattern of DNA in somatic cells is faithfully preserved by a maintenance methylase activity (Wigler *et al.*, 1981).

A DNA methyltransferase has been cloned from mouse cells (Bestor *et al.*, 1988) and this protein has a C-terminal domain similar to that of bacterial methyltransferases. The N-terminal domain is not essential for methyltransferase function. If the N-terminal domain is removed then methyltransferase activity is still detected (Bestor *et al.*, 1988). The N-terminal domain has been shown to contain a targeting sequence which allows cell-cycle dependent association with replication foci (Leonhardt *et al.*, 1992). Such a targeting function would be expected since the hemimethylated DNA resulting from replication requires methyltransferase activity to allow for the clonal propagation of methylation levels.

Li *et al.* (1992) have shown that a 3 fold drop in genomic ⁵methyl-cytosine levels, due to mutation of the methyltransferase gene leads to embryonic lethality, demonstrating that methylation performs an essential function during development. This lethality is not observed in embryonal stem (ES) cells, which with a similar drop in genomic methylation levels exhibit no alteration in growth or morphology. Methylation therefore appears to be required for development and not for growth in culture.

1.3.2 De novo methylation

When DNA is introduced into mammalian cells, the methylation pattern of the introduced DNA depends on the cell line into which it is placed. *De novo* methylation of DNA introduced into embryonal cells and transgenic mice (Kolsto *et al.*, 1986; Frank *et al.*, 1991) has been shown to follow the gene's normal pattern. If an *in vitro* methylated *APRT* transgene is studied, the CpG island region of the gene becomes demethylated, whereas the 3' end of the gene retains methylation and is *de novo* methylated at additional sites (Frank *et al.*, 1991). If a mouse *Thy1.1* gene is microinjected into fertilised eggs it is observed that the CpG island stays unmethylated while the flanking DNA becomes methylated (Kolsto *et al.*, 1986). Embryonal cells in culture also possess *de novo* methylation activity. If constructs are transfected into embryonal cells then the methylation pattern of the introduced DNA will be modified, if necessary, to a pattern appropriate for the sequences introduced. CpG island genes transfected into embryonal cell types will display unmethylated CpG islands and methylated non-CpG regions,

while non-CpG island genes will display no demethylated regions (Frank *et al.*, 1991; Cedar and Razin, 1990). This demethylation of CpG islands in embryonal cells is a rapid process and can be detected within 48hrs after transient transfection (Frank *et al.*, 1991). In differentiated higher eukaryotic cells, the presence of *de novo* methylation activity depends on the cell line used (Szyf *et al.*, 1989). When DNA fragments encoding the mouse steroid 21-hydroxylase (*C21*) gene are transfected into the mouse adrenocortical tumour cell line Y1, sequences within the *C21* gene become methylated while flanking vector sequences or *neomycin* gene sequences that were cotransfected do not become methylated, suggesting that the *de novo* methylase activity may be sequence specific. *De novo* methylase activity was not observed in Y1Kin8 cells, L cells or L8 myoblasts in the above or other studies (Szyf *et al.*, 1989; Pollack *et al.*, 80). *De novo* methylation activity has also been demonstrated in human fibroblasts (Joblanka *et al.*, 1987), rat 2 fibroblasts (Burnett and Gallimore, 1985) and Adenovirus type 12 transformed hamster cells and rat brain tumour cells (Vardimon *et al.*, 1980). Variation in *de novo* methylase activity presumably reflects differences in presence of methyltransferases and demethylases or in the molecules which regulate their activity. *De novo* methylation has also been demonstrated in cells recovering from 5-Aza cytidine treatment (Gasson *et al.*, 1983) If SAK 8 cells; resistant to dexamethosone, are treated with 5-AzaC then dexamethosone sensitive subclones can be obtained. These clones can revert to dexamethosone resistance within 10 weeks of treatment. It was shown that this change in dexamethosone sensitivity was closely correlated to the methylation level of the cells after 5-AzaC treatment. Methylation levels were low after treatment and gradually recovered back to their original levels as cells regained resistance to dexamethosone. *De novo* methylation has also been shown to affect endogenous DNA sequences. In animals CpG islands are normally free of methylation even if transcriptionally silent, with the exception of the inactive X chromosome where CpG islands become methylated. This is not the case in NIH3T3 cells and L cells (Antequera *et al.*, 90), where over half of the CpG islands appear to be heavily methylated. It was shown that this *de novo* methylation had affected non-housekeeping genes, since several tissue specific genes were shown to have methylated CpG islands, while the housekeeping gene's CpG islands remained methylation free. These epigenetic mutation events lead to the repression of many genes in cultured cell lines (Antequera *et al.*, 1989; Holliday, 1987).

1.3.3 Gene expression.

That methylation does not appear to play a direct role in gene expression has been demonstrated in various pieces of work. The δ - crystallin genes of chick embryo lens

cells become hypomethylated only after expression of the gene has started (Sullivan *et al.*, 1989), and the major chick vitellogenin gene also does not undergo demethylation prior to expression (Bruch *et al.*, 1983). In eutherian mammals methylation appears to be involved in the stable repression of genes on the inactive X-chromosome (Grant and Chapman, 1988), although not in the initial inactivation process. An inactive X-chromosome *Hprt* gene becomes methylated after transcriptional inactivation of the gene. Methylation is also involved in the repression of non-CpG -island genes, but only if the gene is expressed from a weak promoter (Boyes *et al.*, 1992). Artificial methylation of constructs of human γ -globin (Busslinger *et al.*, 1983), β -globin (Yisraeli, *et al.*, 1986) or rat α -actin (Yisraeli *et al.*, 1988) leads to the repression of these genes in cell lines which express these genes weakly. In high expressing cell lines this repression is not observed: the methylated α -actin gene is expressed normally in rat myoblasts. The methylated E2A promoter was fully active in the presence of adenovirus transactivator E1a (Langer *et al.*, 1986). That high levels of expression can overcome repression by methylation has also been shown for mouse and human α -globin genes (Boyes *et al.*, 1992). Constructs of these genes methylated at low levels are not transcribed. The addition of an SV40 enhancer overcomes this methylation repression and allows transcription to occur *in vitro*. Repression due to high levels of methylation can not be overcome by the presence of an enhancer. This repression is an indirect affect since addition of methylated competitor DNA will allow transcription to occur from methylated templates (Boyes *et al.*, 1991). The proteins implicated as the mediators of this indirect repression are the MECP proteins (Methylated cytosine binding proteins). Levels of MECP1 have been shown to be related to the degree of repression observed (Boyes *et al.*, 1991), whilst no direct evidence that MECP 2 is involved in this repression process has been obtained (Meehan *et al.*, 1993). MECP 1 and 2 have been shown to bind in a sequence independent manner to methylated cytosine residues. MECP 2 has a requirement for only one residue to be present while MECP 1 requires a minimum of 12 methylated cytosines to mediate efficient binding, although weak binding to CpG poor genes has been demonstrated, which is easily displaced by transcription (Boyes *et al.*, 1992).

1.3.4 Methylation and chromatin structure.

It has been shown that chromatin containing methylated DNA is less accessible to nucleases and endonucleases suggesting it has a less "open" configuration compared to unmethylated chromatin. In isolated nuclei digested by the methylation insensitive restriction endonuclease *MspI*, which cleaves the sequence CCGG, the enzyme cuts the chromatin predominantly in the unmethylated regions, rather than at all of its potential

cleavage sites (Antequera *et al.*, 1989). Preferential cleavage of unmethylated chromatin in nuclei is not demonstrated by restriction enzymes that do not have a CpG dinucleotide in their recognition site. It has subsequently been shown that CpG islands exhibit a disrupted chromatin configuration (Tazi and Bird, 1990). They are depleted for histone H1, compared to the rest of the genome and histones H3 and H4 are highly acetylated within CpG islands. In addition the central region of a CpG island is nucleosome free.

The MECP proteins, especially MECP 1, which requires multiple methylated cytosines for efficient binding, would appear the most likely candidates for the protection of methylated CpGs to nucleases, by directly binding to these sites and preventing access to nucleases (Boyes *et al.*, 1992).

Methylated minichromosomes have also been shown to be in an inaccessible conformation. After replication a methylated minichromosome has been shown to become inaccessible to the V(D)J recombinase, and also to exogenous endonucleases (Hsieh *et al.*, 1992). It would appear therefore that, methylated DNA perhaps through interaction with the MECPs, is packaged into a conformation which is less open to endogenous cellular enzymes as well as exogenously added ones.

1.4 CHROMATIN

1.4.1 The nucleosome

Despite the great variety and complexity of eukaryotic chromosomes, they are at one level remarkably similar, and simple. Chromosomes are made up of chromatin; DNA and protein in roughly equal mass. This is formed into repeating units; nucleosomes, which contain 4 basic histone proteins, but interactions with other non-histone proteins are required for appropriate packaging and function. The nucleosome is the only well understood level of chromatin packaging. Nucleosomes have been studied using nuclease digestion, to cleave chromatin into nucleosomal subunits. The nucleosomes released contain between about 160 (yeast) and 240bp (Sea cucumber) of DNA (Van Holde, 1989). The nucleosome size is species specific, but can also be tissue specific within a species. The nucleosome can however be digested down to a stable "core particle" which contains 1.75 loops of DNA, equivalent to 146bp of DNA in any species or tissue. This core particle is effectively invariant in all eukaryotic cells; except for some vertebrate sperm, where protamines replace histones for chromatin packaging. The "core particle" is defined as a 146bp of DNA wrapped around an octamer of two

each of the histones H2A, H2B, H3 and H4. The particle is roughly cylindrical with a diameter of 11nm and a height of 5.5nm. Histone H1 binds the linker DNA outside of the core particle, and has been shown to be involved in stabilising the nucleosome. Chromatin lacking histone H1 will not package nucleosomes into higher order solenoid structures "30nm fibres" (Thoma *et al.*, 1979) while the addition of H1 to H1 depleted chromatin will allow the formation of higher order structures (Allan *et al.*, 1981). This activity of stabilising the nucleosome and producing higher order structures, allows H1 to act as a general repressor of transcription, by preventing transcription complex formation.

1.4.2 Histone structure

Core histones are bound to the DNA by non-covalent forces. Histones are basic proteins and vary in their respective lysine content. Their overall positive charge is distributed unevenly. The C-terminus of histones show the properties of typical globular proteins, while the N-terminus is more basic and "tail-like". The core histones are amongst the most highly conserved of all known proteins. The arginine rich H3 and H4 are the least variant. All known histone H4 sequences contain exactly 102 residues while all except one sequenced H3 proteins contain 135 residues; the sole exception is the mouse H3 which contains 134 residues (Van Holde, 1988). Only one amino acid change is detected between pea and calf H4 sequences and this is a conservative substitution. Histones may carry post-translational modifications, such as ubiquitination, phosphorylation and acetylation and may be related to differing structural requirements of the proteins throughout the cell cycle. Monoclonal antibodies against acetylated histone H4 will detect the same modified site in virtually all species tested. This constancy must reflect the critical role of H3 and H4 in nucleosome structure. However this is contradicted by findings that in yeast, the histone H4 tails are dispensable, with little affect on growth or viability (Kayne *et al.*, 1988; Durrin *et al.*, 1991). Histones H2A, H2B and H1 show no such restriction and can vary considerably.

The largest variation is seen in Histone H1. Divergence of H1 variants may be so considerable, that they may be easily classified as distinct proteins. Histone H5 the major histone H1 variant in avian reticulocytes, was long thought to be a distinct histone protein, but sequence comparison appears to suggest that it is only a very highly diverged variant. These histone H1 variants may affect local regions of chromatin to keep them in a state of readiness for expression, replication or compacting the chromatin to exclude the region from some interactions. Both fission yeast and budding yeast, have not as yet, been found to have H1 homologues (Yanagida, 1990). This

may either be due to proteolysis of the protein during extraction, a level of divergence that means that a homologue may not be easily recognised, or that the genomes of the yeasts can function and segregate normally without such a protein. However a protein with an "H1-like" activity appears to be present, since "30nm fibres" can be detected in *S. cerevisiae* (Widom *et al.*, 1991).

1.4.3 Histone modification

Histones do not play a purely passive or obstructive role in chromatin metabolism. It appears that histone modifications, mainly acetylation and phosphorylation, play roles in chromosome structure, and transcription. It has been shown that H1 phosphorylation increases and phosphorylation of H3 occurs as cells enter mitosis (Csordas *et al.*, 1990), and is removed as the cells enter G1. These phosphorylation changes are associated with chromosome compaction, and are proposed to be necessary for condensation to occur (Gurley *et al.*, 1978; Matthews and Morton-Bradbury, 1978).

Acetylation of core histones is universal, and occurs at the lysine residues in the N-terminal tail. There are 4 sites of acetylation in the tail region of H4. It has been shown using rabbit polyclonal antibodies and protein sequencing, that the usage of these sites varies from species to species (Turner, 1992). In humans and bovines lysine 16 is the major site of mono-acetylation, while in cuttlefish it is lysine 12. In *Drosophila* lysines 8 and 12 are the major sites. As the acetylation state increases from the mono- to tetra-acetylated isoforms there is a non-random usage of potential sites. *Tetrahymena* and cuttlefish sites are used in an invariant order (7, 5, 11, 15 and 12, 5, 16, 8 respectively), with the result that each isoform consists of essentially a single molecule. In human and bovine cells more flexibility in site usage occurs, especially in more acetylated isoforms. Histones H3 and H2 also show a flexible but non-random usage of acetylation sites (Thorne *et al.*, 1990). This pattern of differential usage presumably reflects the relative activities of the acetylases and deacetylases in each species. The N-terminal tail of the core histones does not appear to be an intrinsic part of the core nucleosome, as can be seen from cross-linking studies, and also from the fact that *S. cerevisiae* can tolerate the loss of the H4 tail (Kayne *et al.*, 1988). However the N-terminal of H4 is required for the appropriate silencing of the mating type genes. Since in H4 tail deletion mutants, both *HML α* and *HMR α* genes are constitutively expressed.

1.4.4 Histones and gene expression.

If a gene is translocated from its normal chromosomal region, to a position in, or adjacent to, a heterochromatic region, then it may become inactivated. This inactivation

can occur in different cells, at different times during development, leading to a mosaic or variegated phenotype. This phenomenon is known as position effect variegation. *Drosophila* has been intensively utilised in studies of this phenomenon, and mutants which can suppress or enhance variegation have been isolated. *Suvar(3)7*, a gene which suppresses variegation when mutated has been cloned and sequenced (Reuter *et al.*, 1991). It has been proposed that *Suvar(3)7* may interact with histones to stabilise heterochromatin, by virtue of the polyacidic stretches located between its zinc-fingers (Turner, 1991). Variegation can also be suppressed by the deacetylase inhibitors sodium butyrate or propionate (Mottus *et al.*, 1980). The suppresser of variegation *Suvar(2)101*, is also associated with an increase in the steady state acetylation levels (Dorn *et al.*, 86), suggesting that non-acetylated histones are required for the formation of heterochromatin. Acetylated chromatin by contrast is enriched in transcribing sequences (Hebbes *et al.*, 1988). CpG islands (Tazi and Bird, 1990), show a nucleosome free region which gives way to highly acetylated "unstable " nucleosomes. Histones H3 and H4 are found to be tetra-acetylated in CpG islands (Tazi and Bird, 1990). These regions are also highly depleted of histone H1. CpG islands show this character in all tissues, independent of expression.

Antibodies to histones H1 and H5 have been used to immunoprecipitate chromatin. In chicken erythrocytes the β -globin gene and the H5 gene are depleted in these histones compared to inactive genes (Kamakaka and Thomas, 1990). Psoralen cross-linking has been used to show that, in Friend cells, ribosomal DNA exists in two different chromatin states. One population of sequences, shows a nucleosome packaged character, while the other shows a lack of nucleosome packaging. These two different populations of structures coexist in the same cell and are maintained throughout the cell cycle (Sogo *et al.*, 1989). Altered nucleosome structures have also been demonstrated in *Drosophila* using chemical cross-linking (Nacheva *et al.*, 1989). These experiments demonstrate that after induction, the chromatin structure of the *Hsp 70* gene becomes less dependent on the globular cores of the core histones. The main attachments of the histones to the DNA, in such a highly induced gene, appears to be mediated through the histone tails. This change in attachment of the nucleosomes is taken to show that during transcription the nucleosomes split to allow the polymerase to progress. In contrast to histones remaining attached to the DNA during transcription. Clark and Felsenfeld (1992) have demonstrated that nucleosomes are repositioned behind a transcribing polymerase suggesting that attachments are broken during transcription. These contrasting results can though be unified if they are the results of studying different aspects of the same process. If the nucleosome is temporarily held onto the DNA by the histone tails during polymerase progression then reattachment to the open

DNA immediately behind the polymerase may be kinetically favourable.

1.4.5. Nucleosome positioning

Due to the non-covalent nature of the DNA-protein interactions involved in nucleosome structure it could be argued that the underlying DNA sequence plays a major role in the determination of the position of nucleosomes. Direct histone DNA interactions, nucleosome exclusion or boundary affects due to permanent protein DNA interactions, such as scaffold attachment, could lead to a nucleosome being positioned on a particular sequence. A short repeated sequence with a nucleosome positioned in each repeat would give the appearance of phased nucleosomes, where the pattern follows the repeat length of the sequence. Alternatively the nucleosomes could be regularly spaced along the DNA, with no apparent relationship to the underlying sequence from cell to cell. Both these situations might occur in different regions of chromosomes.

That the bulk of nucleosomes in a cell are randomly situated with regard to DNA sequence, was shown by reassociation studies of DNA from digested chromatin from *Drosophila* and rat liver cells. Trimer nucleosome DNA that was denatured, reannealed and digested with S1 nuclease produced a smear of fragments, not the specific 200, 400 and 600 bp bands expected for positioned nucleosomes (Prunell and Kornberg, 1978). This study however does not address whether positioning of nucleosomes may occur at some specific sites in genomes.

The 172 bp monomer sequence forming the α -satellite of the African green monkey, has been shown to have a phased nucleosomal pattern (Zhang *et al.*, 1983). Multiple forms of phasing patterns have been found with a nucleosome positioned at different sites in the sequence for each pattern. Multiple forms of phasing have also been reported in the mouse major satellite (Zhang and Horz, 1984). Whether these results reflect the true nucleosomal positioning in the cell or are an effect of the sequence specificity of the nucleases used remains debatable. *Xenopus laevis* has 8 *t-RNA* genes occurring in a 3.2kb cluster repeated 100 times. In mature erythrocytes where the genes are inactive, positioned nucleosomes are located along the repeated DNA in a regular fashion. In contrast no positioning was detected in two tissues actively transcribing the genes (Humphries *et al.*, 1979). The *Drosophila Hsp26* locus has a positioned nucleosome located at -140 to -300 relative to the transcriptional start site (Elgin, 1988). Flanking this site are two heat shock response elements which could be brought into close proximity by looping of the DNA around the nucleosome and may allow the elements to work more efficiently. In mice it has been shown that the major β -globin

gene in L-cells is covered by a continuous array of phased nucleosomes from -3000bp to +1500bp relative to the cap site. In erythroid cells where the gene can be transcribed the same phasing exists but is interrupted from positions -200 to +500, where a non nucleosomal structure occurs (Benezra and Carbon, 1986). Some of the clearest evidence for nucleosome positioning comes from yeast (Bloom *et al.*, 1982). In the centromeres of *S.cerevisiae* there appears to be an array of twelve positioned nucleosomes flanking the 250bp centromeric core. Plasmids containing the sequences, minus the central core, showed the same nucleosomal positioning. The *S. cerevisiae* acid phosphatase gene *PHO5*, displays four positioned nucleosomes around a hypersensitive region containing an Upstream Activation Sequence (UAS). Upon activation of the gene, these four nucleosomes are specifically removed (Almer *et al.*, 1986). These nucleosomes are required for efficient suppression of the gene (Almer *et al.*, 1991). It would appear that although positioned nucleosomes may account for a very low percentage of nucleosomes, where this positioning occurs it is important to the proper functioning of the underlying sequences.

1.4.6 Nucleosome repeat length

Although core nucleosomes are 146bp in all studied eukaryotes, the amount of linker DNA is variable. Since nucleosome repeats do not appear solely to be determined by the underlying DNA sequence (Prunell and Kornberg 1978), it is interesting to ask whether nucleosome repeats are determined by their histone content or by other, as yet unknown, factors which may affect nucleosome spacing. Specific sequences within a cell may have a different nucleosomal repeat from that of the surrounding chromatin. While *Xenopus* bulk chromatin exhibits a nucleosomal repeat of 187bp (Humphries *et al.*, 1979), the *Xenopus* 5S genes exhibit a repeat of 178bp (Gottesfeld, 1980). Mouse cells show an average repeat of 200bp, while in tissue culture they have a repeat of 190bp (Van Holde, 1989). It has been shown that all vertebrate tissue culture cells display a nucleosome repeat of 185 -190bp (Van Holde, 1989). This change in repeat length in cultured cells presumably is a result of the silencing of certain non-essential genes, involved in determining nucleosomal spacing, during cell line formation. *S. pombe* has a nucleosome repeat of between 155-160bp (Nakaseko *et al.*, 1988; Polizzi and Clarke, 1991). The *ura4* and *ade6* genes of *S. pombe* after partial micrococcal nuclease digestion, display a nucleosomal repeat of 155-160bp as do the centromeric flanking repeats. As noted earlier neither budding nor fission yeast appear to have a histone H1 protein, this may be reflected in this shorter repeat. Studies involving core histone reconstitution onto poly[d(A-T)] poly[d(A-T)] tracts and then the addition of H1 variants found that the H1 variants produced order in the nucleosomes present. The

repeat formed was largely dependent on the initial average nucleosome packing density before addition, rather than on the H1 variant added (Stein and Mitchell, 1988), suggesting that H1 plays no role in determination of nucleosome repeat length. Specificity of nucleosome repeat length it can be assumed, must be largely dependent on the interactions involved in chromatin formation and maturation following replication.

1.4.7 Chromatin replication.

During S-phase not only must the DNA be replicated, but the chromatin structure must be replicated as well. How nucleosomes are segregated during replication has been a matter of much debate. As the replication fork moves through chromatin the nucleosomes could remain attached to the leading strand or dissociate from the DNA during fork progression, to reform on the replicated DNA. Initial evidence, analysing nucleosome segregation in the absence of protein synthesis showed that old nucleosomes preferentially segregated to the leading strand during replication (Siedman *et al.*, 1978; Cusick *et al.*, 1984). Replication of chromatin, reconstituted from purified histones, exhibits a conservative pattern of nucleosome segregation (Bonne-Andreas *et al.*, 1990). This conservative fashion of nucleosome segregation has been challenged. Nucleosomes when cross-linked to the DNA, and studied using electron microscopy, can be seen to be missing from the region directly behind the progressing replication fork (Sogo *et al.*, 1986). Dispersive segregation of nucleosomes is also proposed by Cuisick *et al.*, (1984) who recovered nascent chromatin from either ^3H or ^{32}P -CTP labelled SV40 virus in African green monkey cells. This nascent chromatin hybridised to both separated strands of SV40, demonstrating that nucleosomes did not preferentially remain on one strand during replication. Cycloheximide, which inhibits new protein synthesis and other metabolic processes, has been used to block new nucleosome formation. Incorporation studies of new histones into replicating chromatin has demonstrated that new H3/H4 tetramers are distributed to both strands after replication and that they can be associated with old H2A/H2B dimers (Jackson, 1990). These studies suggest that nucleosomes are not segregating conservatively to one strand after replication but possibly reform onto the nascent DNA after passage of the replication fork.

Assembly of new nucleosomes on replicated DNA appears to follow a multistep process. *In vitro* systems using complex cellular extracts, have now allowed the mechanisms of nucleosome assembly to be partly elucidated. Nucleosome assembly occurs a short way behind the replication fork (Sogo *et al.*, 1986), with the initial assembly of a H3/H4 complex on the DNA, followed by the addition of H2A/H2B

dimers. In *Xenopus* egg extracts, assembly is mediated by two factors, N1 which bind and transfer H3/H4 tetramers to the DNA, while nucleoplasmin binds H2A/H2B dimers (Kleinschmidt and Martinson, 1984; Dilworth *et al.*, 1987). This process in *Xenopus* extracts is not dependent on replication. It is unclear whether this chaperone function of these proteins is specific to embryogenesis in *Xenopus*. A protein CAF-1 has been isolated from mammalian extracts (Smith and Stillman, 1989, 1991) which promotes chromatin assembly in a replication-dependent fashion. CAF-1 targets the deposition of newly synthesised histones H3 and H4 to replicating DNA. This protein utilises newly synthesised H3 and H4, which unlike old H3 and H4 cannot self assemble on DNA. This H3/H4 intermediate appears to be a stable intermediate in the nucleosome assembly process. The next step is the conversion of this intermediate to the mature nucleosome by the addition of H2A/H2B dimers. This reaction appears to occur in a replication independent fashion, and occurs slowly after replication. (Almouzni *et al.*, 1990; Fotodar and Roberts, 1989) This stepwise mechanism for the assembly of nucleosomes may account for the change in nuclease sensitivity of new "immature" chromatin after replication. The maturation of nascent chromatin, from a highly nuclease sensitive state, to that of mature chromatin, takes approximately 10 -20 minutes in mammalian cells (Cusick *et al.*, 1983; Klempenaar *et al.*, 1980). The exclusion of nucleosomes from CpG islands would appear to occur during the chromatin maturation process. During the formation of nascent chromatin the histones and transcription factors will compete for binding, and due to the generally high affinities of transcription factors, they may exclude nucleosome formation, in these promoter regions.

1.5 HIGHER-ORDER STRUCTURE

1.5.1 Nucleosome organisation.

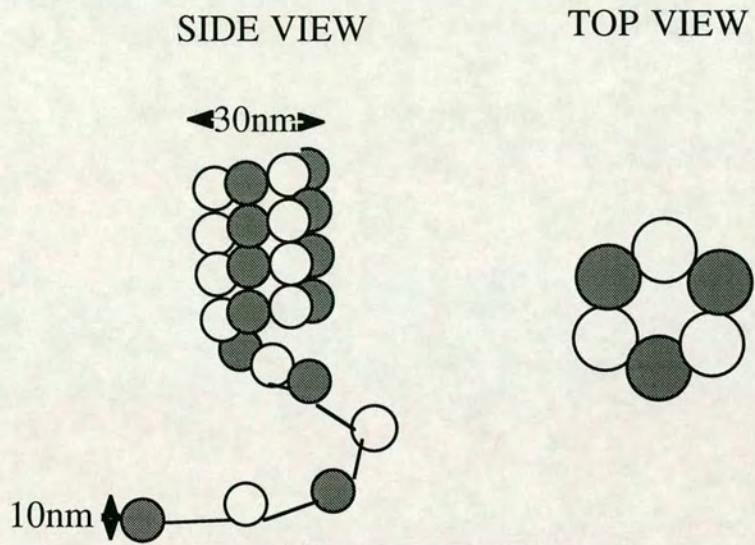
To produce a metaphase chromosome, nucleosomes must be packaged into higher-order structures. The next level up from the nucleosome is the "30nm fibre", which has been shown to be dependent on the presence of histone H1 (Thomas, 1979; Mc Ghee *et al.*, 83) The structure of the "30nm fibre" as with all levels of higher order structure is fraught with controversy.

There are three basic models for the packaging of the "30nm fibre" as illustrated in Figure 1.4. The solenoid model of Finch and Klug, (1976) proposes that the nucleosomes are wound to form a single helical path which constitutes the solenoid. The linker DNA is wrapped in a helical path between the nucleosomes. Butler, (1984) proposes a modified solenoid model, where the linker DNA is kinked into the central

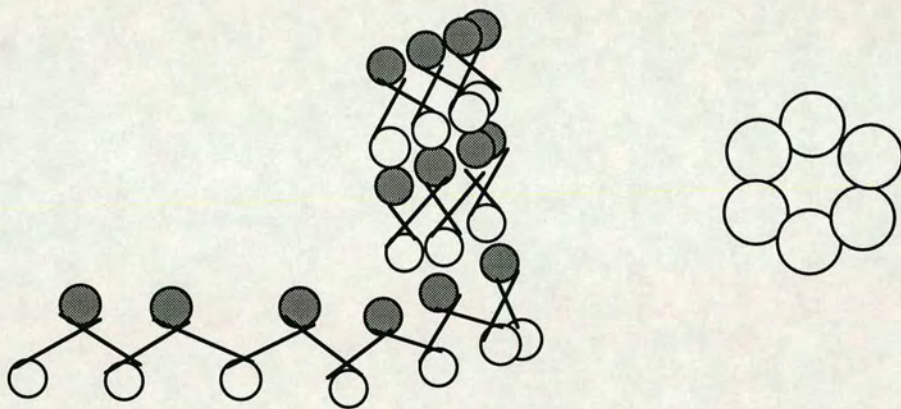
core of the solenoid. This modification allows the solenoid to have a constant diameter and mass per unit length. The second model is the twisted ribbon model proposed by Woodcock *et al.*, (1984). This model was inspired by the zig-zag appearance of nucleosomes at low ionic strength. The model is a flat ribbon consisting of two parallel stacks of nucleosomes, connected by the linker DNA. This ribbon is wrapped as if round a cylinder to produce a helical structure. Woodcock proposed that the model would have a constant diameter and that interactions among the nucleosomes would be conserved. The third model is the cross-linker model. Non-sequentially arranged nucleosomes are connected by the linker DNA traversing the central core. The nucleosomes in this model are double helically arranged along the length of the fibre (Williams *et al.*, 1986). This model proposes a filled central core, diameters which vary with linker length and a variable mass per unit length. These models at least allow certain criteria to be tested, and are represented in figure 1.4.

The solenoid and twisted ribbon models have been supported by observations of chromatin fibres (Widom and Klug, 1985). X-ray diffraction studies have suggested that nucleosome interactions are conserved (Widom and Klug, 1985; Koch *et al.*, 1987). Hydrodynamic studies have also shown that the fibre has a constant diameter, and that nucleosome interactions are conserved between fibres (Mc Ghee *et al.*, 1983). Woodcock *et al.*, (1991) present electron microscopy (E.M) evidence that nucleosome interactions are conserved, and that fibres are of a constant diameter. They claim that any variance observed is due to variable affects of the binding of stains to the fibres. They also propose that their evidence supports the twisted ribbon model of nucleosome packaging. Statistical surveys of fibres in E.M. have shown that the fibre diameter varies as a function of the linker DNA length. *Thyone briareus* (sea cucumber) has a nucleosome repeat of 227bp and a diameter of 43.5nm when packaged, while *Necturus maculosus* has a repeat of 194bp and has a fibre diameter of 32.0nm (Athey *et al.*, 1990). In earlier studies using different buffers Alegre and Subirana (1989) found that mouse thymus DNA, 195bp repeat, had a diameter of 20.1nm, chicken erythrocytes, 208bp, had a diameter of 26.8nm while sea cucumber sperm, 227bp, was 29.3nm. 3-D reconstruction of fibres from E.M. studies has shown that there is a lack of regularity in the interactions of nucleosomes and that the fibres exhibit a zig-zag pattern of nucleosomes (Subirana *et al.*, 1983, 1985). Models proposing a hollow centred fibre are also not supported by the finding that chromatin fibres have filled centres (Smith *et al.*, 1990). This was calculated using density studies from scanning electron microscopy (S.E.M.) experiments. Although the bulk of the evidence appears to support the cross-linker models of nucleosome packaging, the arguments are far from being resolved.

Figure 1.4 MODELS OF HIGHER ORDER NUCLEOSOME PACKAGING



a: Solenoid model of Finch and Klug. 1976.



b: Double helical ribbon model. Worcel et al 1981.



c: Cross centre model. Williams et al, 1986.

Nucleosomes and the "30nm fibre" can account for approximately a 50-fold ratio of packaging. Higher orders of compaction must be present for the 10000 fold packaging ratio necessary to fit DNA into the metaphase chromosome. What these higher orders of structure are and how they are formed and packaged remains an area of high conjecture.

1.5.2 Chromosomal loops

It is now widely accepted that the "30nm fibre" is gathered into loop domains, which may be helically coiled to produce the metaphase chromosome (Rattner and Lin, 1985). How these loops are formed and what relevance they play in the metabolism of the cell remains unclear. That chromatin is formed into loops was shown visually by dehistonising metaphase chromosomes (Paulson and Laemmli, 1977) with high salt extraction. The presence of supercoiled loops in interphase nucleoids has also been demonstrated (Cook and Brazell, 1976a). These methods show that DNA extends in loops from a central core in chromosomes, and from a "cage like" structure in the interphase nucleoids respectively. These loops are topologically constrained since the addition of an intercalating agent such as ethidium bromide alters the sedimentation coefficient of the nucleoids and contracts the DNA halos (Cook and Brazell, 1975). The investigation of loop domains and their attached structures has proven to be difficult, due to the variety of biochemical methodologies used to investigate them. The major forms of analysis are described in Table 1.1. The potential for artefactual attachment, or loss of attachment of DNA sequences during experimentation, raises argument as to which procedure most closely represents the *in vivo* state.

1.5.3 Loop size.

Due to the differences in the nuclear structures retrieved from the loop formation methodologies it is interesting to discuss the different loop sizes, and the proteins and sequences recovered from the procedures. A major problem with comparing the different methods is that they have not been undertaken on the same cell types. Scaffold stabilisation studies have largely concentrated on the yeasts and *Drosophila* Kc cells, while nucleoskeletal studies have largely been done on HeLa cells. A direct comparison of the different protocols for the analysis of loop size has, been undertaken in HeLa cells (Jackson *et al.*, 1990). After labelling with ³H-thymidine the apparent loop size averages are calculated by comparing the relative sizes and amounts of the released and attached fractions. Table 1.1 shows the findings of Jackson *et al.* (1990).

Non-stabilising methodologies give a considerably larger loop size average than the stabilising preparations. Neither system shows any appreciable change in loop size

TABLE 1.1 : METHODS FOR NUCLEAR ATTACHMENT ANALYSIS.

STRUCTURE		PROCEDURE	FUNCTIONAL ASSOCIATIONS	HeLa Loop size	REFERENCE
nucleoids		detergent lysis 2M NaCl extraction digestion of DNA	replication transcription	123kb	Cook <i>et al.</i> , 1984
nucleoskeleton		agarose encapsulation detergent lysis digestion electroelution physiological conditions	replication transcription non transcribed sequences	80kb	Jackson, and Cook, 1985, 1986
scaffolds		nuclear isolation Cu 2+/ heat stabilisation LIS extraction	in vitro topo II binding consensus complex structures	15kb	Mirkovitch <i>et al.</i> , 1986 Gasser and Laemmli, 1986
matrices	type 1	detergent lysis Cu ²⁺ stabilisation 2M NaCl extraction	various systems Lamins A, B, C	48kb	Lebkowski <i>et al.</i> , 1982
	type 2	detergent lysis 2M NaCl extraction	replication transcription		

LOOP SIZES IN HeLa CELLS TAKEN FROM JACKSON *et al.*, (1990)

average during the cell cycle.

1.5.4 Transcription and replication occurs on nuclear cages.

Nuclear cages / type II matrices are obtained by extracting histones from nuclei with 2M NaCl washes. What remains is a fibrous structure, with loops of DNA emanating from this "cage like " structure. Nuclear RNA hybridised to these structures preferentially hybridises to the DNA at the base of the loops in the cage like structure (Jackson *et al.*, 1981). Influenza virus is an RNA virus; infecting negative strands are first transcribed and the resulting positive strands are then replicated into new virion RNA. After labelling infected chicken fibroblasts with ^3H uridine, it was shown by hybridisation that both negative and positive strands were closely associated with the cage. In contrast a control rhadovirus, whose reproduction is cytoplasmic was not associated.

Replication also appears to be associated with this nuclear cage structure. (Dijkwel *et al.*, 1979; Mc Cready *et al.*, 1980) When cells are pulse labelled with ^3H -thymidine prior to matrix isolation and unattached DNA removed by nuclease digestion, the bulk of the labelled DNA remains associated with the cage/matrix. If labelling time is increased and the matrices observed by E.M., the labelled DNA can then be detected in the skirt of unattached DNA. These experiments, carried out by high salt extraction of nuclei, suggest that replication and transcription occur at discrete sites within nuclei, and that there is some form of association between the polymerases and the nuclear architecture. It could be suggested that these results are induced by the artificially high salt conditions. Support for a structural association of the replication machinery also comes from light microscopy of pulse labelled nuclei (Nakamura *et al.*, 1986; Nakayasu and Berezney, 1989; Mills *et al.*, 1989). When replicating nuclei are pulse labelled with either bromo-deoxyuridine (BrdU) or biotinylated deoxy-uridine 5' triphosphate (dUTP), sites of replication can be visualised microscopically. Optical sections through labelled nuclei show replication is confined to a few hundred foci, each of which must consist of several hundred replication forks, that appear to remain clustered during S-phase. Since there is no evidence that nuclear DNA exists as arrays of parallel bundles of fibres it must be assumed that since these foci remain clustered that the surrounding DNA is spooled through these fixed replication sites. These experiments demonstrate that transcription and replication occurs at sites attached to the nuclear cage.

1.5.5 The nucleoskeleton, transcription and replication

The structural association of replication and transcription is substantiated by work

involving physiologically buffered systems. Cells are encapsulated in agarose beads, permeabilised by the addition of detergent and then the unattached chromatin removed by electrophoresis after nuclease digestion of chromatin (Jackson and Cook, 1985). The remaining nucleoskeletal structure can then be assayed for residual nuclear activities and also for attached sequences. Electron microscopy of these residual structures reveals that the remaining chromatin is attached to some form of nucleoskeleton. It is proposed that intermediate filaments may be part of this supporting structure (Jackson *et al.*, 1988).

This residual chromatin and nucleoskeleton can be shown to preferentially contain active genes and also the replicative and transcription activities of the cell. When cells are labelled with ^3H -uridine prior to nucleoskeletal preparation, 92% of the nascent RNA transcripts remain attached, despite only as little as 2% of the chromatin remaining after extraction (Jackson and Cook, 1985). If RNA polymerase activity is assayed after unattached chromatin extraction, by labelling transcripts with ^{32}P UTP, approximately 80% of the activity is attached to the nucleoskeleton regardless of which nuclease digestion is used prior to extraction. If bulk DNA is compared to nucleoskeleton attached DNA in *polyA*+ RNA hybridisation, then nucleoskeletal DNA is seen to be enriched for transcribed sequences, as observed with cages and type II matrices (Jackson and Cook, 1985). Specific sequences have also been shown to be enriched. Ribosomal sequences are retained on the nucleoskeleton (Dickinson *et al.*, 1990). It is known that ribosomal sequences exist in two different chromatin conformations (Kamakaka *et al.*, 1990). Accessibility studies suggest that 30% of ribosomal sequences are relatively inaccessible to nucleases and therefore presumably inactive (Weisbrod *et al.*, 1982). After extraction around 70% of ribosomal sequences are attached to the nucleoskeleton, this presumably correlates to the active fraction (Dickinson *et al.*, 1990b). If replication is studied with the addition of ^3H -thymidine, 90% of the DNA polymerase activity is retained in the nucleoskeleton (Jackson and Cook, 1986). This activity was assayed during the cell-cycle and was shown to be S-phase cell specific. Replication foci can be directly visualised upon the nucleoskeleton after appropriate labelling procedures demonstrating that this function is indeed associated with the nucleoskeleton (Hozak *et al.*, 1993)

These studies suggest that the nucleoskeleton, and therefore loop formation, involve an equilibrium of different attached sequences at different times in the cell cycle as sequences are transcribed and replicated. It is also implied that, opposed to conventional wisdom, replication and transcription activities occur in association with the

nucleoskeleton at fixed sites at any particular moment. This is not to say that they are literally fixed in space in the nucleus during the cell- cycle.

1.5.6 Scaffold attachment.

Studies of the nuclear scaffold lead to an opposed view of loop formation. In LIS extracted nuclei a reproducible set of proteins and sequences are recovered in the attached fraction (Mirkovitch *et al.*, 1984). This reproducibility only occurs if the cells or isolated nuclei have undergone heat shock or divalent cation stabilisation, before extraction. These treatments have been shown to alter the solubility characteristics of proteins, by as yet unknown mechanisms. Nuclear complexes can also be formed by other cellular insults such as arsenite or amino acid analogues which have been shown to induce heat shock genes. Cellular responses to stress appear therefore to be a major part of nuclear scaffold formation (Littlewood *et al.*, 1987).

Scaffold attachment regions (SARs) have been mapped in many cell lines, around various genes. The *Drosophila* histone gene cluster binds to the scaffold in the H1-H3 spacer region, with two major sites of attachment (Mirkovitch *et al.*, 1984). SARs have also been located 5' and 3' of the *Drosophila Adh*, *SgS-4* and *Ftz* genes. These SARs appear to co-map with upstream regulatory regions of the genes concerned, and form loops of between 4.5 and 13Kb. These sites all contain the *in vitro* topoisomerase II (topo II) consensus cleavage sequence (Gasser and Laemmli, 1986). In *S. cerevisiae* SARs appear to mainly co-map with putative origins of replication, autonomous replicating sequences (ARS), and also with the centromeric sequences (Amati and Gasser, 1988). ARS associated scaffold binding in *S. cerevisiae* has also been demonstrated for the E silencer region of the *HMR-E* silent mating locus (Abraham *et al.*, 1984). The 3' ends of the histone *H4* gene and the *HO* gene of *S. cerevisiae* have ARS sequences which are scaffold attached (Bouton and Smith, 86; Kearsley, 1984). Not all SARs in *S. cerevisiae* are ARS sequences, as can be shown by the physical separation, by restriction digestion, of the ARS and SAR functions of the *ftz* SAR (Amati *et al.*, 1990). An 800kb region of the *Drosophila* genome has been analysed for ARS function, in *S. cerevisiae* and SAR attachment (Brun *et al.*, 90). This study showed that the ARS function was always adjacent to, or superimposed on, SAR sequences. This close association suggests some common sequence or structural features shared between these sequence classes. These ARS sequences were only a subclass of the total SAR population.

1.5.7 Scaffold binding conservation.

Scaffold binding is conserved across species boundaries. *Drosophila* SARs if added to rat liver scaffold preparations, bind these heterologous scaffolds (Izzuralde *et al.*, 1988). *S. cerevisiae* and *S. pombe* SARs bind scaffolds from the heterologous species, *Drosophila* SARs also bind to yeast scaffold preparations (Amati and Gasser, 1988). These reconstitution studies also demonstrated that SARs vary in their affinities for their own and heterologous scaffolds. The SAR regions were found to be attached independently of the activity state of the gene, they surround (Gasser and Laemmli, 1986). In *Drosophila*, SARs bind to both interphase and metaphase scaffold preparations. This implies that scaffold attachment may occur permanently throughout the cell-cycle, and also in a non-tissue specific manner (Mirkovitch *et al.*, 1988). As well as containing the *in vitro* topoII consensus cleavage site (Gasser and Laemmli, 1986), it has been observed that SARs contain other recognisable motifs. The presence of A and T boxes has been found in these regions. These boxes are characterised as being approximately 10 bp in length, and were initially proposed as potential orientational sequences for the SAR. Intact SAR regions are highly (70%) A-T rich, and statistically only the T box is found preferentially in SARs (Amati and Gasser, 1990). It is proposed that this box may be involved in SAR attachment by causing helix distortion.

Scaffold attachments have also been mapped flanking the Chinese hamster *DHFR* gene (Dijkwel and Hamlin, 1988), human interferon *B* gene (Bode and Maas, 1988), human apolipoprotein *B* gene (Levy-Wilson and Fortier, 1989), chicken lysosyme gene (Stratling and Dolle, 1986), human β -globin gene (Jarman and Higgs, 1988), and human telomeric sequences (De Lange, 1992). The above SARs are also A-T rich, but not all of them contain the *in vitro* topoII consensus, and not all *in vitro* topoII consensus sequences fall within SARs (Jarman *et al.*, 1988). SARs in these systems appear to be associated with the limits of the chromatin domains of the genes concerned.

1.5.8 Topoisomerase II and scaffold attachment.

Topoisomerase II is the major protein of the metaphase scaffold (Gasser and Laemmli, 1986; Earnshaw *et al.*, 1985). 60% of topoII remains associated with the metaphase scaffold after LIS extraction. Topo II has been shown to be essential for proper chromosome condensation and segregation during mitosis in *S. pombe* (Uemura *et al.*, 1987) and for condensation in higher eukaryotes (Adachi *et al.*, 1991). Antibody labelling of topoII in metaphase chromosomes has shown the protein to be associated with the axial core of chromosomes. The requirement for this protein in chromosome condensation and segregation and also its association with SARs implies that topoII is a candidate for a major loop anchoring protein however, topo II levels are found to be

high only in proliferative cells, in non-proliferating cells such as terminally differentiated chicken erythrocytes, topoII levels have been shown to be less than 300 molecules per cell (Heck and Earnshaw, 1986). Recovery of topoII cleavage intermediates *in vivo*, has allowed the analysis of topoII's true cleavage site (Kas and Laemmli, 1992). Chemical recovery of intermediates shows that topoII cleavage occurs *in vivo* at a G-C rich consensus sequence, rather than at the A-T rich *in vitro* consensus. This G-C rich cleavage core is generally flanked by A-T rich sequences. Interestingly, it was found that the *Drosophila* H1-H3 spacer was cleaved *in vivo*, in the linker region between the positioned nucleosomes found in this region, rather than as a broad region of binding found after LIS extraction.

1.5.9 Scaffold loop boundaries.

When genes are transposed to new sites by rearrangement or transfection their expression may be affected by the influences of the new chromosomal environment. It has been proposed that since SARs are permanently attached during the cell cycle they should protect genes within their domains from effects of surrounding chromatin. Constructs made from the chick lysosome 5' SAR flanking a reporter gene, result in stable levels of reporter gene expression, free from position dependent effects (Steif *et al.*, 1989; Phi-van *et al.*, 1990). The *Drosophila Hsp70* genes are bounded by two specialised chromatin structures (*scs*, *scs'*) which are proposed to be associated with the boundaries of the gene domain and sites of topoII binding (Udvardy *et al.*, 1985). These potential SARs have been shown to confer position independent stable levels of expression on a reporter gene. Transgenic fly lines produced with a construct containing the *scs* flanking a reporter *white* gene result in position independent expression if *scs* sequences are 5' and 3' of the reporter gene (Kellum and Schedl, 1991). These *scs* sequences have also been found to inhibit enhancer activity when placed between promoter and enhancer. This is not true of all SARs since the SAR from the *Drosophila Hsp 70* intergenic spacer has no such enhancer blocking ability. (Kellum and Schedl, 1992). Due to the linear differentiation observed in *Drosophila* polytene chromosomes, it is attractive to propose that band boundaries might be associated with scaffold attachment. It has been shown however that mapped *Drosophila* SARs on polytene chromosomes are not coincident with polytene band boundaries (Mirkovitch, 1986). Recently a protein B52 associated with boundaries of transcriptionally active chromatin in *Drosophila* has been characterised (Champlin *et al.*, 1991). This protein contains a motif, conserved in RNA binding proteins and single stranded DNA binding proteins, called RNA recognition motif (Query *et al.*, 1989) and also a regular repeat of serine and arginine or lysine, and is consequently very basic and reminiscent of the

histone displacing protamines. It is unknown whether this protein is associated with the scaffold.

1.5.10 Scaffold proteins.

Although topoII is the major protein isolated from the metaphase scaffold, can it be the major loop formation protein in cells such as chicken erythrocytes where topoII is present at less than 300 molecules per cell? Other proteins have also been isolated from scaffold preparations. RAP-1 fractionates in *S. cerevisiae* scaffold preparations and binds the silencer *E* and *I*, and promoter *P* elements of the *HML* loci. RAP-1 is involved in loop formation at these loci and is presumed to play a role in HML silencing through loop formation. (Hoffman *et al.*, 1989) ACBP, ARS consensus binding protein has been purified from scaffold preparations (Hoffman and Gasser, 1991). This protein shows sequence specific single stranded DNA binding to the T-rich strand of the ARS consensus. This protein is proposed to be involved in replication initiation. ARBP, attachment region binding protein has been isolated from chicken scaffolds. (Von Kreis *et al.*, 1991). This protein has been shown to bind preferentially to the chick lysosyme SAR and to some SARs from *Drosophila*, mouse and human genes.

Evidence suggests that sites are bound in a cooperative manner in loop formation. Nuc-2 protein from *S. pombe* scaffolds is required for mitotic disjunction. This protein has motifs which are proposed to be involved in DNA binding and scaffold attachment. Nuc-2 contains 10 repeating units called tetratricopeptide repeat (TRP) domains which form snap helices which may stabilise and aggregate with other nuc-2 like proteins upon DNA binding. A serine rich DNA binding domain interrupts the snap helices and appears to bind A-T rich DNA *in vitro* (Hirano *et al.*, 1990) This protein shows similarity to *S. cerevisiae* proteins SK13 and SSN6 and the *Aspergillus nidulans* bimA protein. A family of such proteins could form part of the nucleoskeleton aggregation with their snap-helix motifs (Hirano *et al.*, 1990). Many other proteins are observed by Western analysis to be retained on the complex scaffold structures (Cardenas *et al.*, 1990), but what their roles are in these structures, remains to be resolved.

Whether the two main methods of analysing nuclear attachment reflect the same *in vivo* loop structures appears unlikely. They may reflect two different loop forming pathways utilised in the cell. The matrix/nucleoskeletal protocols may reflect functional associations of DNA for loop formation and be the major association during interphase, while the scaffold procedures may reflect the later stages of packaging required to condense chromosomes.

1.5.11 Packaging loops into chromosomes

E.M. studies of human and mouse metaphase chromosomes suggest that the chromatin loops are formed into a 200-300nm fibre which is then helically coiled in metaphase (Rattner and Lin, 1985). HeLa metaphase chromosomes depleted of histone H1 and then labelled with an anti topo II antibody, show a helical staining of the topoII protein along the chromosome. It is also observed that the coils of sister chromatids have an opposite helical handedness (Boy de la tour and Laemmli, 1988). At present we know very little about these highest levels of packaging.

1.6 LINEAR DIFFERENTIATION OF CHROMOSOMES.

Metaphase chromosomes are not homogeneously packaged along their length, linear differentiation of regions is observed with a variety of techniques. After compaction of chromosomes, constrictions may be observed along their length, inhibitors of replication may induce fragility in specific regions, and various staining procedures allow visualisation of differential structure or activity, as a banding pattern, along the chromosomal length.

1.6.1 Secondary constrictions

A condensed chromosome exhibits a primary constriction associated with centromeric sequences, (discussed later) but may also exhibit secondary constrictions. Most secondary constrictions are the result of the activity of the rDNA sequences found in these regions. Ribosomal gene expression occurs in the nucleus at a visible domain termed the nucleolus. Nucleoli are clearly visible in the interphase nucleus. Their presence is the result of the high activity levels of the rDNA sequences, and subsequent ribosome formation in these domains. As cells enter mitosis the nucleoli diminish until only the secondary constrictions at the site of rDNA sequences are apparent. These constrictions are presumed to result from inhibition of higher packaging levels due to the high transcriptional activity of the rDNA genes. Occasionally, especially after cold treatment, heterochromatic, or less frequently, other regions fail to condense and form weakly stainable constrictions (tertiary constrictions; negative heteropycnosis).

1.6.2 Chromomeres

Linear differentiation of chromosomes may also be observed in meiotic cells where the chromosomes are only partially contracted or in *Diptera* cells containing polytene chromosomes, where the replication of DNA leads to multiple arrays of chromosomes aligned laterally in perfect register. These polytene chromosomes have a banded

appearance due to the differential packaging levels along their length. These highly compact bands are known as chromomeres. Chromomeres can be disrupted by the induction of the gene or genes contained within them. Heat shock can be observed to induce the chromomeres containing the heat shock genes to decondense and puff out as high levels of transcription occur, these puffs are known as Balbiani rings. In oocytes of many species the meiotic chromosomes, are roughly 200 times less compact than in mitosis (Bostock and Sumner, 1978), and differential condensation can be observed, with highly condensed regions, chromomeres, are interspersed by interchromomeric regions. These chromomeres are associated with loops which contain expressing genes as assessed by autoradiography after radiolabelling transcripts (Reiger *et al.*, 1976).

1.6.3 Fragile sites

Chromosomal gaps as opposed to constrictions, may be induced by inhibiting replication at potential fragile sites (Laird *et al.*, 1987). ³H thymidine labelled *Microtus* chromosomes which are subjected to premature chromosome condensation, induced by cell fusion (Sperling, 1974), show chromosomal gaps corresponding to late replicating regions. In *Drosophila* "weak sites" in polytene chromosomes can be induced by the replication inhibitor fluorodeoxyuridine (Hagle, 1972). Fragile sites, for example human Xq27, which are cytogenetically observed as an absence of higher packaging can be induced by shortening the G2 phase of the cell cycle, for example by cell fusion or by addition of caffeine after folate deprivation (Yunis, 1981). A number of other agents influence the frequency of fragile sites in higher eukaryotes including aphidicolin, which inhibits DNA polymerase α , fluorodeoxyuridine, which inhibits thymidine synthetase, and BrdU. All of these agents appear to induce fragility by inhibiting or delaying DNA replication, leading to a short G2 phase. It therefore appears that fragile sites generally are late replicating regions and that inhibiting replication or induction of premature condensation leads to a lack of packaging in these regions.

In the case of the fragile site at Xq27 in the human genome a mutation has been detected which is associated with the expression of the fragile phenotype. A gene FMR-1, which encodes an mRNA expressed in the brain, has been shown to contain a trinucleotide repeat (CGG) domain which has been expanded in affected individuals (Verkerk *et al.*, 1991). Methylation of the CpG island for this gene is also associated with the fragile phenotype.

1.6.4 Chromosome banding.

Various techniques have been developed which demonstrate reproducible linear

differentiation along mammalian mitotic chromosomes. These banding techniques use stains, fluorochromes and base analogs. Banding patterns, are believed to reflect differences in the packaging, composition and activity of euchromatin in the different regions. Banding techniques are applicable to mammals, and some are applicable to reptiles, birds, plants and certain fish (Bickmore and Sumner, 1989). G-banding, (Geimsa staining partially trypsin digested metaphase chromosomes) produces a recognisable pattern of negative and positive bands. These banding patterns, are believed to reflect some difference in the packaging of euchromatin in the different regions. The positive and negative band appearance is reversed for R banding where chromosomes are partially heat denatured prior to staining. The other forms of banding largely correlate to this G-banding pattern. Q-banding using A-T specific fluorochromes e.g., DAPI, quinacrine or Hoescht 33258 shows a direct correlation to G banding. Q-negative banding uses G-C specific fluorochromes such as chromomycin or mithromycin, and shows an inverse correlation to G-banding. Replication banding using BrdU, either early or late in S-phase, reveals a banding pattern similar to G or R banding respectively. A problem of resolution occurs with the number of bands observed. In mid-prophase human chromosomes display over 2000 bands (Yunis, 1981), but many fuse and are masked as condensation occurs, to produce the 300 bands observed at prometaphase (Sumner, 1990). A correlation exists between regions of chromosomes containing genes. G-negative (R), early replicating G-C rich bands contain >80% of all genes as assessed by cytogenetic localisation and mapping of CpG islands (Bickmore and Sumner, 1989). These negative bands also correspond to early replicating regions and late condensing regions during mitosis and to inter chromomeric regions of meiotic chromosomes. The distribution of long and short interspersed repeats (Lines and Sines respectively) also correlate to these banding patterns, as assessed by *in situ* hybridisation. Lines are preferentially located in G-positive bands, Sines in G-negative bands. Both sets of sequences account for around 10% of the genome, and may play some part in producing banding patterns.

1.7 CENTROMERES

1.7.1 Primary constrictions

Mammalian centromeres appear as the primary constrictions within metaphase chromosomes. This constriction, is apparently caused by a lack of helical packaging of the 250nm fibre, in these domains (Rattner and Lin, 1985, 1987). This presumably allows for the correct interactions to occur between centromeric sequences and proteins involved in kinetochore structure and function. *In situ* hybridisation with mouse α -satellite sequences suggests that only partial decondensation of the centromeric region

may occur during interphase. (Rattner and Lin, 1987; Manueldis *et al.*, 1982)

Centromeres are essential for proper chromosome segregation at division. They are the sites at which mitotic or meiotic spindles attach to the kinetochore and also the site of sister chromatid association until the onset of anaphase. The kinetochore, an ultrastructurally defined plate-like structure observed at the outer surface of the centromere, is not only the microtubule attachment site but may also contain the mechano-chemical motor responsible for the movement of the chromosomes towards the spindle poles during anaphase.

1.7.2 Centromere proteins.

Several centromere specific proteins (CENPs) have been identified by autoimmune sera from patients with the CREST syndrome of progressive systemic sclerosis. CREST sera detect 4 proteins which are localised to centromeres; CENP-A, B, C and D. CENP-B is the best characterised, and is an 80Kd protein which can bind human alphoid repeats (Matsumoto *et al.*, 1989). The human and mouse CENP-B cDNAs are 96% homologous, in agreement with the conservation of the DNA sequences which may be bound by the respective proteins in mouse and man. CENP-B appears to be permanently bound to centromeric regions as immunofluorescence detects CENP-B as a localised signal in interphase nuclei (Cooke *et al.*, 1990) CENP-B is localised to the alphoid sequences beneath the kinetochore plate, (Cooke *et al.*, 1990) and its levels vary between centromeres due to differences in sequence composition and size of α -satellites regions. CENP-A (17kD) appears to be a centromere specific histone related to histone H3. Its distribution in the centromeric domain is unknown (Palmer *et al.*, 1987). CENP-C (140kD) appears to be associated exclusively with functional kinetochores i.e. binds only active centromere of dicentric chromosomes (Rothfield *et al.*, 1985; Earnshaw *et al.*, 85). CENP-D (50kD) was initially detected in rodent extracts and then shown to be present in human extracts by Western analysis (Kingwell and Rattner, 1987; Earnshaw *et al.*, 1985). The relationship between these proteins and kinetochore structure and function was investigated by injecting purified IgGs from CREST sera into cells. It was found that these IgG preparations disrupted mitotic events in these cells (Bernat *et al.*, 1990). If introduced before the S/G2 transition then the assembly of the kinetochore was disrupted, and introduction during G2 lead to formation of a fragile kinetochore and mitotic arrest in metaphase (Bernat *et al.*, 1991). The common protein detected between the different IgG preparations was CENP-B, suggesting that it is central to centromere function. Two murine monoclonals raised against CENP-B however failed to affect mitosis when injected into cells, although whether functionally important epitopes were bound is questionable (Bernat, 1990).

CENP-B fusion proteins have recently been overexpressed in human cells and fail to inhibit mitosis. Expression of deletion forms of CENP-B in human cells has shown that a 158 amino acid domain is responsible for centromere localisation. Binding of CENP-B proteins, deleted for all but this signal sequence, to the alphoid sequences, fails to inhibit mitosis (Pluta *et al.*, 1992). It may be that the IgG preparations used for microinjection, neutralised low copy number proteins which could not be detected by Western analysis. That CENP-B is not essential for kinetochore formation is suggested by the lack of detectable CENP-B at the Y chromosome centromere.

1.7.3 Centromeric DNA

Although centromere specific proteins have been detected, it is also important to analyse the DNA sequences that they interact with. 5% of the human genome is made up of families of alphoid repeats, which localise to the centromeres. These repeats of α -satellite DNA are arranged in arrays of different subclasses spanning contiguous centromeric regions. An alphoid subclass has been shown to be a site of anticentromeric antibody binding. This subclass contains a 17bp consensus motif, that has the ability to bind CENP-B *in vitro* and this motif is named the CENP-B box (Matsumoto *et al.*, 1989). Interestingly the CENP-B box is absent from the Y chromosome alphoid.

In mouse there are two predominant families of repeat associated with centromeric regions. The major satellite is a 234bp monomer, which appears to be evolved from smaller ancestral units. The minor satellite is a 120bp monomer and is believed to be evolved from the major satellite (Wang *et al.*, 1988). Minor satellites contain the same CENP-B box found in some human alphoid repeats. In *Mus musculus* and *Mus spretus* both major and minor satellite forms are found. In *Mus caroli* only the major satellite sequence is observed by *in situ* hybridisation, and it is not confined to the centromere but is found in the chromosome arms. Major and minor satellites of *Mus musculus* and *Mus spretus*, show localisation to the centromeres of all chromosomes except the Y. In *Mus musculus* the minor satellite is localised to the kinetochore region, while in *Mus spretus* the minor satellite is dispersed throughout the centromeric domain (Rattner *et al.*, 1990). These differences in distribution and presence of the satellite subclasses complicates conclusions drawn as to their importance in kinetochore formation, and these results were obtained by *in situ* hybridisation, it is possible to question whether, some sites are protected from hybridisation and thus are not detected.

1.7.4 Yeast centromeres

As compared to the large and highly complex nature of higher eukaryotic

centromeres, those of the yeasts are fairly simple and more accessible to analysis. The *S. cerevisiae* centromere is contained within 125bp of defined sequence that contains no repeats, but is made up of 3 elements; a 78-86 bp A-T rich (>90%) region flanked by two conserved domains. Single base pair changes or short deletions can completely inactivate centromeric function. *S. cerevisiae* centromeres are not chromosome specific, and may retain some form of spindle attachment throughout the cell cycle, since the spindle is present for most of the cell cycle (Bloom and Carbon, 1989). These centromeres are flanked directly by transcribed genes. *S. cerevisiae* may not represent a good model for the study of mammalian centromere structure and function since the spindle appears to be permanently attached throughout the cell cycle, lack of chromosome condensation and closed mitosis (no nuclear membrane breakdown) .

By contrast the mitotic cycle of the fission yeast *S. pombe* is more similar to that of other higher eukaryotes. *S. pombe* centromeres are also larger and more complex than those of *S. cerevisiae* and contain repetitive DNA sequences. *S. pombe* has three centromeres which are 40kb-*cen1*, 80Kb-*cen2*, 110Kb-*cen3* (Murakami *et al.*, 1991). This is intriguing since the size of centromeres seems inversely related to chromosome size. Each centromere consists of three regions, a central core(cc), flanked by innermost repeats (imr), which in turn are surrounded by repeating elements. The central core regions of *S. pombe* centromeres are approximately 5kb and contain unique sequences , although *cen1* and *cen3* share the same highly homologous sequence tm.. Tm is assumed to be non-essential for centromere function since it is absent from *cen2*. These central core sequences are not transcribed and exhibit a non-nucleosomal structure in micrococcal nuclease digests (Pillozzi and Clarke, 1991). If these core sequences are transferred to *S. cerevisiae*, on a minichromosome, they adopt a normal nucleosomal pattern, demonstrating that specific proteins are required to produce this structure (Polizzi and Clarke, 1991; Chirashige *et al.*, 89)

Innermost repeat sequences flank the central cores. Three different classes of imr are found within the *S. pombe* genome with each represented twice, flanking their respective centromere cores. These repeats do not cross hybridise between centromeres apart from their *t-RNA* gene content. *t-RNA* genes show a high degree of clustering in these centromeric regions (Kuhn *et al.*, 1991; Takahashi *et al.*, 1991), 10% of *t-RNA* genes are localised to 1.8% of the genome. Three of these *t-RNA* genes are unique in the genome, and are therefore presumably expressed. Previous to the localisation of these genes it was thought that no transcription appeared to come from centromeric regions (Clarke and Baum, 1990). These imr regions exhibit a disrupted nucleosomal pattern with micrococcal digestion (Polizzi and Clarke, 1991).

FIGURE 1.5 STRUCTURE OF *S. pombe* CENTROMERES

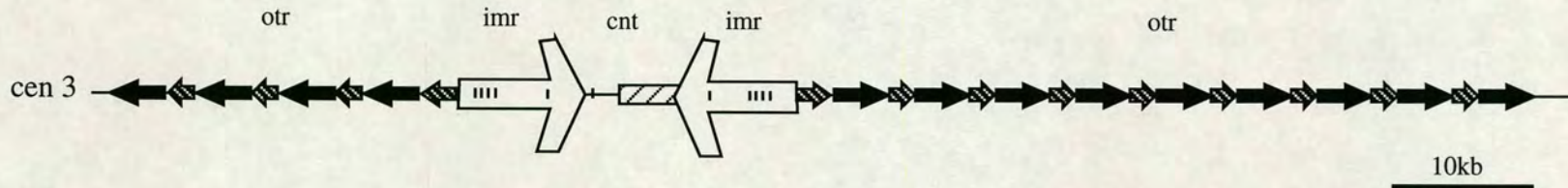
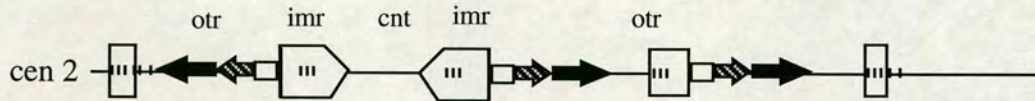
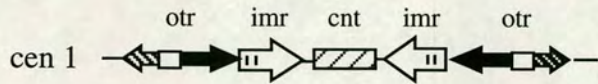


Figure 1.5. Shows the structure of the centromeric regions of *S. pombe*. The arrows in the outer repetitive regions (otr) represent the common repetitive elements dg and dh (filled and hatched arrows respectively). The large open arrows represent the innermost repeat regions (imr) which are centromere specific. The centre sequence (cnt) is also non homologous between centromeres apart from the boxed tm sequences of cen 1 and cen3. the vertical lines represent tRNA genes. The distal parts of imr2 are duplicated in otr2 indicated by the open boxes.

50

These central regions are then surrounded by repeats. Depending on the source of report, these repeats are termed either dg, dh (Nakaseko *et al.*, 1986, 1988; Chikashige *et al.*, 1989) or K,L,B and J (Clarke and Baum, 1990), of which the former two are equivalent to K and L respectively, while B and J are effectively ignored in this terminology. These repeats combine in different patterns to produce centromere specific arrays. dg and dh form a 6.6kb repeating motif, which is repeated twice in *cen1*, three times in *cen2*, with a truncated 4.8kb version repeated 13 times in *cen3*. This truncation is due to a shorter 1.6kb dh sequence, while a longer 3.8kb version is located beside the innermost repeat. These repeats show a normal nucleosomal configuration. It appears that dg and dh are required for meiosis 1 whilst imr is required for mitosis and meiosis (Murakami *et al.*, 1991). How these centromeric sequences would interact with the proteins of higher eukaryotes is unknown, although the lack of sequence similarity between *S. pombe* centromeres and mouse centromeric regions would lead to the conclusion that no specific or structural interactions would necessarily occur. These relatively complex centromeres may prove to be a better model system for those of higher eukaryotes, than those of budding yeast..

1.8 TELOMERES

The ends of the linear chromosomes of eukaryotes would be susceptible to fusion or degradation if, the ends were not protected by telomeres. Telomeres however in the short term are not essential elements of a chromosome. *Drosophila* mutants have been created which lack telomeres upon specific chromosomes (Levis, 1989; Biessmann and Mason, 1988). These terminal deletions were created by destabilisation of a subtelomeric P-element (Levis, 1989) or X irradiation in a *mu 2* background (Biessmann and Mason, 1988). The terminal sequences of the mutated chromosomes were located within known gene sequences. The lack of telomeres was not lethal to the cells, which continued to divide for several divisions. However of these ends was found to occur over time. The senescence of these cell lines shows that, although, telomeres are not essential for cell viability, they are important for chromosome integrity.

Telomeric sequences have been sequenced from many species and have been found to consist of tandem arrays of short repeats. The human and possibly all vertebrate telomeres repeats consist of TTAGGG (Cross *et al.*, 1989), while *S. pombe* repeat consists of T₁₋₂ ACA₀₋₁C₀₋₁G₁₋₆ (Sugawara *et al.*, 1986). Telomeres vary greatly in size between different species, *S. pombe* telomeres have been demonstrated to be roughly 300 bp in length, (Allshire *et al.*, 1988) while human telomeres comprise 10kb (Allshire *et al.*, 1989) of sequence on average and those of mouse can be as large as



150 kb (Kipling and Cooke, 1990; Starling *et al.*, 1990). Telomeres are not of a constant size. Harley *et al.* (1990) has shown that in primary human cell lines telomeres decrease in length during time in culture and have calculated that on average 50bp of telomeric repeat is removed at each subsequent division. A budding yeast strain containing a mutation of the *Est -1* (ever shorter telomeres) gene also displays a phenotype of decreasing telomeric length leading to eventual death of the cells (Lundblud and Blackburn, 1989). *Est-1* has been postulated to be a component of the complex involved in telomere maintenance. The length of telomeres appears to be maintained by an equilibrium of telomeric sequence addition and loss. It has been demonstrated that long telomeres are found upon the chromosomes of cells from sperm and foetal tissue while blood from ageing donors and rapidly turning over tissues such as the colonic mucosa contain shorter telomeres (Hastie *et al.*, 1990; Harley *et al.*, 1990; Cooke *et al.*, 86). Decreasing telomeric length has also been observed in certain cancers which may reflect the number of divisions the cancer cells have undergone. This loss of telomeric sequence suggests that somatic tissues have lost the capacity for maintenance of telomere length. It has been proposed that this loss of telomeric sequence may lead to senescence of cells (Greider, 1990). Immortalised cell lines stabilise their telomere lengths at a shorter length than that found in tissues (De Lange *et al.*, 1990; Kipling and Cooke, 1990), as exemplified by the comparison of telomeres from C127 cells and sperm (Kipling and Cooke, 1990).

Telomere addition has been found to occur by the addition of telomeric repeat sequence subunits. This telomeric sequence addition is catalysed by a ribonucleoprotein enzyme telomerase. Telomerase is an unusual ribonucleoprotein reverse transcriptase, which polymerises the addition of telomeric sequences to the G-rich strand, under the guidance of a complementary RNA template (Greider and Blackburn, 1989; Shippenlantz *et al.*, 1990). Telomerase activities have been demonstrated *in vitro* in ciliates (Blackburn, 1991) and human HeLa cell free extracts (Greider and Blackburn, 1989; Morin, 1989). The RNA component of telomerase has been shown to act as a template for the addition of telomeric sequences (Shippenlantz *et al.*, 1990; Greider and Blackburn 89). This was most elegantly demonstrated by Yu *et al.* (1990) who demonstrated that mutations within the template sequence of the RNA species lead to a corresponding change in the telomeric sequences added. *Tetrahymena* or oxytricha telomeres introduced into yeast on linear plasmids are modified by the addition of yeast telomeric sequences (Szostak *et al.*, 1982; Shampay *et al.*, 1984; Pluta *et al.*, 1984). In mammalian cells breakage of human chromosome 16 (Wilkie *et al.*, 1990) is followed by addition of telomeric sequences and in Chinese hamster cells Farr *et al.* (91) has demonstrated that introduced linear constructs which cause chromosome breakage upon

integration have telomeric sequences added to their ends when located at the end of a chromosome. These findings demonstrate that the addition of telomeric sequences to chromosomal ends may be a general phenomena. It should also be noted that in certain instances above (Matsumoto *et al.*, 1987; Wilkie *et al.*, 1990; Kamper *et al.*, 1989) that no telomeric or subtelomeric sequences were located at, or in the proximity of, the end of the chromosome to which the telomeric sequences were added. Thus demonstrating that there are no extended sequence requirements for this function to occur. Such activities may be important in consideration of the transgenome structure of the F7.2 cell line.

In this project I have set out to investigate the basis of the novel chromatin structure observed within the *S. pombe* sequence domains of the hybrid cell lines established by Allshire *et al.* (1987). I have concentrated on three main areas, methylation, nucleosome structure and chromatin loop formation which may affect chromosome structure and lead to observable differences at metaphase between these *S. pombe* transgenomes and murine genomes. It is hoped that this analysis undertaken will lead to a better understanding of mammalian chromosome structure. Our visible assay, of differential chromosome packaging, should allow correlations to be derived from comparisons of molecular characteristics of the *S. pombe* and murine domains and their relative affects on chromosome structure.

CHAPTER 2. MATERIALS AND METHODS

2.1 Mammalian Cell Culture

2.1.1 Thawing of cells stored in liquid nitrogen

All cell lines used in this project were stored in liquid nitrogen in foetal calf serum supplemented with 10% DMSO. Cell lines were recovered from their frozen state by rapidly thawing them in a 37°C water bath and then resuspending them in 9ml of the appropriate medium in a sterile 12ml centrifuge tube. Cells were pelleted for 5 minutes at 1000g. The cell pellet was then resuspended in 5ml of the appropriate media and transferred to a 25cm³ tissue culture flask (Nunc).

2.1.2 Maintenance of cell lines in culture

Two types of cell line were used in this project and each required different culture conditions.

Mouse C127

This is a mouse mammary tumour cell line. The cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco Biocult) supplemented with 10% foetal calf serum (FCS) (Gibco Biocult) which had been inactivated for 30 minutes (mins). Cells were grown as an attached culture at 37°C with a 10% CO₂ atmosphere, in tissue culture flasks (Nunc).

Hybrid mouse *1S. pombe* lines F1.1, F7.2, DC.1, DC.2

These lines are cultured similarly to C127 cells, with the addition of 400µg/ml of G418 (Gibco).

2.1.3. Transfer and harvesting of cells.

Confluent cultures were washed with sterile PBS at 37°C and released by addition of a 1:10 trypsin: versene solution, at 37°C for 5 minutes. 10ml of the appropriate medium were then added and cells counted using a haemocytometer. Cells were then harvested by centrifugation for 5 minutes at 1000g, or split appropriately to new sterile flasks and the appropriate medium added. Where procedures vary from this, this will be specifically mentioned.

2.1.4 Chromosome preparation.

Metaphase chromosome spreads were prepared from colcemid blocked cells. After harvesting cells were swollen in hypotonic media (0.07M KCl) for 20 minutes at room temperature, before fixation in 3:1 methanol: acetic acid for 20 minutes. Cells were resuspended in fresh fix and spread by dropping onto cleaned microscope slides. After drying slides were stored in a vacuum dessicator to age before use in *in situ* hybridisation.

2.2 Bacterial Cell Culture

2.2.1 Media and Additives

All media were sterilised by autoclaving.

L-Broth and Agar

Per litre : 10g tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, 2.46g MgSO₄ pH 7.2. The agar contains 15g agar in addition.

Terrific Broth

To 900ml of dH₂O add 12g bacto-trytone, 24g bacto yeast-extract, 4ml glycerol after autoclaving, and just prior to use add 100ml of sterile 0.17M KH₂PO₄, 0.72M K₂HPO₄.

Media Additives

When appropriate, antibiotics were added to the media. Ampicillin was used at 25 - 50µg/ml; tetracycline at 12 - 15µg/ml.

2.2.2 Bacterial Strain Used

E. coli strain DH1: supE44, hsdR17, recA1, endA1, gyrA96, thi-1, rel A1.

2.2.3 Bacterial cell culture.

After streaking cells on the appropriate selective solid media the plates were grown inverted over night at 37°C.

Colonies to be grown up were picked from the appropriate plate and grown in the selective liquid medium overnight at 37°C.

Any variation to these procedures will be mentioned in the relevant place.

2.3 Yeast cell culture

2.3.1 *S. pombe* strain used.

S. pombe Int-5: cdc2.33 leu1.32 ura4.294 h⁻

This strain was created by the insertion of an SV₂neo/*ura4* construct into the *ura4* locus of strain ED628 (Allshire *et al.*, 1987).

2.3.2 Solutions

Edinburgh minimal medium (EMM): 14.7mM potassium hydrogen phthalate, 15.5mM Na₂HPO₄, 93.5mM NH₄Cl, 2%w/v glucose, 20ml/l Salts (50 x stock), 1ml/l Vitamins (1000 x stock), 0.1ml/l Minerals (10000 x stock). For the growth of *S.pombe Int-5* strain L-leucine was added to 0.5mg/ml. After autoclaving a few drops of preservative (1:1:2 chlorobenzene/ dichloroethane/ chlorobutane) is added.

YE medium: 0.5% (w/v) Oxoid yeast extract, 3% (w/v) glucose.

Solid media were produced by the addition of 2% Difco bacto-agar

Salts (50 x stock): 0.26M MgCl₂.6H₂O, 4.99mM CaCl₂.2H₂O, 0.67M KCl, 14.1mM Na₂SO₄

Vitamins (1000x stock): 4.2mM pantothenic acid, 81.2mM nicotinic acid, 55.5mM myo-inositol, 40.8mM biotin.

Minerals (10000 x stock): 80.9mM boric acid, 23.7mM MnSO₄, 13.9mM ZnSO₄.7H₂O, 7.4mM FeCl₂.6H₂O, 2.47mM Molybdic acid, 6.02mM KI, 1.6mM CuSO₄.5H₂O, 47.6mM citric acid.

2.3.3 Culture of *S. pombe Int-5* cells.

S. pombe Int-5 cells were maintained on plates, on solid YE media, at 4°C. Fortnightly restreaking was required to maintain a good stock. *S.pombe Int-5* was also maintained as glycerol stocks -70°C, in YE which had been adjusted to 30% glycerol.

S.pombe Int-5 cells were grown on YE solid plates at 25°C.

S.pombe Int-5 cells were grown in EMM to a density of 10^7 cells/ml at 25°C.

2.4 Manipulation of Plasmid DNA

2.4.1 Small scale plasmid preparations

The Alkaline Lysis Method

The method used was a modification of the methods described in Sambrook, Fritsch and Maniatis (2nd edition) Volume 1, p.1.25 - 1.28.

2.4.2 Large Scale Plasmid Preparation

The Alkaline Lysis Method

This method is a modification of that in Sambrook, Fritsch and Maniatis (2nd edition) Volume 1, p. 1.38 - 1.39. The purification procedure was that of equilibrium centrifugation in CsCl-ethidium bromide gradients also in Sambrook, Fritsch and Maniatis Volume 1, p. 1.42 - 1.43 and extraction by organic solvent (p. 1.46).

2.5 Screening of the cosmid library

2.5.1 Plating of the library.

A cosmid library was gratefully received from Prof. Yanagida (Kyoto, Japan). The library was a partial Sau 3AI digest of *S.pombe* genomic DNA in the cosmid vector pSS-10 (Nakaseko *et al.* 1986). 25µl of cosmid library stock was used to infect 125µl of a fresh over night culture of *E. coli* DH1 cells in L-broth. This mixture was then allowed to infect for 30 minutes at 25°C. 50ul of this infection was then plated out on each of two 20x20 cm ampicillin selective plates. These plates were allowed to absorb the fluid, then grown inverted at 37°C for 30 hours. 1000 colonies from these plates were subsequently picked onto two gridded 20x20 cm nylon membranes, to be used as mastercopies of the library. After streaking on the nylon membranes these clones were also dipped into 96 well microtitre plates. The nylon membranes were then grown inverted on 20x20 cm agar plates at 37°C over-night. The microtitre dishes were also grown overnight at 37°C, with gentle shaking and subsequently altered to 30% glycerol and stored at -70°C for future analysis.

2.5.2 Replicating the cosmid library.

The master copies of the library were placed colony side up on Whatman 3MM paper. Another 20x20 cm piece of nylon membrane, pre dampened, was then overlaid on the master copy. This was then covered by a piece of Whatman 3MM paper, and a rolling pin gently rolled over this sandwich to transfer the colonies. A syringe needle dipped in ink was then used to pierce alignment holes through both pieces of nylon. The two nylon membranes were then separated and placed on agar plates. The master copy was then grown for 6hrs, and a second duplicate of the mastercopy produced as above. The replicas were grown overnight to allow colony growth. DNA from the colonies was then immobilised on the membranes as in 2.10.2.

2.6 Preparation of DNA from cultured cells.

2.6.1 Preparation of DNA from mammalian cells.

After harvesting of cells 3mls of lysis buffer (150mM NaCl, 100mM Tris pH 8.0, 100mM EDTA, 0.5% SDS) were added per 2×10^7 cells while mixing. 10 μ l/ml of 10 μ g/ μ l RNase A were added and then incubated at 37 $^{\circ}$ C for 1hour. 50 μ l/ml of 5 μ g/ μ l pronase were added and then incubated at 37 $^{\circ}$ C for at least 4hours. An equal volume of phenol was then added and mixed well. Aqueous and phenolic layers were then separated by centrifugation at 2.5K rpm for 10 minutes at room temperature. The aqueous layer was then removed and mixed with an equal volume of phenol/chloroform / isoamyl alcohol (25/24/1), and mixed well and then spun at 2.5K rpm for 10 minutes at room temperature. The aqueous layer was again removed and mixed with an equal volume of chloroform/ isoamyl alcohol (24/1). This was again mixed well and centrifuged as above. The aqueous layer was again removed and this time mixed with 0.5 volumes of 7.5M NH₄OAc and 2 volumes of absolute alcohol. The precipitate was then spooled out using a sealed glass pastette and allowed to air dry. This was then washed for 15 minutes in 75% ethanol and allowed to air dry. The end of the pastette was then placed in an 1.5ml eppendorf tube and 1ml of T.E. added. The tube was then aggitated at 4 $^{\circ}$ C until the DNA was in solution.

2.7 Preparation of DNA from *S.pombe Int-5* cells.

These preparations were produced by following a procedure of Dr Fantes (Edinburgh, Scotland)

Protoplasting buffer: 50mM citrate-phosphate pH 5.6, 1.2M sorbitol, 40mM EDTA with 0.2% (v/v) β -mercaptoethanol added just before use.

Triton Buffer: 50mM Tris HCl pH 7.5, 50 mM EDTA, 1% Triton X100.

A 1l culture of *S.pombe Int-5* was grown in EMM at 25°C to a density of 10^7 cells/ml. The culture was spun for 5 minutes at 3000 rpm. The cell pellet was taken up in 50ml of 50mM citrate-phosphate, 1% (v/v) β -mercaptoethanol, and cells repelleted. The cell pellet was then taken up in 15ml of protoplasting solution. A 10ml solution of protoplasting solution containing 50mg of Novozyme (Novo Biolabs) and 10 mg Zymolyase 100T (Seikagaku Kogyo Co. Ltd.) was incubated at 37°C until most cells appeared spherical under phase contrast microscopy. The solution was gently layered onto a 20ml sucrose cushion (0.8M Sucrose, 0.2M KCl) and spun at 1500rpm for 5 minutes. The solution was gently aspirated from the pellet which was then resuspended in 25ml of triton buffer and left for 2-3 minutes. This was then spun at 300g for 10 minutes and the supernatant carefully poured off. Nuclei were then pelleted, from this supernatant, at 8000g for 20 minutes at 4°C. The supernatant was then discarded and the nuclei lysed by the addition of 3ml of (100mM Tris. HCl pH7.5, 100mM NaCl, 50mM EDTA, 1%SDS). 15ul of RNase A (20 μ g/ μ l) were added and the solution incubated for 30 minutes at 37°C. 150ul of 5ug/ul Proteinase K were added and the solution incubated at 37°C for at least another 4 hours. The isolation of DNA from this solution was then identical in procedure to that detailed above (Section 2.6).

2.8 Manipulation of DNA by enzymes

Restriction endonuclease digests were carried out according to the protocols issued by the manufacturers using buffers A, B, L, M and H supplied by Boehringer Mannheim. Genomic DNA was digested overnight with approximately 5 - 10 units of enzyme per μ g of DNA. Plasmids and cosmids were digested for 2 hours. The reactions were stopped by the addition of 1x stop mix from a 10x stock (0.2M EDTA pH 8, 15% Ficoll and Orange G dye).

2.9 Separation of DNA fragments by gel electrophoresis.

TAE buffer per litre of 50x stock : 242g Tris base, 57.1ml glacial acetic acid and 100ml 0.5M EDTA pH 8.0

TBE buffer per litre of 1x stock : 54.6g Tris base, 27.5g Orthoboric acid, 4.65g EDTA.

Horizontal agarose gel electrophoresis was used for the separation of DNA fragments of up to 50kb in size. The concentration of agarose used was between 0.5 and 2% made up in 1x TAE buffer or TBE buffer. After running the gels were stained

for 15 minutes in a 0.5 μ g/ml solution of EtBr. The gels were photographed after two 15 minute washes in dH₂O.

Specific fragments of DNA used for cloning or radiolabelling were also isolated by electrophoresis. However the agarose used was of low gelling temperature (LGT) in order that the DNA fragment required (cut out of the main body of the gel under long wave UV light) could be purified using the method of Burmeister and Lehrach, 1989. The gel slice was melted at 65°C and then EDTA pH 8.0 added to 5mM and NaCl to 100mM. The gel was cooled to 37°C and 2 units of agarase (Sigma) added per 100ml of gel. This was incubated for >1hr and then the DNA was phenol/chloroform extracted, precipitated and then resuspended in TE before use.

2.10 Transfer of DNA to Nylon and Nitrocellulose Membranes

2.10.1 Southern transfer of DNA from gel to membranes

DNA separated on agarose gels was transferred onto reusable nylon membrane (Hybond-N) or nitrocellulose (0.2 or 0.45 μ m pore diameter) (Schleicher+ Schuell) using the method of Southern, 1975. 0.2 μ m pore diameter nitrocellulose was used for blotting nucleosome length DNA, the small diameter of the pores allows more of the very small, <200bp, DNA fragments to be successfully retained on the filter as compared to 0.45 μ m nitrocellulose. Prior to the transfer the gel was photographed under UV light with a ruler so that the DNA transferred to the membrane could be sized at a later date. The UV illumination also introduces thymidine dimers into the DNA which makes the transfer of large DNA molecules easier. The gels were denatured in 1.5M NaCl, 0.5M NaOH for two 30 mins periods at room temperature, washed in water and then neutralised in 3M NaCl, 0.5M Tris at pH 5.0 for two further periods of 30 mins. DNA was transferred by capillary blotting with 20 x SSC. The gel was placed onto Whatman 3MM wetted with 20 x SSC and the membrane wetted in 2 x SSC was placed on top with no air bubbles trapped in between it and the gel. Two pieces of 3MM wetted in 2 x SSC were placed on top of the membrane and then a stack of paper towels put on top. A glass plate was rested on the top and the gel allowed to blot o/n. When the transfer was complete, the membrane was washed in 2 x SSC and air dried. The DNA was covalently bound by exposing the membrane DNA-side down to UV light for 5 mins for Hybond-N, or by baking in a vacuum oven at 80°C for 2 hours for nitrocellulose.

To reuse the membrane any hybridised probe was removed by submerging the filter in 300ml of boiling water and allowing this to cool for 3 hours. The removal of the probes were then tested by overnight autoradiography.

2.10.2 Immobilisation of DNA from bacterial colonies onto Hybond-N

Bacterial colonies for screening were grown on nylon membranes, as described in Section 2.5. To release the DNA from these colonies the filters were lifted from the agar plates and placed colony side up onto Whatman 3MM soaked in a 1% solution of SDS for 5 minutes, to lyse the cells. The filter was then transferred to Whatman 3MM presoaked in denaturing solution, for 5 minutes DNA side upper most, followed by transfer onto a neutraliser soaked piece of Whatman 3MM for a further 5 minutes and then washed in 2 xSSC. Care was taken to ensure that thorough washing had removed all cell debris from the filter. The filter was then allowed to air dry and then DNA covalently linked to the membrane by exposure to UV light for 5 minutes.

2.11 DNA Probes

Where necessary inserts were isolated from plasmids for use as hybridisation probes as described by Allshire and Bostock (1986). The 0.7kbrNTS probe was a 0.7kb HindIII/ Bam HI fragment from the plasmid pSp 329 (Toda *et al.*, 1984). The *ura4*, *ade6* and *stb* probes were all generous gifts of J. Kohli (Allshire *et.al.*, 1987). *Ura4* was a 1.8kb Hind III fragment from the plasmid pURA 4. A 2.5kb Xho I fragment from M13 RFAS 59 served as the *ade6* probe. The *stb* probe was a 1.3kb Eco RI fragment from the plasmid pURA 3 1.3. The SV₂neo probe was a Bam HI fragment as described by Southern and Berg, (1982). The dh probe was derived from plasmid pSS 242 as a 2.3kb Bgl II fragment. A 4.6 kb Eco R1 fragment was isolated from *pcy1* for the *cyh1* probe. *His1* was isolated as a 6kb fragment from pMN1. A 1.3 kb Hind III fragment from KS (*sds 21*) was used as the *sds21* probe. The *dis2* probe was a 2.4 kb Hind III partial insert in the plasmid opuc 19. The *nuc 2* gene was contained in the plasmid pUC (*nuc 2*) 2-3. The dh, *cyh1*, *his1*, *sds21*, *dis2* and *nuc2* plasmids were all kind gifts from Professor M. Yanagida (Kyoto). The *nus 21* probe was an 8kb Eco RI fragment in the plasmid pNSU 21 (Sugawara unpublished). The *ars1* probe was a 960 bp Eco RI fragment from the plasmid pIRT-2 subcloned from pFYM 1 (Losson and Lacoute, 1983). The tm probe was obtained as a preisolated insert from Dr Allshire (Edinburgh). The *nda3* probe was the kind gift of Dr Fantes (Edinburgh).

A rat Spi-2 cDNA from pSPFL 2.1 (Inglis and Hill, 1991) which cross hybridises to the murine Spi-2 complex was used as the *con* probe, a generous gift from Dr Inglis (Edinburgh). The *pax6* probe was obtained as 1.6kb Eco RI insert within the plasmid pSM-I (Hill *et al.*, 1991). The *Hox-8* plasmid λ 26 was also the kind gift of Dr Hill (Edinburgh) (Monaghan *et al.*, 1991). The *sl* and *Hprt* plasmids were both the kind gifts of Dr Jackson (Edinburgh).

2.12 Radiolabelling of DNA

2.12.1 Random prime reaction.

DNA probes were labelled by random priming using the method of Feinberg and Vogelstein (1984). DNA was boiled for 5 mins to denature it before use. Random hexanucleotides anneal to the denatured DNA and act as primers for complementary strand synthesis. This is catalysed by Klenow enzyme from the 3'-OH ends of the primers in the presence of $\alpha^{32}\text{P}$ -labelled dCTP and unlabelled dATP, dGTP and dTTP. A commercially available random priming kit from Boehringer Mannheim was used.

25ng of denatured DNA was added to a reaction mixture containing dATP, dGTP and dTTP each at 25mM, random hexanucleotides and buffer. 2 units of Klenow enzyme and 30 μCi $\alpha^{32}\text{P}$ dCTP (3,000Ci/mmol, 10mM/ml, Amersham) were also added. The reaction (in a volume of 20ml) was incubated at 37°C for >1hr. The percentage incorporation of the label into the DNA was estimated by TCA precipitation of counts on a Whatman GF/A filter (10% TCA quantitatively precipitates oligonucleotides of >20 bp). Proteins and unincorporated nucleotides were removed from the probe by centrifugation through a Sephadex G-50 column. The DNA is excluded from the column and passes through first. The probe was denatured once more, immediately before use.

2.12.2 End labelling of DNA.

5 μg of mammalian cell line DNAs and 50ng of *S.pombe* DNAs were digested with the appropriate restriction endonucleases (Hpa II and Msp I). 1 μg of these reactions was then subjected to gel electrophoresis to check for completion of the reactions. To the rest of the reaction were added 50 μCi $\alpha^{32}\text{P}$ dCTP and 10 units of Klenow (Boehringer Mannheim). Reactions were allowed to proceed for 1 hour at 37°C. Unincorporated counts were removed from the reaction by elution through a column of Sephadex G50.

This reaction is in principle the same as the random prime above (2.10). Klenow adds the $\alpha^{32}\text{P}$ dCTP to the 3-OH of the cleavage site, partially filling the sticky end. Since for *Hpa II* and *Msp I* the cleavage is C/CGG then only the radio nucleotide need be added.

2.13 Nucleic Acid Hybridisation

Church and Gilbert Buffer

7% SDS, 0.5M NaPO₄ at pH 7.2. This hybridisation solution is described by Church and Gilbert, 1984. It was used for most random primed DNA probes.

Denhardt's Buffer

0.5% SDS, 0.1% sodium pyrophosphate, 5x SET/Denhardt's solution (Stock solution is 20x: 0.4M Tris pH 7.8, 3M NaCl, 20mM EDTA, 0.4% Ficoll, 0.4% Polyvinylpyrrolidone (PVP), 0.4% BSA.) Denatured sonicated salmon sperm DNA was added as a competitor at 100mg/ml. This solution was used with certain random primed DNA probes and always when using nitrocellulose membrane.

All filters were double sealed into plastic bags and pre-hybridised for >30mins at 68°C. After addition of denatured probe all hybridisation incubations were overnight at 68°C.

After hybridisation, the filters were rinsed twice in 2x SSC, 0.1% SDS at 68°C for 10 mins each and then more stringently washed twice with 0.1x SSC, 0.1% SDS at 68°C for 30mins. The filters were then placed in plastic bags in preparation for autoradiography.

2.14 Photography

Agarose gels were placed on a UV transilluminator and photographed using a Polaroid MP4 camera fitted with a red filter using Kodak T-MAX Professional film 4052 with an exposure time of 15 secs. The film was developed in an X1 X-OGRAPH automatic X-ray film processor.

2.15 Autoradiography

2.15.1 Fixation of agarose gels.

To enable autoradiography of end labelled DNA fragments after gel electrophoresis the gel were treated as follows. the gel was washed for 5 minutes in 7% acetic acid,

then rinsed 3 times in deionised water. The gel was then placed on a piece of Whatman 3MM and dried on a gel drying apparatus prior to autoradiography.

2.15.2 Autoradiograph development.

The autoradiography in this project was carried out using X-ray film (Fuji Medical or Kodak X-OMAT) in light tight cassettes containing intensifying screens. Filters or dried gels were first exposed o/n at -70°C and then they were further exposed for different lengths of time and/or temperatures. The X-ray films were developed using an X1 X-OGRAPH automatic X-ray film processor.

2.15.3 Densitometry of Autoradiographs.

Autoradiographs of the dried gels were given to Dr Daryl Green who carried out the densitometric analysis.

2.16. Mammalian nuclei preparation and micrococcal nuclease digestion.

2.16.1 Mammalian nuclei preparation.

I have used a modified method of Hewish and Burgoyne (1973). Cells were harvested by scraping, from tissue culture flasks, and washed twice in ice cold PBS. Cells were then resuspended in 10ml of lysis buffer: 10mM Tris.HCl pH7.5, 10mM NaCl, 5mM MgCl_2 , 1mM PMSF, 0.5% Triton X-100 (Sigma), This solution was then passed through a luer lock needle, until the cells were broken and the nuclei had been released as monitored by phase contrast microscopy. Nuclei were pelleted by centrifugation at 3000 rpm for 5 minutes 4°C . Nuclei were then washed three times in lysis buffer lacking Triton X-100 then resuspended at 10^7 /ml in storage buffer: 25% glycerol, 50mM magnesium acetate, 0.1mM EDTA, 5mM DTT, 50 mM Tris. HCl pH 8.0, and stored at -70°C for up to 1 week.

2.16.2 Mammalian micrococcal nuclease digestion.

Nuclei were pelleted by centrifugation at 13000 rpm 4°C for 10 minutes then resuspended at 10^7 /ml in Digestion buffer: 15mM Tris. HCl pH 7.5, 60mM KCl, 200mM NaCl, 3mM MgCl_2 , 0.5mM DTT, 0.25M sucrose, 1mM CaCl_2 . For certain reactions nuclei were resuspended at 2×10^7 /ml. To 100 μl of this suspension 10 μl of 200u/ml micrococcal nuclease (Pharmacia) were added, and digestion allowed to continue for various lengths of time, at 37°C . An equal volume of stop solution: 1M NaCl, 20mM EDTA, 1% SDS, was added at the appropriate time point, to terminate the reaction. 2 μl of 50mg/ml Proteinase K were added to each reaction and then incubated

at 37°C for 4 hours. DNA was then purified by addition of an equal volume of phenol, mixing well and centrifuging at 13000 rpm for 5 minutes. The aqueous layer was then recovered and mixed well with an equal volume of phenol/chloroform/ isoamyl alcohol, 25/ 24/ 1 and centrifuged. The aqueous layer was again recovered and mixed with equal an volume of chloroform/ isoamyl alcohol (24/1). The aqueous layer was this time mixed with 2.5 volumes of absolute alcohol, and precipitated at -20°C for 30 minutes. DNA was pelleted by centrifugation at 13000rpm 4°C for 20 minutes. The DNA was then washed in 75% ethanol and dried under vacuum. The DNA was then taken up in an appropriate volume of TE and analysed by gel electrophoresis.

2.17 Yeast nuclei preparation and nuclease digestion.

Method 1

A fresh 10^7 cells/ml culture, in EMM, of *S.pombe Int-5* was protoplasted as described in Section 2.7. These protoplasts were then lysed in mammalian lysis buffer 2.16.1 with added 1mM PMSF. Cells were allowed to lyse for 5 minutes on ice. Nuclei were then pelleted as in Section 2.7 and washed 3 times in the above buffer without Triton X-100. Nuclei were resuspended at 10^9 /ml in nuclease digestion buffer (2.16.2) with added 1mM PMSF. 50µl aliquots were made up to 100µl volumes and digested and DNA purified as described in Section 2.16.2 .

Method 2

The procedure was as above, except that the nuclei were resuspended at 10^9 /ml in nuclease digestion buffer from Chikashige *et al.* (1989) SPC: 20mM PIPES, 0.1mM CaCl₂, 1M sorbitol, 1mM PMSF. 5ul of 200u/ml micrococcal nuclease (Pharmacia) were added to each 100ul reaction, and reactions stopped and DNA purified as described in Section 2.16.2.

Method 3

Same as method 2 above, except that the protoplasts were lysed for 5 minutes 4°C with 20mM PIPES, 0.05mM EDTA, 1mM PMSF, 1% Triton X-100. Nuclei were pelleted at 8000 rpm 4°C for 15 minutes, and washed twice in buffer minus Triton X-100.

Method 4

This procedure was modified from Polizzi and Clarke (1991). A $1\text{L } 10^7$ cells/ml culture of *S.pombe Int-5* was centrifuged at 3500K for 5 minutes. The cells were then protoplasted as in section 2.7. Nuclei were released by resuspending the proplasts in lysis buffer: 20mM PIPES, 18% Ficoll, 0.5mM MgCl₂, and stirred at 4°C for 15 minutes. This solution was then layered onto a cushion of 20mM PIPES, 20% glycerol, 8% Ficoll, 0.5mM MgCl₂, 1mM PMSF, and centrifuged at 20000 rpm for 40 minutes at 4°C. Nuclei were resuspended at 10^9 /ml in 20mM PIPES, 0.1mM CaCl₂, 0.5mM MgCl₂, 1mM PMSF. Micrococcal nuclease (Pharmacia) was then added and the reactions carried out and DNA recovered as above.

Method 5

The same procedure as method 4 except that after the nuclei were pelleted through the Ficoll cushion the nuclei were taken up in SPC and digested as in method 2.

2.18 Manipulation of cells in agarose beads.

2.18.1 Encapsulation of cells in agarose beads

Mammalian cells were harvested as described in Section 2.1.3. After harvesting as described in Section 2.1.3, cells were washed once in PBS and pelleted at 1000g for 5 minutes. 1 volume of 2.5% agarose (Sigma type VII) made up in PBS, precooled to 39°C, was added to 4 volumes of cells (2×10^7 ml) and mixed well, by pipetting. This solution was then added to a round bottomed flask containing 2 volumes of paraffin oil (Boots) prewarmed to 39°C. This flask was then sealed with parafilm and shaken vigorously, in a flask shaker, for 30 seconds until a creamy emulsion was formed. Beads were then formed by rapidly spinning the flask on ice water for 5 minutes. The bead solution was then poured into a centrifuge tube and the remaining solution recovered with two 1/3 volume washes of PBS. The beads were then recovered by centrifugation at 2500g for 1min. This step is repeated with disruption of the paraffin/ aqueous interface until the beads had all pelleted. The supernatant is then carefully removed by aspiration, and the pellet washed twice with ice cold PBS.

2.18.2 Permeabilisation and digestion of cells in agarose beads.

After washing the agarose beads were spun at 2500g for 1 min, and taken up in 10 volumes of pH 7.4 buffer plus Triton X100 0.5%. pH 7.4 buffer: 130mM KCl, 10mM Na₂HPO₄, 1mM MgCl₂, 1mM DTT, 1mM Na₂ATP and pH'd to pH 7.4 with K₂HPO₄. Lysis was allowed to occur for 15 minutes on ice, and repeated three times, with fresh pH 7.4 buffer plus Triton X100. Beads were pelleted between each lysis

step for 1 min at 2500g. After final lysis the beads were pelleted and washed three times in 10 volumes of ice cold pH 7.4 buffer. Beads were then taken up in an equal volume of pH 7.4 buffer and digested with the relevant amount of restriction enzyme at 32°C for 1 hour. After digestion the beads were placed on ice until loaded onto a gel.

2.18.3 Electrophoresis of unattached chromatin.

A 2% agarose (Sigma type II) solution in dH₂O was mixed with an equal volume of 2X E-buffer, preheated to 50°C, and the gel formed in a vertical gel former with 3mm spacers, leaving approximately 3cm for sample loading. E. buffer : 78.6mM KCl, 10mM Na₂HPO₄, 9.4mM KH₂PO₄, 1mM MgCl₂, 1mM DTT, 1mM Na₂ATP, pH 7.4 with 1M K₂HPO₄. No combs were used for most of the experiments, except during the trial runs. After cooling the beads were loaded onto the gel. The gels were electrophoresed in 1 x E buffer at 4°C for 16 hr at 30V in a vertical gel apparatus. The buffer was recirculated, using a peristaltic pump, to prevent ionic shift, and maintained at 4°C by passing the tubing through an ice bucket.

After running the buffer in the upper tank was syphoned off. The gel was then carefully removed from the tank, and the beads recovered by pastetting off from the top of the gel. The beads were then digested in an SDS 0.1%/ proteinase K 50µg/ml solution, at 37°C for 4hrs to digest chromatin proteins. The gel was removed from the plates and chromatin proteins were digested in an SDS/ proteinase K solution for 4hrs. After digestion a slice of the gel was stained with EtBr for 15min, and washed in dH₂O for 15 minutes. The gel slice was then photographed as described in Section 2.14.

2.18.4 Recovery from beads.

After protein digestion the beads were melted at 69°C for 1hr. An equal volume of phenol at 69°C was then mixed with the molten bead solution and centrifuged at 2500K for 10min. The upper aqueous layer was recovered and mixed well with an equal volume of phenol: chloroform 1:1. Phenolic and aqueous layers were then separated by centrifugation at 2500k for 10 minutes. The aqueous layer was recovered and mixed well with an equal volume of chloroform: isoamyl alcohol 24:1, and centrifuged at 2500 K for 10 minutes. The aqueous layer was recovered and mixed with 2.5 volumes of absolute alcohol, and the DNA precipitated for 20min at -70°C and centrifuged at 13000g for 25 minutes. The supernatent was removed and the pellet washed in 70% ethanol. After repelleting, the remaining alcohol was removed under vacuum, and the pellet taken up in TE pH7.4. This DNA solution was passed through an "Elutip-D" (Schleicher + Schuell) according to the manufacturers instructions.

2.18.5 Recovery of loop DNA from gel

After proteinase K treatment the gel was aligned with the EtBr stained fraction and the DNA containing region cut out. This DNA containing gel slice was then fragmented and placed into a length of dialysis tubing and 20ml of 1xTBE added. The gel fragments were then pushed to one end of the tube, and the tube electrophoresed overnight at 50V in 1x TBE to elute the DNA from agarose. The current was then reversed for two minutes to remove the DNA from the walls of the tubing. The contents of the tube were then filtered through muslin, so that the DNA solution could be obtained. This fraction was then passed through an "Elutip-D" (Schleicher + Schuell) according to the manufacturers instructions to obtain a concentrated clean DNA solution.

2.18.6 Analysis of DNAs

DNAs recovered from these fractions were cut to completion with the appropriate enzyme and analysed by gel electrophoresis as described in Section 2.9, and blotted as described in Section 2.10.1.

2.19 Manipulation of *S.pombe* in agarose beads.

2.19.1 Encapsulation of *S.pombe* cells in agarose beads.

5×10^9 cells, grown in EMM, were used for each experiment. *S.pombe* cells were encapsulated the same way as described in Section 2.18.1 with the exception that the cells were taken up in 1M sorbitol before addition to the agarose.

2.19.2 Lysis of *S.pombe* cell wall in agarose beads.

S. pombe cells have to have their cell wall partially degraded before nuclei can be produced. This was achieved by the action of the enzyme zymolyase which punctures holes in the cell wall, and should therefore allow the diffusion of enzymes and small molecules required in subsequent steps. Beads were resuspended in 10ml of 1M sorbitol, 0.5mM spermidine pH 7.4. 4mg of zymolyase were then added and incubated at 30°C for 3 hours. Beads were then centrifuged 3500rpm for 1 minute and then washed twice in 1M sorbitol, 0.5mM spermidine.

2.19.3 *S.pombe* nuclei preparation in agarose beads.

Beads were lysed as in Section 2.18.2, with the exception that the pH7.4 buffer was made up with 1M sorbitol. Washes were also carried out in pH7.4 buffer with added 1M sorbitol.

2.19.4 Digestion of *S.pombe* chromatin in agarose beads.

Chromatin was digested as above Section 2.18.2, with the exception that 1M sorbitol was added to the pH7.4 buffer.

2.19.5 Recovery of DNAs from beads and gel.

This was achieved as described in Section 2.18.

CHAPTER 3. ANALYSIS OF CONTENT AND STABILITY OF *S. pombe* SEQUENCES IN MOUSE / *S. pombe* HYBRID CELL LINES.

3.1 INTRODUCTION

As discussed in Chapter 1, hybrid cell lines were established from the fusion of *S. pombe* protoplasts and mouse C127 cells. The fused cells were then grown on G418 which selected for the presence of an expressing SV2neo gene. The SV2neo gene had been introduced into the *ura4* locus of *S. pombe* chromosome III prior to fusion (Allshire *et al.*, 1987). It was observed cytogenetically, by *in situ* hybridisation (Perry *et al.*, unpublished observations), that one of the hybrid cell lines, F1.1, contained a chromosomal constriction at the site of integration of the *S. pombe* sequences. Similar constrictions have also been observed in other integrated hybrid cell lines (Section 1.1), at the site of *S. pombe* sequence integration. Since these chromosomal constrictions are common to all, so far determined, sites of extensive *S. pombe* sequence integration, into mouse chromosomes, it can be assumed that this affect is not due to position affects at sites of integration, but is intrinsic to the interactions of the foreign *S. pombe* DNA regions with the chromosomal proteins of the mouse cell.

The major cell line used in this study is the F1.1 cell since as this was the first integrated hybrid cell line investigated and discovered to contain an anomalous chromosomal region.

The size of the *S. pombe* insert in F1.1, as estimated from cytological studies, is assessed to be in excess of 5-10 Mb. However this is only a close approximation since, as this region appears to be packaged differently from the surrounding mouse chromatin, a direct comparison of the observed length of insert to the length of a known region of mouse genome would not deliver an accurate comparison.

In order to analyse the interactions at these *S. pombe* inserts which may be responsible for the unusual appearance of these regions at metaphase, it is necessary to investigate which parts of the *S. pombe* genome have entered these cells. This chapter addresses which *S. pombe* sequences have entered the mouse genome during establishment of the F1.1 cell line. Only one gene (SV2neo) from the *S. pombe* *Int 5* genome is required under the selective conditions used. Regions other than the SV2neo gene however are likely to be present co-integrated at the same site. The likelihood of co-integration is greatly affected by linkage to the selected marker and therefore chromosome III markers are the most likely to be found, however non-syntenic markers may be present (Bickmore *et al.*, 1989). In detailed studies of DNA introduced by PEG induced fusion it has been

Fig 3:1 PHYSICAL MAP OF THE *S.pombe Int -5* GENOME

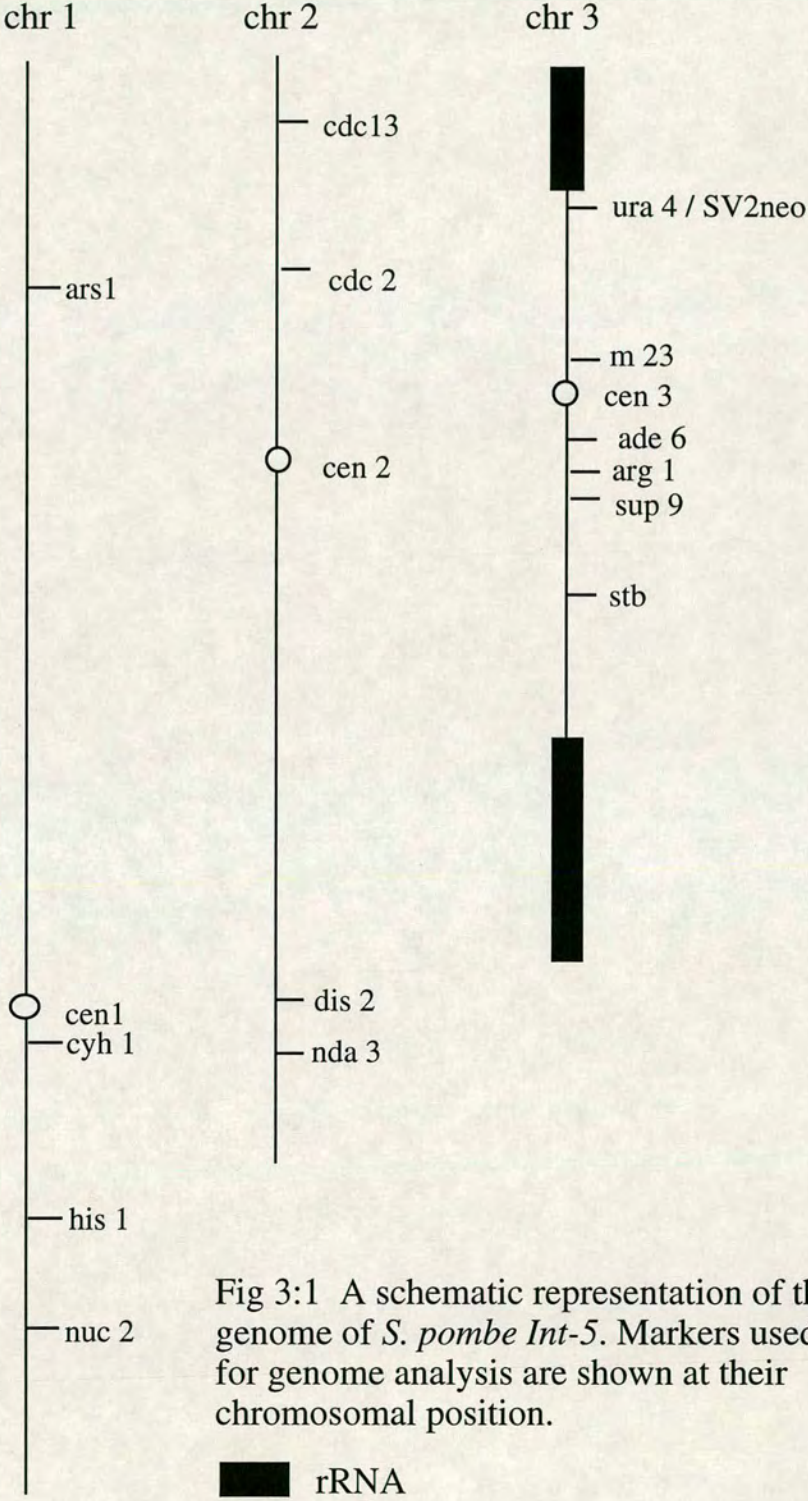


Fig 3:1 A schematic representation of the genome of *S. pombe Int-5*. Markers used for genome analysis are shown at their chromosomal position.

■ rRNA

References:
 Fan et al., 1988
 Kohli, personal communication
 Maier and Bickmore, unpublished observations

observed that large regions of introduced DNA may enter the recipient genome at a single copy level and also be free of detectable rearrangements (Bickmore *et al.*, 1989). Transfectant cell lines established by CMGT procedures were demonstrated to vary in the level of rearrangements found within their transgenomes. Since these hybrid cell lines were established by PEG fusion of *S. pombe* protoplasts and C127 cells, by analogy it may be expected that the *S. pombe* transgenomes would be largely unrearranged relative to the donor yeast chromosomes. This assumes that the fusion of yeast and mouse cells is similar in nature that of the vertebrate to vertebrate cell fusions which have been analysed previously.

Some analysis has also been undertaken on a hybrid cell line, F7.2, within which the *S. pombe* sequences, as assessed by *in situ* hybridisation, have not integrated into the mouse genome (Perry *et al.*, unpublished observations) but are presumed to be maintained as extra chromosomal units while the cells are maintained under selective conditions. The F7.2 cell line serves as a control for cis-acting effects upon the F1.1 transgenome of adjacent mouse chromatin. This is particularly relevant for the chromatin studies of chapter 6.

3.2 ANALYSIS OF THE *S. pombe* CONTENT OF HYBRID CELL LINES

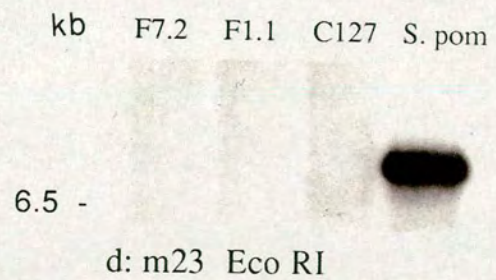
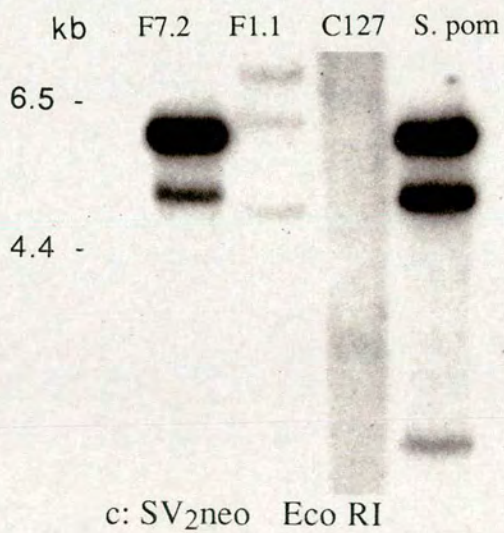
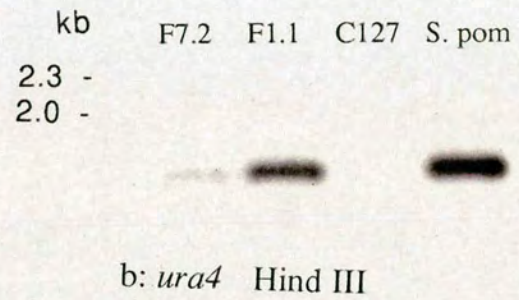
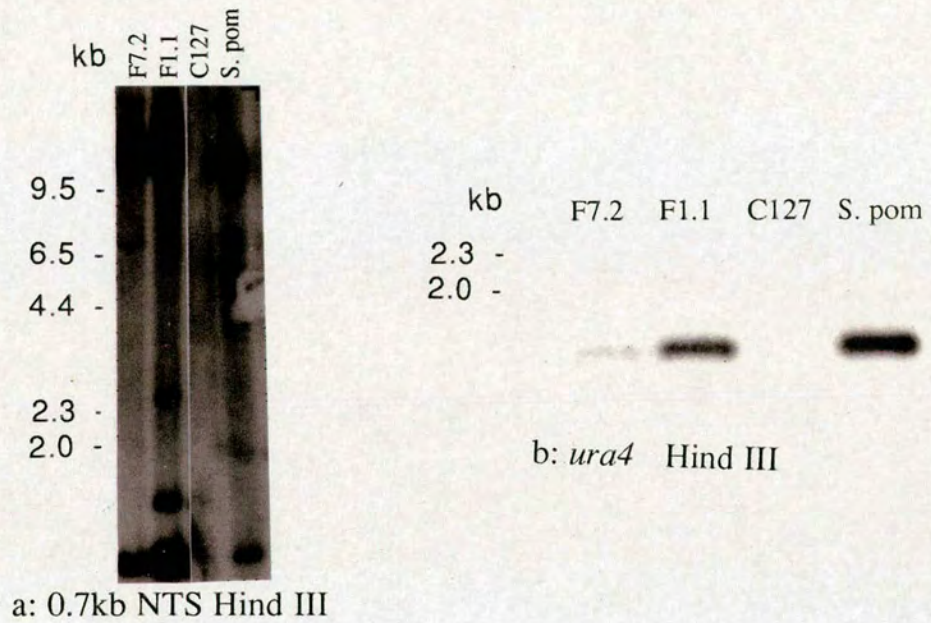
3.2.1 Probes used from *S. pombe* *Int 5* genome

As demonstrated in Figure 3.1 the *S. pombe* genome consists of 3 chromosomes. The SV2neo gene had been inserted into the *ura4* locus, which is located in the short arm of chromosome III, to produce the *Int 5* strain used as the donor for the fusions (Allshire *et al.*, 1987). The markers which I have used in this study are spread throughout the *S. pombe* genome, although the major emphasis has been placed upon markers for chromosome III, since these markers are physically linked to the selected gene and are therefore more likely to be co integrated into the cell with the SV2neo construct. *S. pombe* chromosome III markers which are absent from the hybrid cell lines, could only be lost due to deletions of chromosome III material prior to transgenome establishment.

3.2.2 *S. pombe* chromosome III marker content of hybrid cell lines F1.1 and F7.2

The rDNA complex which constitutes approximately 1/3 of *S. pombe* chromosome III can be detected using a probe for the non transcribed spacer region (0.7kb rNTS). This probe detects all of the ribosomal subunits of *S. pombe*. Hind III digests of F1.1, F7.2, C127 and *S. pombe* DNAs were probed with 0.7kb rNTS (Figure 3.2). The *S. pombe* lane demonstrates the hybridisation expected for these rDNA sequences, with a major

Figure 3.2: The *S.pombe* chromosome III marker content of the F1.1 and F7.2 hybrid cell lines



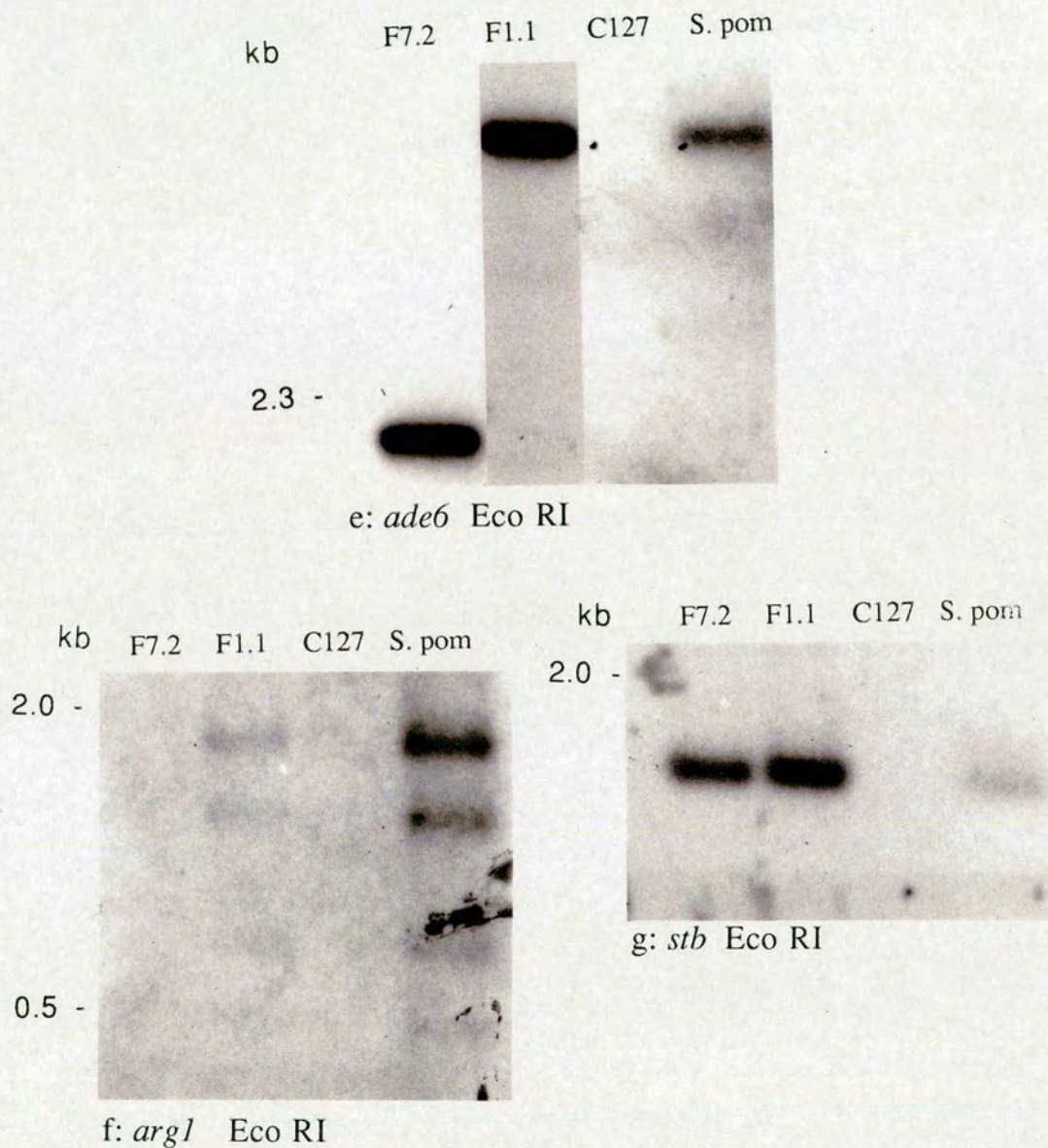


Figure 3.2: 10 μ g of mammalian cell line DNAs from F1.1, F7.2 and C127 and 50ng of *S. pombe Int-5* (*S. pom*) DNA were digested with the restriction enzymes as indicated. DNAs were subjected to gel electrophoresis and then transferred to nitrocellulose. The indicated probes were then hybridised to these filters and the hybridisation analysed by autoradiography.

band of approximately 10.5 kb and a small <1kb band. The extrachromosomal F7.2 hybrid cell line also demonstrates this pattern, showing that these sequences are largely unrearranged within the extrachromosomal element in this cell line. However there are two weakly hybridising bands at approximately 15kb and 6.6kb, suggesting that rearrangement of some of the copies of rDNA may have taken place. The F1.1 cell line also shows the hybridisation pattern of the *S. pombe* rDNA complex although in this cell line, there appears to be a larger amount of rearrangement than that seen in F7.2 since minor hybridisation bands are visible at 1.8 and 2.5kb in addition to the major 10.5 and 1kb bands. It may be unsurprising that over 1.5Mb of repeated DNA would enter a cell without some level of rearrangement. The ribosomal sequences in *S. pombe* are adjacent to the telomeric sequences and therefore this rearrangement may be related to the integration event during establishment.

In Figure 3.2 b F1.1, F7.2, C127 and *S. pombe Int 5* DNAs have been digested with Hind III and hybridised with a probe from the *ura4* region. No hybridisation is detected from the C127 DNA and hybridising bands of identical size are seen in the *S. pombe* and both hybrid cell line lanes. Therefore no rearrangements of the *ura4* locus have occurred during transgenome formation. The lower level of signal in the F7.2 lane is due to unequal loading of the DNAs.

A probe for the SV₂neo construct has been used to probe DNAs cut with Eco RI in Figure 3.2c. A smear of hybridisation is observed within the C127 lane, however no specific bands of hybridisation are detected within the mouse DNA and therefore the hybridisation within the other lanes can be considered specific to the SV₂neo constructs in these DNAs. In *S. pombe* DNA it can be seen that three major bands of hybridisation are detected. In F7.2 the lower of the three bands (2.8kb) is deleted. The other two bands do not appear to be altered in size but the lower of the two seems reduced in intensity, so a small specific rearrangement has occurred. In F1.1, all the SV₂neo hybridising bands detected are highly rearranged. This rearrangement however cannot be associated with a loss of activity of the construct concerned, since the cells grow under G418 selection.

An anonymous DNA marker m23 from close (<200kb) to the *S. pombe* chromosome III centromere has been used to probe DNAs, cut with Eco RI in Figure 3.2d. No hybridisation to this probe is detected in either F1.1 or F7.2 hybrid cells. Since this probe is close to the chromosome III centromere it is possible that the centromere is also deleted from these cell lines. Since this sequence is absent from both the hybrid cell lines analysed it could be hypothesised that this sequence may be poisonous to the mouse cells. This could be further clarified by analysing other hybrid and transfectant cell lines.

In Figure 3.2e an *ade6* probe has been hybridised to Eco RI digests of the F1.1, F7.2, C127 and *S. pombe* DNAs. *Ade6* is located in chromosome III on the arm opposite that of the *ura4* selectable locus. F1.1 contains a non rearranged *ade6* hybridising band suggesting that this region has entered these cells intact. In the case of F7.2 a new band is detected at approximately 2.2 kb, suggesting that this region has undergone rearrangement during establishment of this cell line.

A probe for *arg1* region has been used to probe Eco RI digests of the relevant DNAs in Figure 3.2f. In the *S. pombe* lane four major bands of hybridisation are detected. These bands are not detected in the C127 control lane. The F7.2 lane also does not contain these bands demonstrating that these sequences are absent from this hybrid. The F1.1 lane contains all four bands (lowest band visible on original autorad) in an unrearranged manner.

A probe to the *stb* region of chromosome III has been used to probe Eco RI digests of the relevant DNAs in Figure 3.2g. Identical bands are detected in *S. pombe*, F1.1 and F7.2 lanes, suggesting that this region has entered both cell lines unrearranged.

3.2.3 *S. pombe* chromosome I and II marker content of hybrid cell lines F1.1 and F7.2.

Markers for *S. pombe* chromosome I and II might not be expected to be part of the cell lines under investigation since they are not initially associated with the selectable SV2neo marker in *S. pombe* *Int5*. The markers for chromosome I and II could only become established within the autonomous element of the F7.2 cell line if they become attached to the SV2neo selectable marker during establishment or in F1.1 through co integration.

Hybridisation to the *cyh1* gene on chromosome I is shown in Fig 3.3a. This region is present in neither of the two hybrid cell lines. In relation to the results obtained later with the dh centromeric repeat (Section 3.2.4) it is interesting that the *cyh1* probe is absent since this is proximal to the centromere of chromosome I and would suggest that cen I is missing since *cyh1* is <1Mb from the centromere.

A probe for the *S. pombe ars1* sequence was used in Figure 3.3b. This sequence isolated by Losson and Lacoute (1983) has been mapped close to *rad4* on chromosome I (Maier and Bickmore, unpublished observations). It can be seen that the *S. pombe* and F1.1 lanes both contain a band of ~4kb in DNAs digested by Hae III. The F7.2 lane and the C127 control lane show no hybridisation. This demonstrates that chromosome I sequences have been established in the F1.1 insert, but appear to be absent from the f7.2 extrachromosomal element.

Figure 3.3: *S.pombe* chromosomes I and II marker content of the F1.1 and F7.2 hybrid cell lines

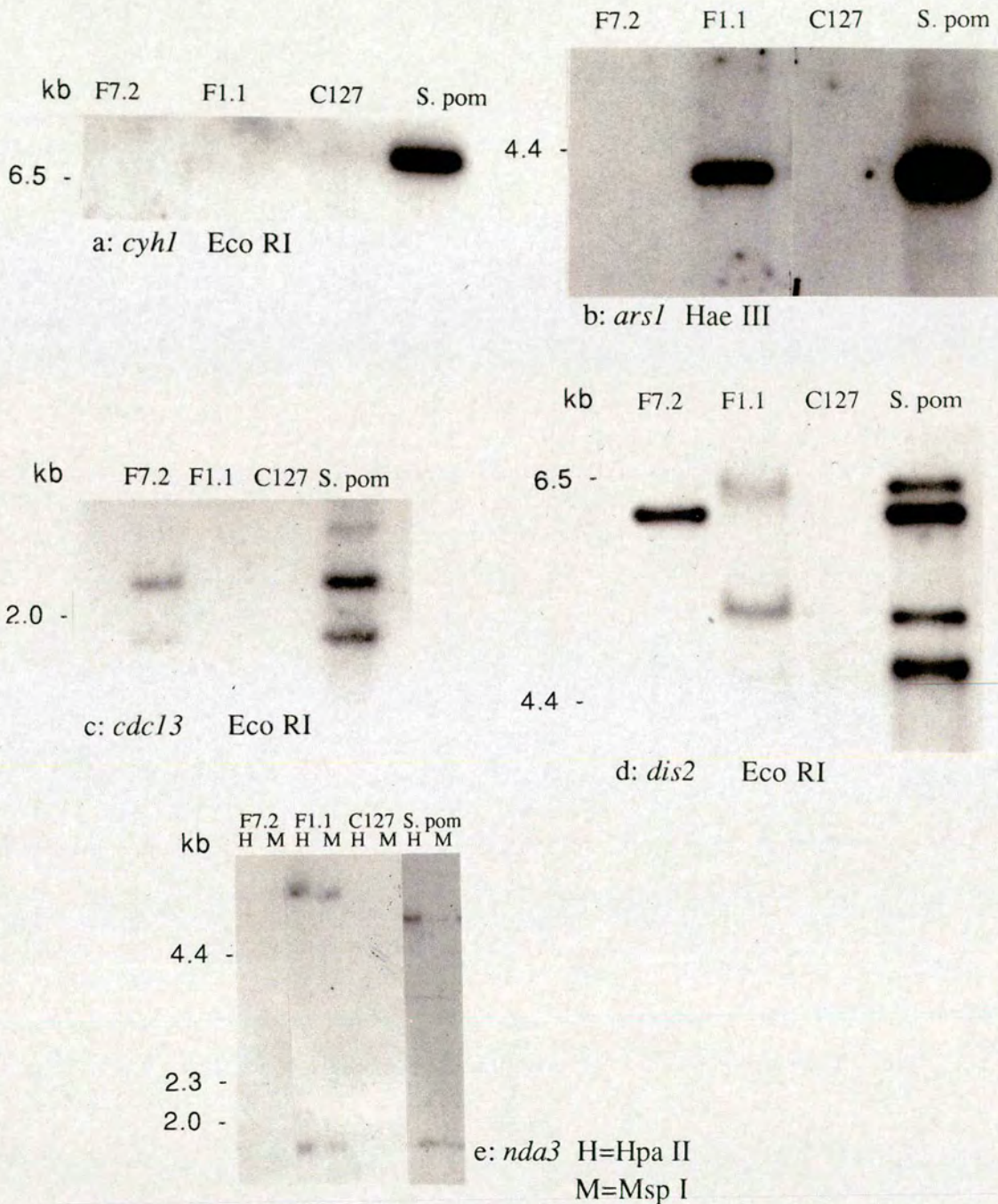


Figure 3.3 10µg of mammalian cell line DNAs from F1.1, F7.2 and C127 and 50ng of *S. pombe* Int-5 (*S. pom*) DNA were digested with the restriction enzymes as indicated. DNAs were subjected to gel electrophoresis and then transferred to nitrocellulose. The indicated probes were then hybridised to these filters and the hybridisation analysed by autoradiography.

It was also found that *his1* and *nuc2* regions of the short arm of chromosome I (Table 3.1) were also absent from these hybrid cell lines (data not shown) as was *cyh1*. The absence of all three of the above regions would suggest that little or no material from this end of chromosome I has entered the hybrid cell lines and that only material from the long arm of chromosome I is present in F1.1.

A probe to the *cdc13* gene in the short arm of chromosome II was used in Figure 3.3c. Two major bands of hybridisation, at 1.8 and 1.6kb, after digestion with Eco RI, are present in *S. pombe* DNA these bands are not detected in the C127 control. The F1.1 cell line shows no hybridisation to this probe. *cdc13* is present unarranged in F7.2 cell line. *cdc13* is the only non chromosome III material detected in the F7.2 cell line.

In Figure 3.3d a plasmid containing the *dis2* gene has been used for hybridisation. It can be seen that in the *S. pombe* lane four major bands of hybridisation are detected. In F7.2 a hybridising band is detected but this is due to the vector sequences of the plasmid cross hybridising to sequences within the SV2neo construct. In the F1.1 cell line it is seen that two bands of hybridisation are detected. The larger of these hybridising bands 6.4kb is the cross hybridisation to the vector sequences in the SV2neo construct. The lower of the two bands (4.8kb) is the *dis2* gene which can be seen to be of the same size as its corresponding *S. pombe* band.

In Figure 3.3e a probe to the *nda3* gene was used for hybridisation. The F7.2 and C127 tracks contain no detectable hybridisation. The F1.1 lane contains the same 1.8kb hybridising band as does the *S. pombe* track, but the larger of the two bands, 4.0kb in *S. pombe*, is seen to be about 4.3 kb in F1.1, this may suggest that this region has undergone rearrangement during establishment.

The *cdc2* region of chromosome II is absent from both hybrid cell lines (Table 3.1). The absence of this region would imply that the rearrangement which incorporated the *cdc13* region into the F7.2 element was limited to the distal region of chromosome II containing this gene. Therefore markers from one arm of chromosome II (*dis2* and *nda3*) are present in F1.1 and no material from the other arm is detected in this cell line.

3.2.4 Analysis of *S. pombe* centromeres and telomeres in hybrid cell lines F1.1 and F7.2.

It would be of interest to analyse the yeast centromeric content of the transgenome of F1.1, since although no homology would be predicted, these sequences contain potential to affect the observed structure of chromosomes (Rattner and Lin, 1985) and the cytological aspects of this are the focus of different studies. *S. pombe* telomeric content of the F1.1 transgenome is also of interest since previous studies have found other telomeric

Figure 3.4 :The *S. pombe* centromere and telomere content of the F1.1 and F7.2 hybrid cell lines

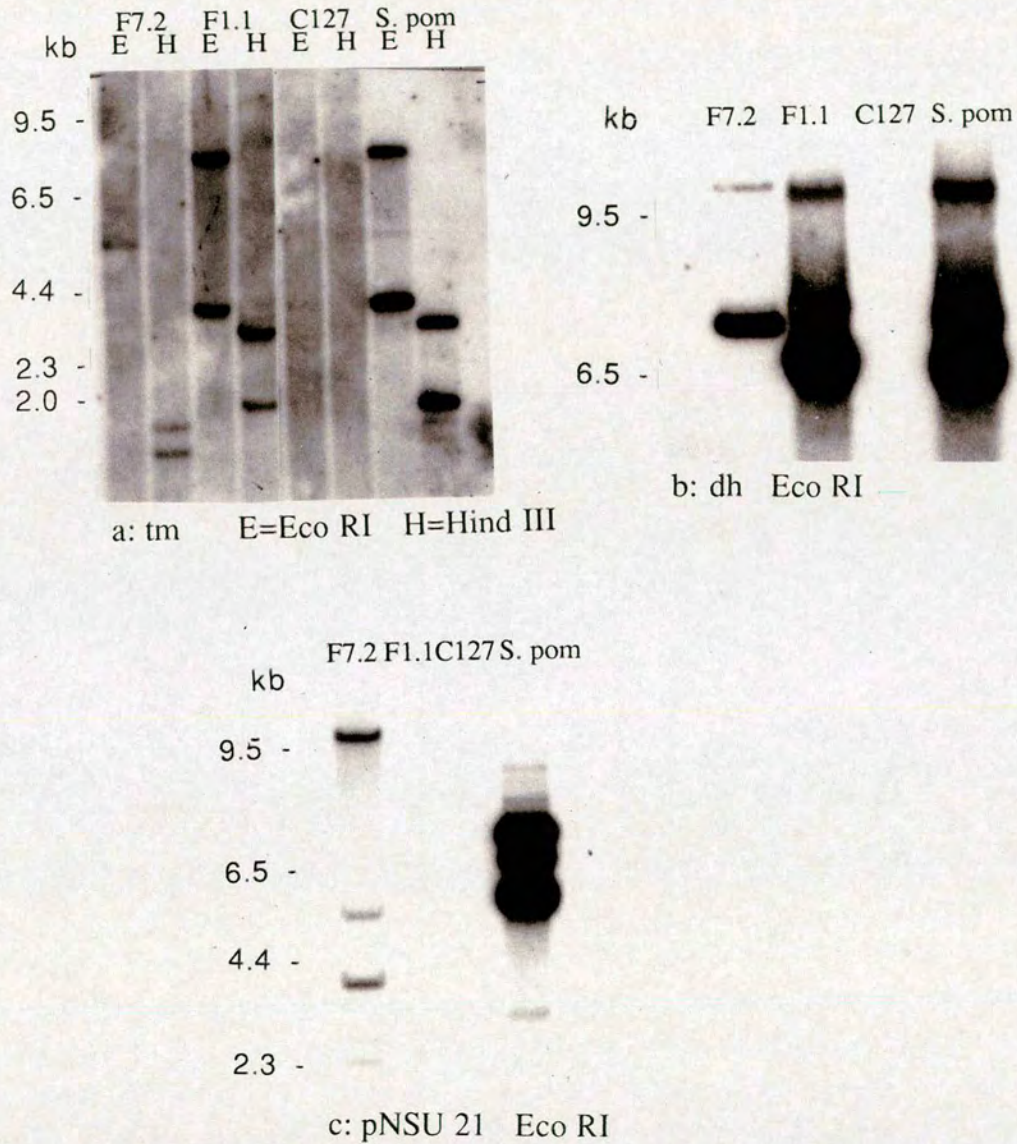


Figure 3.4: 10µg of mammalian cell line DNA and 50ng of *S. pombe* *Int-5* DNA were digested with the restriction enzymes as indicated. DNAs were subjected to gel electrophoresis and then transferred to nitrocellulose. The indicated probes were then hybridised to these filters and the hybridisation analysed by autoradiography.

sequences to be missing from hybrid cell line transgenomes (Huxley and Gnirke, 1992). The lack of these sequences from the transgenomes may be related directly to the process of transgenome formation in integrated transgenome cell lines.

In Figure 3.4a a probe to the *S. pombe* centromeric core sequences, tm, is used for hybridisation. This probe is found in the centromeres of *S. pombe* chromosomes I and III, although not in chromosome II (Takahashi *et al.*, 1991). The probe has been hybridised to Eco RI and Hind III digests of the F1.1, F7.2, C127 and *S. pombe* DNAs as indicated. F1.1 shows an identical hybridisation pattern to that of *S. pombe*. The hybridising bands are identical to those described by Takahashi *et al.* (1992) with 8kb Eco RI and 3.5kb Hind III fragments originating from cen 3, with 4kb Eco RI and 2kb Hind III fragments from cen I. Therefore it would appear that both cenI and cenIII copies of the tm sequence are present within the F1.1 transgenome. In the F7.2 the hybridising bands correspond to weak bands visible in *S. pombe* *Int5*. It may be that this hybridisation is due to cross hybridisation to t-RNA sequences which are found within the tm probe, and are highly clustered within the centromeric regions (Takahashi *et al.*, 1991; Kuhn *et al.*, 1991). It would therefore appear that F7.2 does not contain the centromeric sequences of chromosome I or more importantly chromosome III, suggesting that rearrangements have taken place in the centromeric region of chromosome III prior to establishment of the F7.2 transgenome.

A probe to the *S. pombe* centromeric repeat sequences dh has been used in Figure 3.4b. This sequence forms, in conjunction with the dg repeat the centromeric repeat structures which are important in defining the centromeres of the *S. pombe* chromosomes (Fig 1.5)(Nakaseko *et al.*, 1987). The dg sequence was found to cross hybridise to the mouse genome (data not shown) and therefore only the dh sequence has been used in these studies. The *S. pombe* DNA shows a pattern of hybridisation expected of Eco RI digestion. It can be seen that there is a band of ~14 kb which is diagnostic of chromosome I and a band of 6.6kb which is characteristic of chromosome 3 (Murakami *et al.*, 1991). Chromosome II is detected by a variety of bands between 7 and 9kb which can also be observed (Chikashige *et al.*, 1989). The difference in the relative hybridisation strengths of each of these bands is due to the copy number at the respective centromeres; cen 3 contains 13 copies of dg and dh while cen1 contains 2 copies and cen 2, 3 copies. In F7.2 DNA it appears that dh sequences are present that may derive from cen 1 and cen 2 although the upper band of ~16kb may be rearranged. A band of 7.4kb is probably from centromere II. No chromosome III dh sequences are detected in F7.2. The rearrangement of the cenI derived dh band and absence of the cen3 repeats from F7.2 is consistent with the observation that cenI and III tm sequences are also absent from F7.2 (Fig 3.4a). In the

F1.1 lane it would appear that most of the dh hybridising bands are represented. The only band apparently aberrant is one corresponding to a chromosome II band of 7.5kb. This suggests that complex rearrangements have been involved in producing this cell line and that one side of cen2 is missing, since some, not all, cen2 bands are present. That the chromosome I dh repeats and the cen1 tm sequence are present in F1.1 is surprising since *cyh1* is absent from this cell line (Fig 3.3a), *cyh1* is closely linked genetically and physically to cen1. This would suggest a breakpoint in the rearrangements lie between the centromere and the *cyh1* gene .

In Figure 3.4c an *S. pombe* telomeric sequence (nsu21) containing probe has been used for hybridisation. The sequences contained within this 8kb probe are the subtelomeric sequences believed to be found at 5 of the *S. pombe* telomeres apart from the long arm of chromosome 3 which does not contain such subtelomeric sequences (Sugawara, personal communication). The F1.1 cell line does not contain any such telomeric sequences. It is interesting to speculate that the loss of telomeres through rearrangement prior to establishment related to the retention of non-syntenic markers in F1.1 since the DNA may be cut and ligated by endogenous nucleases prior to integration. It is of interest that the F7.2 hybrid cell line contains *S. pombe* telomeric sequences, though these sequences are highly rearranged in this cell line. It would be interesting to investigate whether these sequences were still terminally positioned in this extrachromosomal element or if the rearrangements have given them an internal position. This could be tested by Bal 31 exonuclease digestion, if nsu21 sequences in F7.2 are terminally positioned then a Bal 31 time course digestion will result in a decrease in size of the hybridising band, since the Bal 31 digests from a free DNA end such as a telomere (Cooke et al., 1986). I consider it unlikely that these *S. pombe* sequences act as functional telomeres even if terminally located, due to the sequence divergence described between *S.pombe* and murine telomeres (Kipling and Cook, 1992). It is more likely that telomerase activity in the mouse cells would add mouse telomeric repeats onto this F7.2 autonomous element in the same manner as that observed with mammalian chromosomes breakage (Wilkie *et al.*, 1990).

A probe to the, as yet unmapped, *sds 21* gene showed that this region was absent from both cell lines studied. Until mapped this probe provides little information as to the nature of the rearrangements involved in both hybrid cell lines.

3.3 STABILITY OF HYBRID CELL LINE TRANSGENOMES.

Cytological studies show that F1.1 (Fig 1.1) and F7.2 differ in the nature of their transgenomes (Wright and Perry, unpublished observations); F1.1 has an integrated transgenome, while the transgenome of F7.2 appears to be extrachromosomal. This was confirmed by a study of the comparative stabilities of these two transgenomes (Fig 3:5). F1.1 and F7.2 were grown on and off G418 selection for approximately 40 generations. The level of transgenome retained, and therefore its stability, was compared in DNA samples from these populations over this time course by hybridisation to the *S. pombe* 0.7kb *rNTS* probe.

Fig 3.5a demonstrates that for the F1.1 cell line, the 0.7kb *rNTS* probe detects no change in level of the transgenome either on or off G418 selection. Any differences between tracks is due to differences in DNA loading as revealed by hybridisation to a mouse control probe (Fig 3:5b). Therefore the *S. pombe* sequences in F1.1 are stably integrated into a mouse chromosome, thus confirming the cytogenetic data on this transgenome.

As previously seen by *in situ* hybridisation Figure 3.5 indicates that the F7.2 transgenome appears to be extra chromosomal in nature. Instability of the F7.2 transgenome is demonstrated by the fact that the 0.7kb *rNTS* signal decreases with time, under non selective conditions but is maintained under selective conditions (Fig 3:5a). Thus the *S. pombe* transgenome is mitotically unstable. This is consistent with an extrachromosomal element that does not contain a functional centromere or that is unable to replicate once per cell cycle. A lack of centromere function is not unexpected since an *S. pombe* centromere probably bears very little sequence similarity to mouse centromeres, and would therefore not be expected to function in mouse cells. Figures 3:4 a and b demonstrate that the F7.2 cell line contains little of the *S. pombe* centromeric repeat *dh* and the *tm* centromeric core sequence is absent, suggesting that F7.2 in fact carries very little *S. pombe* centromere sequences anyway .

3.4 Discussion

In this chapter I have shown that the transgenomes of F1.1 and F7.2 hybrid cell lines are non-linear in character, and also contain non syntenic material. A complex pattern of retention of the *S. pombe* sequences has been found in F1.1 and F7.2 (Table 3:1). The F1.1 insertional hybrid exhibits a very complex pattern of markers from different regions of the *S. pombe* genome. Chromosome III markers are mostly retained but with regions of loss close to *cen3* and at telomeres (Table 3:1 c). This implies that some breakage and religation of *S. pombe* sequences has occurred prior to integration or during stabilisation of the transgenome. Markers from 1 arm of chromosome I and 1 arm of chromosome II are also detected. The complexity of the rearrangements involved is exemplified by the fact that the centromere of chromosome I appears to be present, as defined by *dh* hybridisation

Figure 3.5 Stability of *S. pombe* sequences in the hybrid cell lines F1.1 and F7.2

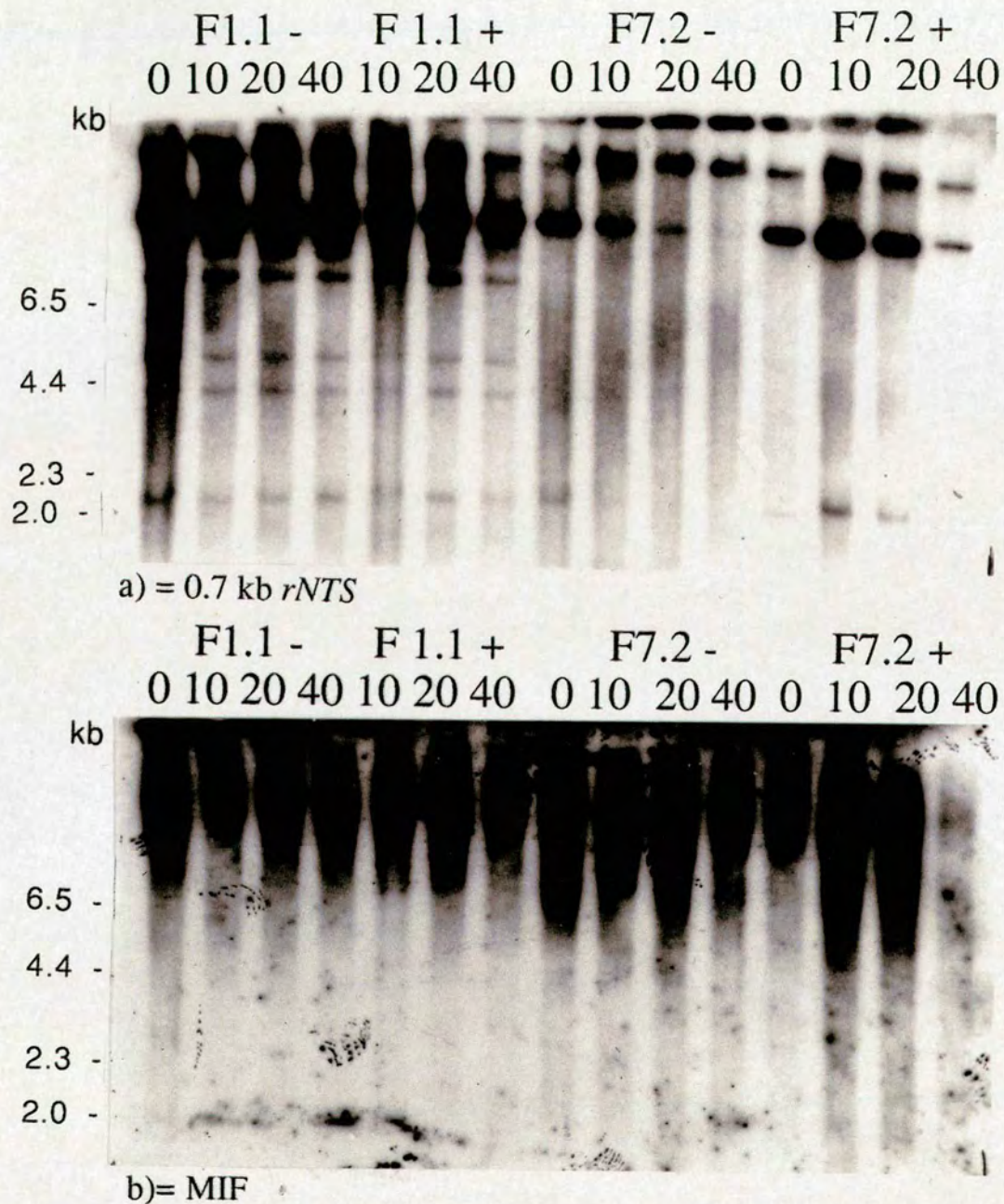


Figure 3.5

F1.1 and F7.2 cell lines were grown on (+) and off (-) G418 selection for up to 40 generations. DNAs were made from cells at selected time points. Numbers above lanes correspond to generation number. Approximately 10 μ g of each DNA was digested with Hind III and transferred to nitrocellulose after gel electrophoresis. The blot was hybridised to *S. pombe* 0.7kb *rNTS* (a) and then to MIF (b). MIF is a murine line sequence and is used as DNA loading control.

TABLE 3:1 *S. pombe* SEQUENCE CONTENT OF HYBRID CELL LINES

a: CHROMOSOME 1

PROBE	<i>ars1</i>	<i>cyh1</i>	<i>his1</i>	<i>nuc2</i>
F 1:1	+	-	-	-
F 7:2	-	-	-	-
Fig 3:3	b	a	NS	NS

NT = NOT TESTED NS = DATA NOT SHOWN + = PRESENT IN CELL LINE - = ABSENT FROM CELL LINE R = REARRANGED IN HYBRID

b: CHROMOSOME II

PROBE	<i>cdc13</i>	<i>cdc2</i>	<i>dis2</i>	<i>nda3</i>
F 1:1	-	-	+	+R
F 7:2	+	-	-	-
Fig 3:3	c	NS	d	e

c: CHROMOSOME III

PROBE	<i>rDNA</i>	<i>ura 4</i>	<i>SV2neo</i>	<i>m 23</i>	<i>ade 6</i>	<i>arg1</i>	<i>sup9</i>	<i>stb</i>
F 1:1	+R	+	+R	-	+	+	-	+
F 7:2	+	+	+	-	+R	-	NT	+
Fig 3:2	a	b	c	d	e	f	Allshire et al., 1987	g

d: CENTROMERIC UNASSIGNED

PROBE	<i>dh</i>	<i>tm</i>	<i>sds 21</i>	<i>nusu 21</i>
F 1:1	+	+	-	-
F 7:2	+	-	-	+R
Fig 3:4	b	a	NS	c
CHR	all	1+3		

74

but the proximal probes *cyh1* and *his1* are absent. It would appear that a significant proportion of the *S. pombe* genome has entered the F1.1 cell line, ~7Mb. The F7.2 cell line also exhibits a complex pattern of formation for its transgenome in the same manner as F1.1. Unlike F1.1, F7.2 contains little non syntenic material with only one gene marker *cdc13* present (figure 3.3c) from chromosome II and certain *cen2* dh repeats being present but no chromosome I material is detected in F7.2. Joining of non syntenic material has occurred before integration as shown by the presence of chromosome I and II markers in the F1.1 insert (Table 3.1b). This presence of non syntenic material may reflect the variable pattern of *S. cerevisiae* DNA integration after fusion with mammalian cells (Pavan et al., 1990; Pachnis et al., 1990; Huxley et al., 1991), and may indicate some fundamental difference between yeast/ vertebrate and vertebrate/ vertebrate heterokaryons.

I have also shown that the transgenomes of the two hybrids differ in nature as was suggested by their cytogenetic profiles. The F1.1 transgenome being stably integrated and therefore stably retained when grown off selection, while the F7.2 transgenome exists as an unstable extrachromosomal element that is lost from the population once selection is withdrawn, suggesting it is unable to segregate correctly at mitosis. Incorrect segregation in this instance would imply either a replication defect or a no centromeric function.

The complex pattern of retention of *S. pombe* DNA in these cell lines implies a complex pattern of rearrangements has occurred during the formation of these hybrid cell lines. These rearrangements contrast with the generally unrearranged nature of transgenomes produced from fusion of mammalian cells (Bickmore et al., 1989). Human DNA sequences introduced in YACs from *S. cerevisiae* background into mammalian cells, by PEG induced fusion, are found to exhibit slight degrees of rearrangement (Pachnis et al., 1990; Pavan et al., 1990; Huxley et al., 1991). The *S. cerevisiae* DNA which is cointegrated in these latter studies has not been well characterised and therefore cannot be directly compared to our own cell lines. The transgenome content of F1.1 and F7.2 reflect a pattern more akin with that of a CMGT type establishment. The rearrangements detected may reflect some fundamental difference between the fusion processes of these vertebrate / fungal hybrids, and those of the more common vertebrate / vertebrate fusions. This may be because the nuclei of the *S. pombe* protoplasts are unstable within the mouse cell or because the introduced yeast DNA is intrinsically less stable in a murine environment. If the *S. pombe* nuclei are unstable in the mouse cell, enzymes involved in DNA degradation may gain access to the DNA as has been suggested to occur during CMGT (Simmons et al., 1978). It would be interesting to study the direct products of the fusion process to see if the *S. pombe* nuclei disassemble upon entry into the mouse cell or whether the *S. pombe* nuclei retains their structure. Ward, (1986), presents electron micrographs of *S. cerevisiae*-hamster cell heterokaryons which show clearly that the *S. cerevisiae* sphereoplast within the hamster cytoplasm is in an apparently intact state. It is not known however whether

these heterokaryons enter mitosis synchronously or how transgenome formation progresses. If there is a lack of synchrony the yeast nuclei may disassemble at a different time to that host nuclei leading to access of introduced sequences to degradative nucleases. Alternatively yeast sequences may be inherently unstable in mammalian cells. This may be due to lack of appropriate DNA modifications such as DNA methylation at time of entry, which may lead to increased accessibility to enzymatic activity. The differing sequence content of yeast DNA as opposed to mammalian DNA may also cause rearrangement to occur.

Knowledge of which sequences are found within the F1.1 and F7.2 hybrid cell lines allows the analysis of specific regions of the transgenomes to be undertaken at a more detailed level and also allows us to extract larger regions surrounding these markers so that a detailed analysis of an extended region can be undertaken in mammalian cells.

CHAPTER 4. SCREENING OF A *S. pombe* COSMID LIBRARY

4.1 Introduction

The analysis in the previous chapter suggests that the F1.1 hybrid cell line contains a significant amount of the *S. pombe* genome (Chapter 3). The interactions of the *S. pombe* sequences in this insert with the proteins of the mouse cell are presumably responsible for the constriction observed in F1.1 (Figure 1.1). It is therefore desirable to obtain large contiguous regions of *S. pombe* DNA, to allow a more complete picture of the interactions in this DNA to be gained.

To obtain large regions of the *S. pombe* genome a cosmid library (a gift of Prof. Yanagida, Kyoto) was screened for regions of the genome known to be in the F1.1 cell line from the analysis described in Chapter 3. Cosmids were selected as a source of large regions of the *S. pombe* genome since they may contain between 33 and 45kb of inserted sequence. The cosmid vector used was the shuttle vector PSS10 (Nakaseko *et al.*, 1986). This vector contains the *bla* gene, which allows selection for ampicillin resistance in *E. coli*, and also the *LEU2* gene which allows complementation of *leu1* mutations in *S. pombe*. Approximately 40kb *Sau* 3A partial fragments from an *S. pombe* genomic digest were inserted into the *Bam* *H1* site of this vector. Deletions caused by recombination between any repeated sequences contained within the cosmid can be minimised by growing the cosmids in recombination deficient *E. coli* strains such as DH1.

4.2 Screening of the *S. pombe* cosmid library.

As described in Section 2.5, the *S. pombe* cosmid library was plated out as replica copies of each half of the library. 1000 cosmid clones were screened. 1000 clones (42Mb) should allow screening of roughly 2.5 *S. pombe* genome equivalents (17.5Mb) and thus most sequences in the *S. pombe* genome should be represented in this population. This size of library permitted ease of screening, and location of clones within the population.

The probes used for screening the library were 0.7kb *rNTS*, *ura4*, *ade6* and *stb*. The strongest colonies were selected, for further screening. All probes used in the screening, successfully provided cosmid clones apart from the *stb* clone.

This primary screening of the cosmid library provided many cosmids requiring further analysis. The *0.7kb rNTS* probe identified 5 clones, *ura4* 3 clones and *ade6* 1 clone.

The 5 clones obtained with the *0.7kb rNTS*, implies that the ribosomal sequences are greatly under represented this population. The ribosomal sequences account for 10% of the *S. pombe* genome (1.2Mb), yet only account for 0.5% of the cosmid library. This low ribosomal content of the library may be due to the fact that ribosomal sequences exist as tandem repeats, which may be prone to deletion during culture in *E. coli*.

4.3 Secondary screening of cosmid clones.

The clones selected from the primary screening of the cosmid library were picked, cultured and DNA prepared for secondary screening. The cosmids were cut with restriction enzymes to allow sizing and analysis of the clones.

4.3.1 rDNA containing cosmids.

The ribosomal cosmids crm 2, 3, 4 and 5 were picked purely as a test of screening the library. The ribosomal sequences contained in these cosmids cross hybridise with mouse ribosomal sequences in the hybrid cell lines thus preventing their use in analysing the F1.1 and F7.2 cell lines (data not shown).

All cosmids selected with the *0.7kb rNTS* probe contain a 7kb Eco R1 band hybridising to the *0.7kb rNTS* (Fig 4.1). We cannot establish whether this is the correct size of Eco RI fragment since not all of the Eco RI cleavage sites in the *S. pombe* rDNA domains have been mapped (Toda et al., 1984). Ribosomal cosmids 2 and 5 (figure 4:1a) contain Eco R1 fragments of approximately 5.5kb hybridising to *0.7kb rNTS*, in addition to the 7kb fragment common to all 4 cosmids. Similarity of the EcoR1 digestion pattern (Fig 4.2 a) from these two cosmids would suggest that there is positional identity between these cosmids, with Crm 2 containing only a subset of the sequences from Crm 5. Cosmids 3 and 4 also contain a similar pattern of Eco R1 bands, suggesting that they may be overlapping, however the *S. pombe* rDNA complex is estimated to be 150Mb, we have only isolated ~180kb of this DNA in cosmids and thus the isolation of overlapping clones seems unlikely and it is more likely that they contain rDNA copies that are very similar to each other.

4.3.2 Cosmids selected for the *ade6* locus

FIGURE 4.1 : SECONDARY SCREENING OF *S. pombe* rDNA COSMIDS.

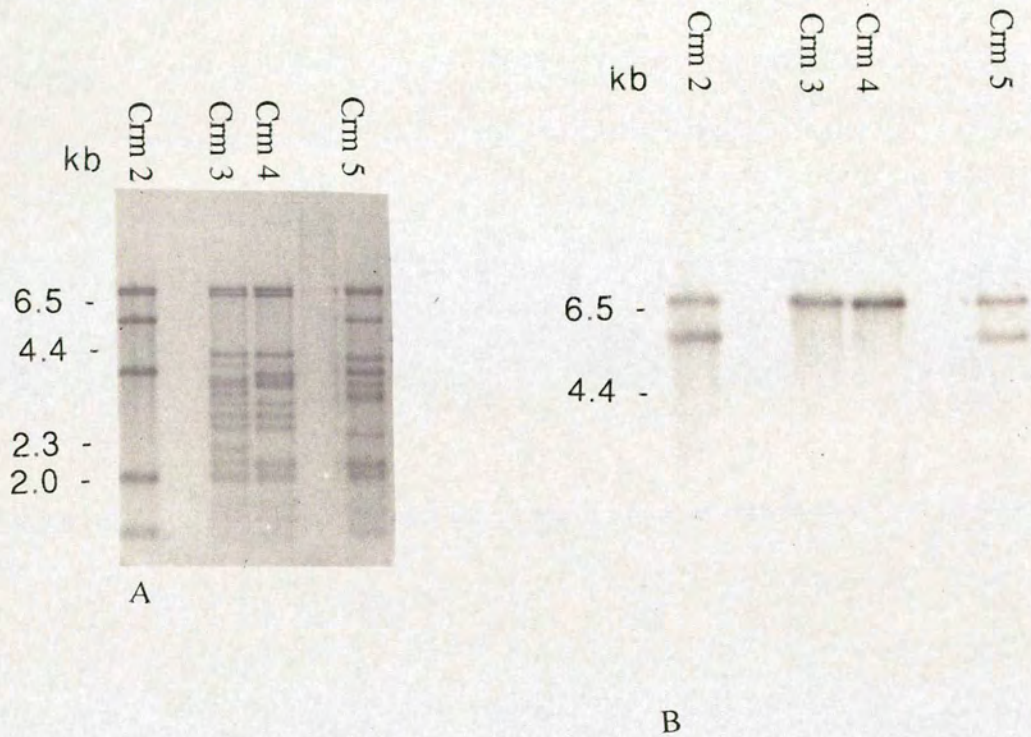


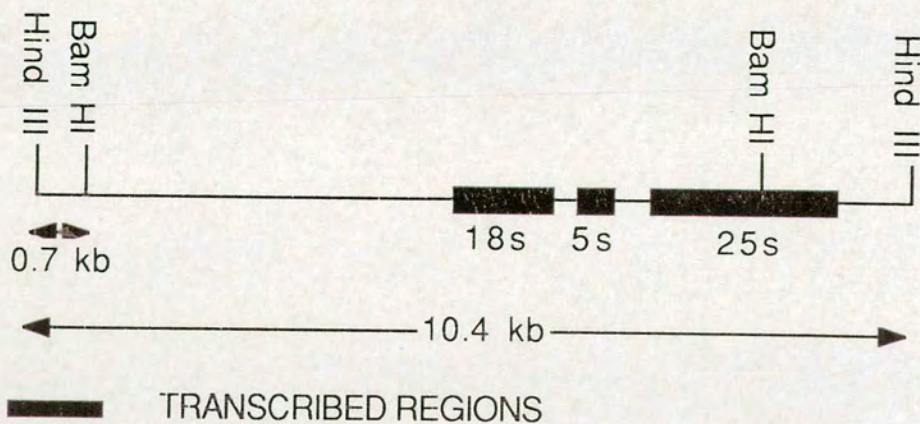
Fig 4.1 : A - Ethidium Bromide staining of Eco RI digests of cosmids Crm 2-5.

Approximately 2.5 μ g of DNA per digest. B- autoradiograph of digests from A

probed with 0.7kb rNTS. C- schematic diagram of *S. pombe* ribosomal subunit.

(Toda *et al.*, 1984)

c. RIBOSOMAL SUBUNIT



Only one *ade6* hybridising cosmid was recovered, Cade6-I. This cosmid is 43.4kb in size (Table 4:1), and contains two bands of hybridisation to the *ade6* probe in Eco R1 digests (Fig 4.2b). These hybridising bands are 3.3 and 0.5 kb in length. Figure 4.2c shows that the *ade6* locus of *S. pombe* does not contain any mapped Eco R1 sites in the gene sequence (Kohli, personal communication). Hybridisation of the *ade6* probe to *S. pombe* DNA cut with Eco RI revealed a hybridising band of 3.3 kb (Fig 3.2e). It would therefore appear that Cade6-I contains an intact *ade6* locus.

4.3.2 Cosmids from the *ura4* locus.

Three cosmids were obtained which hybridised to a probe for the *ura4* gene. In Eco R1 digests Cura4-I and -II both contain a *ura4* hybridising 6.5kb band (Figure 4:3b). These cosmids are also in the maximal range for insert size (Table 4:1). Cura4-I and Cura4-II appear to be overlapping as their Eco RI profiles are very similar, although there is a 3.4 kb band from Cura4-I not present in Cura4-II and a 2.8kb band in Cura4-II not present in Cura4-I, presumably these bands are from opposite ends of the cosmids. Cura4-III contains a different *ura4* hybridising band (3.8kb) from Cura4-I and II and the Eco RI profile is markedly different from both of these cosmids. These differences between the cosmids Cura4-I, -II and -III cosmids suggest that Cura4-I and II are highly overlapping, with the majority of their sequences from one side of the locus. Cura4-III must overlap these cosmids but the majority of the Cura4-III sequences must come from the region on the opposing side of the *ura4* gene to those of Cura4-1 and -II or else it may contain rearranged sequences. Since the flanking sequences to this locus are not well mapped it is difficult to know the orientation of Cura4-I and II relative to the genomic *ura4* locus. We have confined most of our subsequent analysis to Cura4 -II.

4.4 Cosmids tested on cell lines.

The cosmids obtained from the screening of the cosmid library were to be used for the analysis of *S. pombe*/ mouse hybrid cell lines. Since this involves using regions of one genome to analyse samples of a mixed genomic content it is important that the *S. pombe* cosmids used do not cross hybridise to the mouse genome. Cosmids were therefore tested for cross hybridisation to the mouse genome by Southern blotting.

The ribosomal cosmids were not tested since the ribosomal sequences of both *S. pombe* and mouse have previously been shown to cross hybridise (Allshire, personal communication). The Cade6-I and Cura4-II cosmids were tested on panels of hybrid cell line DNAs plus C127 and *S. pombe*.

FIGURE 4.2 : SECONDARY SCREENING OF Cade6-I

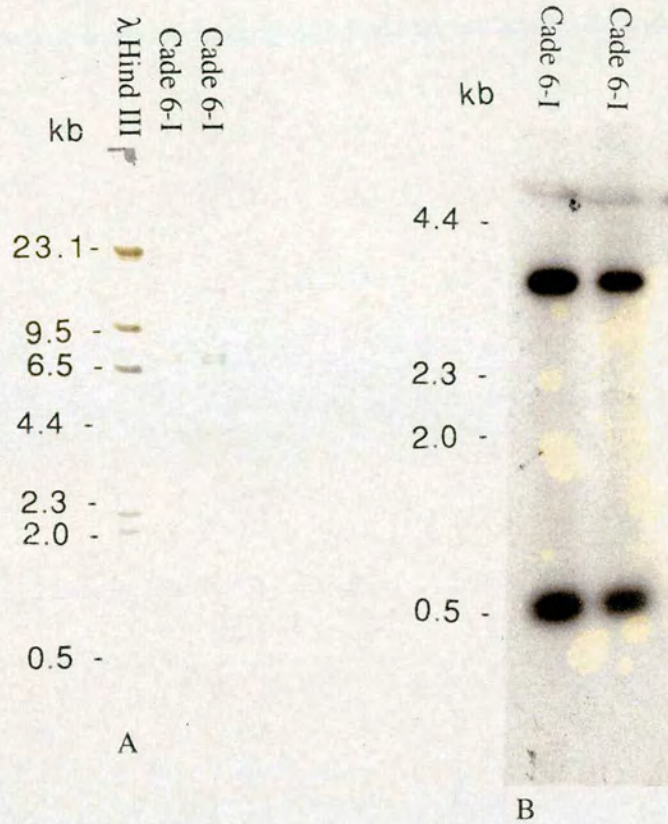


Fig 4.2 : A- Ethidium bromide staining of Eco RI digests of Cade6-I. Approximately 2.5µg of DNA per digest. B- Autoradiograph of digests from A hybridised to a 2.5kb Xho I/Bam HI *ade6* probe. C - Schematic diagram of the *ade6* region of *S. pombe* chromosome III (Kohli, personal communication).

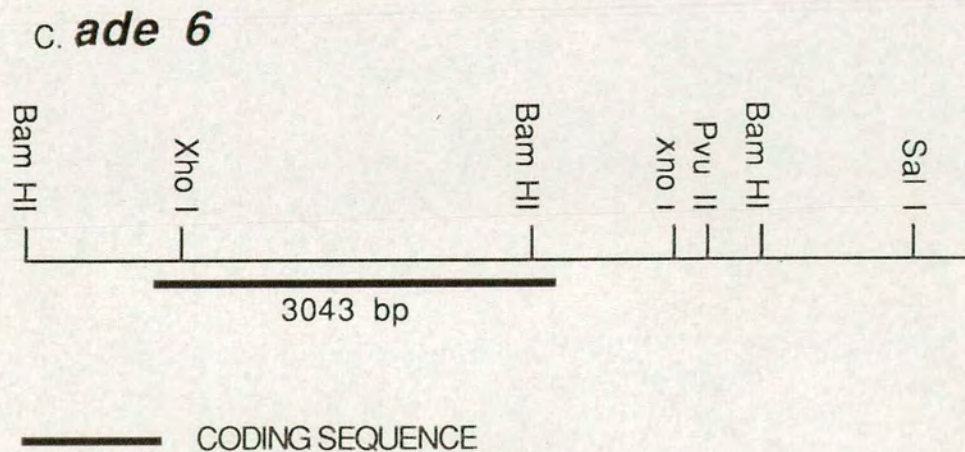


FIGURE 4.3 : SECONDARY SCREENING OF *S. pombe ura4* COSMIDS.

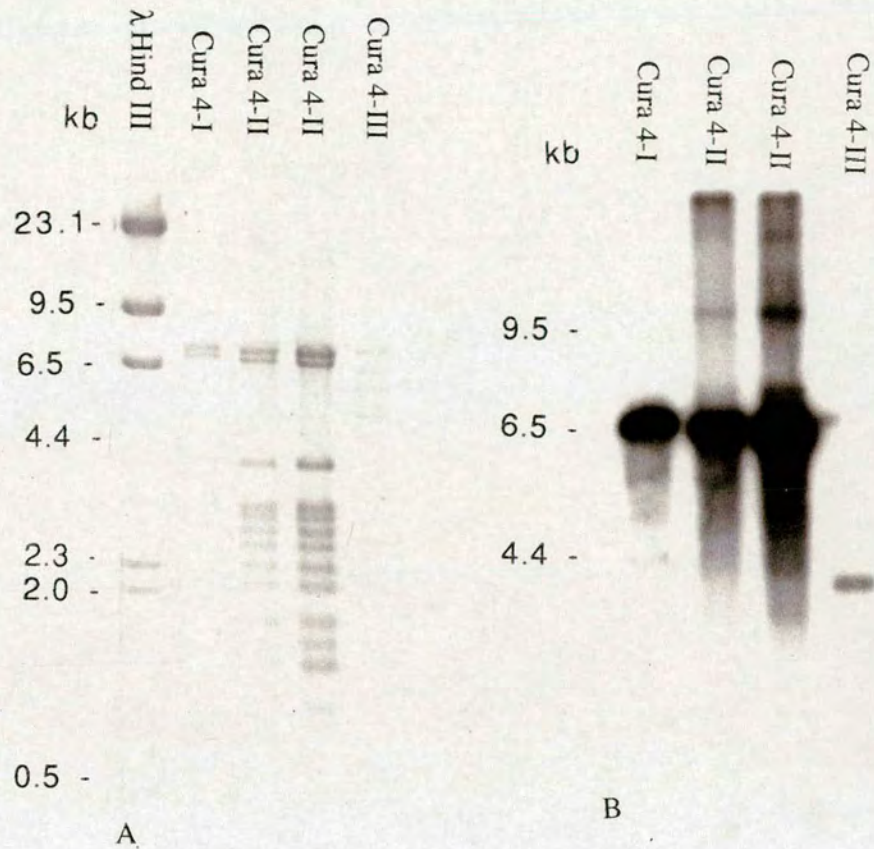


Fig 4.3 : A - Ethidium bromide staining of Eco RI digests of Cura 4-I to III.

Approximately 2.5 μ g of DNA per digest. B- Autoradiograph of digests from A hybridised to a 1.8kb Hind III *ura4* fragment. C- Schematic diagram of the *ura4* gene (Allshire *et al.*, 1987).

C. **URA 4**

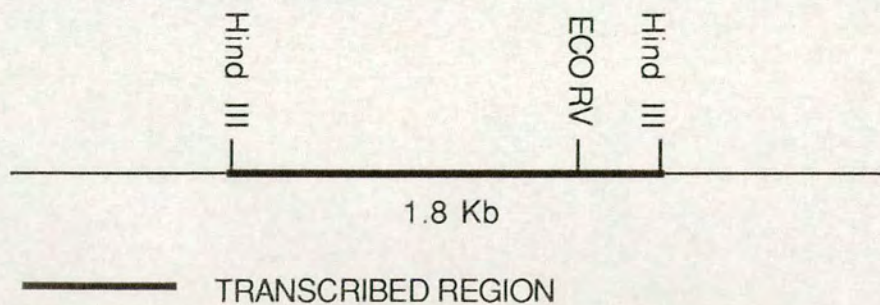


TABLE 4.1 SIZES OF SELECTED COSMIDS FROM AN

S. pombe COSMID LIBRARY

COSMID	SIZE (kb)	HYBRIDISING BANDS (kb)	PROBE
Crm II	24.2	7.0, 5.5	<i>0.7 kb rNTS</i>
Crm III	43.9	7.0	
Crm IV	40.4	7.0	
Crm V	42.6	7.0, 5.5	
Cade 6-I	43.4	3.4, 0.5	
Cura 4-I	43.1	6.5	<i>ura4</i>
Cura 4-II	42.4	6.5	
Cura 4-III	43.4	6.5	

TABLE 4.1

SIZE: total size of cosmid, vector plus insert.

HYBRIDISING BANDS correspond to the bands which hybridise to the probe appropriate for cosmid type.

Figure 4.4 shows that the Cade6-I cosmid does not cross hybridise to the C127 mouse genome, thus allowing this cosmid to be utilised for future analysis. The hybridisation pattern of this cosmid to F1.1 is very similar to the hybridisation pattern of *S. pombe* DNA, suggesting that substantial regions of this DNA have been retained intact during the formation of F1.1. In the Hind III tracks only one band differs in the *S. pombe* and F1.1 hybridisations (8kb in F1.1, 5kb in *S. pombe*). All the other cell lines analysed show more extensive rearrangement of the DNA hybridising to Cade6-I. The F7.2 hybridisation is limited to only a few bands, suggesting substantial rearrangement of this region has occurred in F7.2. Two transfectant cell lines have also been studied, DC.1 and DC.2. These cell lines were established by the introduction of naked *S. pombe* DNA to C127 cells and selected for resistance to G418 (Allshire *et al.*, 1987). These cell lines have been analysed for comparison of the introduced material between transfectant and fusion derived cell lines. In the DC.1 and DC.2 cell lines some of the *S. pombe* Cade6-I hybridising bands are absent, suggesting that neither of these two hybrids formed by DNA transfection have retained the *ade 6* locus in its entirety

Hybridisation of Cura4-II to the DNAs from hybrid cell lines is shown in Figure 4.5. As seen with Cade6-I (Figure 4.4) the hybridisation is very similar in pattern between the F1.1 cell line and yeast DNAs. The major difference is a band lacking from the F1.1 hybridisation at ~6.5 kb but present in *S. pombe Int 5*. This suggests that as with the *ade 6* locus this region has entered the F1.1 cell line with only slight rearrangement occurring. The F7.2 cell line is highly rearranged for this region with a new band being formed of ~18 kb, and only the bands between 5.8 and 7.5 kb remaining from the pattern seen in *S. pombe*. DNA from the hybrid cell line F7.1, a hybrid cell line containing a large autonomous *S. pombe* element, (Allshire *et al.*, 1987) exhibits a pattern of hybridisation within which all of the bands present in the yeast hybridisation are present but it also includes some new novel bands (e.g. 16kb band in F7.1). The above pattern of hybridisation suggests that entry of more than one copy of the locus into the cell, or duplication of the locus which entered the cell followed by rearrangement of the other copy/s. The transfectant cell lines DC.1, DC.2 demonstrate extensive rearrangement of the *ura4* region.

All the cell lines studied contain extensive amounts of DNA from the *ura4* region as would be expected since this is the DNA which flanks the selectable marker for the establishment of the cell lines. In contrast the *ade6* region, ~1Mb from the *ura4* locus exhibits extreme variance in the amount of DNA from this region which is retained in the cell lines. This pattern is not unexpected, since the further away the sequence is

FIGURE 4.4 : Cade6-I CONTENT OF FUSION AND TRANSFECTANT CELL LINES.

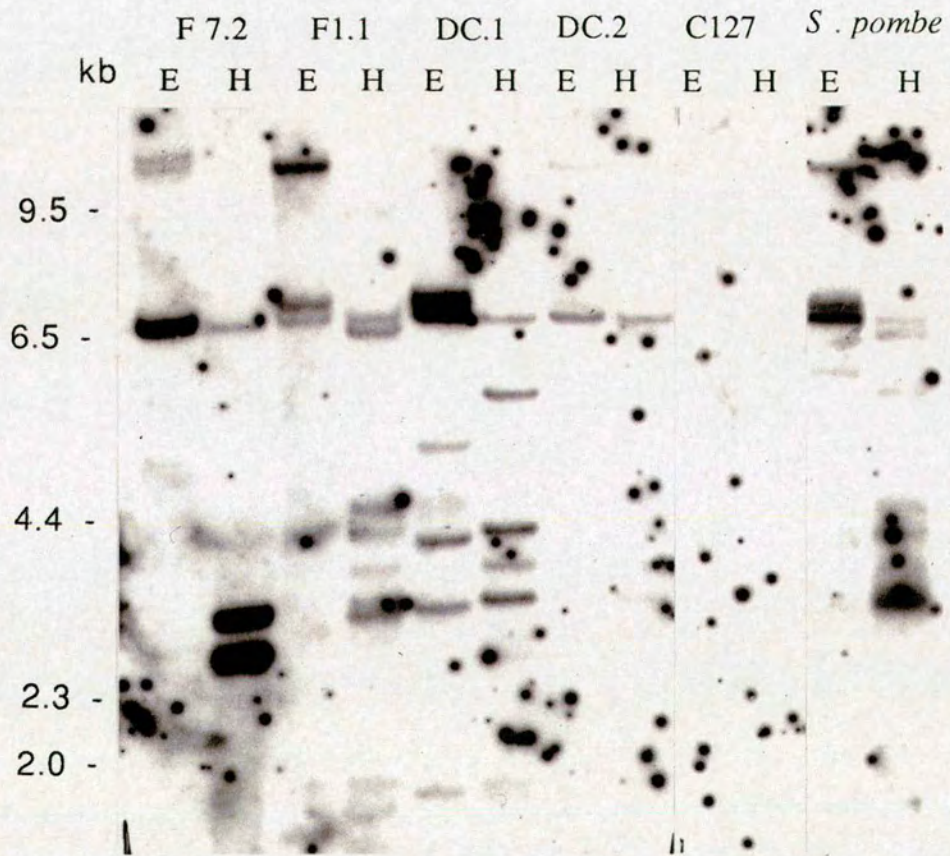


Fig 4.4 : 10 μ g of DNA from cell lines F1.1, F7.2, DC.1, DC.2 and C127 and 100ng of yeast (*S. pombe*) DNA were digested with either Eco RI (E) or Hind III (H) as indicated. Digests were electrophoresed and then DNA transferred to a nylon membrane. Cade6-I DNA which had been random primed was used to probe the DNAs.

FIGURE 4.5 : Cura4-II CONTENT OF FUSION AND TRANSFECTANT CELL LINES.



Fig 4.5 : 10 μ g of cell line DNA from F1.1, F7.2, F7.2, DC.1, DC.2 and C127 and 100ng of yeast (*S. pombe*) DNA were digested with Eco RI. After gel electrophoresis, the DNA was transferred to a nylon membrane. Random primed Cura4-II DNA was used to probe these digests.

from the selectable marker the less likely it is to be co-integrated or be part of the same extrachromosomal element.

The F1.1 and F7.2, yeast - vertebrate fusion, cell lines demonstrate a pattern of co-retention within the recipient cell which is largely dependent on synteny with the selectable marker (Chapter 3). The DNA which has entered the mouse cells during formation of these hybrid cell lines has also been rearranged. In previous yeast - mammalian cell fusions it has been found that rearrangements are fairly common (50%), although usually slight, within the transgenomes formed (Pachnis *et al.*, 1990; Pavan *et al.*, 1990; Huxley *et al.*, 1991). These transgenomes however are largely derived from human sequence based YACs, and may not therefore reflect the situation found with introduced *S. pombe* sequences. The yeast DNA which was found to co integrate in these studies has not been extensively studied and therefore little comparison can be made.

Obtaining these cosmids should aid analysis of the F1.1 cell line, since their extensive sequence content will allow investigation of large regions of the yeast transgenome.

CHAPTER 5 METHYLATION OF *S.pombe* SEQUENCES IN THE F1.1 HYBRID CELL LINE.

5.1 INTRODUCTION.

CpG DNA methylation is a DNA modification found in vertebrates and is present throughout the vast bulk of mammalian genomes (Bird, 1987). The function of this modification in the vertebrate genome has long puzzled investigators (Bird, 1992). Its fundamental importance in vertebrate development has been demonstrated by the creation of methyltransferase mutations in mice by homologous recombination (Li *et al.*, 1992). The phenotype of these mice, (which have less than 30% of normal methyltransferase activity) is embryonic lethality. These experiments finally reveal that methylation is of vital importance to vertebrate development. The embryonic lethality may be caused by the low methyltransferase activity preventing the silencing of specific genes at particular times during development. Although these experiments suggest that methyltransferase is required for development, the cultured embryonal stem (ES) cells carrying the same mutation can tolerate the 70% drop in methyltransferase activity and the accompanying three fold drop in genomic methylation levels, with no apparent phenotype. If methylation is involved in regulation of gene expression, the unchanging developmental status of ES cells in culture may not be an appropriate environment to assess its effects. The level of methylation provided by a 30 % level of methyltransferase activity may be sufficient to maintain appropriate cellular functions. These recent findings suggest at least that methylation is not essential for genome viability, chromosome condensation or segregation.

The question remains as to why vertebrates exhibit global methylation of their DNA? It has been shown that despite the rarity of CpG dinucleotides in the genome and a dedicated repair mechanism, over one third of all point mutations causing human genetic diseases are caused by mutation of CpG to TpG (Cooper *et al.*, 1983). The embryonic lethality of certain methyltransferase mutations (Li *et al.*, 1992) suggests that although DNA methylation increases the genetic load its importance in development must be great enough to counterbalance this methylation may also function to minimise the effects of retroviruses and other transposable elements within the genome. It appears that methylation plays an indirect role in gene expression mechanisms, and may also be involved in chromosome packaging, through the activity of two proteins, methylated CpG binding proteins (MECP I and II) which bind methylated cytosine residues in a non sequence specific manner (Lewis *et al.*, 1992). MECP I has been shown to bind sequences containing a minimum of 12 methylated CpGs and to mediate indirect repression of transcription through binding of

methylated promoters.(Meehan *et al.*, 1989; Boyes and Bird, 1992) The extent of repression depends on the strength of the promoter and also on the level of CpG in the sequence (Boyes and Bird, 1992). A role for MECP I in gene repression is implied, since extracts with low MECP I levels inefficiently repress methylated genes (Boyes and Bird, 1991). MECP II has recently been cloned and characterised (Lewis *et al.*, 1992). Unlike MECP I, MECP II only requires one methyl C residue for binding. Immunocytochemistry with antibodies directed against MECP-II show that it colocalises with levels of methyl CpG in the genome. MECP II may also be involved in the protection of methylated regions from nucleases since it is present at high levels in brain cells where protection is high, and present in low levels in PC13 cells where it is well established *in-vitro* that the level of protection of methylated regions is also low. Methylated regions of chromatin are more resistant to cleavage by exogenous nucleases than unmethylated regions (Antequera *et al.*, 1988). It has also been shown that after replication a methylated region containing the VDJ recombinase recognition sequence becomes protected from VDJ recombinase cleavage through the creation of a chromatin configuration which was not present prior to replication even if the sequence was previously methylated (Hsieh *et al.*, 1992). The MECP proteins have been proposed to be involved in the higher order packaging of chromatin to create these protected methylated regions (Meehan *et al.*, 1987).

Since the *S. pombe* genome has no CpG methylation, if the yeast insert in F1.1 preserves this methylation free state, no MECP proteins will bind in this region. This might contribute to the constriction observed in the *S. pombe* sequence region of the F1.1 hybrid chromosome.

However cell lines have been shown to accumulate methylation during their life in culture. This accumulation can affect virtually all non-essential sequences including genes not required for life in the tissue culture flask (Antequera *et al.*, 1990). However, not all established mammalian cell lines contain *de novo* methylase activity (Szyf *et al.*, 1989) They may not express the appropriate methyltransferase or they may express high levels of a demethylase. Vertebrate cells in culture could silence the methylase function with no apparent affect to the cell, the experiments of Li and Bestor (1992) suggest that methyltransferase is at best required at low levels in cultured cells and may be entirely non essential.

We do not know whether C127 cells contain *de novo* methylase activity. If they do then we might expect to see progressive methylation of of the *S. pombe* insert in F1.1. If the lack of methylation of the yeast DNA in F1.1, and consequent absence of methyl-binding proteins, contributes to the cytological appearance of the yeast insert in F1.1 then

we might expect to see changes in this cytology accompanying the process of progressive methylation of the DNA during passage in culture. If there is no increase in methylation of the yeast transgenome this could be the prime cause of constriction.

In this chapter I have analysed the methylation status of yeast DNA sequences in F1.1 at various passage stages of culture and compared this to the cytological appearance of chromosomes at these time points.

5.2 Quantitative analysis of methylation of cell lines.

In Chapter 3 I analysed which regions of the *S. pombe* genome have entered the C127 genome to form the F1.1 hybrid cell line. This now allows me to examine any changes in the methylation status of different regions of the yeast transgenome in the hybrid cell line.

CpG methylation is most easily studied by cutting DNA with the isoschizomers Hpa II and Msp I. These restriction endonucleases cleave the same recognition sequence CCGG. Msp I will cleave the sequence irrespective of the methylation status of the second C, while Hpa II will only cleave this sequence if the second C is unmethylated. Therefore methylation of this site can be detected by running digests of DNA cut with each enzyme side by side during electrophoresis. After transfer to hybridisation membrane the methylation status of discrete regions can be determined by hybridisation with appropriate probes. Methylation of CCGG sequences results in Hpa II hybridising bands that are longer than those produced by Msp I.

Since Antequera et al (1990) have shown that endogenous levels of methylation can increase during passage of tissue culture cells I have characterised the overall methylation of the cell lines under study. 5ug of mammalian DNAs, from the F1.1, F7.2 and C127 cell lines and 50ug of *S. pombe* DNA were cleaved with the isoschizomers Hpa II and MspI. The ends of the fragments generated by these enzymes were then partially filled in by end labelling with $\alpha^{32}\text{PCTP}$ to allow quantitation of the fragments. The end labelled reactions were then resolved by gel electrophoresis after which the gel was fixed, dried and subjected to autoradiography (Section 2.16). The autoradiographs from these reactions were then subjected to densitometry, by Dr D Green, to allow profiles of the weight average of fragments derived from these digests to be generated.

Figure 5.1 shows the resulting autoradiographs and the densitometry profiles of the various cell lines. The absence of methylation in the *S. pombe* genome is immediately apparent from the similarity of the Hpa II and Msp I profiles of the autoradiograph and also of the densitometry profile (Lanes 7 + 8). By comparison, as expected, all three of the

FIGURE 5.1 COMPARISON OF OVERALL CELLULAR METHYLATION LEVELS BETWEEN CELL LINES AND IN YEAST

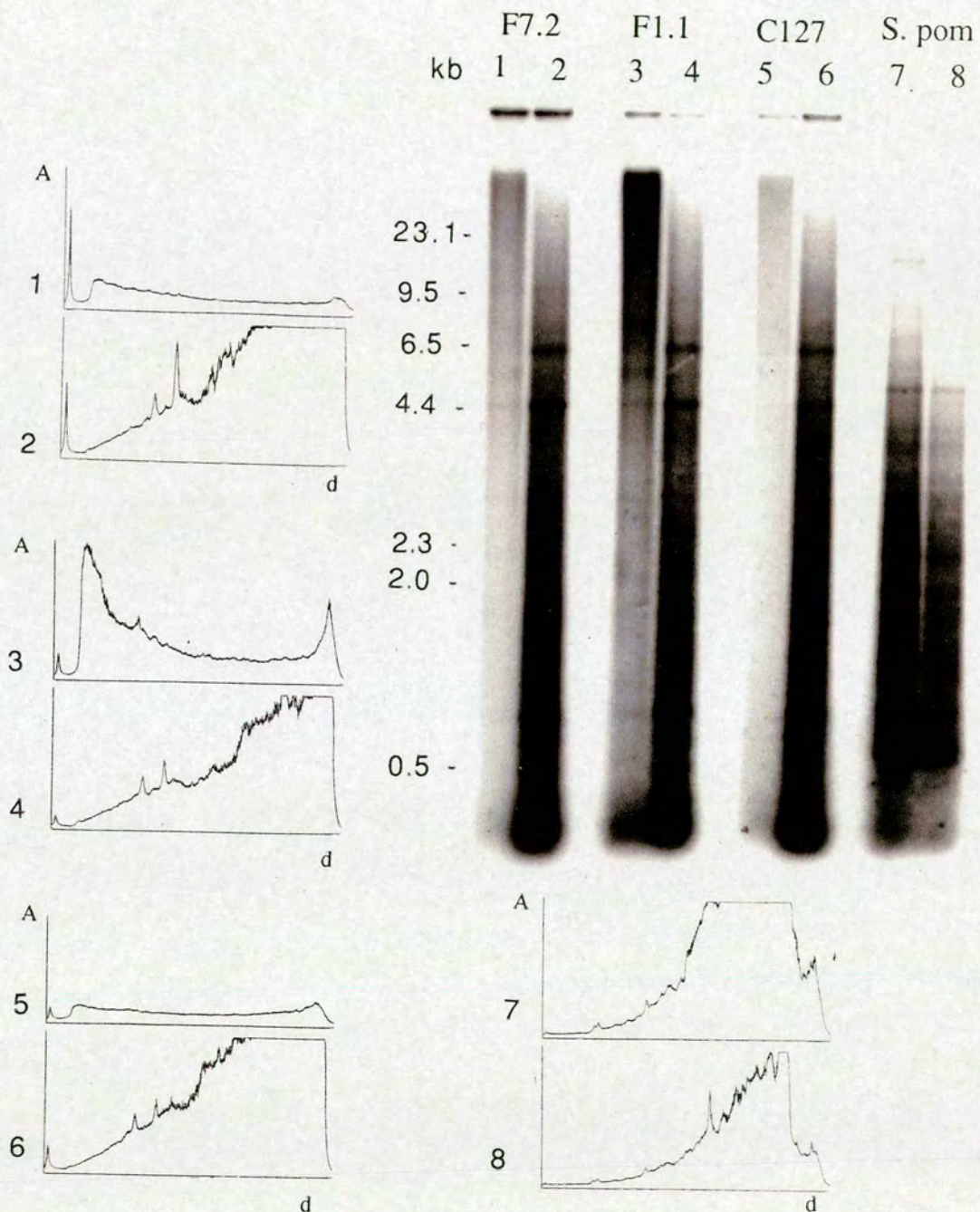


Fig 5.1 DNAs from passage number 30 of F7.2, F1.1, C127 and from *S. pombe* were cleaved with Hpa II (tracks 1, 3, 5 and 7) or Msp I (tracks 2, 4, 6 and 8) and fragments end labelled with $\alpha^{32}\text{P}$ dCTP. Aliquots were run out by gel electrophoresis. The gel was subsequently fixed, dried and autoradiographed. Autoradiographs were subjected to densitometry to produce profiles of fragment distribution for each reaction. Profile numbering corresponds to lane numbering. d= distance from origin. A= absorbance.

mammalian genomes (F1.1, F7.2 and C127) used in this study exhibit CpG methylation. This is apparent from the large difference between the profiles of Hpa II and Msp I digests of these cell lines. The Msp I reactions show a single curve profile with increasing numbers of molecules at decreasing size ranges. By contrast the Hpa II profiles show two discrete peaks of density, one containing large >15kb fragments, while the other contains small <500bp size fragments, demonstrating that mammalian genomes are largely methylated with small discrete unmethylated domains, which correspond to CpG islands. Quantitative differences can be observed between the reactions for each of the mammalian genomes under study, this largely being caused by the amount of sample added. If the Hpa II and Msp I peaks are compared for any particular cell line, no significant qualitative differences can be seen between F1.1, F7.2 and C127 DNAs. The hybrid cell lines were grown for 30 generations, after recovery from liquid nitrogen. Whilst this is too short a time to see any general accumulation of methylation within a cell line it does confirm that the process of fusion to the yeast spheroplasts and establishment of the resulting hybrid cell lines has not changed the underlying methylation pattern of the C127 genome and that the *S. pombe* content of the hybrid cell lines F1.1 and F7.2 does not detectably affect the overall methylation levels of the DNA. These results however do not tell us about the methylation levels that may exist on the *S. pombe* sequences themselves in these hybrid cell lines since they account for a tiny percentage (> 0.1%) of the DNA analysed.

5.3 Methylation levels in the *S. pombe* insert of F1.1

To assess the methylation status of the *S. pombe* insert in the F1.1 cell line, DNAs from F1.1, C127 and *S. pombe* cells were cut with Hpa II, Msp I, and Bam HI in conjunction with either of the former enzymes. These digests were resolved by gel electrophoresis, and then transferred to filters. *S. pombe* probes shown to be present in the *S. pombe* insert of F1.1, (Chapters 3 and 4), were used to probe these digests.

Initial methylation studies were done on DNAs from cells at early passage stages (<passage 30 from cells initially stored in September 1986). When digested with Hpa II, Msp I and probed with *S. pombe* sequences no methylation of the *S. pombe* insert in F1.1 was apparent. Fig 5.2 demonstrates this quite clearly. *Nda3*, a probe from the distal end of *S. pombe* chromosome II, was used to probe Hpa II and Msp I digests of F1.1 DNA harvested from cells at passage 16. As expected no methylation of *nda3* is apparent in the *S. pombe* genome. In F1.1 DNA the hybridisation pattern to the *nda3* probe is identical to that of the *S. pombe* genome, i.e. there are no new bands in the Hpa II track not apparent in the Msp I digested DNAs. Thus no methylation of this locus has occurred in F1.1 cells

FIGURE 5.2 METHYLATION STUDY OF THE *nda3* REGION FROM F1.1 CELLS AT PASSAGE 16.

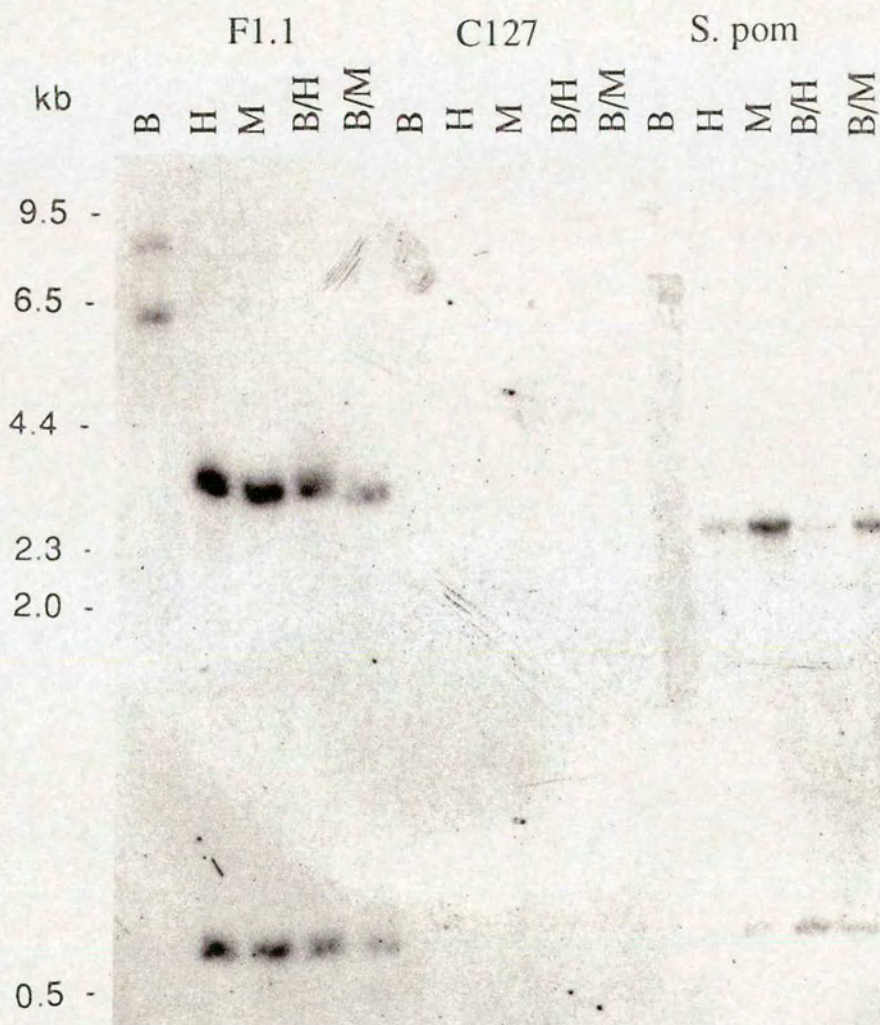


Fig 5.2 10 μ g of F1.1 and C127 cell line DNAs and 50ng of *S. pombe Int-5* (*S. pom*) DNA were cleaved with B- Bam HI, H- Hpa II, M- Msp I or combinations of these enzymes as indicated. After gel electrophoresis DNA was transferred to filters and probed with *nda 3*. Passage number of F1.1 cells =16

at these early passage stages and this is also true of other probes used on F1.1 passage 16 cells (data not shown).

The absence of methylation of the *S.pombe* insert in these F1.1 cells may have several causes. Firstly, the C127 cell line may not contain a *de novo* methylase activity, as has been found to be the case in many other cell lines (Szyf *et al.*, 1990), or alternatively the hybrid cell line sampled at passage 16, may be too young, for enough generations to have passed for methylation to have accumulated to high enough levels to detect by Hpa II/ Msp I hybridisation.

After cosmid clones for the *ade6* and *ura4 S. pombe* loci were obtained, (Chapter 4), these clones could be used to assess methylation over more extensive regions of the *S. pombe* transgenomes. These cosmids were not used to analyse F1.1 methylation until after passage 30, which is considerably later than the passage 16 cells used for the *nda3* result in Fig 5.2.

Cade6-I has been hybridised to Hpa II and Msp I digests of F1.1, C127 and *S. pombe Int 5* DNAs in Figure 5.3. This hybridisation was carried out on F1.1 cells harvested at passage 30. In contrast to the findings in Fig 5.2, the F1.1 lanes in Fig 5.3, show a different pattern of hybridisation to that of *S. pombe* DNA. In the F1.1 Hpa II digested lanes there are bands not present in the Msp I digests. This implies that at some of the sites Hpa II is unable to cleave its appropriate sequence, because the sites are methylated. It therefore appears that between passage 16 (Fig 5.2) and passage 30, methylation has reached detectable levels within the *S. pombe* insert of F1.1. There are however still a large number of bands in common between the Hpa II and Msp I digests at this passage and, indeed, none of the Msp I bands are completely absent from the Hpa II lanes. This means that not all sites are methylated completely, and that the methylation pattern is not homogenous throughout the cell population. This methylation detected in the F1.1 *S. pombe* insert at passage 30 indicates that C127 cells do indeed contain a *de novo* methylase activity.

In Figure 5.4 the Cura 4-II cosmid has been hybridised to Hpa II and Msp I digests of F1.1 DNA from cells at passage 50. There has been a large change in the methylation levels of the F1.1 *S. pombe* insert. The C127 lanes show some slight hybridisation which is not observed in any other hybridisations using this cosmid, which must therefore be assumed to be due to plasmid contamination. These contaminant bands do not correspond to any bands in either *S. pombe* or F1.1 digests and therefore do not affect the interpretation of the results gained from the F1.1 hybridisation. In the F1.1 tracks, hybridising bands in both

FIGURE 5.3 METHYLATION STUDY OF THE Cade6-I REGION
OF THE F1.1 *S. pombe* INSERT AT PASSAGE 30.

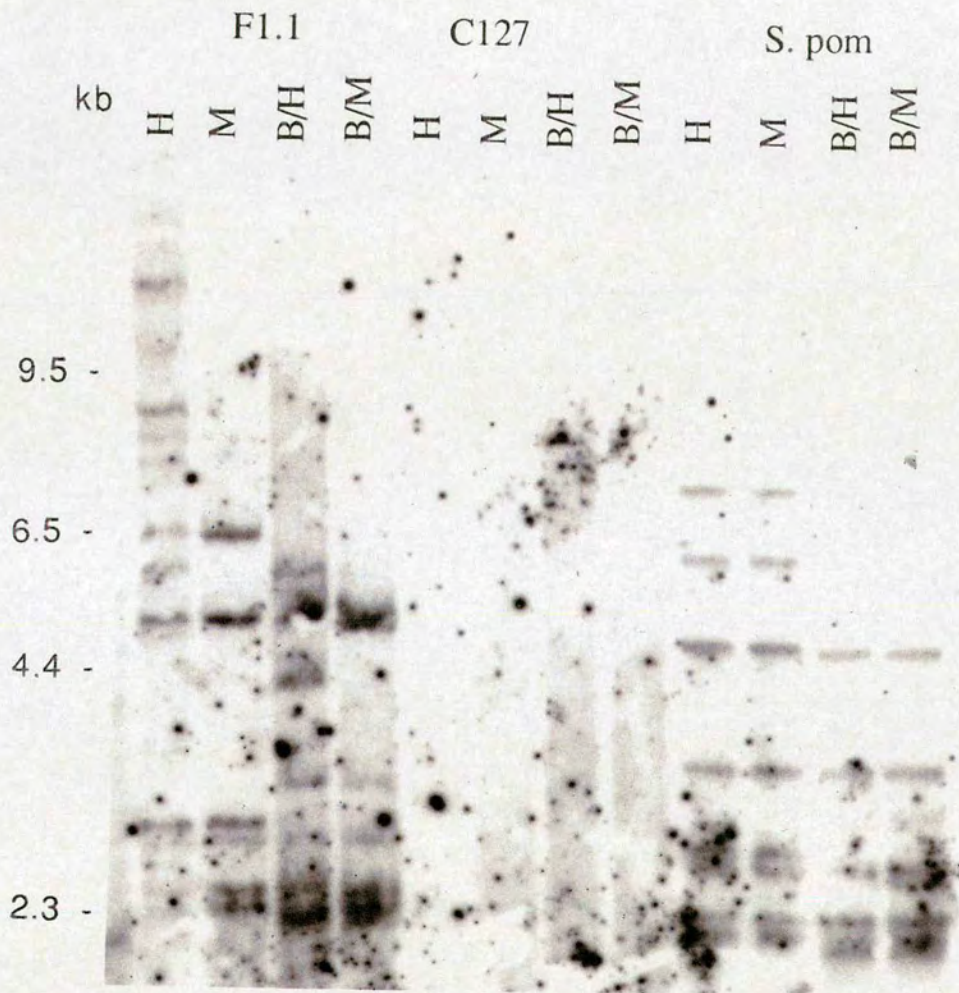


Fig 5.3 10 μ g of F1.1 and C127 cell line DNAs and 50ng of *S. pombe* (*S. pom*) DNA were cleaved with B- Bam HI, H- Hpa II, M- Msp I or combinations of these enzymes as indicated. After gel electrophoresis DNA was transferred to filters and probed with Cade6-I random primed DNA.

FIGURE 5.4: METHYLATION STUDY OF THE Cura4-II REGION
OF THE F1.1 *S. pombe* INSERT AT PASSAGE 50.

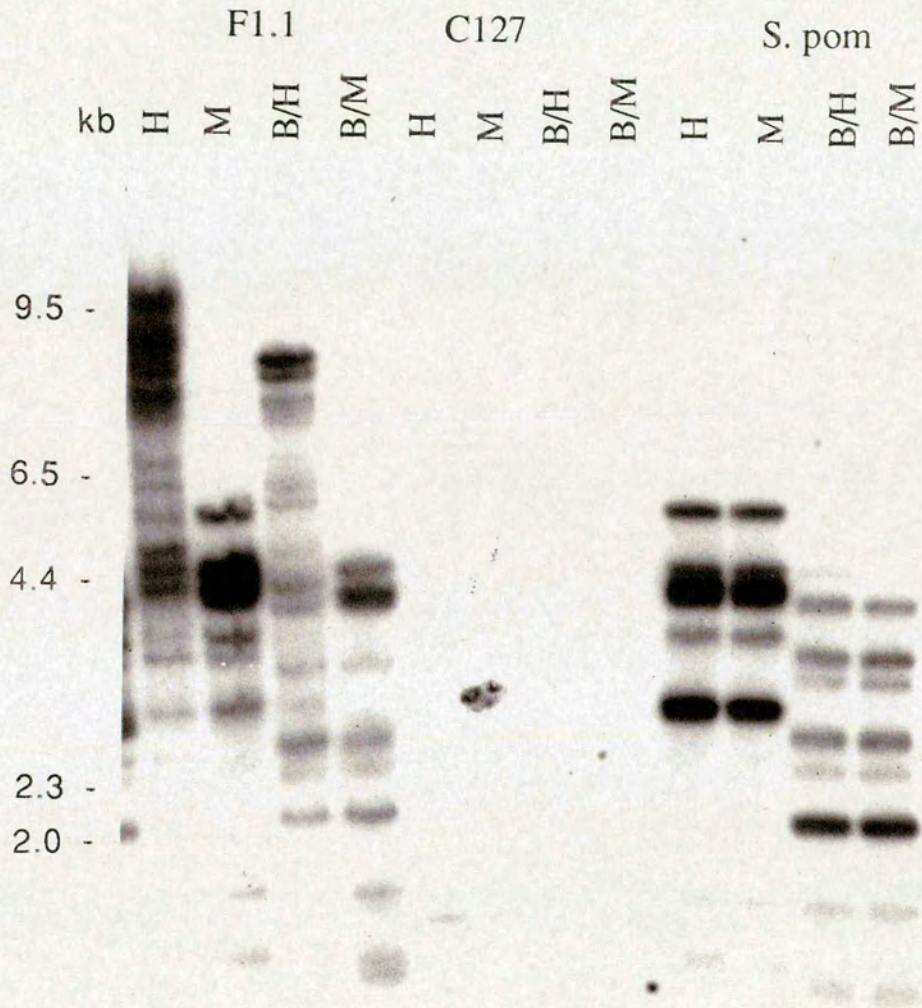


Fig 5.4 10 μ g of F1.1 and C127 cell line DNAs and 50ng of *S. pombe* (*S. pom*) DNA were cleaved with B- Bam HI, H- Hpa II, M- Msp I or combinations of these enzymes as indicated. After gel electrophoresis DNA was transferred to filters and probed with Cura4-II random primed DNA.

the Hpa II and Bam HI/ Hpa II digests are considerably larger than those in the corresponding Msp I lanes. Indeed, some of the Msp I bands (eg. largest band (~4.5kb) in the Bam HI / Msp I) appear to be completely absent from the corresponding Hpa II digest. This indicates that by passage 50 the *S. pombe* insert in F1.1, at least at and around the *ura4* locus, is very highly methylated.

The results in Figures 5.2, 5.3 and 5.4 suggest that a large degree of CpG methylation is gradually accumulating throughout the *S. pombe* insert of the F1.1 hybrid cell line during passage in culture. Appearance of methylation in a previously unmethylated region tells us that C127 cells contains a *de novo* methylase activity. Studies of methylation in cell lines have shown that methylation also accumulates on endogenous sequences gradually during time in culture (Antequera *et al.*, 1990) but this is a much slower process than I have observed for the *S. pombe* transgenome in F1.1. Since methylation of the *S. pombe* insert in F1.1 is increasing, the MECP proteins may be able to bind the yeast DNA in later passage cells. Since these proteins have been implicated in chromosome structure, increasing methylation may lead to a change in the appearance of the insert. It is therefore important to be able to directly correlate methylation levels with the cytology of the cell line to show directly whether the morphology of the F1.1 insert changes with increases in methylation levels. This may also tell us something about the role of methyl CpG binding proteins in mitotic chromosome morphology.

5.4 Time course study of methylation in the F1.1 insert.

Since methylation had been observed in the *S. pombe* sequences of the F1.1 hybrid at various time points during culture, questions can be asked as to the nature of the accumulation of methylation in these regions. By analysing cells sampled at different passages it might be possible to assess whether methylation accumulates stochastically within this insert or in discrete bursts.

To enable me to study methylation within this insert, the earliest cells available were retrieved from storage (November 1986). These cells were slightly older than those of the initial study (September 1986), and therefore the subsequent passages do not correspond to those in Section 5.3. These cells were grown in culture and DNA made from different passage stages. Metaphase chromosome spreads were also prepared from F1.1 cells at distinct passages from which DNA was also made, to allow parallel studies of DNA methylation and cytology to be carried out.

F1.1 DNAs prepared from cells at different passages were cleaved with either Hpa II or Msp I in conjunction with Bam HI. In Figure 5.5, DNA from F1.1 cells at passages 2, 9 and 20 and C127 and *S. pombe* DNAs have been hybridised with the probe *con* (contrapsin), a mouse probe which detects a family of proteinase inhibitors. This probe was used as an internal control for methylation of mouse sequences, to indicate whether there was any general increase of methylation within the mouse genome. Although as shown in Section 5.2, no general increase in methylation of the F1.1 cell line was observed over that of C127 cells, however we wished to ascertain whether there were any temporal changes in methylation of the mouse DNA. No hybridisation is detected in *S. pombe* DNA. The Bam HI/Hpa II lanes and Bam HI/Msp I lanes are identical between all F1.1 DNAs analysed at different passages. The F1.1 lanes are also identical to those of the C127 digests demonstrating that no new methylation has accumulated within these sequences during formation of the fusion hybrid. Therefore the *de novo* methylase activity within F1.1 cells, appears to act primarily only on the introduced yeast DNA and not on the endogenous mouse genome. Preferential activity of methyltransferase within the *S. pombe* domain of F1.1 might be a reflection of a more open chromatin configuration within this region of the F1.1 genome giving preferential access of the DNA methyltransferase to the yeast transgenome.

Figure 5.6 shows hybridisation of *S. pombe* rDNA sequences, using the 0.7kb rNTS probe, to the same digests as shown in Fig 5.5. The F1.1 lanes demonstrate a small amount of methylation even within the earliest passage (P2) analysed. This is not unexpected, since these cells were older than the initial passages studied in Fig 5.2 where no methylation at the *nda3* locus was detected. This methylation gradually increases with time, as seen from the increase in strength of the larger bands in the Hpa II lanes and the decreasing strength of the 1kb Msp I hybridising fragment, with increasing passage number. Indeed by passage 20 the 1kb Msp I band is almost completely absent from the Hpa II lane, indicating that almost all of the cells in the population have methylated these sites. There appears to be no reason to assume that this methylation is not being accumulated stochastically over time, since the increase appears to be steady.

Figure 5.7 demonstrates the accumulation of methylation over time in culture, within the region of the F1.1 transgenome detected by Cade6-I. Methylation is again detectable at a low level in F1.1 cells at the initial passage P2. As the cell line is passaged the methylation within the *S. pombe* insert increases. By passage 20, some of the Msp I hybridising bands (0.6 and 3.0kb) are absent from the Hpa II lanes, in a similar manner to the results in Figure 5.6.

FIGURE 5.5: TIME COURSE METHYLATION STUDY OF MOUSE
con SEQUENCES IN F1.1

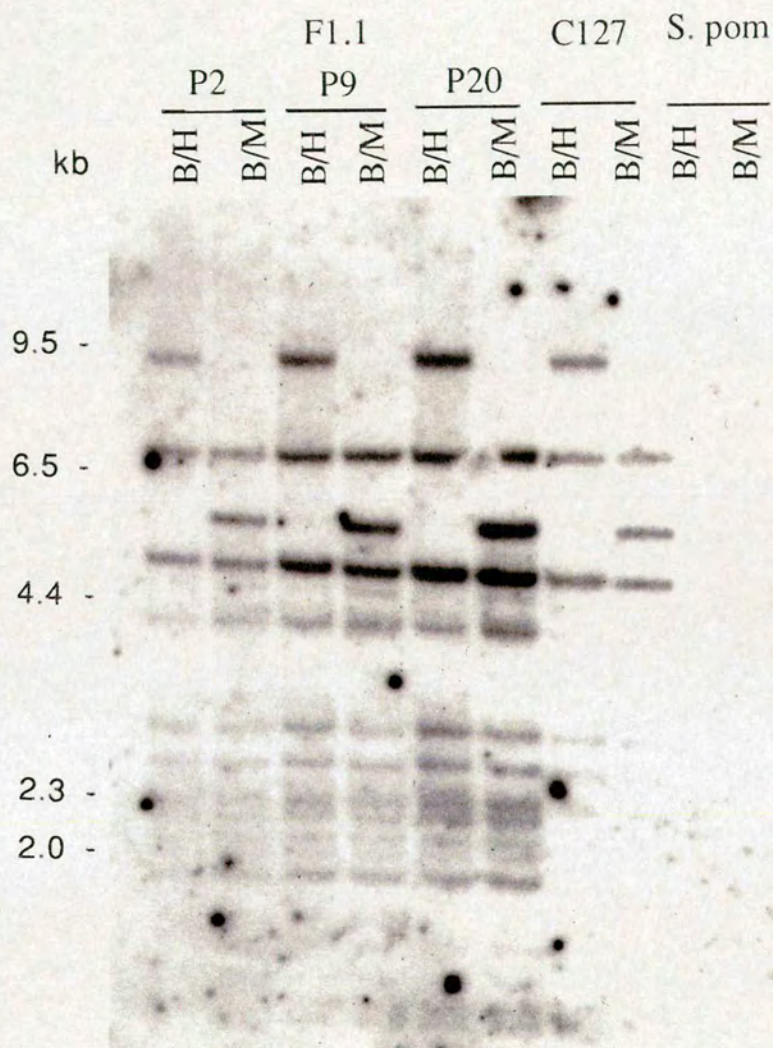


Fig 5.5 DNAs were made from appropriate passages (P) of F1.1 cells. 10 μ g of F1.1 and C127 cell line DNAs and 50ng of *S. pombe* (*S. pom*) DNA were cleaved with Bam HI (B) followed by either Hpa II (H) or Msp I (M) as indicated. After gel electrophoresis DNA was transferred to filters and probed with *con* random primed sequences

FIGURE 5.6 TIME COURSE METHYLATION STUDY OF *S. pombe*
0.7kb rNTS SEQUENCES IN F1.1

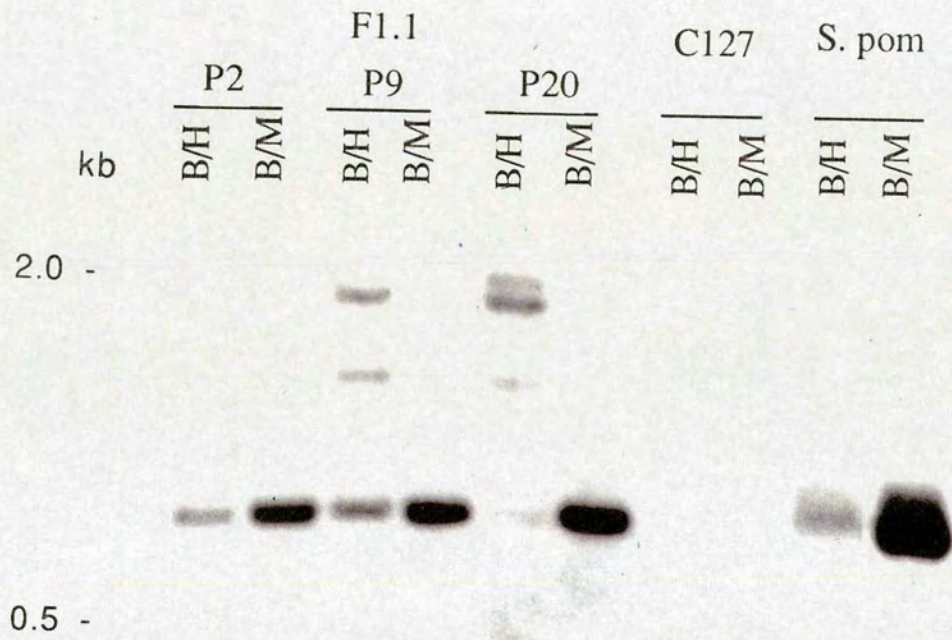


Fig 5.6 DNAs were made from appropriate passages (P) of F1.1 cells. 10 μ g of F1.1 and C127 cell line DNAs and 50ng of *S. pombe* (S.pom) DNA were cleaved with Bam HI (B) followed by Hpa II (H) or Msp I (M) as indicated. After gel electrophoresis DNA was transferred to filters and probed with 0.7kb rNTS random primed DNA

FIGURE 5.7: TIME COURSE METHYLATION STUDY OF *S. pombe*
Cade6-I SEQUENCES IN F1.1

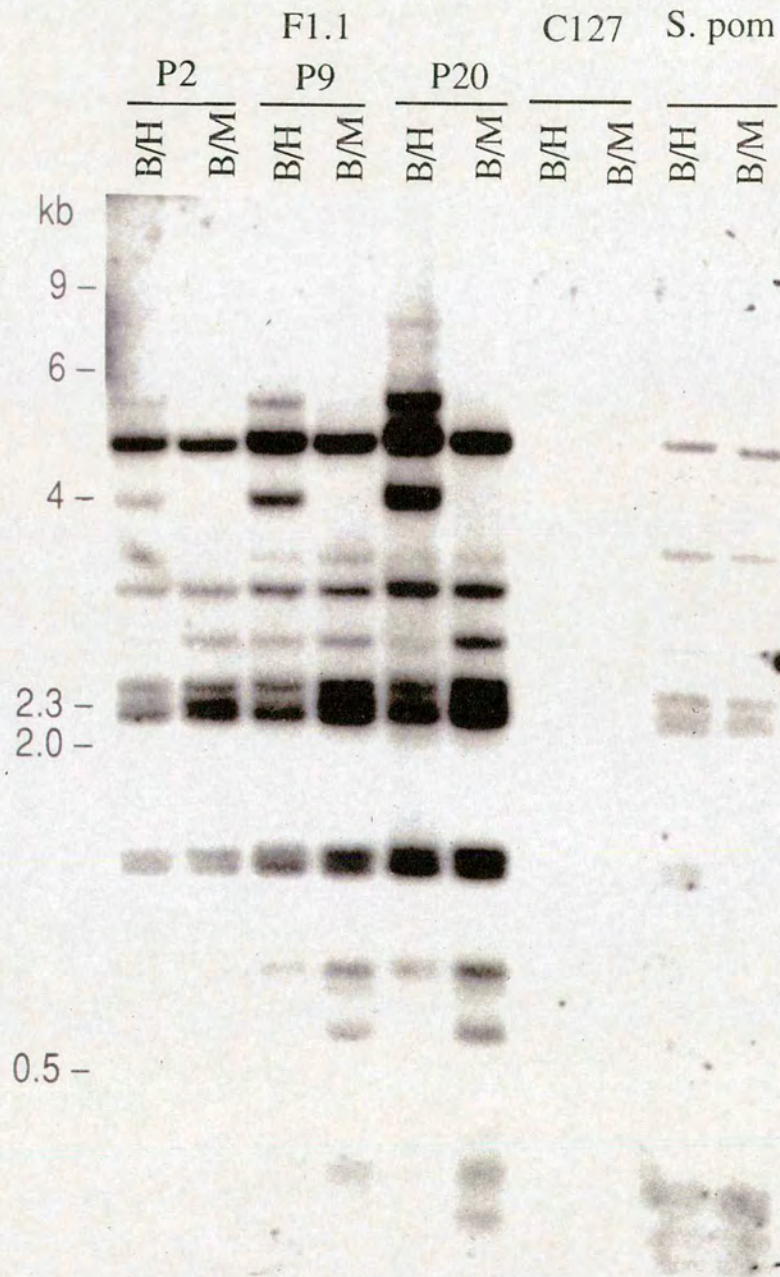


Fig 5.7: DNAs were made from appropriate passages (P) of F1.1 cells. 10 μ g of F1.1 and C127 cell line DNAs and 50ng of *S. pombe* (*S. pom*) DNA were cleaved with Bam HI (B) followed by Hpa II (H) or Msp I (M) as indicated. After gel electrophoresis DNA was transferred to filters and probed with Cade6-I random primed DNA

Figure 5.8 shows the methylation in the Cura4-II region for F1.1 cells at passages P2, P10 and P20. As in Figures 5.6 and 5.7 it can be seen that at passage 2 there is a small amount of methylation, that gradually increases with time, so that at later passages P10 and P20 more of the Hpa II hybridisation can be found in larger bands. No methylation is observed in the corresponding region of the *S. pombe* genome and no cross hybridisation to the C127 genome is observed.

Figure 5.9 demonstrates the amount of methylation accumulated in passages P2, P20, P40 of F1.1 cells within the centromeric repeat sequence dh. This sequence is part of the repeated sequences which surround all three *S. pombe* centromeres. As shown in Chapter 3 the F1.1 cell line may contain, in an unarranged state, the full *S. pombe* centromere I and III regions. This was suggested by the presence of dh hybridising bands characteristic of *S. pombe* centromeres from chromosomes 1 and III in F1.1 (Figure 3.5b) and also from the similarity of the hybridisation pattern of the centromere core sequence of chromosome I and III (tm) in F1.1 and *S. pombe* DNAs. There may therefore be approximately 35 copies of dh in F1.1, occupying approximately 140kb of trasgenome. Figure 5.9 shows that these sequences are unmethylated in the *S. pombe* genome and that they do not hybridise to any discrete bands in the C127 genome. In F1.1 there is a small amount of methylation in cells at P2. Methylation accumulates to such an extent that by passage P20 the majority of the Hpa II hybridising bands are no longer resolved on 0.8 % agarose gels. This may not reflect a larger increase in methylation in this region, compared to the single sequence regions so far studied, but is more of a reflection of the sequence bias of this area which is AT rich, therefore providing very few opportunities for cleavage by Hpa II. By passage P40 the methylation has increased to such an extent that the hybridising fragments do not enter the gel, suggesting that most of the available sites have become methylated. We conclude that there is an extensive and progressive methylation of these centromeric regions in F1.1.

The increasing *de novo* methylation of the *S. pombe* sequences in F1.1 cells suggests that methylation might continue to accumulate until the majority of the available sites are methylated. Alternatively a state might be reached where demethylation balances methylation, at which point F1.1 methylation would reach a steady level. To investigate this F1.1 cells were cultured extensively so that populations of cells up to passage 150 could be sampled. Figure 5.10 shows an analysis of the methylation within the Cura4-II region in these cells. Methylation, already detected within the region at P16, steadily increases through P40 and P150. By this late passage stage most of the hybridising bands have moved to the top of the gel and only a few bands corresponding to Msp I bands are

FIGURE 5.8: TIME COURSE METHYLATION STUDY OF *S. pombe*
Cura4-II SEQUENCES IN F1.1

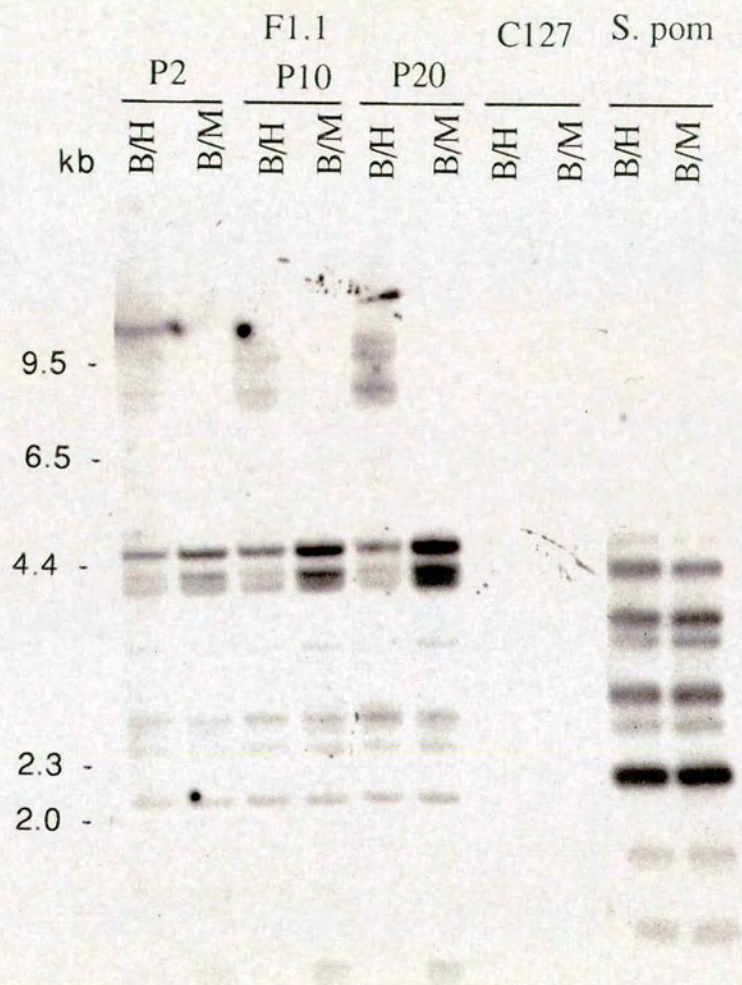


Fig 5.8 : DNAs were made from appropriate passages (P) of F1.1 cells. 10µg of F1.1 and C127 cell line DNA and 50ng of *S. pombe* DNA were cleaved with Bam HI (B) followed by Hpa II (H) or Msp I (M) as indicated. After gel electrophoresis DNA was transferred to filters and probed with Cura4-II random primed DNA.

FIGURE 5.9: TIME COURSE METHYLATION STUDY OF *S. pombe* dh SEQUENCES IN F1.1

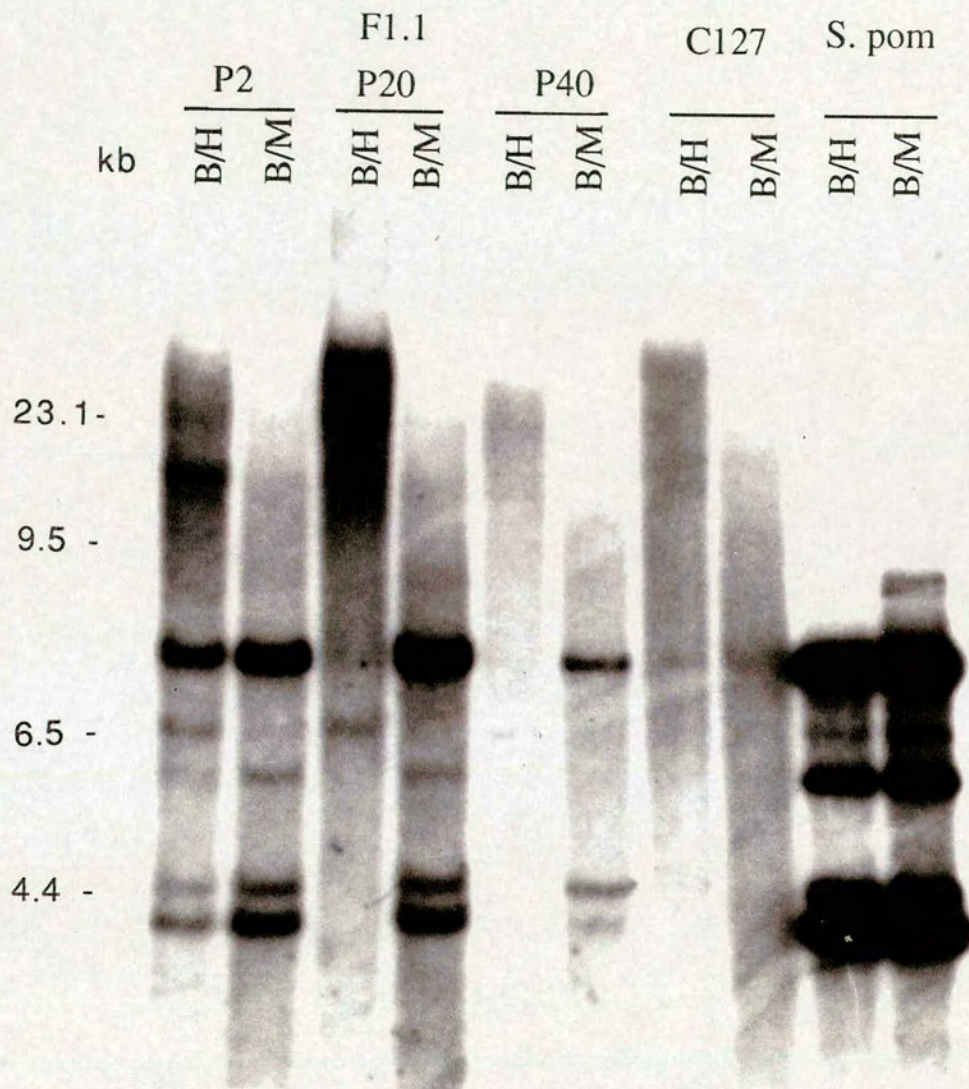


Fig 5.9 DNAs were made from appropriate passages (P) of F1.1 cells. 10 μ g of F1.1 and C127 cell line DNA and 50ng of *S. pombe* (*S. pom*) DNA were cleaved with Bam HI (B) followed by Hpa II (H) or Msp I (M) as indicated. After gel electrophoresis DNA was transferred to filters and probed with dh random primed DNA

FIGURE 5.10: EXTENDED TIME COURSE METHYLATION STUDY
OF *S. pombe* Cura4-II SEQUENCES IN F1.1

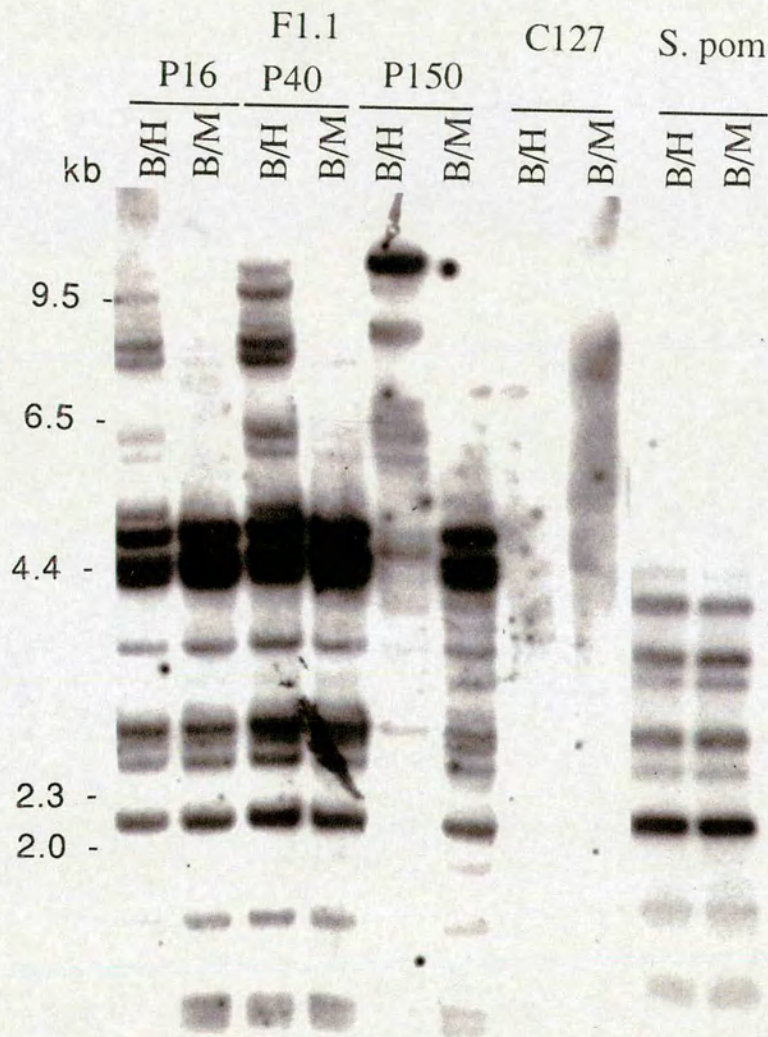


Fig 5.10 DNAs were made from appropriate passages (P) of F1.1 cells. 10 μ g of cell line DNA and 50ng of *S. pombe* DNA were cleaved with Bam HI (B) followed by Hpa II (H) or Msp I (M) as indicated. After gel electrophoresis DNA was transferred to filters and probed with Cura4-II random primed DNA

detectable. This gross shift in the level of methylation within the F1.1 insert leads us to assume that all available sites are becoming methylated and that this may proceed to saturation if these cells could be followed for long enough. It is difficult to compare the levels of methylation within this *S. pombe* insert to those of the surrounding mouse genome. The levels of methylation within the mouse sequences are presumably constrained by selection for functioning house keeping genes necessary for growth in tissue culture. It is possible that since the *S. pombe* insert is purely a piece of passenger DNA, that no controls on methylation levels of it within F1.1 cells will occur.

5.5 Cytology of the *S.pombe* insert during increasing methylation.

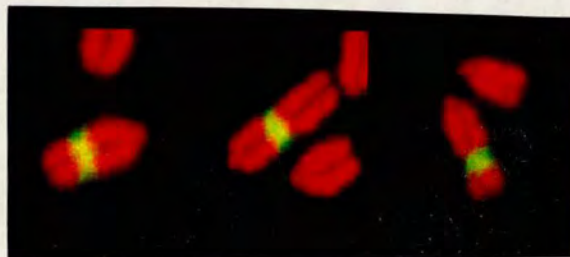
The initial studies of the F1.1 cell line by *in situ* hybridisation of chromosomes revealed that the chromosome containing the *S. pombe* sequences was thinner, or waisted at the site of *S. pombe* integration (Fig1.1) These cytological studies used early passage cells. As I have shown here, these cells contained no or only very slight levels of methylation within the yeast sequences. It has been proposed (Meehan *et al.*, 1989) that methylation may play a role in the higher order packaging of mammalian chromosomes. Therefore a lack of methylation in yeast DNA in these early passage cells may be a contributory factor to the striking morphology seen in F1.1. The *S. pombe* insert in F1.1 is very large (possibly >10Mb) and therefore the increasing methylation levels I have observed during prolonged culture may affect the appearance of the chromosome through increased binding of proteins such as MECP 1 and 2.

To answer this question F1.1 cells at passage stages described in Figure 5.5 to 5.10 were analysed by *in situ* hybridisation with biotinylated total *S. pombe* DNA. This hybridisation was done mainly by Dr Perry and co workers. The results in Figure 5.11 demonstrate that from passage P2 (Figure 5.11 a) to P40 (Fig 5.11 b) there has been no apparent change in the appearance of the *S. pombe* waisted region in F1.1, even though a substantial increase in the methylation levels the *S. pombe* region has occurred in these cells. Therefore although methylation has been proposed to be a part of chromosome packaging mechanisms (Meehan *et al.*, 1989) we have been unable to observe any correlation between metaphase chromosome structure and DNA methylation in this instance.

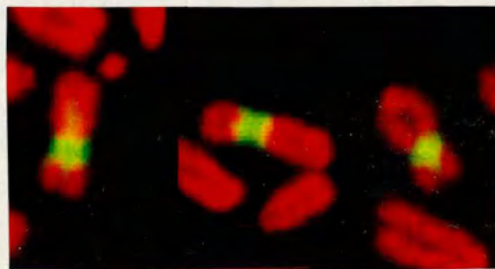
5.6 Conclusions

I have shown that the *S. pombe* insert in F1.1 accumulates methylation to high levels during culture and is thus a good substrate for a *de novo* methylase activity present in the

FIGURE 5.11: CYTOLOGICAL ANALYSIS OF *S. pombe* INSERT OF F1.1 OVER TIME IN CULTURE.



A: P2



B: P40

Fig 5.11: F1.1 cells were taken from time course passages P2 and P40 and metaphase chromosomes prepared from them. These chromosomes were stained with propidium iodide to detect DNA and hybridised in-situ with biotinylated total *S. pombe* DNA. Biotin was detected by avadin-FITC. Images were captured by confocal microscopy (Dr P Perry).

C127 host cell line. Over the same time scale we do not observe any concurrent increase in methylation of endogenous mouse DNA. The high levels of methylation of the F1.1 yeast transgenome observed, may be related to a more open chromatin structure within the *S. pombe* sequence domain than that present within the bulk mouse genome. High levels of methylation within this domain may also result from *S. pombe* DNA not being depleted of the dinucleotide CpG, the substrate for methyltransferase activity. The mouse genome is known to be heavily depleted of these sequences (Bird, 1987) and also selection for housekeeping function may further prevent methylation of mouse DNA. It is interesting to note that increasing methylation around the *ura4* locus does not compromise expression of SV2neo since cells are continually grown on selection.

It is impossible to estimate the rate of methylation occurring within the *S. pombe* sequence from the analyses which have been undertaken. This could be investigated by quantitative PCR analysis of particular sites.

We have been unable to associate the increased methylation of *S. pombe* DNA in F1.1 with any change in metaphase chromosome appearance. It would be interesting to carry out immunocytochemistry studies of this region using the antibodies obtained by Lewis *et al* (1992) to MECP II. Since MECP II is a protein which has been shown to bind in a non sequence specific manner to methylated DNA then we would expect this protein to bind the transgenome if methylated. I would therefore expect that over the time in culture increasing levels of MECP II would be detected at the site of *S. pombe* integration. The MECP proteins have been proposed to be part of the chromatin packaging processes of the cell. It therefore comes as a surprise that while this insert changes from a region of non- to high-MECP potential binding, it is not accompanied by any detectable change in chromosome structure. Although methylation appears to play no part in affecting the structure of this novel chromosomal region, this finding does not preclude methylation from being involved in chromatin packaging. It is possible that methylation plays a more important role in packaging of the interphase chromosome than it does at metaphase. Independence of these functions would allow methylation to be involved with the protection of DNA observed in methylated regions from assault by nucleases (Antequera *et al.*, 1989; Hsieh *et al.*, 1992) while different mechanisms may package all sequences, regardless of methylation for metaphase.

Since this waisted appearance of the *S. pombe* insert in F1.1 is not affected by methylation levels then the structure of this insert must be determined by other processes of chromosome packaging independent of methylation. We therefore need to look at levels

above that of epigenetic modification of the yeast DNA, for an explanation of the morphology we observe in the F1.1 chromosome.

CHAPTER 6. NUCLEOSOME STRUCTURE OF *S. pombe* SEQUENCES IN HYBRID CELL LINES

6.1 Introduction.

Eukaryotes manage to fit their large and complex genomes into the nucleus through a series of chromatin packaging stages which result in a 10^4 packaging ratio in the metaphase chromosome. The first of these is the formation of nucleosomes. Short DNA regions, 146bp, are wound (1.75 loops) around an octamer of proteins called histones (H2a, H2b, H3 and H4). The nucleosome core particle may then be stabilised by the presence of histone H1. Mouse cells contain histone H1 or related subtypes, while in *S. pombe* or *S. cerevisiae* no H1-like protein has been demonstrated. Nucleosomes can be visualised by electron microscopy as beads on a string (Olins and Olins, 1974) and have been shown to be spaced regularly along the DNA. The distance from the start of one nucleosome to the start of the next is found to be cell specific. In mouse tissues this distance varies from 170bp to 220 bp (liver cells) while all mouse cell lines have a repeat of approximately 185bp (Van Holde, 1989). It is not known what accounts for this difference between different tissues or why cell lines have roughly identical repeats irrespective of source. By contrast *S. pombe* chromatin has been shown to have a general nucleosome repeat of approximately 155bp (Yanagida, 1990). This 155bp repeat is found throughout the *S. pombe* chromatin with the exception of the central core centromeric sequences (Polizzi and Clarke, 1991), where a non nucleosomal structure is present.

Nucleosomes although occurring regularly along chromatin, are generally not positioned. Prunell and Kornberg (1978) demonstrated with reassociation studies that the bulk of nucleosomes do not occur at determined sites in the genome. Positioned nucleosomes appear to occur infrequently in the genome, but are important for correct functioning where they occur (Bloom and Carbon, 1982; Almer and Horz, 1986; Benezra et al., 1986).

Chromatin structure has to be replicated at S phase. Although some argument still remains it appears that histones dissociate from the DNA as a replication fork passes through a region (Sogo, 1986; Cuisick, 1984; Jackson, 1990). New nucleosomes are formed on the replicated DNA as H3/H4 tetramers reassociate with the DNA followed by the addition of H2A/H2B dimers to form the nucleosome core particle. This process takes some 10- 20 minutes and may account for the maturation of chromatin after replication (Cuisick et al., 1983; Klempenaar et al., 1980). Many proteins and interactions may be

involved in determining the chromatin structure of the cell including exclusion of nucleosomes from regions such as CpG islands.

It appears that the nucleosome repeat length of chromatin has a direct affect on the size of the '30nm' fibre, formed after packaging. E. M. evidence suggests a statistical correlation exists between the diameter of the '30nm' fibres from different species and their inherent nucleosome repeat length (Alegre and Subirana, 1989; Athey *et al.*, 1990). These findings impinge on models of chromatin packaging. The correlation of diameter and nucleosome repeat length requires that nucleosomes are packaged without direct conservation of relative nucleosome positions in the relevant fibre structure, as is predicted in certain models. The only model presently suggested that allows flexibility in relative nucleosome positioning after packaging is the cross centre model of nucleosome packaging proposed by Williams *et al.* , (1986) (Figure 1.4).

The F1.1 and F7.2 hybrid cell lines provide a good model system for the analysis of the behaviour of large foreign DNA regions in mammalian cells. The nucleosome structure of the yeast transgenome sequences is important to analyse since it will provide insight into how foreign sequences interact with chromatin proteins.

Due to our lack of understanding of chromatin structure, we cannot predict how the *S. pombe* region will be packaged by the endogenous mouse chromatin proteins of the F1.1 cell line. It may be that these regions will merely be packaged into nucleosomes as if a normal piece of the mouse genome. If this occurs, nucleosome structure can not be the cause of the constriction seen in the *S. pombe* insert region of F1.1. The unusual appearance of the *S. pombe* insert could however result if these regions were to maintain an *S. pombe* nucleosomal repeat i.e. 155bp instead of 185 bp, which would in turn lead to a difference in the diameter of the '30nm' chromatin fibre. This might arise if a conservative pattern of nucleosome segregation occurred after replication, such that the initial conservation of *S. pombe* nucleosomal structure could be propagated through cell division. The resulting thinner or disrupted basic chromatin fibre might result in a thinner chromosomal region at metaphase after being propagated through subsequent levels of chromatin packaging. I have therefore investigated nucleosome structure in the *S. pombe* regions of hybrid cell lines, by nuclease digestion of their chromatin.

6.2 Nuclease digestions.

To study the nucleosomal conformation of the *S. pombe* regions in the hybrid cell lines it is necessary to cleave the chromatin in such way as to release the nucleosomal units. This

is achieved by preparing nuclei and digesting the chromatin with micrococcal nuclease. Micrococcal nuclease cleaves the spacer regions linking the nucleosomes. Controlling the amount of micrococcal nuclease digestion so that a partial reaction is obtained provides nucleosomal multimers of all potential lengths (Fig 6.1). The DNA from these digests is then purified and separated by size with gel electrophoresis. I have used micrococcal nuclease since any sequence preference of the enzyme should not bias the results when studying the general nucleosomal structure of the *S. pombe* sequences in the transgenome. After electrophoresis the nucleosome repeat of bulk chromatin can be calculated directly by measuring the difference in size between each subsequent multimer on ethidium bromide stained gels. Repeat lengths of specific regions within the genome can be calculated in the same way after hybridisation of the DNA with the desired probe.

As discussed above, the genomes of *S. pombe* and mouse cell lines have been shown to contain different lengths of nucleosomal repeat in their bulk chromatin. The nucleosome repeat length is defined as the distance, in base pairs, between the end point of a nucleosome, to the end point of the next nucleosome. *S. pombe* has a repeat of approximately 155bp (Chikashige *et al.*, 1989; Polizzi and Clarke, 1991) while mouse cells lines have repeats between 180 and 190bp (Van Holde, 1989). Although there is only 30 bp difference between the *S. pombe* repeat and mouse repeat, they can easily be distinguished since this initial difference is exaggerated in each subsequent nucleosomal multimer to produce distinctive nucleosomal ladders. Figure 6.2 demonstrates this. It can be seen that although the *S. pombe* core particles, C, are identical in size to those of the mouse, by nucleosome multimer C + 3N the difference is quite pronounced. In mouse cell lines the C + 3N nucleosome multimer migrates slower (~710 bp) than the 603 bp marker band from ϕ X174 Hae III whereas the *S. pombe* C + 3N nucleosomes are very close to this size (620 bp). The C + 4N nucleosomes from mouse cell lines migrate at 900bp, those from *S. pombe* migrate at 770bp and therefore these nucleosome multimers run higher and lower respectively than the 872 bp band from ϕ X174 Hae III marker. Therefore under the correct electrophoresis conditions the mouse and *S. pombe* nucleosome ladders can easily be distinguished.

6.3 Analysis of *S. pombe* nucleosome ladders.

S. pombe cells were grown to optimal densities of 10^7 cells per ml and then subjected to several different procedures for purifying nuclei, as detailed in Section 2.16. Method 1 was an attempt to produce *S. pombe* nucleosomal ladders under approximately identical conditions to those used for mammalian cells. This procedure lead to general degradation of the *S. pombe* chromatin producing a smear on EtBr stained gels, and no evidence of a

FIGURE 6.1. SCHEMATIC DIAGRAM OF MICROCOCCAL NUCLEASE DIGESTS

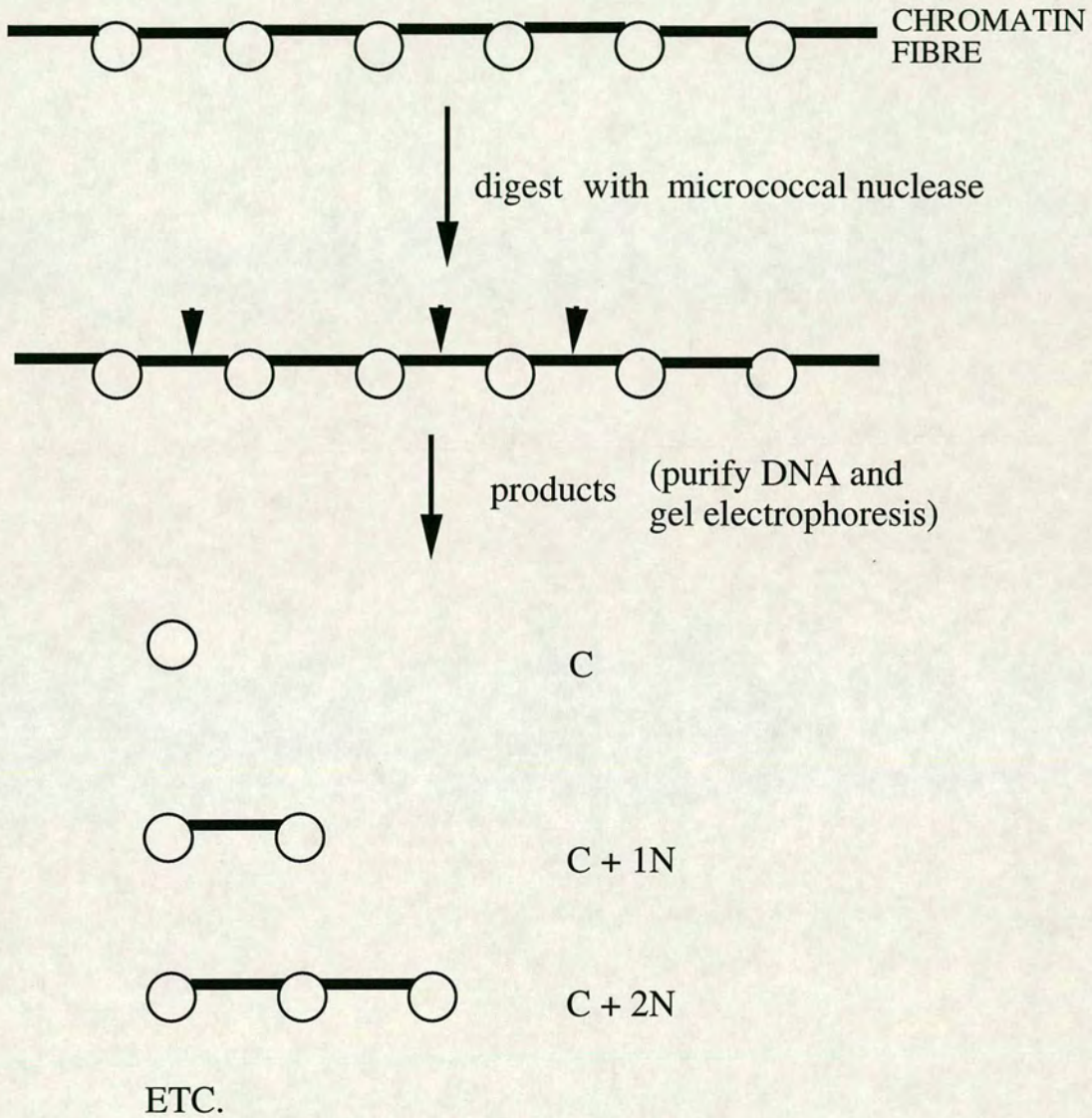


Fig 6.1



NUCLEOSOMES



LINKER DNA



SITE OF MICROCOCCAL NUCLEASE CLEAVAGE

C = NUCLEOSOME CORE PARTICLE

N = NUCLEOSOME REPEAT LENGTH

FIGURE 6.2 SCHEMATIC COMPARISON OF *S. pombe*
AND MOUSE NUCLEOSOME LADDERS

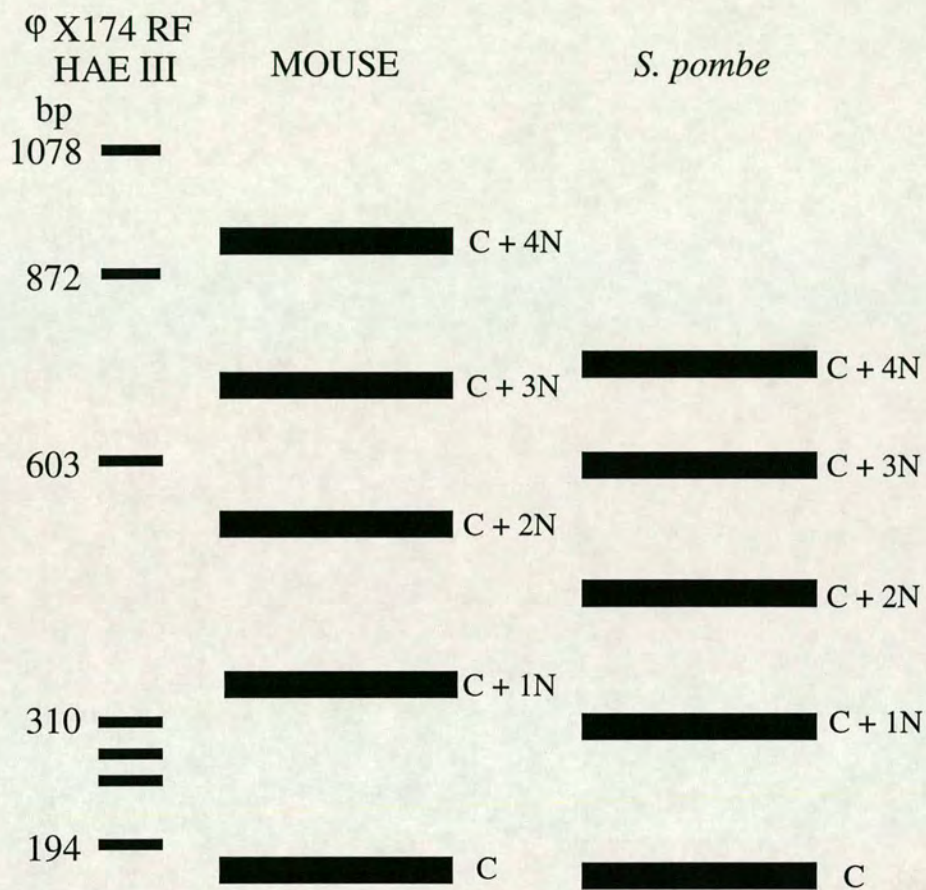


Fig 6.2 : The black bars represent the positions of nucleosomes after gel electrophoresis. C = the core particle which is ~146 bp. N = an average nucleosome repeat length after digestion.

ladder (data not shown). This could either be due to presence of proteases or endogenous nucleases not removed or inhibited during preparation, destroying nucleosome structure. Method 2 was an attempt to preserve chromatin structure by using a PIPES and sorbitol based buffer, a more osmotically stable environment for the yeast spheroplasts commonly used in yeast chromatin manipulations. In my hands this procedure also lead to excessive degradation of the chromatin. The problem was therefore believed to be partly due to coprecipitation of degrading enzymes with the *S. pombe* nuclei. In method 3 the addition of EDTA to the buffer to chelate Mg^{2+} and Ca^{2+} ions present in buffers used in methods 1 and 2, should alleviate the problem, since the EDTA should sequester the free divalent cations and prevent enzyme activation during nuclei isolation. Method 4, modified from the method of Polizzi and Clarke, (1992), appears to partially circumvent this problem by cleaning the nuclei through a glycerol cushion. This produced nucleosome ladders although a certain amount of degradation was still present in controls. Method 5 introducing a buffer from Chikashige *et al.*, (1989) to the digestion stage of method 4, showed little improvement. I therefore decided to use the protocol of Polizzi and Clarke (92), method 4, for preparations of yeast nuclei since this appeared to give the best and most consistent results in my hands.

Figure 6.3 shows the average nucleosome spacing in the genome of *S. pombe Int 5* by digestion with micrococcal nuclease and Ethidium bromide staining of the digestion products after electrophoresis. A time course reaction was used so that at least one of the experimental tracks would provide an analysable distribution of nucleosomal repeats. It is apparent that the DNA in the 0 minutes control lane is virtually undegraded while the 9 minute time point control lane with no enzyme added shows a small degree of degradation, presumably due to the activity of endogenous nucleases and proteases active in the digestion buffer. This degradation of the chromatin explains why the ladders in the experimental lanes are smeared. A small amount of smearing is expected since the micrococcal nuclease will not cleave at the same site in each linker region resulting in differences in the lengths of overhanging linker regions of the digested nucleosomal multimers. It can be seen that the core particle, C, is approximately 146bp in size, the expected size for the core particles of any species. Degradation and digestion products can be seen below these core particles. C+1N is also of the expected size for the *S. pombe* nucleosomal repeat (~300 bp). The subsequent larger nucleosomal multimers are also of the expected sizes, (C +2N, 460bp; C + 3N, 620 bp). A loss of DNA appears to occur over the time course, (3 minute lane compared to 9 minute lane). This is presumably due to the activity of nucleases degrading the chromatin.

FIGURE 6.3 MICROCOCCAL NUCLEASE DIGESTION OF *S. pombe* NUCLEI.

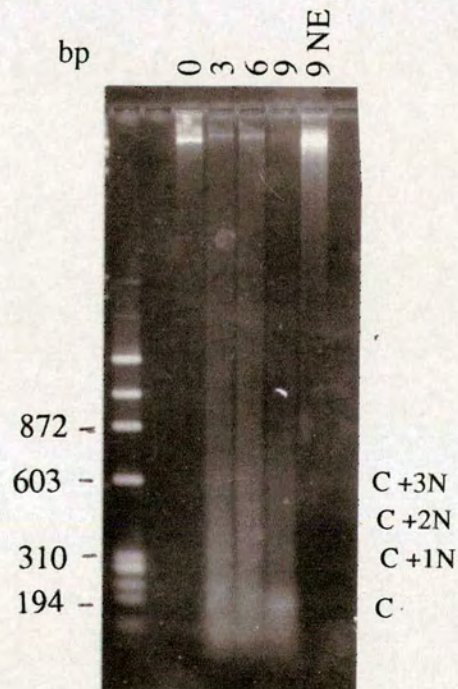


Fig 6.3 : *S. pombe* nuclei were prepared and digested with micrococcal nuclease for varying lengths of time using method 4 as described in Section 2.17. Digests were size fractionated by agarose gel electrophoresis through 2% agarose. The numbers above the lanes correspond to digestion time in minutes. NE- no enzyme added. Size marker is ϕ X174 Hae III. C = core particle. C +N = core particle + N nucleosome repeat units.

It was important to check that specific regions being studied also conform to this canonical *S. pombe* nucleosomal repeat in *Int-5* cells since nucleosome repeats can vary within a cell at specific regions. The dh centromeric repeat was used to probe *S. pombe* micrococcal nuclease digests electrophoresed through 2% agarose and transferred to nitrocellulose with a small pore size (0.2 μ m) to minimise loss of small DNA fragments (Figure 6.4). The no enzyme control is observed to be largely undegraded, demonstrating that the DNA was intact prior to digestion. Unfortunately the control lane, with no added enzyme, exhibits a large degree of degradation, but this does not appear to seriously affect the nucleosome ladder produced. It can be seen that, as expected, dh detects a core particle of approximately 150 bp. The second and third nucleosomes are also of the expected sizes with the fourth nucleosome (620bp) being approximately in line with the 603 bp size marker. This ladder demonstrates that dh has a nucleosome repeat of 155bp, identical to that of bulk chromatin in *S. pombe*. This finding is consistent with the results of Chikashige *et al.*, (1989) who demonstrated a 155 bp nucleosomal repeat in the dg repeat, the other half of the centromeric repeating sequence.

Figure 6.5 demonstrates that several other regions of the *S. pombe* genome adopt the nucleosomal repeat found in bulk chromatin. Figure 6.5a shows the SV₂neo sequences from chromosome III hybridised to micrococcal nuclease ladders from *S. pombe Int 5*. The core particle is again ~150 bp and the subsequent multimers show the appropriate sizes with C + 3N being of 620 bp in size. This is the pattern expected for *S. pombe* genomic sequences, however it is interesting to remember that these sequences are not endogenous to the *S. pombe* genome. The SV₂neo sequences were transfected into the *S.pombe ED628* line to produce the *S. pombe Int-5* line (Allshire *et al.*, 1987). The SV₂neo construct contains bacterial sequences, a neomycin resistance gene, and the late early region promoter of the SV40 virus. These foreign sequences appear to adopt a normal *S. pombe* nucleosomal configuration in the yeast. This demonstrates that these viral and bacterial sequences are recognised and packaged as normal pieces of DNA within the *S. pombe* genome.

Figures 6.5 b and 6.5 c show that the cosmids Cade6-I and Cura 4-II hybridised to nucleosome ladders from *S. pombe*. Both of these large regions of contiguous yeast sequence appear to be packaged in the same nucleosomal pattern as bulk chromatin.

It therefore appears that the *S. pombe* sequences we are using in the analysis of F1.1 are all packaged in the 155bp nucleosome repeat characteristic of *S. pombe* chromatin in (SV₂neo) the strain *Int 5* used in the construction of F1.1. It is of interest that the foreign sequences in this *S. pombe Int 5* line adopt the *S. pombe* repeat. This indicates that

FIGURE 6.4 MICROCOCCAL NUCLEASE DIGESTION OF *S. pombe* NUCLEI PROBED WITH dh SEQUENCE.

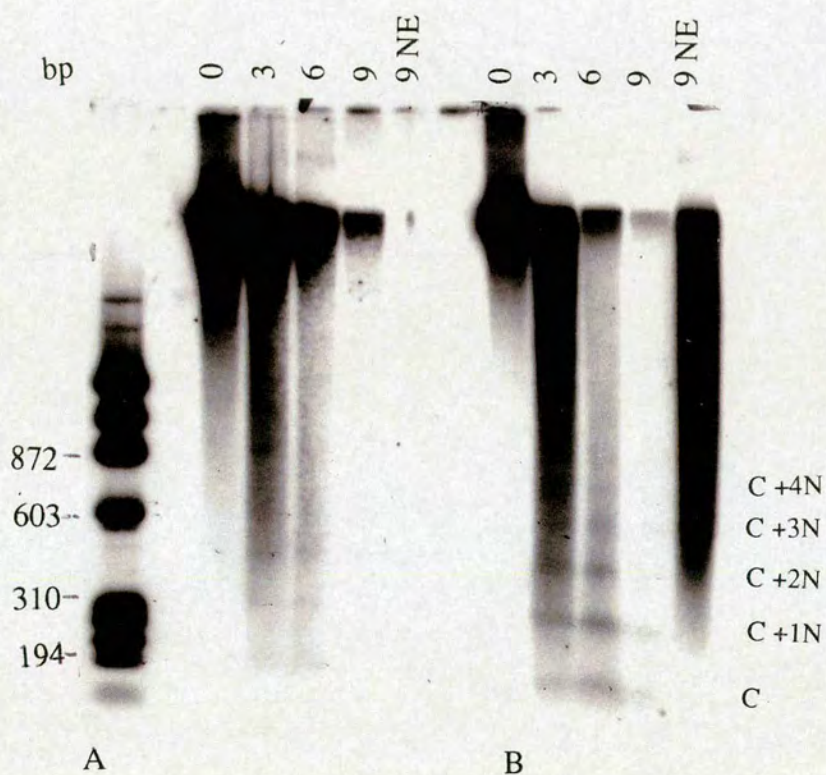


Fig 6.4 : *S. pombe* nuclei were prepared and digested as described in Figure 6.3. Digests were size fractionated by gel electrophoresis, and transferred to 0.2um pore nitrocellulose. The filter was probed with random primed dh sequence. Numbers above the lanes correspond to digestion time in minutes. NE- no enzyme added. Size marker is ϕ X174 Hae III. C = core particle. C +N = core particle + N nucleosome repeat units. Panel A shows a different set of digests to those in B.

FIGURE 6.5. MICROCOCCAL NUCLEASE DIGESTION OF *S. pombe* NUCLEI PROBED WITH SV₂neo, Cade6-I and Cura4-II

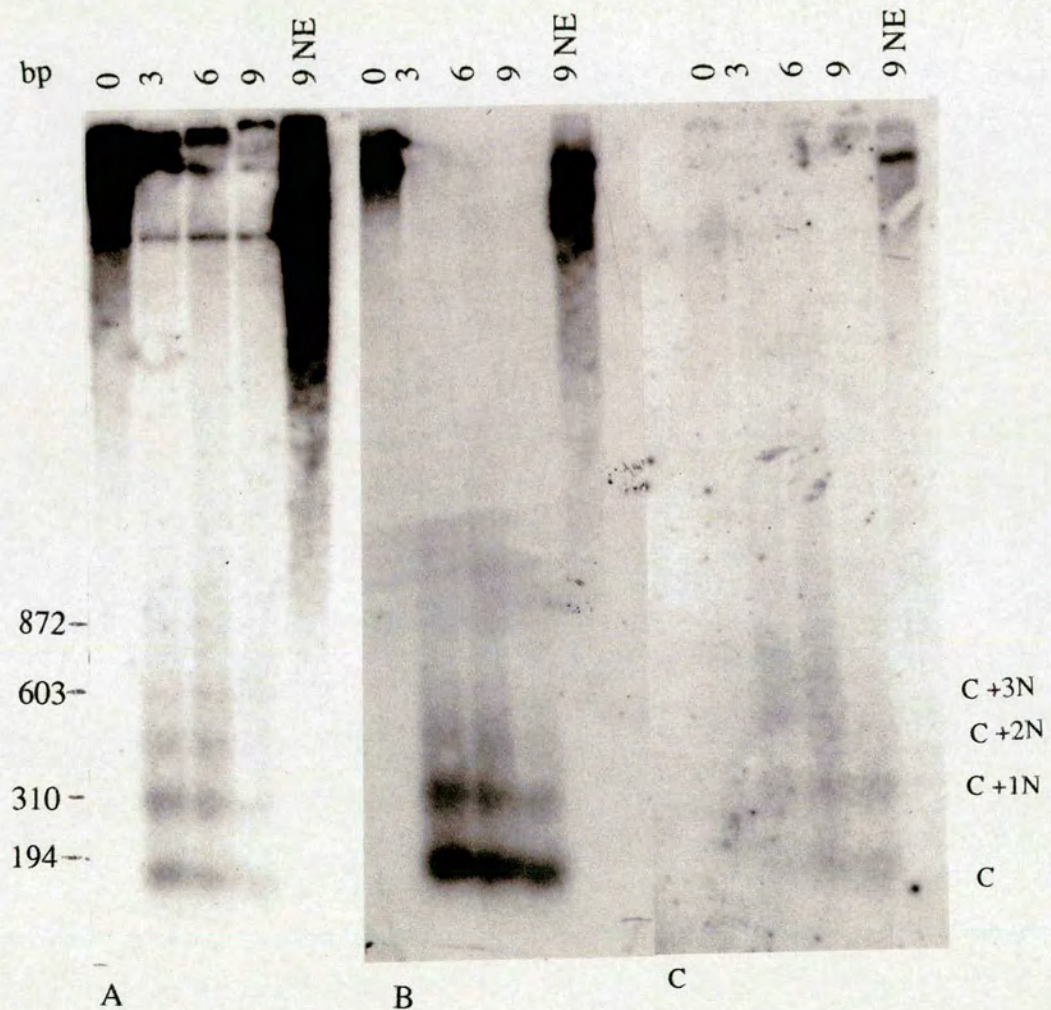


Fig 6.5 : *S. pombe* nuclei were prepared and digested as in Figure 6.3. Digests were size fractionated by gel electrophoresis, and transferred to a nitrocellulose membrane. DNA was probed with random primed A- SV₂neo, B- Cade6-I and C- Cura4-II sequences. Numbers above the lanes correspond to digestion time in minutes. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

nucleosome packaging is not generally sequence dependent and may suggest a similar fate for *S. pombe* sequences in mouse cells. *Int 5* however was created from *S. pombe* strain *ED628* by DNA transfection while the transgenome of F1.1 was produced by fusion, which introduces chromatin, with resident nucleosomes rather than naked DNA.

6.4 Nucleosome structure of *S. pombe* sequences in hybrid cell lines.

Cells from the appropriate mammalian cell lines were harvested and nuclei preparations and micrococcal nuclease digests carried out as described in Section 2.16. Briefly cells were lysed in a 0.5% Triton X-100 (Sigma) solution. The salt concentration was 10mM to preserve nuclear structure, above 0.8M NaCl histones dissociate from DNA (Ohlenbusch *et al.*, 1967). Nuclei were then washed to remove detergent and digested in appropriate buffer. DNA was then purified from the reactions and resolved by agarose gel electrophoresis.

Mouse cell lines have a nucleosomal repeat of approximately 185bp. The corresponding expected sizes of the nucleosome multimers is represented schematically in Fig 6.2. It can be readily seen that the mouse ladder is different from that obtained from *S. pombe* cells. If the *S. pombe* sequences within the hybrid and transfectant cell lines used in these studies adopt the mouse nucleosomal repeat then their nucleosomal ladders should be in perfect register with those of the mouse sequences. The *S. pombe* sequences can only be detected by hybridisation. Hybridisation of specific sequences to nucleosomal reactions produces a much lower signal compared to the signals from conventional restriction endonuclease digestions. As opposed to signal being concentrated into a very tight band of hybridisation, in micrococcal nuclease reactions the sequences to be detected are spread throughout the length of the track, due to the non specific nature of the cleavage of the DNA by micrococcal nuclease. Therefore either long exposures of autoradiographs are necessary to obtain a signal or larger amounts of DNA must be loaded onto the gels to allow the sequences to be detected.

When chromatin from mouse cell lines is digested with micrococcal nuclease, and separated by gel electrophoresis the expected nucleosomal profile is realised. Figure 6.6 shows the EtBr stained gel of nucleosome ladders from the F1.1 cell line. It can be seen that as expected the core particle is ~146bp, as found in all species. Nucleosome C + 1N is approximately 340 bp in size and C + 2N is 530 bp, indicating that the nucleosome repeat in these cells is approximately 190 bp in length. This result demonstrates that the mouse nucleosomal repeat in F1.1 cells is markedly larger than that of *S. pombe* and can be easily differentiated from it.

FIGURE 6.6. MICROCOCCAL NUCLEASE DIGESTION OF F1.1 NUCLEI.

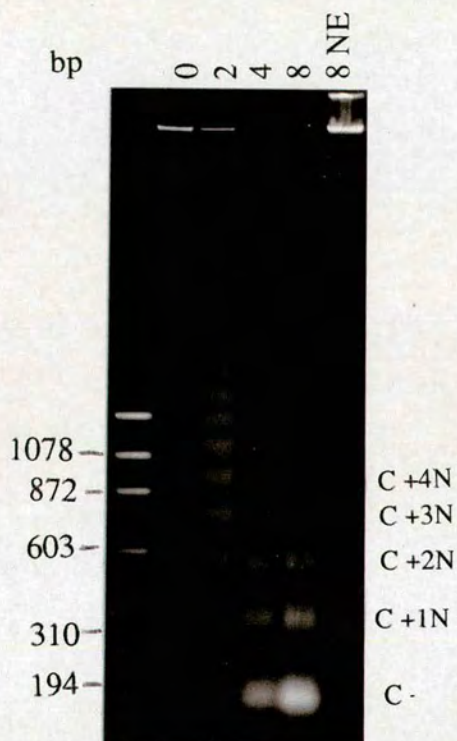


Fig 6.6 : Nuclei were prepared from F1.1 cells and digested as described in Section 2.16. Digests were size fractionated by agarose gel electrophoresis. Numbers above the lanes correspond to digestion time in minutes. Size marker is ϕ X174 Hae III. NE- no enzyme added. C = core particle. C +N= core particle + N nucleosome repeat units.

FIGURE 6.7 MICROCOCCAL NUCLEASE DIGESTION OF F1.1, F7.2 AND C127 NUCLEI PROBED WITH SV₂neo.

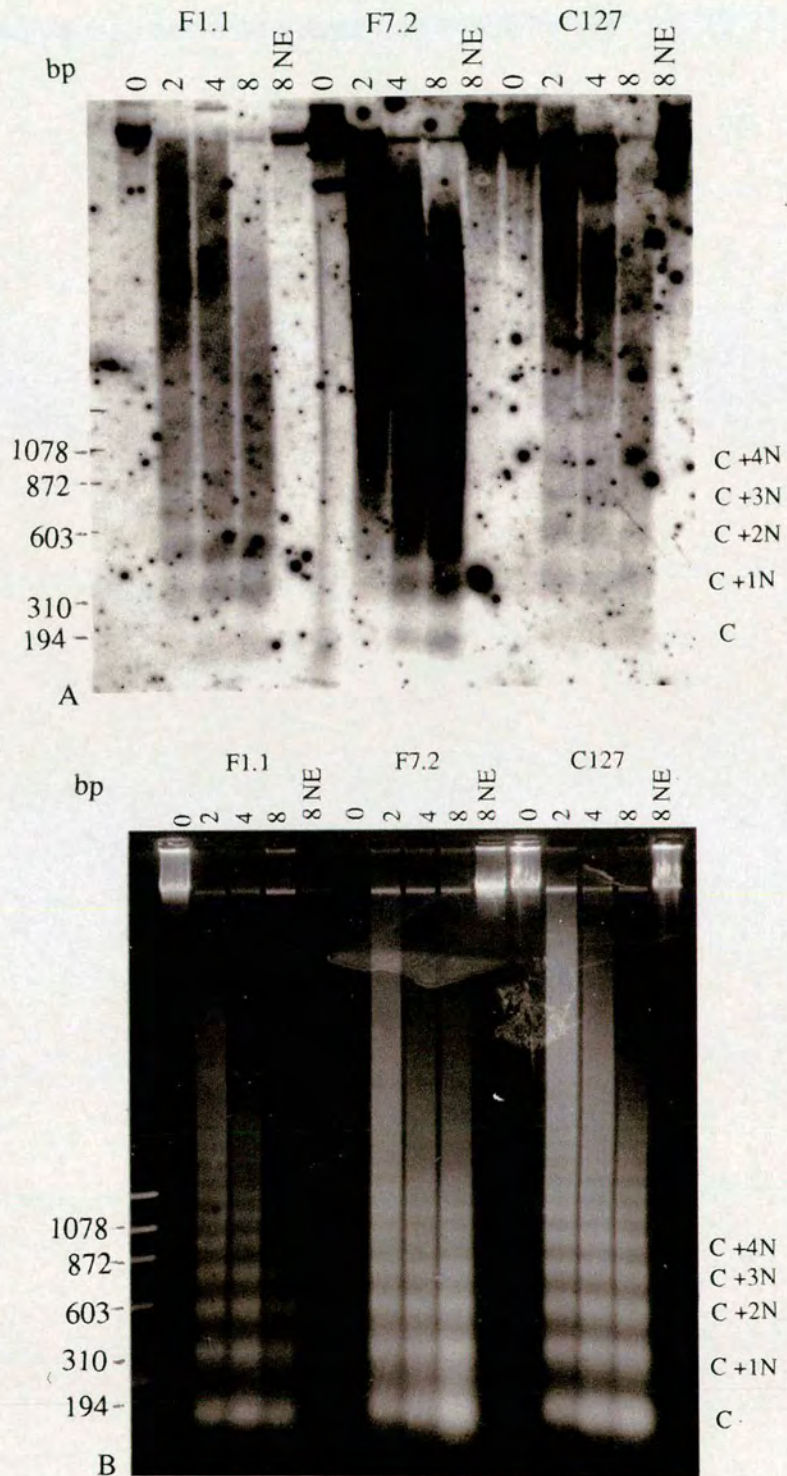


Fig 6.7 : Nuclei were prepared and digested as described in section 2.16. Digests were size fractionated by 2% agarose gel electrophoresis, and transferred to 0.2um nitrocellulose. The filter was hybridised with random primed SV₂neo sequences. The resulting autoradiograph is shown in panel A. B is a photograph of the corresponding EtBr stained gel. Numbers above the lanes correspond to digestion time in minutes. Size marker is ϕ X174 Hae III. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

After establishing that the murine nucleosomal repeat was discernible from that of the *S. pombe* repeat, the important factor was to investigate the nucleosome ladders of the *S. pombe* sequences in the hybrid cell lines.

Figure 6.7 demonstrates the hybridisation pattern of SV₂neo to nucleosome ladders from F1.1, F7.2 and C127 nuclei. It can be seen that there is a high level of cross hybridisation to the C127 genome. This cross hybridisation occurs even after high stringency washing conditions (0.1 x SSC at 68⁰C). This cross hybridisation prevents these sequences from being analysed in the F1.1 and F7.2 cell lines since it is impossible to distinguish the specific and non specific hybridisation signals. Indeed most of the hybridisation appears to derive from the C127 genome since the EtBr profile of the gel (Fig 6.7b) shows that approximately equal amounts of DNA were loaded in each lane. The SV₂neo construct is the only region of the *S. pombe* *Int 5* genome that is known to be expressed in the hybrid cell lines, since these are the sequences which confer resistance to G418. It was therefore disappointing that these sequences were non analysable.

Figure 6.8 shows the hybridisation of the Cura 4-II cosmid to nucleosome ladders from F1.1, F7.2 and C127 cells. In *S. pombe* this region was found to be packaged in the 155bp nucleosome repeat (Figure 6.5). However as seen with the SV₂neo sequences there is non specific cross hybridisation to C127 DNA and this therefore precludes this region from analysis.

A similar situation was also encountered with the Cade6-I cosmid and the *S. pombe* centromeric core sequence tm. The cross hybridisation to the centromeric core tm sequences was particularly unfortunate. In *S. pombe* these sequences are found to be non nucleosomal in structure (Polizzi and Clarke, 1991). Upon transfer to *S. cerevisiae* it is found that these sequences adopt a normal *S. cerevisiae* nucleosomal pattern, suggesting that a specialised chromatin structure is present in *S. pombe* at these sequences. It would have been of great interest to see whether, in these hybrid cell lines, such an unusual chromatin structure was present or not at these loci.

Cross hybridisation to the C127 genome appears to prevent analysis of the nucleosome packaging of large regions of the *S. pombe* sequences in the hybrid cell lines. To try to circumvent this problem the original small gene probes used to select the cosmids from the *S. pombe* library were used (Section 3.2). It was hoped that their smaller size would limit the likelihood of cross hybridisation to C127 DNA, although they provide less information.

FIGURE 6.8. MICROCOCCAL NUCLEASE DIGESTION OF F1.1, F7.2 AND C127 NUCLEI PROBED WITH Cura4-II COSMID.

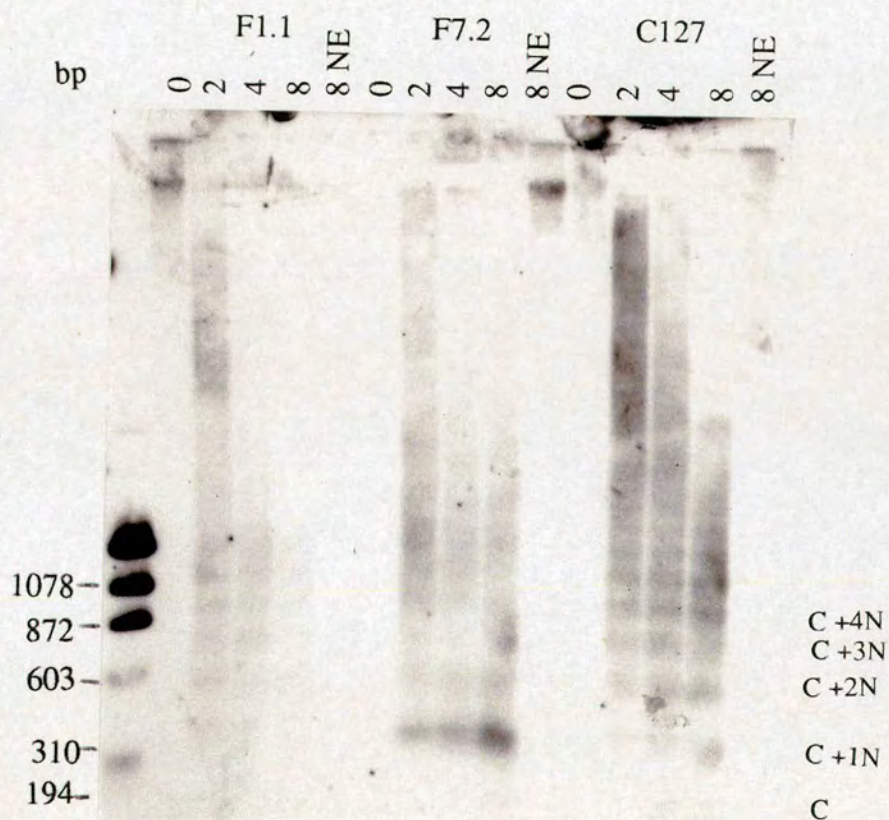


Fig 6.8 : Nuclei were prepared and digested as described in Section 2.16. Digests were size fractionated by 2% agarose gel electrophoresis, and transferred to 0.2um nitrocellulose. The filter was hybridised with random primed Cura4-II cosmid. Numbers above the lanes correspond to digestion time in minutes. Size marker is ϕ X174 Hae III. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

Figure 6.9a depicts the hybridisation of the 2.5 kb Xho I *ade6* fragment to micrococcal nuclease digests of the F1.1 and C127 cell lines (Fig 6.9b). This smaller probe removes the cross hybridisation, to the C127 genome, detected with the larger Cade6-I cosmid. The nucleosome core particle can only be faintly detected due to the small size of these fragments which do not bind well to the 0.2µm pore diameter nitro-cellulose and can easily pass through these holes. Although, due to the large amount of DNA loaded, the nucleosome multimers have migrated more slowly than their true size, it would appear that these *ade6* sequences have adopted the mouse nucleosome repeat. The perfect register of the *ade6* hybridisation bands and the position of the mouse bulk chromatin bands in Fig 6.9b demonstrate this. Retention of an *S. pombe* nucleosome ladder would have led to a loss of register between the hybridisation signal and the position of bulk nucleosomes, as differences between the *S. pombe* and mouse repeat lengths become exaggerated in the large nucleosome multimers.

Figure 6.10a depicts the hybridisation of the *S. pombe ura4* 1.8kb HindIII fragment to nucleosome ladders of F1.1, DC.1 and C127 cell lines. It is noticed that the *ura4* probe shows negligible cross hybridisation to the C127 genome, whereas Cura4-II did (Fig 6.8). Some other section of the Cura4-II cosmid, apart from the *ura4* gene must therefore be responsible for this cross hybridisation. The resulting nucleosomes perfectly match the murine nucleosomal pattern of Fig 6.10b. Individual nucleosomes appear to be slightly larger than expected, C+ 2N ~600bp. This may be due to the amount of sample loaded (20µg) causing drag in migration. The F1.1 and DC.1 cell lines were created by differing procedures, F1.1 was created by introduction of chromatin into C127 cells while DC.1 was produced by introduction of naked DNA. However figure 6.10 shows that both introduction procedures are followed by the adoption of a mouse nucleosome ladder by the introduced DNA. These results augment those of Figure 6.9 which shows that the *ade6* locus also adopts a murine nucleosomal repeat in F1.1 cells.

It is possible that the mouse nucleosome repeat found over *S. pombe* loci in the F1.1 and DC.1 cell lines results from a spreading affect of flanking mouse chromatin into the integrated yeast transgenomes of these cell lines. The *S. pombe* transgenome in the F7.2 cell line however has been shown to be extrachromosomal to the mouse genome and therefore not subject to such spreading affects. It is therefore interesting to investigate the nucleosome packaging of the F7.2 cell line. Figure 6.11 shows the hybridisation of *ura4* to nucleosome ladders of F7.2 and C127 cell lines. The *ura4* sequences adopt the mouse nucleosomal configuration, exemplified by the size of the C+3N and C+4N nucleosome multimers. For the C+3N multimer the mouse repeat size of 720kb is detected rather than

FIGURE 6.9 MICROCOCCAL NUCLEASE DIGESTION OF F1.1 AND C127 NUCLEI PROBED WITH *S. pombe ade6* .

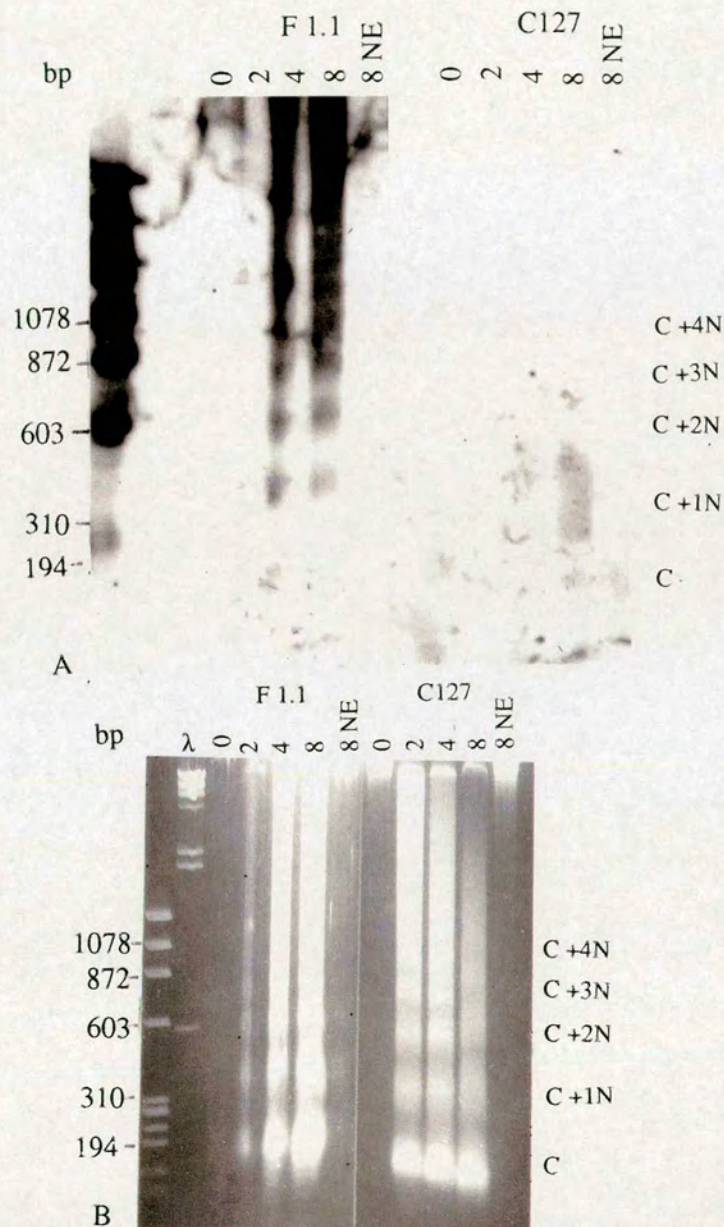


Fig 6.9 : Nuclei were prepared and digested as described in Section 2.17. Digests were size fractionated by gel electrophoresis, and transferred to nitrocellulose. The filter was hybridised with random primed *S. pombe ade6*. The resulting autoradiograph is shown in A. B is a representation of the EtBr stained gel. Numbers above the lanes correspond to digestion time in minutes. Size marker is ϕ X174 Hae III. $\lambda = \lambda$ Hind III marker. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

FIGURE 6.10. MICROCOCCAL NUCLEASE DIGESTION OF F1.1, DC.1 AND C127 NUCLEI PROBED WITH *S. pombe ura4*.

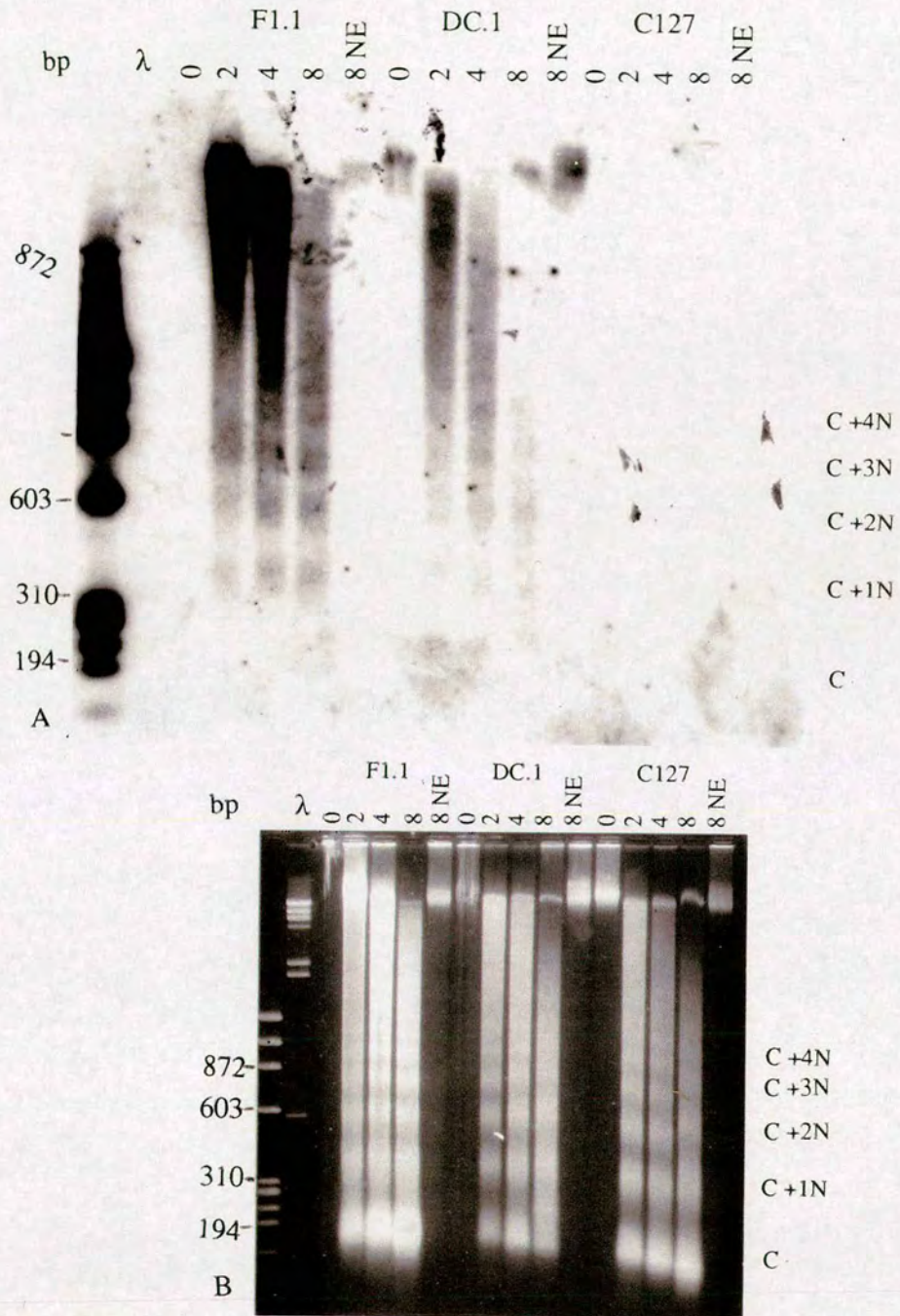


Fig 6.10 : Nuclei were prepared and digested as described in Section 2.17. Digests were size fractionated by 2% agarose gel electrophoresis and transferred to 0.2um nitrocellulose. The filter was hybridised with random primed *S. pombe ura4*. The resulting autoradiograph is shown in A. B is a representation of the EtBr stained gel. Numbers above the lanes correspond to digestion time in minutes. Size marker is ϕ X174 Hae III. λ = λ Hind III marker. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

FIGURE 6.11 MICROCOCCAL NUCLEASE DIGESTION OF F7.2 AND C127 NUCLEI PROBED WITH *S. pombe ura4*.

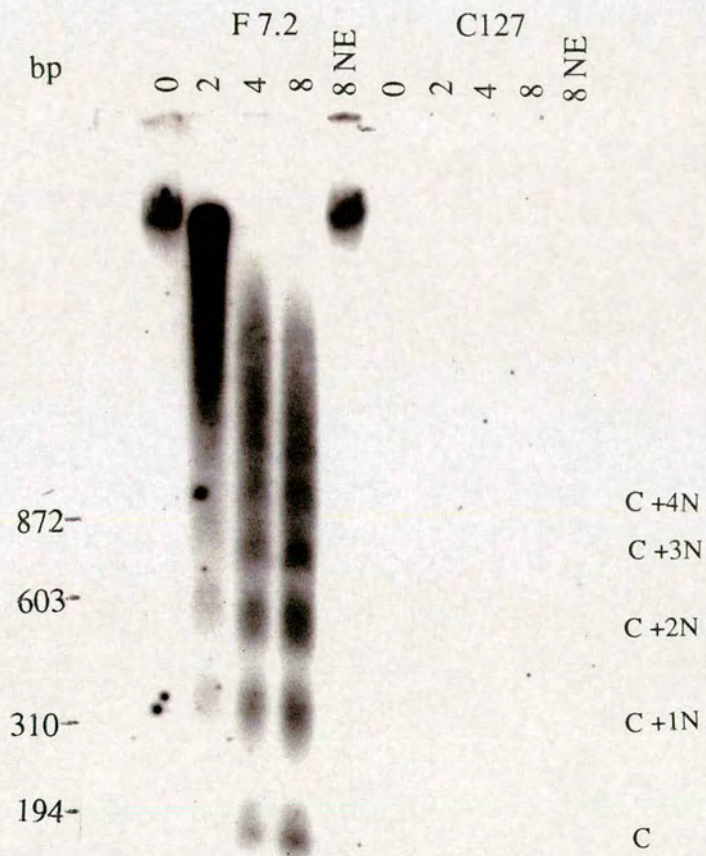


Fig 6.11 : Nuclei were prepared and digested as described in section 2.17. Digests were size fractionated by 2% agarose gel electrophoresis, and transferred to 0.2um nitrocellulose. The filter was hybridised with random primed *S. pombe ura4* . Numbers above the lanes correspond to digestion time in minutes. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

the 610bp expected for an *S. pombe* ladder and with the C+4N multimer the murine size of 910bp is detected as opposed to the yeast size of 760bp. This agrees with results for the *ura4* and *ade6* loci in F1.1 and demonstrates that introduced extrachromosomal chromatin in F7.2 adopts the murine nucleosomal repeat.

The *ura4* and *ade6* *S. pombe* loci are only able to detect small regions of the DC.1, F1.1 and F7.2 transgenomes. We were unable to use larger cosmid based probes to examine longer lengths of DNA as non specific cross hybridisation to the mouse genome was found with them, obscuring the nucleosome repeat of the yeast DNA. The *S. pombe* centromere repeat dh detects 200kb of centromeric DNA from centromeres 1, 2 and 3. Although dh will detect little in F7.2 it appears that over 140kb of *S. pombe* centromeric sequence is present in F1.1. Use of this sequence should allow a large region of the F1.1 transgenome to be analysed.

Figure 6.12 demonstrates the hybridisation detected by the *S. pombe* centromeric repeat sequence dh to micrococcal nuclease digests of nuclei from F1.1, F7.2 and C127 cell lines. No significant level of cross hybridisation is detected by this probe to the C127 genome. The pattern of hybridisation is the same for both F1.1 and F7.2 lanes. Faint hybridisation can be detected at the expected position for the core particles in both cell lines. The C+1N multimer is approximately 340bp in size, the expected size for a mouse nucleosomal repeat. The C+3N and C+4N nucleosomal multimers provide stronger evidence of the adoption of a mouse nucleosome ladder as they are 720bp and 910bp in size respectively, the expected sizes for a mouse ladder. This result therefore demonstrates that the dh *S. pombe* centromeric sequence adopts a murine nucleosomal pattern in mouse cells. dh may detect up to 100kb of sequence in the F1.1 cell line and therefore represent a relatively large region of the insert.

Figure 6.13 depicts the hybridisation of dh to micrococcal nuclease digests of nuclei from the F1.1 hybrid cell line, DC.1 transfectant cell line and C127 parental cell line. Increased amounts of DNA were also loaded (20µg) in each track and therefore the core particles of the expected ~150bp length are seen in the F1.1 lanes. The C+1N, C+2N and subsequent nucleosome multimers also migrate to the expected murine sizes. In the DC.1 lanes 10µg samples were loaded and so the core particle is barely observed. The nucleosome ladder detected by dh in DC.1 is the same as that seen in F1.1 and F7.2 (Fig 6.12) and is characteristic of mouse chromatin, consistent with the results of the *ura4* hybridisation (Fig 6.10), demonstrating that naked DNA and chromatin are not differentiated, with regard to nucleosome structure, after introduction.

FIGURE 6.12 MICROCOCCAL NUCLEASE DIGESTION OF F1.1, F7.2 AND C127 NUCLEI PROBED WITH *S. pombe* dh SEQUENCE.

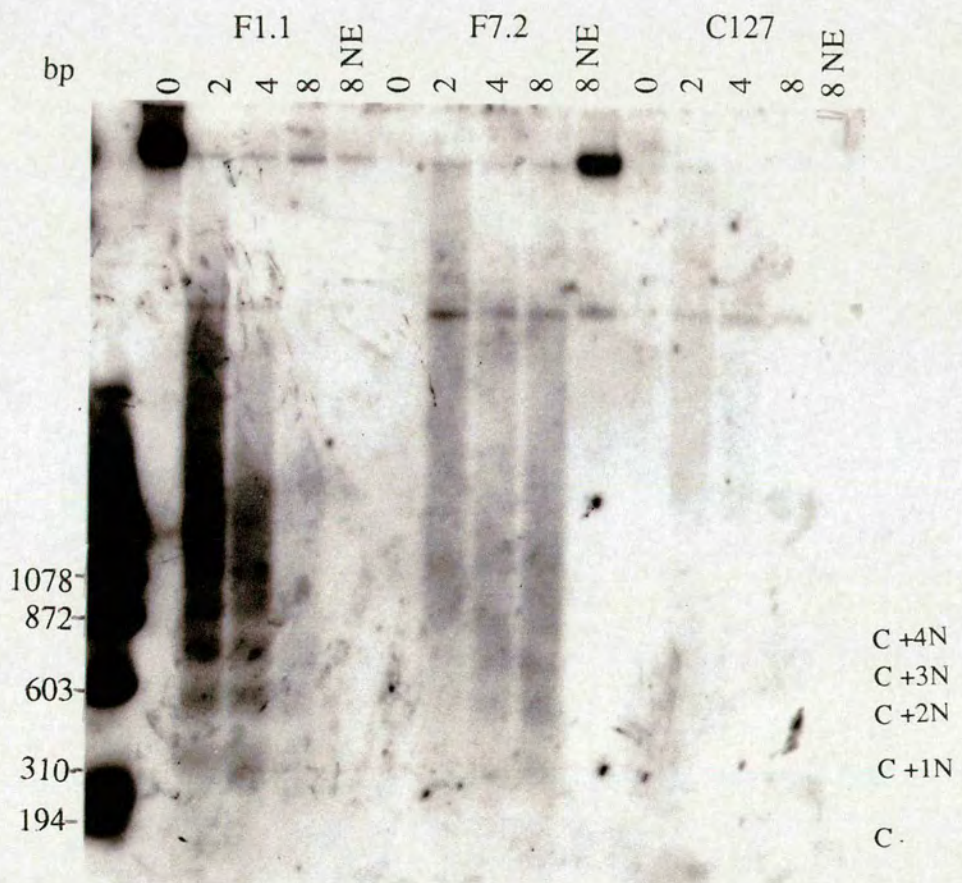


Fig 6.12 : Nuclei were prepared and digested as described in Section 2.16. Digests were size fractionated by 2% agarose gel electrophoresis, and transferred to 0.2um nitrocellulose. The filter was hybridised with random primed dh sequences. Numbers above the lanes correspond to digestion time in minutes. Size marker is ϕ X174 Hae III. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

FIGURE 6.13 MICROCOCCAL NUCLEASE DIGESTION OF F1.1, DC.1 AND C127 NUCLEI PROBED WITH *S. pombe* dh.

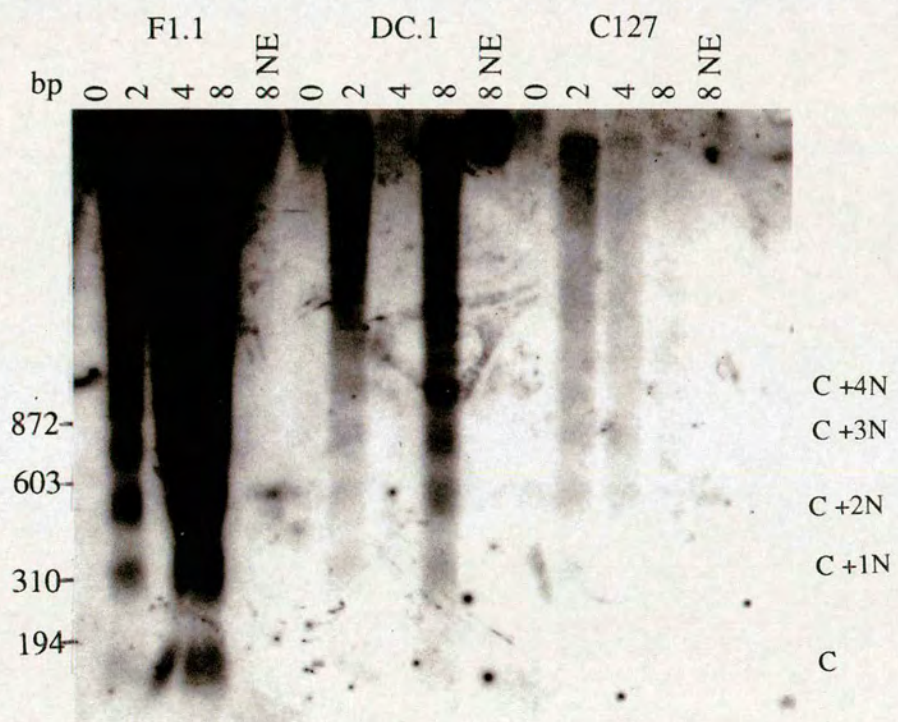


Fig 6.13 : Nuclei were prepared and digested as described in Section 2.16. Digests were size fractionated by 2% agarose gel electrophoresis, and transferred to 0.2um nitrocellulose. The filter was hybridised with random primed dh sequences. Numbers above the lanes correspond to digestion time in minutes. NE- no enzyme added. C = core particle. C +N = core particle + N nucleosome repeat units.

6.4 Conclusions.

In analysing the cause of the unusual chromosome structure observed in the F1.1 cell at the site of *S. pombe* sequence integration, it was felt that nucleosome structure could affect the metaphase structure of this region if the *S. pombe* sequences either displayed a disrupted nucleosomal profile or maintained their original yeast specific 155bp nucleosome repeat. It has previously been demonstrated that the diameter of the "30nm" fibre is dependent upon the nucleosome repeat length of the chromatin (Alegre and Subirana, 1985; Athey *et al.*, 1990). Since the "30nm" fibre is the basis upon which higher chromatin packaging is superimposed, then any differences here between yeast and mouse DNA may be reflected during condensation at metaphase.

It was found that in the DC.1 cell line that all *S. pombe* sequences tested had adopted a mouse nucleosomal repeat (Figs 6.10 and 6.13). DC.1 is a transfectant cell line, produced by the introduction of naked yeast DNA into C127 cells. That the introduced DNA in this line has adopted a mouse nucleosomal repeat suggests that the nucleosome repeat is dependent upon histones and other chromatin accessory proteins rather than the underlying DNA sequence.

In F1.1, a cell line derived by the introduction of chromatin, the yeast transgenome is also found to adopt the mouse nucleosome repeat. Suggesting three possible mechanisms of packaging the F1.1 chromatin. Firstly the introduced yeast chromatin may be stripped of its proteins to produce naked DNA during transgenome formation. Alternatively during the replication of the introduced yeast chromatin in the mouse cell new histones were assembled on the *S. pombe* daughter molecules with no reference to the position of the old molecules on the incoming strands. Lastly the nucleosome repeat could be altered by cis acting effects of the flanking mouse chromatin on the integrated yeast chromatin. The possibility that the mouse repeat could be adopted due to cis acting effects of murine sequences would appear to be negated by the fact that the transgenome of F7.2 is extra chromosomal and therefore unlikely to contain an appreciable amount of murine DNA yet still adopts a mouse nucleosomal repeat. Presuming that the nucleosomes of the yeast transgenomes are retained until replication occurs these findings would suggest that nucleosomes are dissociated from the DNA at replication and replaced by new nucleosomal structures. Such findings therefore have implications for models of nucleosome behaviour at replication.

It is unfortunate that the probes, SV₂neo and tm, exhibit a high degree of non-specific cross hybridisation to the C127 genome. As the SV₂neo sequences are the only sequences

known to be expressed from these *S. pombe* transgene regions, their nucleosome character would have been of particular interest. Adoption of the mouse nucleosome ladder would show that these foreign sequences can adopt the endogenous repeat and still be efficiently expressed. The *S. pombe* centromeric core sequence would also have been interesting to study since in *S. pombe* this sequence exhibits a non nucleosomal structure (Polizzi and Clarke, 1991). Upon transfer into *S. cerevisiae* this sequence is observed to adopt a nucleosomal repeat. It would have been of interest to establish whether this non-nucleosomal structure is maintained in mouse cells.

I have shown that the *S. pombe* sequences in mouse hybrid cell lines adopt a normal mouse nucleosome repeat. Binding of murine histone H1 by these *S. pombe* regions would be of interest to establish. Binding of Histone H1 is known to be required for higher levels of packaging to occur (Thoma et al., 1979), yet no H1 activity has been detected on yeast chromatin. Adoption of the murine nucleosome repeat may suggest that the *S. pombe* transgenomes are indeed bound by histone H1. Histone H1 binding could be detected by immunocytochemistry using an antibody specific to H1.

The results presented in this chapter show that the *S. pombe* sequences within the mouse cell are packaged into nucleosomal arrays indistinguishable from those of mouse chromatin. Since no difference can be found, at this level, between *S. pombe* and mouse sequences the nucleosome packaging of the *S. pombe* sequences would not appear to be responsible for their observed anomalous morphology in the F1.1 cell line. The cause of this structure must therefore be at a higher level of chromatin packaging.

CHAPTER 7. ATTACHMENT OF *S. pombe* SEQUENCES TO THE NUCLEOSKELETON OF THE F1.1 HYBRID CELL LINE.

7.1 INTRODUCTION.

Chromatin packaging above the level of the nucleosome repeat, is poorly understood. It is generally agreed, however, that the chromatin of eukaryotic genomes is gathered into looped domains, and it has been proposed that these looped domains are packaged to produce the metaphase chromosome (Rattner and Lin, 1985). The existence of looped chromatin domains within eukaryotic nuclei was first demonstrated by visibly observing dehistonised chromosomes (Paulson and Laemmli., 1977), under electron microscopy and by production of nuclear matrices (Cook and Brazell, 1975). Both of these experiments demonstrated that chromatin loops emanate from a proteinaceous central structure. That these loops were torsionally constrained, was demonstrated by the change in halo size upon addition of an intercalating dye such as EtBr (Cook and Brazell, 1976). Much controversy surrounds the proposed mechanisms for formation of these looped domains. Two main avenues of investigation have been followed. Scaffold preparations have been produced by the extraction of soluble material from nuclei after nuclear stabilisation and LIS treatment (Mirkovitch *et al.*, 1984). Nucleoskeletal preparations have been achieved by digestion of chromatin under "physiological conditions" and removal of unattached material by electrophoresis (Jackson and Cook, 1985).

Scaffold preparations have proved remarkably popular to investigators and a large body of literature exists using this method. In this procedure nuclei are stabilised by divalent cation and heat treatment. All soluble proteins are then removed from these nuclei with LIS (Mirkovitch *et al.*, 1984). These preparations are therefore relatively easy to carry out and produce consistent results (Gasser *et al.*, 1989). If chromatin is digested, by restriction endonucleases prior to LIS extraction, then sequences not attached to a central scaffold type structure will be removed. The sequences retained in these nuclei are presumed to be the sites of attachment to the nuclear scaffold and also to delimit the domains of chromatin loops. Attachments to a scaffold structure have been mapped surrounding the *Drosophila* genes *Adh*, *Sgs-4*, and *ftz* (Mirkovitch *et al.*, 1984). In *S. cerevisiae* attachments have comapped with autonomous replicating *ARS* sequences (Amati and Gasser, 1988), the *HMR-E* silencer (Abrahm *et al.*, 1984), the *histone H4* gene (Bouton and Smith, 1986) and the *HO* gene (Kearsey, 1984). In higher eukaryotes

attachments have been mapped near the human β -interferon gene (Bode and Maas, 1988), the chicken lysozyme gene (Stratling and Dolle, 1986), and the human β -globin gene (Jarman and Higgs, 1988). Binding of scaffold attachment regions (SARs) is conserved, if the region is tested for attachment to scaffold preparations after transfer to a foreign cell type (Amati *et al.*, 1990). The major sequence similarities of SARs are that many contain *in vitro* topoisomerase II (*topoII*) consensus sequences (Gasser *et al.*, 1986) and that many colocalise with sequences which have potential ARS function in *S. cerevisiae* (Brun *et al.*, 1990). SAR sequences are found to be bound throughout the cell cycle suggesting that these regions may be permanently attached, and therefore have potential as domain forming boundaries and may be significant in metaphase chromosome structure.

The biological significance of these SAR regions as domain forming boundary sequences is, however, open to question. The most frequent single protein found to be associated with these SAR sequences is topo II. Studies on *in vivo* topo II cleavage have demonstrated that, as opposed to the A/T rich *in vitro* consensus sequence, *in vivo* topo II is found to cleave at a G/C rich sequence (Kas and Laemmli, 1992). It has also been found that high topo II levels can be successfully used as a marker for proliferation (Heck and Earnshaw, 1986). In non proliferating cells such as terminally differentiated chicken erythrocytes topo II is found at levels of only 300 molecules per cell and hence it is hard to envisage a role for it as a major structural protein of chromosomes in these cells. Studies have shown that topo II depleted extracts affect entry of nuclei into mitosis (Adachi *et al.*, 1991; Uemera *et al.*, 1987). Blockage of entry into mitosis and its increase in abundance after replication suggest that topo II is functionally required late in the cell cycle. This is supported by its role in chromosome segregation and condensation in *S. pombe* (Rose *et al.*, 1990; Uemera *et al.*, 1987). These findings suggest that topo II is not a major protein involved in loop formation throughout the cell cycle.

It has been suggested that most of the loop attachments observed in dehistonised chromosomes and nuclear matrices are caused by transcription and replication (Jackson and Cook, 1985, 1986). Nuclear cages/ matrices are formed by 2M NaCl extraction of nuclei. The cage like structures at the base of the chromatin loops were shown to be the sites of transcription (Jackson *et al.*, 1984) and replication activities (Dijkwel *et al.*, 1979). Such preparations are open to the criticism that these structures are observed after high salt extraction and therefore may be artefactual.

Manipulations have therefore been carried out under approximately physiological conditions in order to minimise the influence of artefactual results. Cells are lysed, chromatin digested and unattached material removed by electrophoresis, all under

physiological ion conditions (Jackson and Cook, 1985). The structure remaining after extraction is termed the nucleoskeleton. The nucleoskeleton has been shown to contain active genes and the vast bulk of the replicative and transcriptional capacity of the nucleus. 80 % of cellular polymerase II activity and 92 % of nascent RNA transcripts are attached to the nucleoskeleton despite as little as 2% of the chromatin remaining (Jackson and Cook, 1985). 90 % of DNA polymerase activity is also found attached to the nucleoskeleton (Jackson and Cook, 1986). Replication has been shown to occur at fixed sites by the localisation of replicative foci observed in labelled nuclei (Nakamura *et al.*, 1986; Nakayusa and Berezney, 1989; Mills *et al.*, 1989).

Average loop sizes found in eukaryotic genomes may be assessed by comparison of the amount and average size of released and attached sequences. A comparison of loop sizes from scaffold and nucleoskeleton preparations has been carried out in HeLa cells (Jackson *et al.*, 1990). Nucleoskeleton preparations suggest an average loop size of 80kb, while scaffold preparations suggest an average 15kb loop size. It appears that most of the attachments in scaffold preparations are produced during the stabilisation procedure. Prior to this stabilisation stage digestion reveals loops of approximately 100kb. Neither of these procedures produces any significant change in average loop sizes throughout the cell cycle.

Results from physiologically buffered systems suggest that the major attachments involved in loop formation are due to transcription. It is not known if transcription is actually required or whether the presence of initiation or other complexes may be sufficient to allow attachment. Little evidence for purely structural attachments, as proposed for scaffold preparations, has been realised though such structural attachments would not easily be detected from the biochemical analysis of these nucleoskeletons that has been undertaken.

I have undertaken analysis of attachment of the *S. pombe* sequences located within the F1.1 insert to the nucleoskeleton, using the physiologically buffered systems. If a region contains, on average, smaller loops than found in the surrounding chromatin, then this may be reflected at metaphase as a smaller chromosomal region (Figure 1.3). In the F1.1 *S. pombe* insert it is therefore possible that if attachments are closer on average than in the surrounding mouse regions than at metaphase these smaller loops will affect the chromosome's appearance. Smaller average loops in this *S. pombe* region of the F1.1 cell line may result from the compact nature of the *S. pombe* genome. In *S. pombe* genes are tightly packed with very little repeated and junk DNA intervening. This provides the potential for the nucleoskeleton to interact with the *S. pombe* region more frequently than would be expected along mouse sequences. If appropriate transcription complexes are

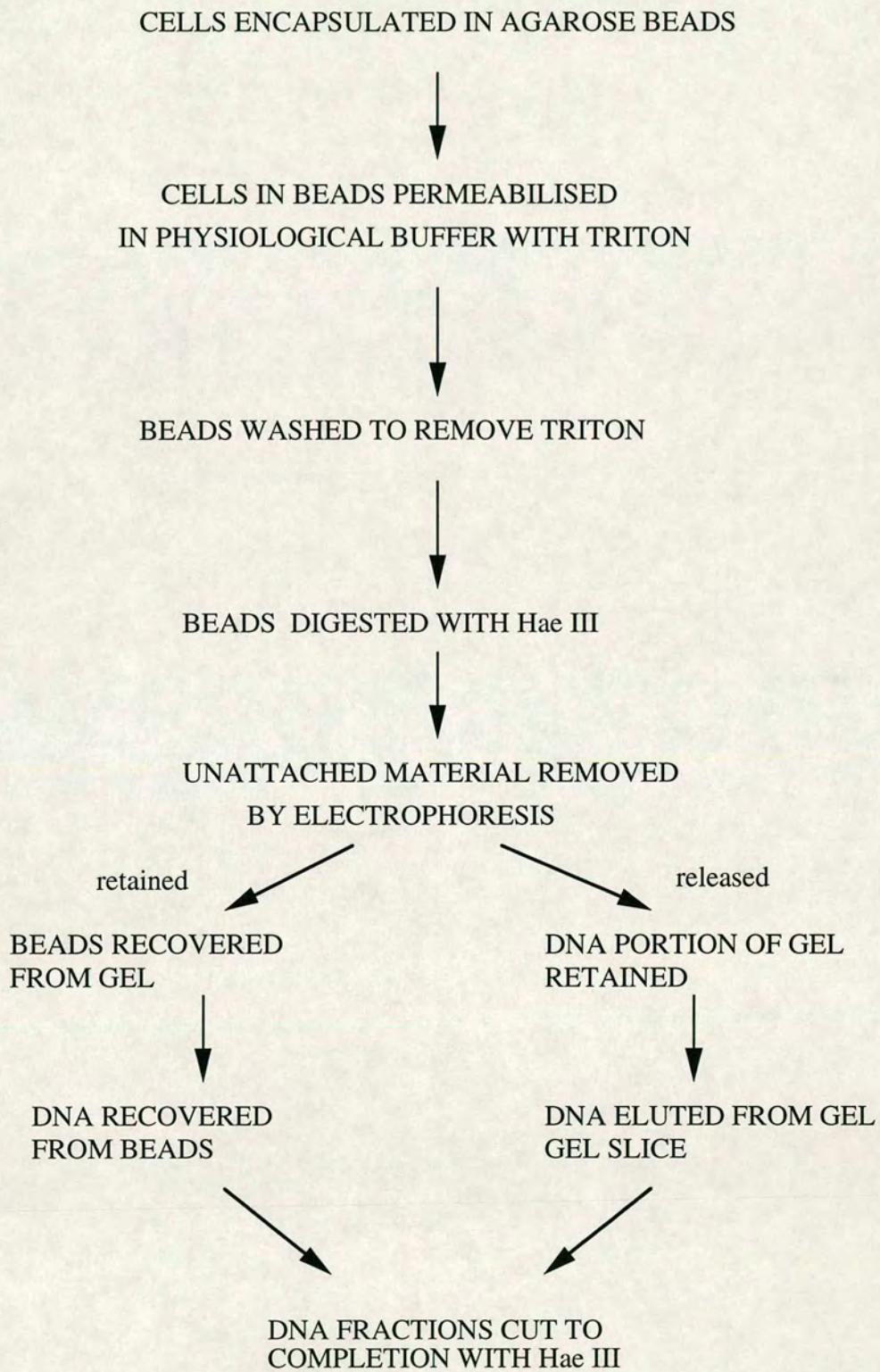
formed, due to conservation of promoter and transcription factor interactions between certain *S. pombe* sequences and mouse proteins (Jones *et al.*, 1988; Kleinschmidt *et al.*, 1990), or if advantageous interactions allow attachment of *S. pombe* sequences, then smaller loops may be formed within this insert region. Alternatively *S. pombe* sequences may not be attached at all resulting in aberrant chromatin packaging.

7.2 Nucleoskeleton procedures.

To investigate the attachment of the *S. pombe* sequences to the nucleoskeleton of the F1.1 hybrid cell line I have modified the procedure of Jackson *et al.* (90). This is shown in Figure 7.1 and described in detail in Section 2.18. Briefly, cells are encapsulated within agarose to form beads, resuspended in a physiologically based buffer and the cells permeabilised by a washing in buffer with added 1% Triton. These permeabilised cells are then digested with the restriction endonuclease Hae III and chromatin not attached to a nuclear substructure removed by electrophoresis of the digested beads. Attached sequences are isolated from the beads by phenol extraction, and the unattached chromatin recovered by electroelution of the DNA from the gel. Impurities were then removed from the DNA by passage through an "Elutip-D" (Schleicher and Schuell). These DNA samples were then redigested to completion with Hae III and size fractionated by gel electrophoresis. DNAs were then transferred to membranes and probed with the required sequences. The analysis undertaken here contrasts to that of Jackson *et al.*, (1990) in that they could only compare attached to total DNA, by isolating the released DNA I can compare released to retained to total, thereby obtaining clearer information as to which sequences are enriched in which fraction.

Hae III digested nuclei after electrophoresis of the unattached DNA have been shown to maintain approximately 5% of the chromatin attached to the nucleoskeleton. After recovery of the sequences from these preparations it would be expected that the attached sequences should be greatly enriched (~20 fold) if interactions are specific. In fact, levels of enrichment of ~3 fold have been found for sequences in these attached fractions suggesting that the interactions involved may be of a dynamic nature rather than permanent (Dickinson *et al.*, 1990) and that sequences contained within the attached fraction are attached at varying levels. Thus sequences may be attached when undergoing transcription, replication or are being processed by other nuclear activities such as DNA repair. Attachment under these conditions would be transient in nature, and the level of attachment within a region would be dependent on the activity of the region within the population of cells analysed. Whether certain sequences within the genome are permanently attached to

FIGURE 7.1 PROCEDURE FOR NUCLEOSKELETON PREPARATIONS



the nucleoskeleton, and therefore much more highly enriched, through activities such as domain boundary functions remains to be seen.

Potential problems with this technique are easily recognised. Since the nature of the attachments appears to be dynamic loss of sequences during electroelution may lead to bias in the results. No fixing or cross linking of the structures in the nucleus occurs during these procedures and during electrophoresis of the unattached chromatin, weak attachments to sequences may be broken and the sequences lost to the released fraction. This can be minimised by keeping the beads at 4°C at all stages apart from digestion to reduce dissociation and protease activity.

7.3 Digestion of chromatin in permeabilised cells.

Since digestion of nucleoskeletons initially occurs on chromatin templates, these reactions will not go to completion, due to protection of a proportion of the potential cleavage sites by nucleosomes and other chromatin proteins. A high degree of partial cleavage with reactions could lead to bias in the results if little of the DNA is removed. Fig 7.3 demonstrates how enzyme concentration greatly affects the level of DNA released from these preparations. Lanes 8 and 9 demonstrate a normal Hae III complete digest carried out on F1.1 and C127 DNAs respectively. It can be seen that the bulk of the DNA is cut to smaller than 2kb fragments as would be expected from a 4 base cutter like Hae III. Lanes 1-7 show DNA from a bead preparation which has been cut with varying levels of Hae III, treated with proteinase K and electrophoresed. Lane 1 demonstrates uncut DNA from a bead preparation which can be seen to be intact and too large to enter the gel. Lanes 2, 3 and 4 show total (unextracted) DNA from bead preparations cut with Hae III at concentrations of 100, 300 and 600 u/ml respectively. The resultant partial digests, caused by protection from cleavage of the chromatin template, can be seen by comparing these lanes with the digests of lane 8 and 9, the average fragment size of the digests has been drastically increased due to the protection from cleavage. The amount of digestion greatly increases as the enzyme concentration is increased from 100 u/ml (Lane 2) to 600 u/ml (Lane 4), with the decrease in the average fragment size corresponding to the increase in enzyme concentration. No significant increase in digestion is achieved above this 600 u/ml level suggesting that all available sites are cleaved at this level of digestion. That the chromatin within these digests is intact can be seen from the nucleosome ladder apparent at the bottom of these digestion lanes. Lanes 5-7 show the DNA retained on the nucleoskeleton after extraction of unattached chromatin, from the digests of lanes 2-4. Comparison of these digests reveals that as the level of digestion increases the amount of attached DNA decreases as would be expected. It is apparent that, retained fragments are

FIGURE 7.2 DIGESTION OF NUCLEOSKELETON PREPARATIONS BY HAE III

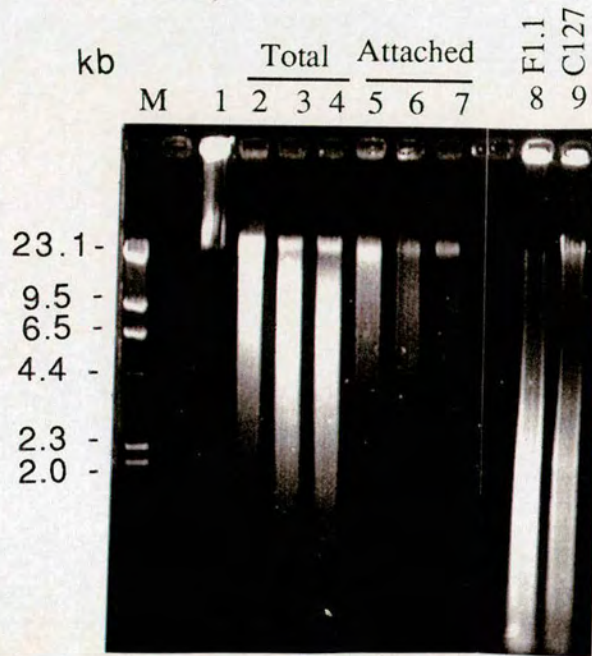


Fig 7.2 : F1.1 cells were harvested and permeabilised as described in Section 2.18. Fractions were digested with Hae III at varying concentrations; Lane 1, 0 u/ml lanes 2+5, 100 u/ml, lane 3+6, 300 u/ml and lane 4+7, 600 u/ml. Lanes 2-4 contain total DNA, lanes 5-7 contain attached DNA. These fractions were then subjected to electrophoresis. The attached chromatin, lanes 5-7 correspond to lanes 2-4 respectively. Hae III digests of total DNA for F1.1 and C127 are shown. Total = digests of total bead DNA. Attached = DNA remaining after extraction.

biased towards larger fragments. This is because as a fragment increases in size it is statistically more likely to contain an attachment site. All future digests were carried out at 600 u/ml of Hae III since this appears to maximise the amount of unattached chromatin released.

After digestion of permeabilised cells and extraction of the unattached material, the released and retained sequences were recovered. These fractions were then redigested to completion, electrophoresed and transferred to hybridisation membrane for further analysis. Figure 7.3 demonstrates two different digests that were used for subsequent analysis. Both of these preparations were the result of digestion with 600 u/ml of Hae III. The digests reflect the normal distribution of Hae III fragments as compared to the digests of total DNA samples. These digests were blotted and used for analysis of retention. In panel B the retained fraction contains only 25 % of the level of DNA present in the released fraction, as assessed by densitometry.

7.4 Attachment of mouse sequences to the nucleoskeleton.

We expect the loop size in mouse to be, on average, 80kb as found in human HeLa cells (Jackson *et al.*, 1990). It would be expected that an average mouse cosmid would contain only one attachment site. Mouse cosmids, due to the frequency of repeated sequences within the genome, proved however difficult to use. I have therefore used cDNAs for analysis of mouse sequences, though these may be expected to be overrepresented in the retained fraction since transcription and attachment are related. Figure 7.4 demonstrates the difference in attachment characteristics seen with 4 mouse cDNAs. The hybridisations were carried out using whole plasmids as probes rather than purified inserts. The most striking feature of all these hybridisations is the enrichment of plasmid hybridising sequences, which derive from the SV₂neo construct, within the attached fraction. There appears to be a very high level of retention of these sequences implying they are strongly attached to the nucleoskeleton. In Figure 7.4 A and D the hybridisation to plasmids sequences is strongest in the retained fraction lane, since the fractions contain equal amounts of DNA (Figure 7.3A) the relative enrichment is high. In Figure 7.4 B and C the plasmid hybridisation can again be seen to be enriched within the retained fraction since in fig 7.3b the retained fraction contains only four times less DNA than the released fraction. The differing pattern of enrichment may be due to the fact that the constructs used contain differing plasmid sequences. That the plasmid derived sequences are so highly retained is of interest since they are located in the region of the SV₂neo gene which is the only known expressing region of this transgenome. These plasmid sequences show that enrichment of sequences within the retained fraction can be observed.

FIGURE 7.3 COMPLETE HAE III DIGESTION OF NUCLEOSKELETON PREPARATIONS AFTER RECOVERY

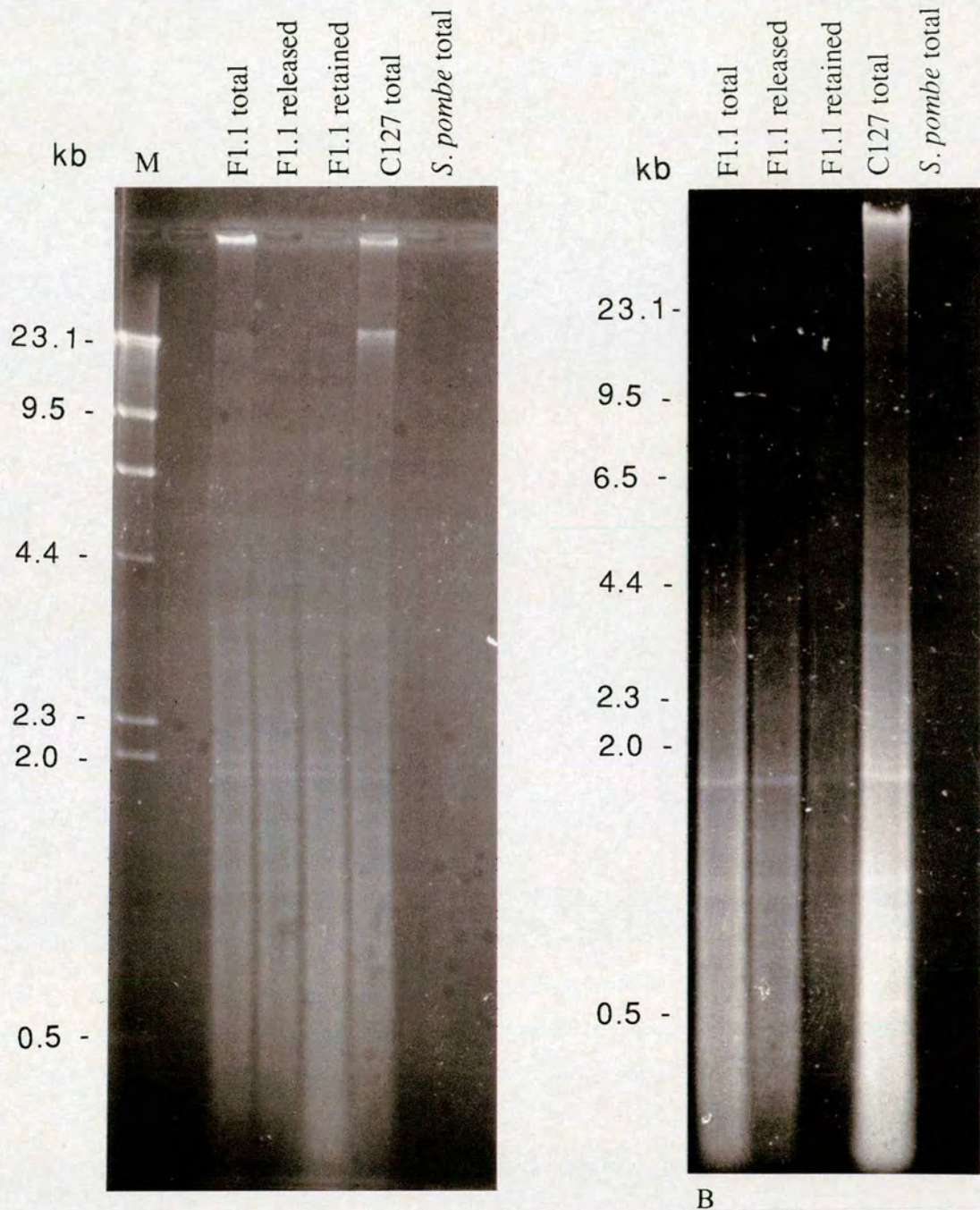


Fig 7.3: Released and retained fraction DNAs, from F1.1 nucleoskeleton reactions were recovered and cut to completion with Hae III. 5 μ g of sample DNAs were then subjected to gel electrophoresis as were 5 μ g of F1:1 and C127 and 50ng of *S. pombe* total DNAs digested with Hae III. A and B correspond to digests of different nucleoskeleton preparations.

FIGURE 7.4. ANALYSIS OF ATTACHMENT OF MOUSE SEQUENCES TO THE F1.1 NUCLEOSKELETON.

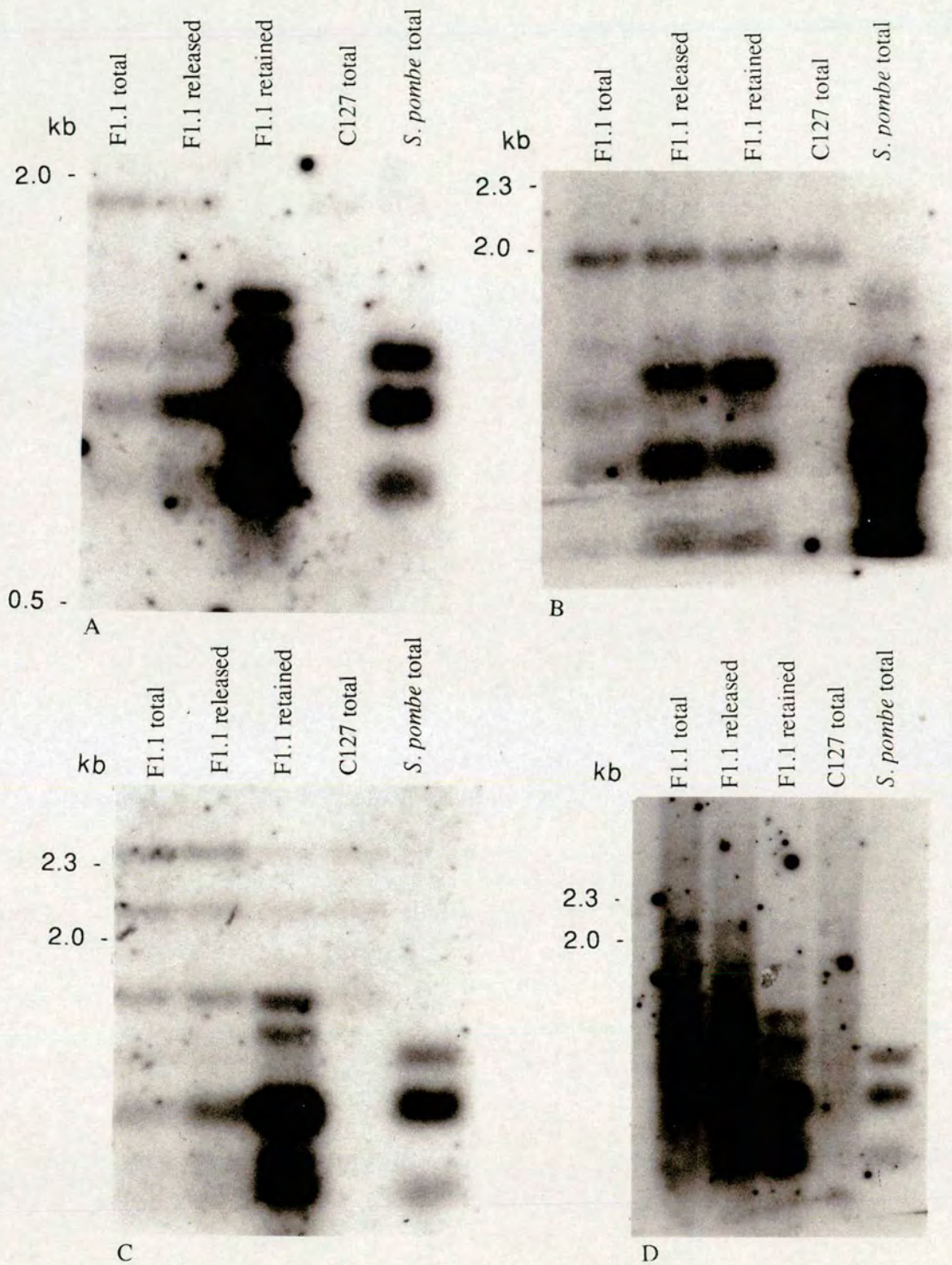


Fig 7.4: Complete Hae III digestions of DNA both released from and retained on the F1.1 nucleoskeleton and F1.1, C127 and *S. pombe* total DNAs were probed with mouse sequences. A: 1.2kb Hind III/Eco RI *Steel* locus insert in pBluescript. B: 1.0kb Eco RI *Hox 8* insert in pUC 19. C: 1.6kb Eco RI *Pax 6* insert in pBluescript. D: 2.3kb Pst I *Hprt* insert in pBluescript. A and D correspond to reactions in Fig 7.3A. B and C correspond to reactions in Fig 7.3B.

Figure 7.4a shows the hybridisation of a plasmid containing a 1.2kb Hind III/ Eco RI insert from the mouse Steel (*Sl*) locus to nucleoskeleton digests shown in Fig 7.3a.. The product of the steel locus is involved in differentiation of various cell lineages in haemopoiesis and in melanocytes (Williams *et al.*, 92). Specific hybridisation to the mouse sequences is seen at approximately 1.9kb. The hybridisation band can be seen in the total DNA digests of F1.1 and C127 DNAs and also in the released DNA fraction from F1.1. No hybridisation is seen within the attached fraction of F1.1 DNA. This is in marked contrast to the plasmid hybridising bands which originate from the SV2neo region of the F1.1 yeast transgenome which are stronger in the nucleoskeleton retained fraction than either total and released fractions, even though Fig 7.3a shows the fractions contain equal quantities of DNA. This result suggests that the mouse *Sl* coding region is not attached to the nucleoskeleton. We do not know whether the *Sl* gene is expressed in F1.1 or C127 cell lines and therefore no correlation can be made between this lack of attachment and expression.

Fig 7.4B demonstrates the hybridisation of a plasmid containing a 1kb Eco RI insert from the mouse *Hox 8* gene, to the reactions shown in fig 7.4b. *Hox 8* is a gene which has been shown to be involved in limb development, and its state of expression in the cell lines tested is unknown (Monaghan *et al.*, 1991). It can be seen that in this case the hybridisation (1.9 and 1.4kb) can be found in both released and retained fractions. This suggests that any attachment of these sequences to the nucleoskeleton is not permanent or is fairly weak and a fraction of the material is released during the procedure.

Fig 7.4 C shows the hybridisation of a 3.1kb Eco RI insert from the mouse *Pax 6* gene to nucleoskeleton preparations shown in Figure 7.3b. *Pax 6* is a gene involved in development of the central nervous system, eye and face (Hill *et al.*, 1991). Hybridising bands are seen in both released and retained fractions implying that these sequences are retained on the nucleoskeleton.

Fig 7.4 D shows the hybridisation of a plasmid containing a 2.3kb Pst I insert from the mouse hypoxanthine phosphoribosyl transferase gene (HPRT). This gene is involved in the salvage of nucleotides for DNA replication and is considered to be a housekeeping gene that is expressed in most cell lines. Although a poor hybridisation, it can be seen that the mouse specific hybridising bands, (1.8 and 2.3 kb), present in the total DNA and released fraction lanes, can only be seen faintly in the retained fraction. This suggests that these sequences may only be weakly associated with the nucleoskeleton. A 1.6kb band is barely apparent within the retained fraction, although it can be seen in both total and released fraction lanes in the original autoradiograph. This suggests that this band is located

further from the site of attachment than the 1.8 and 2.3kb bands. This hybridisation was carried out on the same reactions (Fig 7.3A) as those of the *Sl* hybridisation (Fig 7.4A) and reveals that differences in the relative levels of attachment can be detected.

Figures 7.3 and 7.4 show that sequences maintained upon the nucleoskeleton or released from it can be differentiated by hybridisation to nucleoskeleton preparations and that sequences are not attached to the nucleoskeleton with a uniform affinity.

7.5 Attachment of *S. pombe* sequences to the nucleoskeleton of F1.1

Attachment of the *S. pombe* DNA sequences to the F1.1 nucleoskeleton has also been examined. It could be envisaged that either the F1.1 yeast transgenome may fail to associate with the mouse nucleoskeleton, or it may be associated very frequently because of the high gene density of the yeast genome. If the *S. pombe* sequences are attached more frequently along their length then this would produce smaller loops in this region than in the surrounding mouse genome.

As demonstrated in Figure 7.4 the plasmid sequences contained within the *S. pombe* insert, in the region of the SV₂neo gene, appear to be very highly attached to the nucleoskeleton of the F1.1 cell. This implies some form of functional or structural attachments specific for these plasmid sequences. The SV₂neo sequences are the only sequences within the *S. pombe* insert which are known to be expressed, as these sequences confer the G418 resistance used to select this lineage. Fig 7.5 a and b demonstrate the attachment of the SV₂neo sequences themselves to the nucleoskeleton of the F1.1 cell line. Figure 7.5 a hybridisation to the reactions shown in Figure 7.3a which have also been hybridised to the mouse *Sl* (Fig 7.4a) and HPRT (Fig 7.4d) sequences which were seen to be unattached and lightly attached, respectively. The SV₂neo sequences in Figure 7.5 appear to be attached to the nucleoskeleton, as shown by the presence of hybridising bands in the retained fraction lanes. Although the absolute level of enrichment is difficult to assess, it appears that these sequences are retained relatively highly. Retention of these sequences on the nucleoskeleton is consistent with the idea that transcribed sequences are associated with the nucleoskeleton. It can be seen that, along with hybridisation to the *S. pombe* sequences, this probe also detects some faint cross hybridisation to the C127 genome. The sequences to which this cross hybridisation is directed however do not appear to be attached to the nucleoskeleton, as they are absent from the retained fraction lane and so make a good contrast to the yeast sequences. Fig 7.5b shows SV₂neo hybridised to reactions used for the *Pax 6* hybridisation (Fig 7.3b and Fig 7.4B). As with the other preparation (Fig 7.5A), it can be seen that SV₂neo is maintained upon the nucleoskeleton.

FIGURE 7.5. ANALYSIS OF ATTACHMENT OF *S. pombe* SEQUENCES TO THE F1.1 NUCLEOSKELETON.

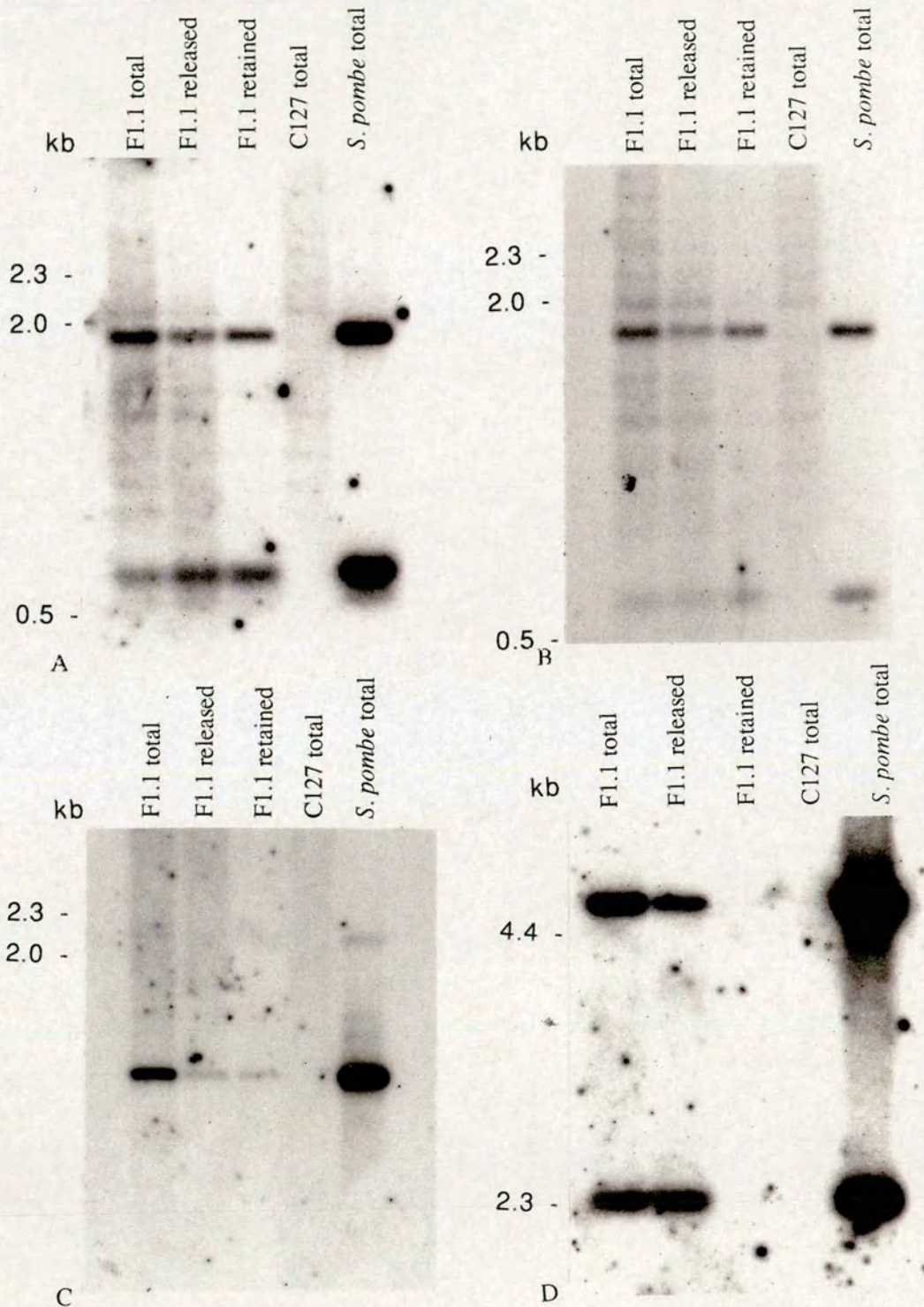


Fig 7.5: Complete Hae III digestions of DNA both released from and retained upon the F1.1 nucleoskeleton and F1.1, C127 and *S. pombe* total DNAs were probed with *S. pombe* sequences. A and B: SV2neo sequence. C: *ars1* sequence. D: *ade6*. A and C correspond to reactions in Fig 7.3A. B corresponds to reactions in Fig 7.3b. D does not correspond to either of the panels in Fig 7.3.

Figure 7.5 c demonstrates the hybridisation of a 1.2 kb Eco RI fragment containing the *S. pombe arsI* (Losson and Lacoute, 1983) sequence to the nucleoskeleton reactions shown in Figure 7.3A. Hybridisation is detected in both retained and released fractions, indicating that this sequence is associated with the nucleoskeleton to some extent. *ArsI* is a sequence in *S. pombe* which shows similar functions to the *S. cerevisiae* ARS I, which can function as an origin of replication. ARS sequences in *S. cerevisiae* have been shown to be associated with scaffold attachment (Amati and Gasser, 1988).

Figure 7.5d demonstrates the hybridisation of the 3.2 kb *ade6* probe to nucleoskeleton preparations. Whilst hybridisation is apparent in the released fraction, no signal is observed in the retained fraction suggesting that these *ade6* sequences are not maintained at all upon the F1.1 nucleoskeleton.

It would appear therefore that *S. pombe* sequences within the F1.1 cells also vary in their attachment to the mouse nucleoskeleton. The one sequence known to express in the F1.1 cell line, SV2neo, has been shown to be associated with the nucleoskeleton as is *arsI*, whilst the *ade6* gene appears to be free of attachment. It has been suggested that attachment may be mediated through various nuclear activities, the most predominant of which is transcription (Jackson and Cook, 1986). Although this may be largely true of endogenous sequences, what type of interactions may occur in a large region of foreign sequence remains to be proven.

7.6 Attachment over an extended region of *S. pombe* sequences to the F1.1 nucleoskeleton.

The analysis of attachment of *S. pombe* sequences to the nucleoskeleton of the F1.1 cell line, shown in Figure 7.5, is only of specific small regions. Analysing a continuous region of the transgenome should allow the frequency of attachments to be analysed. If chromosome loops are smaller within this *S. pombe* region, for example less than 20kb, then a cosmid of 40kb in length should contain several attachment sites. A mouse cosmid of similar size would only be expected to contain one or less sites of attachment, if mouse loops are indeed 80kb in length.

Figure 7.6 shows the hybridisation of the Cura4-II cosmid to the F1.1 nucleoskeleton fractions shown in Figure 7.3 b. Comparison of the retained and released fractions shows that many of the bands detected by this cosmid are present at some level within the retained fraction. Different levels of attachment of the bands within this cosmid can be seen. Two major bands of hybridisation (1.8 and 2.4kb) appear to be much more strongly retained

upon the nucleoskeleton than others in this region. This is apparent from relative comparison of the bands in the released and retained fractions with the strongest attachment to DNA in the retained fraction even though this contains the least amount of DNA. That bands are also preferentially released is seen by the enrichment of bands within the released fraction. Bands of hybridisation at 3.2, 2.4 and 2.2kb can be seen to display a relative enrichment within the released fraction. A higher level of representation within the released fraction suggesting that these sequences are very infrequently associated with the nucleoskeleton.

Since the Cura4-II cosmid was isolated by selection for the *ura4* gene, it would be interesting to map the *ura4* gene sequences, within this cosmid and analyse their attachment in Figure 7.6. Figure 7.7 shows the Cura4-II cosmid cut with various enzymes and hybridised with the *ura4* gene probe. The Hae III digestion in Figure 7.7a can be seen to be virtually identical to the Cura4-II hybridisation pattern detected in Figure 7.6. When this cosmid is probed with *ura4* gene sequences, Figure 7.7 b, the bands detected are of 2.4 and 1.8 kb in size. If the hybridisation of Cura4-II to the retained fraction of Figure 7.6 is compared to the pattern of hybridisation of the Cura4-II to the *ura4* gene sequences, it is seen that the most highly retained sequences within the cosmid are those of the *ura4* gene sequences themselves. Strong specific interactions between the mouse nucleoskeleton and the *ura4* gene may imply that the transcription machinery of the mouse cell can specifically interact with the sequences of the *ura4* gene. That associations with other sequences within the cosmid are less frequent or weaker than those found with the *ura4* sequences suggests a wide variation in the levels of interactions found within this region.

7.7 Conclusions

In this chapter I have shown that nucleoskeleton preparations as devised by Jackson and Cook (1986) can be used to detect whether a sequence is attached to the nucleoskeleton or not. Jackson *et al.*, (1990) only analysed the retained fraction and relied on comparison of hybridisation signal to that of total DNA to determine levels of enrichment. I have analysed both released and retained nucleoskeleton fractions. Which should allow a clearer picture of differential attachment to be gained than by comparison of only the retained fraction against total DNA.

I have detected attachment of the *S. pombe* sequences to the F1.1 nucleoskeleton thus yeast sequences are able to make specific associations with the proteins of the mouse nucleoskeleton. As seen with the mouse sequences, varying levels of attachment are detectable. The SV₂neo sequences have been shown to be attached to the nucleoskeleton

FIGURE 7.6 ANALYSIS OF ATTACHMENT OF Cura4-II SEQUENCES TO THE NUCLEOSKELETON OF F1.1 CELLS.

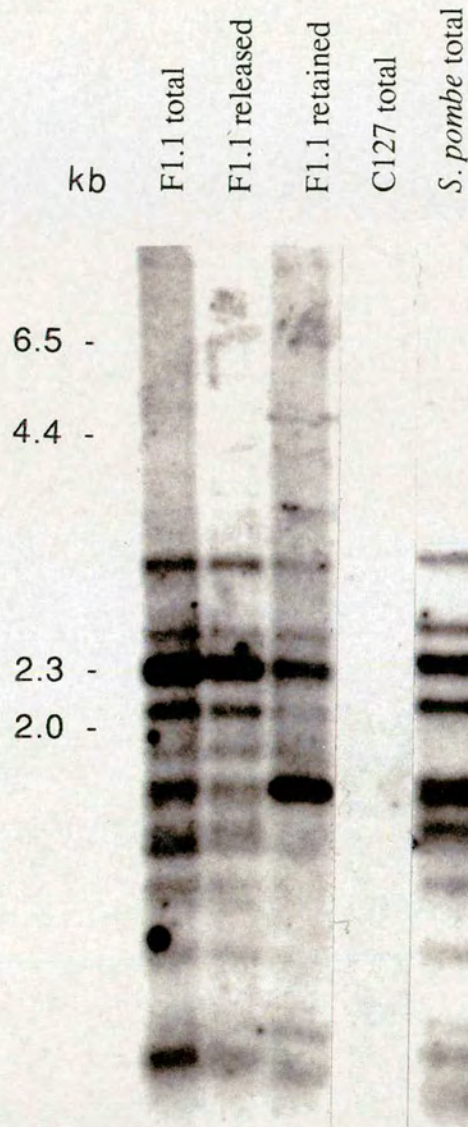


Fig 7.6 : The complete Hae III digests of the F1.1 nucleoskeleton preparation shown in fig 7.3B, were probed with the Cura4-II cosmid.

FIGURE 7.7 Cura4-II COSMID CUT WITH VARIOUS ENZYMES AND PROBED WITH THE *ura4* GENE SEQUENCE.

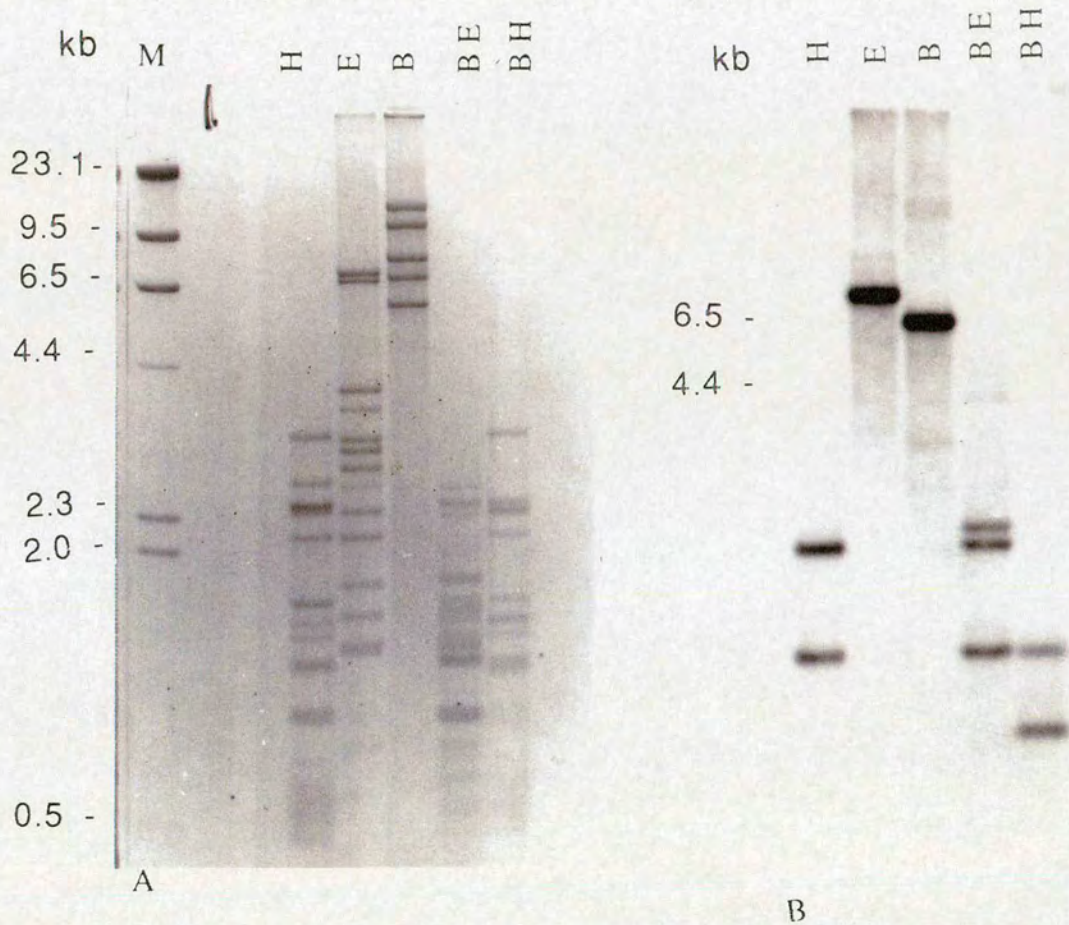


Fig 7.7 : 4ug of Cura4-II DNA were cleaved with Hae III (H), Eco RI (E), Bam HI (B) and combinations thereof, as shown. Digests were subjected to gel electrophoresis and transferred to membranes. Digests were then probed with *ura4* . A is a photograph of the ethidium bromide stained gel corresponding to the autoradiograph in B.

while the *ade6* sequences do not appear to be attached. The *arsI* sequence was also found to be attached. That the level of attachment to the nucleoskeleton varies within the *S. pombe* insert is exhibited by the hybridisation of the Cura4-II cosmid (Fig 7.6). In particular two bands in Fig7.6 appear to be highly retained on the nucleoskeleton. The association of two restriction fragments from a 40kb cosmid could be extrapolated to produce a loop size of ~20kb. That various sites within the Cura4-II cosmid appear to be attached may imply a loop size in this yeast region smaller than this tentative ~20kb figure. This is several fold more frequent than expected on average for the mouse DNA itself (80kb). This proposal requires further analysis over other extended regions of the *S. pombe* insert in the F1.1 cell line, to see if this level of attachment was a general property of large *S. pombe* regions within this and other hybrid cell lines. The smaller loop size postulated for this yeast domain may therefore be the cause of the abnormal cytology observed within this region at metaphase (Fig 1.3). If the smaller loops present within this region are packaged, at metaphase, in a similar manner to the surrounding mouse loops, then it may be expected that this region would appear constricted in comparison. This situation may be comparable to the constricted appearance of centromeres in mammalian cells (Rattner and Lin, 1985).

The high frequency of attachment postulated for the yeast transgenome of the F1.1 cell line may be related to the high gene density of the *S. pombe* genome. This hypothesis is supported by Fig 7.7, which shows that the highly retained fragments in Fig 7.6 correspond to the coding region of the *ura4* gene. The bulk of nucleoskeleton attachments are believed to result from transcription, however, it appears that active transcription may not be a necessity. Jackson *et al.* (unpublished observations) have demonstrated that poisoning for transcription may be sufficient to allow attachment to be detected. Analysis of attachment of plasmid sequences demonstrated that promoter sequences could be associated with the nucleoskeleton without transcription occurring from the associated gene, suggesting that formation of the majority of a transcription complex may be sufficient to allow attachment. If transcription of certain *S. pombe* genes occurred within this yeast transgenome or if *S. pombe* promoters could be recognised sufficiently to form a preinitiation complex then the high gene density may allow for the postulated small loop size. To test the *S. pombe* transgenome for transcription, nuclear run on experiments were attempted. Unfortunately limiting time prevented any data from being gained from these experiments. If more time had been available then it may have proven fruitful to test for transcription by either Northern analysis or by PCR of isolated RNA, although these latter two possibilities require that relatively stable RNA is formed.

Preliminary studies were carried out to determine whether nucleoskeleton preparations could be obtained from *S. pombe* cells. These experiments were singularly unfruitful. It was believed that either the conditions used were incorrect for *S. pombe* nuclei or that excessive protease activity was present, which precluded any worthy analysis. These studies therefore require extensive further analysis. Other laboratories have also reported difficulty in these experiments.

I have been able to demonstrate that *S. pombe* sequences appear to be attached to the nucleoskeleton of the mammalian nucleus. These studies suggest that loops formed within the *S. pombe* region may be smaller on average than in the surrounding mouse genome. It has been proposed that these loops are directly packaged down to produce the metaphase chromosome (Rattner and Lin, 1987). It is suggested that smaller loops may lead directly to the thinner *S. pombe* chromosomal regions observed in the F1.1 cell line.

Chapter 8 . Discussion.

8.1 Aims of project.

The work described in this thesis involved the analysis of the DNA content, modification and chromatin structure of *S. pombe* sequences in *S. pombe* mouse hybrid cell lines. Cytological constrictions within the *S. pombe* sequence domains of hybrid cell lines within which integration of *S. pombe* sequences had occurred, (F1.1, F48C2, F48A1), were apparent after condensation of the chromatin during mitosis (Figs 1.1 and 1.2). Since the appearance of the constrictions was independent of integration site, these constrictions were believed to be intrinsic to the interactions of the *S. pombe* sequences with mouse chromosomal proteins. It was therefore felt that by studying the interactions of the *S. pombe* sequences with mouse proteins, or by analysing certain DNA modifications associated with these interactions, that it would be possible to elucidate the cause of the constrictions and hence highlight aspects of chromatin structure important in chromosome morphology.

Higher order packaging levels of chromatin have largely proved refractory to analysis. The hybrid cell lines produced by Allshire *et al.*, (1987) provide a powerful resource for the analysis of the interactions which may affect chromosome morphology. Since the constrictions are observed within the *S. pombe* regions at metaphase this provides a basis for correlating metaphase morphology with specific protein/DNA interactions within this domain.

Apart from offering insight into interactions important for chromosome morphology this study also addresses questions about the chromatin structure adopted by transgenomes. Knowledge of the interactions that occur between transgenomes and the host chromatin packaging processes is important, since the chromatin structure of a gene appears to be related to its efficient regulation therefore knowledge as to what interactions with the host proteins occur within transgenomes will be an important consideration for transgenic and gene therapy studies.

8.2 Hybrid cell line content

To allow analysis of the interactions within the *S. pombe* sequence domain of the F1.1 cell line it is, first of all, essential to identify which *S. pombe* sequences are present. This was achieved by testing a battery of probes on digests of F1.1, F7.2,

DC.1, DC.2, C127 and *S. pombe* DNAs. This analysis (Table 3.1) demonstrated that a complex pattern of rearrangement of the *S. pombe* sequences has occurred prior to establishment of both F1.1 and F7.2 hybrid cell lines. In both of these cell lines, most of the chromosome III markers have entered the cell lines as would be expected from their initial linkage to the selected SV2neo marker, although deletion of some markers has occurred. In contrast to the similar chromosome III content F1.1 and F7.2 retention of markers from *S. pombe* chromosomes 1 and 2 in F1.1 and F7.2 is very different. In F7.2 only one marker from Chromosome II is present, no chromosome I material was retained. In F1.1 dh sequences from centromere I and *ars1* from one arm of chromosome I are present, while the proximal markers *cyh1* and *his1* from the other arm of chromosome I are both absent. A similar pattern emerged from chromosome II with the centromere and markers from one arm (*nda3* and *dis2*) being present but markers from the other arm were absent. This suggests that approximately 7Mb of yeast genome may have formed the F1.1 transgenome.

The presence of non syntenic DNA in the F1.1 transgenome suggest ligation of incoming *S. pombe* chromosomes prior to integration. Although the presence of non syntenic material within the cell lines is not surprising (Bickmore *et al.*, 1989), the prevalence of rearrangements within the transgenomes was unexpected. Of the 11 markers present in the F1.1 cell line, 25% (3 of 11) exhibit rearranged bands. A similar level of rearrangement is found within the F7.2 hybrid (3 out of 8 markers). These rearrangements are exemplified by the cosmid hybridisations shown in figures 4.5 (Cade6-I) and 4.6 (Cura4-II). This Data shows that microdeletions and rearrangements are present within very short 35kb regions of *S. pombe* sequence within these hybrid cell lines. These cosmids have also demonstrated that the DNA transfectant cell lines (DC.1 and DC.2) contain highly rearranged transgenomes. The degree of rearrangement seen appears to be greater than that seen for the YAC transgenomes formed by PEG fusion of *S. cerevisiae* and mammalian cells (Pavan *et al.*, 1989; Pachnis *et al.*, 1990; Huxley and Gnirke, 1991). These studies however, did not characterise the cointegrated yeast genomic DNA, which our hybrid cell line content may more closely reflect, since primary interest in these experiments lay in the human derived sequences carried on YACs in these *S. cerevisiae* cells.

Microdeletions and rearrangements are not typical of the transgenome structures normally derived from the fusion of vertebrate cells together (Bickmore *et al.*, 1989). Such degrees of rearrangement are however typical of transgenomes derived from

CMGT experiments where the DNA, as chromatin, is introduced by chemical precipitation and not within the protection of a nuclear membrane (Porteous, 1987). In typical fusion experiments between vertebrate cells it is believed that a heterokaryon is formed and the nuclei of the fused cells synchronise and enter mitosis simultaneously with all chromosomes aligning on a single metaphase plate (Harris, 1970). Containment within the nucleus protects the incoming DNA from cytoplasmic degradation and allows the establishment of transgenomes free from microdeletions. In the case of CMGT and DNA mediated gene transfer the DNA and chromatin is exposed to the degradative enzymes in the cytoplasm, followed by religation of this cut DNA to form complex transgenomes (Porteous, 1987).

Ward *et al.* (1986) have demonstrated that heterokaryons are also formed after *S. cerevisiae* and mammalian cell PEG induced fusion. Instability of these heterokaryons, or a different mechanism of hybrid establishment, may allow degradative enzyme access to the introduced DNA in these instances. There is also a possibility that the yeast DNA sequences are inherently unstable within the mouse cell and that the rearrangements are occurring after initial establishment. Inappropriate packaging of the *S. pombe* transgenome is a possibility cause for such instability of sequences within this region. Little evidence for inappropriate packaging of these *S. pombe* sequences has been obtained however the F1.1 transgenome seems stable after extensive passage in culture (Fig5.10).

The *S. pombe* centromeric sequences (tm and dh) appear to have entered the F1.1 cell line in a relatively intact state (fig 3.4 a, b), apart from the apparent lack of centromere 2 dh hybridising bands. In contrast little centromere DNA has become established within the F7.2 extra chromosomal transgenome. It is interesting to speculate that the centromeric sequences may be excluded from this autonomous transgenome if they were, for some reason, affecting the transmission of the autonomous element. A partially functioning heterologous centromere could interfere with partitioning of the transgenome at metaphase and produce a slower growing clone through mis-segregation if, for example, the *S. pombe* centromeric sequences interact inappropriately with the host mitotic apparatus. Since these cell lines are largely chosen for analysis by their ability to grow efficiently the loss of centromeric sequences could be a reflection of the rearrangements which occurred allowing this cell line to grow more rapidly than a clone bearing centromeric sequences. However we do not know if there is sequence similarity between *S. pombe* (Murakami *et al.*,

1991) and mouse centromeres and the lack of *S. pombe* centromeric sequences in F7.2 may purely be a reflection of random rearrangement during establishment.

The *S. pombe* telomeric sequences show a pattern of retention in the F1.1 and F7.2 cell lines which may be very much related to the different processes of establishment of these integrated and autonomous transgenomes. *S. pombe* telomeric sequences are not found in the F1.1 transgenome while the F7.2 cell line contains rearranged *S. pombe* telomeric sequences (Fig 3.4c). It has previously been shown (Huxley and Gnirke, 1991) that upon integration of *S. cerevisiae* DNA into mammalian genomes, the yeast subtelomeric Y' sequences are very rarely present. This and our finding that the integrated F1.1 *S. pombe* transgenome also lacks telomeric sequences suggests that loss of ends of molecules may be related to the integration of such transgenomes and the coligation of non-syntenic material at the integration site present in F1.1 and also in the *S. cerevisiae* experiments (Huxley and Gnirke, 1991). The F7.2 autonomous transgenome has not lost the telomeric sequences and it would be interesting to establish whether they are still telomeric and whether they have acted as a template for addition of mouse telomeric sequences.

Obtaining information as to which regions of the *S. pombe* genome had entered the hybrid cell lines allows specific regions to be analysed for modifications and interactions with the mouse nuclear proteins. I have undertaken to study epigenetic modifications of the *S. pombe* sequences, their nucleosomal structure in the murine environment and their chromosomal loop formation characteristics.

8.3 Methylation.

Methylation was target of interest in the analysis of the novel chromosome morphology of the *S. pombe* insert in F1.1. MECP proteins (Meehan *et al.*, 1988; Lewis *et al.*, 1992) which bind methylated DNA in a non sequence specific manner had been isolated, which appeared to be involved in packaging chromatin into a less accessible configuration than that of non methylated chromatin (Antequera *et al.*, 1989). An initially unmethylated *S. pombe* genome was introduced into the methylated genome of the mouse cell during formation of the transgenomes (Allshire *et al.*, 1987). Not all cultured cell lines contain a *de novo* methylase activity (Szyf *et al.*, 1990), thus it was postulated that if such a large insert (>10Mb) were to remain non methylated then lack of binding by chromatin packaging proteins, such as the MECPs, could lead to presentation of an aberrant morphology at metaphase.

As demonstrated in chapter 4 the F1.1 cell line exhibits dramatic *de novo* methylase activity upon the *S. pombe* insert region. Since no overall difference in methylation is detected between the parental C127 cells and the hybrid cell line it is concluded that this *de novo* methylase activity is specific to the insert region. The *con* mouse probe shows no increase in methylation (Fig 5.1) over a time scale within which the *S. pombe* insert region dramatically increases methylation (Fig 5.6). This increase in methylation is not correlated with any change in the morphology of the F1.1 insert at metaphase (Fig 5.11). This implies that methylation is not involved in the formation of this morphological structure, or that the methylated cytosines in this region are prevented from interacting with the MeCPs. The latter situation seems implausible but could be tested with antibodies which detect MECP 2 (Lewis et al., 1992).

The finding that methylation, at least in this instance, is not a factor in apparent chromosome morphology implies that methylation does not play a major role in the level of metaphase chromosome packaging we are analysing. The evidence presently available on the role of DNA methylation suggests it is an indirect effector of gene regulation through its interactions with the MECP proteins. Mice deficient for DNA methyltransferase have been found to have a recessive embryonic lethal phenotype, while the corresponding cultured ES cells, which have a 70% drop in total genomic 5MeC, levels, seem to be normal in growth and morphology (Li et al., 1992). This suggests that methylation plays a role in the fine tuning of gene regulation, but is not of itself important in a structural role in the chromosome. Such an argument is strengthened by our finding that no affect on metaphase chromosome packaging was observed despite a dramatic increase in the methylation status of the insert in F1.1.

Methylation however has been postulated to be an important determinant of mammalian chromosome structure. Meehan et al (1993) suggested that MECP 1 may be involved in guiding nucleosome formation and that this chromatin may progress to a heterochromatic structure involving association with MECP 2. Such a mechanism would largely preclude MECP 1 from higher order packaging affects, although MECP 2 could be involved at this level. The cytidine analogue 5aza-cytidine causes chromosome undercondensation especially in pericentric heterochromatin (Mitchell, 1992; Schmid et al., 1984). It is proposed that methylation may be important structurally within satellite DNA regions (Mitchell, 1992), since the majority of DNA methylation is present here and MECP2 is detected by immunocytochemistry (Lewis

et al., 1992), and also throughout the genome since ⁵Aza-cytidine treatment has been shown to expand G-bands. However Haaf and Schmid, (1989), find that ⁵Aza-cytidine induced undercondensation of *Microtus* facultative and constitutive heterochromatin occurs independently of ⁵Aza-cytidine incorporation, suggesting that the affects observed are not mediated through interference of the normal methylation pattern of the chromatin, but through some other affect of this cytidine analogue.

Since methylation was not found to be the major cause of the novel metaphase structure of the F1.1 insert, we must look to another level of chromatin packaging for an explanation of this phenomenon.

8.4 Nucleosome repeat length

The lowest level of chromatin packaging is assembly into nucleosomes. It has been shown that nucleosome repeat length, the distance along the DNA between the start of subsequent nucleosomes, is species specific. *S. pombe* has a repeat of ~160bp (Polizzi and Clarke, 1991; Nakaseko *et al.*, 1988) while mouse cells have a repeat of ~185bp (Van Holde, 1989). It is unclear how these repeats are determined. Due to the conserved nature of the core histones it would seem that accessory factors aid the nucleosome placement in each species. This is suggested by the fact that most known cultured mouse cell lines have a repeat of approximately 185bp, while mouse tissues vary in repeat length from 195 bp (liver bulk chromatin) to 210bp (bone marrow) (Van Holde, 1988). Little is known about what these factors could be, although, H1 and H1 variants may well play a role. In this regard it is interesting to note that no H1 like activity has been detected in yeast chromatin.

Nucleosomes are the basic level of chromatin packaging upon which all other levels are superimposed. It therefore seems reason to speculate that differences at this level may affect higher packaging levels. Statistical analysis of chromatin fibres has shown that variation in the nucleosome repeat length is reflected in the diameter of the "30nm " fibre (Alegre and Subirana, 1989; Athey *et al.*, 1990). Variation in this fibre may be exaggerated during higher levels of packaging and eventually be reflected in the apparent morphology of the chromosome region concerned. A differing nucleosomal repeat in a region of ~7Mb, may allow an observable alteration in morphology as seen in the F1.1 and other hybrid cell lines with integrated yeast transgenomes.

I have shown, by hybridisation to partial micrococcal nuclease digests, that *S. pombe* chromatin has a repeat of ~160bp (Fig 6. 2, 3 b, c) agreeing with previous findings (Nakaseko *et al.*, 1988; Polizzi and Clarke, 1991). It is of interest that the SV2neo sequences present in *S. pombe Int-5* are packaged into a repeat typical of *S. pombe* bulk chromatin. The SV2neo sequences are not endogenous *S. pombe* sequences, they were introduced through transformation to produce the *Int 5* strain. In yeast it would appear that introduced DNA will adopt the endogenous chromatin structure, agreeing with previous studies of chromatin structure in yeasts after transformation. *S. pombe* centromeric core tm sequences transferred to *S. cerevisiae* are found to adopt the a nucleosome ladder characteristic of bulk *S. cerevisiae* chromatin rather than the nuclease protected structure adopted by these sequences in *S. pombe* (Polizzi and Clarke, 1991). Bernardi *et al.*, (1992) demonstrate that the *URA 3 S. cerevisiae* gene when transferred to *S. pombe* and the *ade6 S. pombe* gene when transferred to *S. cerevisiae* both fail to retain their previous chromatin structure. In the DNA transfectant cell line DC.1 I have found that the introduced *S. pombe* sequences have adopted the murine nucleosomal repeat (Fig 6: 10,11).

In F1.1 and F7.2 the *S. pombe* DNA has been introduced as chromatin, not naked DNA. This allows us to test if replication displaces and replaces nucleosomes without regard to the previous chromatin structure or whether a conservative pattern of nucleosome segregation allows the original pattern to be maintained. In the case of the *S. pombe* sequences introduced as chromatin by PEG fusion, to produce the hybrid cell lines F1.1 and F7.2, (Figures 6.9-13) it is found that a murine repeat is adopted by all of these sequences. These findings imply that after replication nucleosomes are formed upon sequences without regard to the previous nucleosome structure. Such a system of nucleosome establishment post replication has implications for the fate of nucleosomes during replication. A conservative nucleosome segregation pattern, without nucleosome dissolution and separation from the DNA, would not be consistent with these results. A system whereby the previous nucleosomes are dispersed during replication and replaced by a new nucleosome pattern after fork progression would agree with these results and also those of earlier microscopy studies (Sogo *et al.*, 1986). It would be difficult to differentiate this possibility from cis acting effects of surrounding mouse chromatin upon the *S. pombe* sequence region unless an extra chromosomal *S. pombe* sequence element free of murine sequences could be tested. Such an element may be located by in situ hybridisation to detect murine sequences co-localising with *S. pombe* sequence

extrachromosomal elements or by hybridisation of pulse field gels of undigested DNA dependent upon the size of the element, elements which are too large may not resolve into the gel away from the mouse chromosomes. The F7.2 cell line appears to have an extrachromosomal transgenome but this has not been tested for the presence of murine sequences. Mouse telomeric sequences may be expected (Blackburn, 1991; Morin, 1989) to be found on the ends of this element. However I have established in Chapter 6 that the yeast DNA in this extrachromosomal element adopts a nucleosome repeat typical of bulk mouse chromatin.

That nucleosome structure is dependent upon the chromatin proteins of the cell, and not upon the sequence may have implications for expression of transgenes. In instances where a conditional promoter is used, which may be dependent upon the presence of a positioned nucleosome and upon the presence of an appropriate factor to remove the nucleosome to allow expression, as is found in the *S. cerevisiae* *PHO5* gene (Almer et al., 1986; Straka and Horz, 1991) and with the MMTV promoter (Cordingley and Hager, 1988) efficient control may not be possible. When nucleosomes are formed by a differing repertoire of chromatin formation proteins specific nucleosome positioning may not be present upon the appropriate sequence. In the case of structural elements such as centromeres where distinct nucleosome patterns have been detected (Polizzi, and Clarke, 1991) and telomeres (Blackburn, 1991), a lack of specific factors required for formation of functional structures on these elements in transgenomes may limit their use as part of artificial shuttle chromosomes.

We assumed that the chromatin formed in yeast DNA regions is identical in protein composition to that of bulk mouse chromatin, and therefore that nucleosome structure is not related to the observed novel morphology in the F1.1 insert. It may be of interest to study the modification of the histones in the *S. pombe* insert chromatin of the hybrid cell lines. Transcribed and transcribable chromatin is enriched with acetylated histones especially H3 and H4 (Hebbes *et al.*, 1988; Ip *et al.*, 1988, Tazi and Bird 1991). The relationship between histone acetylation and transcription is also illustrated by the fact that position effect variegation may be suppressed by deacetylase inhibitors, such as propionate or butyrate (Mottus *et al.*, 1980). The suppression of position effect variegation suggests that heterochromatinising proteins are inhibited by the presence of acetylated histones (Turner, 1991). Therefore future study of the presence of acetylated core histones within the *S. pombe* insert may allow

a more detailed picture of the nucleosome structure of this domain than can be gained from micrococcal nuclease studies alone. This would be best achieved by the use of antibodies specific for acetylated histones in immunocytochemistry.

DNA methylation was not found to affect the morphology of the F1.1 insert and the nucleosome structure of the region was also found to be indistinguishable from that of the rest of the mouse genome. The distinction between the metaphase appearance of the yeast transgenome in F1.1 and the rest of the mouse chromosomes must therefore lie at levels of chromosome condensation above that of the 30nm fibre. The study of these levels of chromatin packaging are difficult to dissect experimentally and fraught with controversy.

8.5 Nucleoskeleton attachment.

It is widely believed that chromatin is organised into loop domains, tethered to a proteinaceous scaffold or skeleton. It is proposed that these loops are then condensed down to produce the metaphase chromosome (Paulson and Laemmli, 1978; Rattner and Lin, 1988). Dependent upon the nature of the higher order packaging processes, these loop domains may play a significant role in the morphology of metaphase chromosomes. A visible lack of final packaging is suggested to be responsible for the extended structure of centromeres (Rattner and Lin, 1987). The size of the loops within a region may affect the chromosome morphology in that larger loops may produce a wider chromosomal region, while an extended region of small loops may produce an apparent constriction at metaphase. A plausible explanation for the constriction in the *S. pombe* region of F1.1 could be that the yeast transgenome is not attached to the nucleoskeleton at all or is inappropriately attached so that average loop size is different from bulk mouse DNA. These possibilities were addressed by analysis of attachment of the *S. pombe* sequences to the nucleoskeleton of mouse cells.

Nucleoskeleton preparations, which involve the removal of unattached chromatin from the nucleus after nuclease digestion, are prepared under physiological conditions and have been found to contain 90% of the transcriptive and replicative capacity of the cell (Jackson *et al.*, 1986, 1987), demonstrating that appropriate interactions can still be formed. Attachments to the nucleoskeleton have been shown to be related to transcription (Dickinson *et al.*, 90b, Jackson and Cook, 1985). The fact that replication is also associated with attachment would suggest that most nuclear

activities are associated with the nucleoskeleton. Little evidence exists for a purely structural, as opposed to functional form of loop attachment, although such attachments might prove elusive with such assays. Attachments are believed to persist throughout the cell cycle and are thought to be present within the metaphase chromosome.

As opposed to this dynamic nature of chromatin attachment to a nuclear architecture, scaffold studies present a model involving permanent sites of chromatin attachment. Mirkovitch *et al.*, (1984) showed that LIS extraction of nuclei following nuclease digestion produced a constant profile of attached sequences. Analysis of these attachments have found co-localisation with ARS elements in yeast (Amati and Gasser, 1988) and also certain promoter and enhancer elements (Gasser and Laemmli, 1986). Scaffold attached sequences (SARs) have been proposed to be attached throughout the cell cycle and therefore to create closed chromatin domains which operate free from outside influences (Gasser *et al.*, 1989). SARs assayed for suppression of position effects are found not to behave as domain boundary sequences, and do not form a closed domain (Kellum and Schedl, 1992). The major protein found in scaffold preparations is topo II (Earnshaw *et al.*, 1985) and its *in vitro* consensus binding sequence was found to be present in the majority of SARs sequenced (Udvardy *et al.*, 1985). Later studies (Kas and Laemmli, 1992) demonstrate that *in vivo* topo II cleaves at a different consensus sequence from that *in vitro*.

The most significant problem with analysis of nuclear attachment is that of artefacts. Nucleoskeleton procedures carried out under physiological conditions would appear to largely limit these problems. Scaffold preparations by contrast use extreme extraction conditions. Littlewood *et al.*, (1987) have shown that scaffold like structures are formed after various cellular insults and Jackson *et al.*, (90) demonstrate that 80% of scaffold attachments are not present prior to LIS extraction. I have chosen to use the nucleoskeleton procedures for my studies of the F1.1 chromosome.

S. pombe sequences may have the potential for attachment to the mouse nucleoskeleton since there are similarities in aspects of the fission yeast's transcription mechanism to that of mammalian cells. It has been shown that *S. pombe* splicing machinery can correctly excise an intron from the SV40 small t antigen transcript (Kaufer *et al.*, 1985). Jones *et al.*, (1988) have also shown that the SV40 promoter in *S. pombe* initiates expression at an identical site to that used in mammalian cells, and

that an AP-1 like factor is present in *S. pombe* which binds at identical regions to that of human AP-1. The *S. pombe* U6 RNA gene is also found to be expressed in human cells and it is suggested that this is due to specific recognition of the gene by factors responsible for expression of the human U6 RNA gene (Kleinschmidt *et al.*, 1990). These findings point to a remarkably high level of conservation of transcription processes between *S. pombe* and mammalian cells. This conservation of functions is in contrast to the lack of conservation between fission and budding yeast, which show a high degree of divergence in recognition of transcription initiation sites (Russell, 1983).

The *S. pombe* genome (1.7×10^7 bp) is 0.5% of the size of the mouse genome (3×10^9 bp). This difference in genome size is presumably not reflected in the number of genes. This relative excess of DNA is reflected in the gene density of each species with *S. pombe* containing more closely packaged genes than is found in the mouse genome. The *S. pombe* region in F1.1, due to its high gene density could be attached to the nucleoskeleton more frequently per unit length than a similar length of murine sequence, due to the conservation of transcriptional machinery. The presence of closer attachments would lead to the formation of smaller chromatin loops within the *S. pombe* sequence insert, compared to average mouse loop sizes, leading to the presentation of a thinner chromosomal structure within this domain at metaphase.

I have shown that differential attachment of sequences to the nucleoskeleton can be demonstrated by analysing both retained and released fractions of these preparations. Previous studies (Dickinson *et al.*, 1990; Jackson and Cook, 1991) have only analysed the retained fractions of nucleoskeleton preparations. Analysis of released as well as retained fractions allows preferential segregation to either of these fractions to be detected rather than purely enrichment within the retained fraction versus total DNA. Figure 7.4 showed that while the murine sequences show differential retention to the nucleoskeleton (*Sl* not retained while others *Hox8*, *Pax6* and *Hprt* are), the *puc9* sequences from the vectors, which detect sequences integrated at the *Int-5 ura4* locus, show high levels of retention. These mouse sequences used are not ideal control sequences, since their expression characteristics are not known and they are cDNAs not genomic DNA and therefore likely to be overrepresented in the attached fraction. Ideally mouse controls for these experiments would be random cosmids but these are difficult to use due to the frequency of repeated sequences. These results suggested that sequences are differentially retained within the beads

dependent upon some intrinsic property. Since attached sequences were also represented within the released fraction at differing levels, attachment is either a dynamic process or attachments are broken during preparation. A certain amount of release could be attributed to proteolysis, but the fact that some sequences are represented purely within the released fraction demonstrates that an artefactual generalised retention of sequences does not occur. In Figure 7.5 I demonstrated that the SV₂neo sequences, located within the *S. pombe* sequence insert and known to be expressed, are retained upon the nucleoskeleton, while the sequences within the mouse genome to which this probe cross hybridises are found purely within the released fraction and are thus unattached. *S. pombe arsI* sequences are also found to be attached to the mouse nucleoskeleton while the *ade6* sequences are not. Therefore differential attachments of *S. pombe* sequences to the mouse nucleoskeleton can be detected on the F1.1 nucleoskeleton.

If the cause of the observed constriction within the F1.1 cell line, at the site of *S. pombe* sequence integration, is postulated to be the smaller average loop size than that found within the bulk of the mouse genome, then it is essential to investigate an extended region of the *S. pombe* insert. Analysis of an extended region is relatively simple for the *S. pombe* sequences, due to a lack of repeated elements in the yeast genome, while the prevalence of repeated elements within the murine genome would interfere with such an analysis. Blocking hybridisation of probes to the murine repeats by addition of cold DNA to the hybridisations would not be satisfactory for analysis of extended regions of the mouse genome as such sequences may be sites of nucleoskeleton attachment themselves.

Analysis of attachment of sequences hybridising to the Cura4-II cosmid to the F1.1 nucleoskeleton (fig 7.6) revealed that *S. pombe* sequences within this region were retained at differing levels upon the nucleoskeleton. The strongest attachment appeared to be to the *ura4* gene sequences themselves (Fig 7.7) with other regions showing differing lower levels of attachment. Certain bands appeared to be relatively enriched within the released fraction. Other bands appear to display a level of attachment intermediate to these levels. These findings suggest that a dynamic process of interactions is occurring within this *ura4* region with certain regions bound more strongly than others. The attachment of the *ura4* gene sequences is stronger than that observed for any of the murine sequences tested, suggesting strong and specific interactions within this region. The use of synchronised cell populations might allow

us to investigate whether any of these attachments are cell cycle dependent or if attachment is relatively stochastic throughout. The level of attachment would suggest that the loop domains within this region are smaller than 35kb since many attachments can be detected within this relatively short region. An average loop size of <35kb would imply that within the F1.1 cell line the *S. pombe* sequence domains average loop size is in the region of 1/2 to 1/3 that expected for the mouse genome. The retention of the *ura4* coding region within Cura4-II cosmid suggests that this may be related to the relatively high gene density of the *S. pombe* genome.

These findings suggest that smaller loops (<35kb) than those expected for the mouse genome 80-90kb may be formed within this *S. pombe* transgenome of F1.1. These smaller loops could lead to the apparent constriction observed within this region during metaphase condensation since this is the only difference we could detect between the yeast and mouse chromatin. It is however difficult at present to discern whether higher levels of packaging are occurring within this region and whether this constriction is due to packaging of smaller loops or a lack of packaging levels, as proposed for centromeric regions (Rattner and Lin, 1988). Other facets of higher packaging levels within the *S. pombe* region could be analysed by detection of topoII patterns within this region at metaphase. Immunocytochemistry of metaphase chromosomes has shown that the marker topo II (Heck and Earnshaw, 1986) presents a helical path of visualisation (Boy de la tour and Laemmli, 1989) along the length of the chromosome which is postulated to represent the path followed by chromatin after higher levels of packaging have occurred. Immunocytochemistry of the F1.1 chromosomes at metaphase with anti topo II antibodies may reveal whether such packaging occurs within the constricted *S. pombe* region.

8.6 Transcription.

Transcription, replication and other biochemical processes are proposed to occur in association with the nucleoskeleton (Cook, 1992). It is therefore of great interest that the SV2neo sequences, *ura4* gene and *ars1* sequences of the F1.1 yeast transgenome are attached to the murine nucleoskeleton. It is known that the SV2neo sequences are transcribed within the F1.1 cell line, since this region conveys G418 resistance to the cells. It has also been known that the SV40 promoter initiates at identical sites in mammalian and fission yeast cells, demonstrating a high level of conservation of the transcriptional machinery between both species (Jones *et al.*, 1988). It is therefore not surprising that this sequence is found attached to the

nucleoskeleton of the mouse cells. What is of more interest, and less easy to explain, is the attachment exhibited by the *ura4* and *ars1* sequences. The very high levels of retention of the *ura4* gene sequences suggests that transcription may be occurring from this gene, though poising for transcription may be sufficient for attachment. Conservation of transcription machinery between *S. pombe* and mammalian cells support this assumption (Jones *et al.* 1988; Kleinschmidt *et al.*, 1990), although opposed to this is the fact that the *ade6* sequences are not retained (Fig7.4). The *ars1* sequence and sites found within the Cura4-II cosmid that appear to be attached at lower levels are more difficult to justify as being retained due to transcription. It would be interesting to know which *S. pombe* sequences are transcribed within the F1.1 cell line. Nuclear run off experiments which extend any RNA molecules being produced within the nuclei were attempted but due to lack of time these experiments were unsuccessful and no useful data were obtained. If the RNA species are stable then either Northern analysis of RNA or PCR of the RNA will also detect the presence of specific RNA species. If the *S. pombe* sequences were transcribed it would be interesting to analyse whether the authentic *S. pombe* transcription initiation sites were utilised or, whether transcription occurred mainly by fortuitous interactions. It would also be interesting to test whether appropriate splicing of *S. pombe* genes, that contain introns, occurs in mouse cells.

8.7 Replication

Replication has also been demonstrated to be associated with the nucleoskeleton. Replication origins in mammalian cells have so far proved highly refractory to investigation, due to the complexity of the systems. Mammalian genomes are too large and complex for 2D gel analysis of replication products and also no genetic system for assaying potential origins of replication exists. The inability to easily detect specific origins of replication may also suggest that in mammalian cells origins of replication have a more complex structure (Mahbubani *et al.*, 1992) than those of lower eukaryotes such as *S. cerevisiae* (Maundrell *et al.*, 1988). In *S. pombe* origins of replication are also found to be more complex than those of *S. cerevisiae* (Beach *et al.*, 1980). Comparison of many *S. pombe ars* elements found them to be no smaller than 0.8kb in size with little internal homology (Maundrell *et al.*, 1988). In the case of the F1.1 cell line where a 10Mb insert of *S. pombe* sequence is located within a mouse chromosome replication of such a large transgenome may prove problematic to the mouse cell if it lacks appropriate origin sequences. That replication must be

initiated within this *S. pombe* region is apparent from its size, a 10Mb insert is too large to passively replicate from origins flanking the insert within the mouse sequences, also since the F7.2 transgenome is autonomous replication must initiate from within these *S. pombe* sequences. It may be that *S. pombe* origins themselves are utilised, but it seems more likely that fortuitous initiation events lead to replication of the insert. Preliminary cytogenetic studies of replication within this *S. pombe* region have been carried out (Perry et al., unpublished observations). By analysing replication through incorporation of BudR into the *S. pombe* region of the F1.1 chromosomes it was found that the *S. pombe* region replicated in the last third of S phase later than most mouse euchromatin, although earlier than certain murine heterochromatic regions. This suggests that replication of the *S. pombe* sequences is under some form of temporal control. Replication of specific *S. pombe* sequences could be measured relative to murine sequences by *in-situ* hybridisation (Selig *et al.*, 1992). This may allow for an accurate measure of the relative timing of replication within the *S. pombe* insert and the nature of initiation sites utilised.

8.8 Metaphase chromosome structure.

Aberrant metaphase morphology of discrete regions has also been found in regions of late or artificially delayed replication (Laird *et al.*, 1987). Chromosomal gaps have been found in *Microtus* chromosomes subjected to premature chromosome condensation, induced by cell fusion (Sperling *et al.*, 1974). After folate deprivation a fragile site in human genomes at Xq27 can be visualised as a chromosomal gap after the addition of caffeine which has the affect of shortening the G2 phase of the cell cycle (Ledbetter *et al.*, 1986). This undercondensation is thought to result because the region failed to be replicated properly during S phase and indeed aphidicolin (Glover *et al.*, 1984) and fluorodeoxyuridine (Hagle *et al.*, 1972) and inhibitors of folate metabolism also lead to the presentation of chromosomal gaps at fragile sites and sites of late DNA replication. The gaps formed at these sites, when studied by electron microscopy appear as thin wisps of chromatin linking two normally packaged chromatin regions (Harrison *et al.*, 1983). It would therefore appear that these gaps are not directly related in structure to those observed with the F1.1 cell line (Fig 1.3). Where we observe clear morphology within the *S. pombe* domain of F1.1, only its diameter is reduced. The apparent morphology of the F1.1 constriction is that of a purely thinner chromosomal region rather than that of a region devoid of higher levels of packaging as would be expected from replication inhibited region.

As Paulson and Laemmli (1978) demonstrated with dehistonised chromosomes, the structure of metaphase chromosomes is derived from the packaging of chromatin loops. Later studies Rattner and Lin (1988) and Boy de la tour and Laemmli (1989) suggest that these loops are aligned radially along a central structure and then packaged in a helical manner, presumably with additional condensation through coiling, during entry into metaphase to produce the observed metaphase chromosome structure. These chromatin loops are presumably related to the nuclear cage attached loops observed by Brazell and Cook (1975), Mc Cready *et al.*, (1980), and also the nucleoskeleton attachments of Jackson and Cook (1985, 1986).

We suggest that smaller loops are present within the *S. pombe* domain of F1.1, than in the mouse genome. This thinner chromosome structure and an apparent constriction being caused after packaging of the smaller radial loops through the subsequent steps of chromosome condensation (Fig 1.3).

8.9 FINAL CONCLUSIONS.

I have shown that the fusions between the yeast *S. pombe* and mouse cells undertaken by Allshire *et al.* , (1987) produced transgenomes which exhibited varying degrees of rearrangement and retention of syntenic and non syntenic markers. A significant proportion of the *S. pombe* genome has entered the F1.1 cell line possibly up to 10Mb, including material from all three *S. pombe* chromosomes. Very few *S. pombe* non-chromosome III markers appear to have entered the autonomous element of the F7.2 cell line by contrast. A distinct difference is observed in the amount of *S. pombe* centromeric sequences which have entered these cell lines. In F1.1 virtually all the *S. pombe* centromeric sequences are found while in F7.2 only a small amount of highly rearranged material appears to have become established. It is unclear at this time whether this is functionally significant. In F7.2 certain *S. pombe* telomere sequences are detected, although their rearranged pattern may suggest that they are no longer telomerically located. The lack of *S. pombe* telomeric sequences from the insert of the F1.1 cell line may be related to the coligation of non syntenic material found within this insert.

I have shown that increasing levels of methylation of the *S. pombe* insert in F1.1 do not relate to its metaphase structure, suggesting that methylation may play no significant role at the level of metaphase chromosome structure we are examining.

I have also found that the nucleosome structure of the transgenomes studied (F1.1, F7.2 and DC.1) is that typical of the host cell. That adoption of the host nucleosome ladder seems to be a general principle associated with transgenomes. This suggests that chromatin structure of transgenomes is determined purely by the host chromatin proteins and not by the underlying sequence, and that any previous nucleosome structure is erased at replication. My findings would not be compatible with models of replication requiring conservative segregation of nucleosomes. That the *S. pombe* sequences adopt the mouse nucleosome repeat would suggest that at the level of "30nm fibre" packaging of the yeast DNA in F1.1 is normal and not related to the novel structure observed, any abnormality must therefore lie at higher levels of chromosome packaging.

Analysis of *S. pombe* sequence attachment to the nucleoskeleton of the mouse cell reveals that frequent and strong attachments may be formed with this *S. pombe* region. A high frequency of attachment to the nucleoskeleton of the mouse cell implies that smaller loops are formed within the *S. pombe* domain than within the murine genome. Therefore there is reason to believe that the constriction observed within the F1.1 cell line is caused by the formation of smaller loops within *S. pombe* domain and this is consistent with morphology observed with E.M. (Fig1.3) assuming that the attachments detected with these experiments are retained in the metaphase chromosome.

These hybrid cell lines appear to be a powerful tool for the analysis of chromosome structure in higher eukaryotes. By marrying the simple genomic structure of yeast with the potential for visualisation provided by mammalian chromosomes, it appears to be possible to investigate elements which may affect chromosome structure. These cells lines provide a resource for analysis of not just chromosome structure but also of replication.

REFERENCES

- ABRAHAM, J., WATSMYTH, K.A., STRATHEN, J.N., KLAR, A.J.S. and HICKS, J.B. (1984).
Replication of mating type information in yeast. *J. Mol. Biol.* 176, 307-331.
- ADACHI, Y., LUKE, M. and LAEMMLI, U.K. (1991). Chromosome assembly in vitro: Topoisomerase
II is required for condensation. *Cell* 64, 137-48.
- ALEGRE, C. and SUBIRANA, J. A. (1989). The diameter of chromatin fibers depends on linker length
Chromosoma 98, 77-80.
- ALLAN, J., COWLING, G. R., HARBORNE, N., CATTINI, P., CRAIGIE, R. and GOULD, H. (1981).
Regulation of the higher- order structure of chromatin by histones H1 and H5. *J. Cell. Biol.* 90,
279- 288.
- ALLSHIRE, R.C. and BOSTOCK, C. J. (1986). Structure of bovine papillomavirus type 1 DNA in a
transformed cell line . *J. Mol. Biol.* 188, 1-13.
- ALLSHIRE, R.C. CRANSTON, G., GOSDEN., J.R., MAULE, J.C., HASTIE, N.D. and FANTES, P.A.
(1987). A fission yeast chromosome can replicate autonomously in mouse cells. *Cell* 50, 391-403.
- ALLSHIRE, R.C., GOSDEN, J., CROSS, S.H., CRANSTON, G., ROUT, D., SUGAWARA, N.,
SZOSTAK. J.N., FANTES. P.A. and HASTIE. N.D. (1988). Telomeric repeat from *T. thermophila*
cross-hybridises with human telomeres. *Nature* 332, 656-59.
- ALLSHIRE, R.C., DEMPSTER, M. and HASTIE, N.D. (1989). Human telomeres contain at least three
types of G-rich repeat distributed non-randomly. *Nucl. Acid Res.* 17, 4611-27.
- ALMER, A. and HORZ, W. (1986). Nuclease hypersensitivity regions with adjacent positioned
nucleosomes mark the gene boundaries of the PHO 5/ PHO 3 locus in yeast. *EMBO. J.* 5, 2681-
2687.
- ALMER, A., RUDOLPH, H., HINNEN, A. and HORZ, W. (1986). Removal of positioned nucleosomes
from yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA
elements. *EMBO. J.* 5, 2689- 2696.
- ALMER, H., CLASSON, B., SAMUELSON, B. and KVARNSTROM, I. (1991). Synthesis of a
phosphonomethyl analogue of 3' deoxy-3-fluorothymidine. *Acta. Chem. Scand.* 45, 766-7.

- ALMOUZNI, G., CLARK, D. J., MECHALI, M. and WOLFE, A.P.(1990). Chromatin assembly on replcating DNA in vitro. Nucl. Acid Res. 18, 5767- 5774.
- AMATI, B. and GASSER, S. M. (1988). Chromosomal ARS and CEN elements bind specifically to the yeast nuclear scaffold. Cell 54, 967- 978.
- AMATI, B. and GASSER, S. M. (1990). Drosophila scaffold attached regions bind nuclear scaffolds and can function as ARS elements in both budding and fission yeast. Mol. Cell. Biol. 10, 5442- 5454.
- AMATI, B., PICK, L., LAROCHE, T. and GASSER, S.M. (1990). Nuclear scaffold attachment stimulates, but is not essential for ARS activity in Saccharomyces cerevisiae: analysis of the Drosophila ftz SAR. EMBO. J. 9, 4007- 4016.
- ANTEQUERA, F., MACLEOD, D. and BIRD, A.P. (1989). Specific protection of methylated CpG in mammalian nuclei. Cell 58, 509- 517.
- ANTEQUERA, F., BOYES, J. and BIRD, A. (1990). High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. Cell. 62, 503-514.
- APARICIO, O.M., BILLINGTON, B. L. and GOTTSCHLING, D.E. (1991). Modifiers of position effect are shared between telomeric and silent mating type loci in S. cerevisiae. Cell 66, 1279-1287.
- ARIEL, M., MC CARREY, J. and CEDAR, H. (1991). Methylation patterns of testis specific genes. Proc. Natl. Acad. Sci. USA88, 2317- 2321.
- ATHEY, B.A., SMITH, M. F., RANKERT, D.A., WILLIAMS, S.P. and LANGMORE, J.P. (1990). The diameters of frozen hydrated chromatin fibers increase with DNA linker length: Evidence in support of variable diameter models for chromatin. J. Cell. Biol. 111, 795- 806.
- BEACH, D., PIPER, M., SHALL, S. (1980). Isolation of chromosomal origins of replication in yeast. Nature 284, 1185-7.
- BENEZRA, R., CANTOR, C. R. and AXEL, R. (1986). Nucleosomes are phased along the mouse β -major globin gene in erythroid and non erythroid cells. Cell 44, 697- 704.
- BERNARDI, F., KOLLER, T. and THORNE, F. (1991). The ade6 gene of fission yeast S. pombe has the same chromatin structure in the chromatin and in plasmids. Yeast 7, 547-58.
- BERNARDI, F., ZATCHY, M. and THOMA, F. (1992). Species specific protein DNA interactions may

- determine the chromatin units of genes in *S. cerevisiae* and *S. pombe*. *EMBO J.* 11, 1177-85.
- BERNAT, R.L., BORISY, G. G., ROTHFIELD, N.F. and EARNSHAW, W.C. (1990). Injection of anticentromere antibodies in interphase disrupts events required for chromosome movement at mitosis. *J. Cell. Biol.* 111, 1519- 1533.
- BERNAT, R.L., DELANNOY, M. R., ROTHFIELD, M.R. and EARNSHAW, W.C. (1991). Disruption of centromere assembly during interphase inhibits kinetochore morphogenesis and function in mitosis. *Cell.* 66, 1229- 1238.
- BESTOR, T., LAUDANO, A., MATTALIANO, R. and INGRAM, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. *J. Mol. Biol.* 203, 971- 983.
- BICKMORE, W.A., MAULE, J. M. , VAN HEYNINGEN, V. and PORTEOUS, D.J. (1989). Long range structure of H-ras selected transgenomes. *Som. Cell. Mol. Gen.* 15, 229- 235.
- BICKMORE, W.A. and SUMNER, A. T. (1989). Mammalian chromosome banding - an expression of genome organisation. *Trends in Genet.* 5, 144- 148.
- BIESSMANN, H. and MASON, J. M. (1988). Progressive loss of DNA sequencers from terminal chromosome deficiencies in *Drosophila melanogaster*. *EMBO J.* 7, 1081- 1086.
- BIRD, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nuc. Acid Res.* 8, 1499- 504.
- BIRD, A.P. and TAGGART, M. (1980). Variable pattern of total and rDNA methylation in animals. *Nuc. Acid. Res.* 8, 1485-97.
- BIRD, A.P. (1987). CpG islands as gene markers in the vertebrate nucleus. *Trends In Genet.* 3, 342- 347.
- BIRD, A.P. (1992). The essentials of DNA methylation. *Cell* 70, 5-8.
- BLACKBURN, E.H. (1991). Structure and function of telomeres. *Nature* 350, 569-73.
- BLOOM, K.S. and CARBON, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* 29, 305- 317.
- BODE, J. and MAASS, K. (1988). Chromatin domain surrounding the human interferon β -gene as defined by scaffold attached regions. *Biochemistry* 27, 4706- 4711.

- BOGGS, S.S., GREG, R.G., BORENSTEIN, N. and SMITHIES, O. (1986). Efficient transformation and frequent single site, single copy insertion of DNA can be obtained in mouse erythroleukemia cells transformed by electroporation. *Exp. Hematol.* 14, 988-94
- BONNE-ANDREAS, C., WONG, M. L. and ALBERTS, B.M. (1990). In vitro replication through nucleosomes without histone displacement. *Nature* 343, 719- 726.
- BOSTOCK, C.J. and SUMNER, A.T. (1978). In *The eukaryotic chromosome*, Amsterdam: North Holland.
- BOUTON, A.H. and SMITH, M.M. (1986). Fine structure requirements for autonomous replication of *Saccharomyces cerevisiae* plasmids. *Mol. Cell Biol.* 6, 2354-63.
- BOY DE LA TOUR, E. and LAEMMLI, U. K. (1988). The metaphase scaffold is helically folded: sister chromatids have predominantly opposite helical handedness. *Cell* 55, 937- 944.
- BOYES, J. and BIRD, A. P. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* 64, 1123- 1134.
- BOYES, J. and BIRD, A. P. (1992). Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl- CpG binding protein. *EMBO J.* 11, 327- 333.
- BRUN, C., DANG, Q. and MIASSOD, R. (1990). Studies of an 800 kilobase DNA stretch of the *Drosophila X* chromosome: comapping of a subclass of scaffold attached regions with sequences able to replicate autonomously in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 10, 5455- 5463.
- BURCH, J.B.E. and WEINTRAUB, H. (1983). Temporal order of chromatin structural changes associated with activation of the major chicken vitellogenin gene. *Cell* 33, 65-76.
- BURNETT, T.S. and GALLIMORE, P.H. (1985). Introduction of cloned human papillomavirus 1a DNA into rat fibroblasts: integration, de novo methylation and absence of cellular morphological transformation. *J. Gen. Virol.* 66, 1063-72.
- BUSSLINGER, M., HURST, J. and FLAVELL, R.A. (1983). DNA methylation and the regulation of globin gene expression. *Cell* 34, 197- 206.
- BUTLER, P.J.G. (1984). A defined structure of the 30nm chromatin fibre which accommodates different nucleosomal repeat lengths. *EMBO J* 3, 2599- 2604.

- CARDENAS, M.E., LAROCHE, T. and GASSER, S.M. (1990). The composition and morphology of yeast nuclear scaffolds. *J. Cell. Sci.* 96, 439- 450.
- CEDAR, H. and RAZIN, A. (1990). DNA methylation and development. *Biochemica et biophysica acta* 1049, 1-8.
- CHAMPLIN, D. T., FRASCH, M., SAUMWEBER, H. and LIS, J.T. (1991). Characterisation of a *Drosophila* protein associated with boundaries of transcriptionally active chromatin. *Genes and Development* 5, 1611- 1621.
- CHIKASHIGE, T., KINOSHITA, N., NAKASEKO, N., MATSUMOTO, T., MURAKAMI, M. and NIWA, O., YANAGIDA, M. (1989). Composite motifs and repeated symmetry in *S. pombe* centromeres: direct analysis by integration of Not I restriction sites. *Cell* 57, 739- 751.
- CHURCH, G.M. and GILBERT, N. (1984) Genomic sequencing. *Proc Natl. Acad. Sci. -USA* 81, 1991.
- CLARK, D.J. and FELSENFELD, G. (1992). Anucleosome core is transferred out of the path of a transcribing polymerase. *Cell* 71, 11-22.
- CLARKE, L. and BAUM, M. P. (1990). Functional analysis of a centromere from fission yeast: A role for centromere specific repeated DNA sequences. *Mol. Cell. Biol.* 10, 1863- 1872.
- COOK, P.R. and BRAZELL, I. A. (1975). Supercoils in human DNA. *J. Cell. Sci* 19, 261- 279.
- COOK, P.R. and BRAZELL, I. A. (1976a). Characterisation of nuclear structures containing superhelical DNA. *J. Cell. Sci.* 22, 303- 324.
- COOK, P.R. and BRAZELL, I. A. (1976b). Conformational constraints in nuclear DNA. *J. Cell. Sci* 22, 287- 302.
- COOK, P.R. (1989). The nucleoskeleton and the topology of transcription. *Eur. J. Biochem.* 185, 487- 501.
- COOKE, C.A., BERNAT, R. L. and EARNSHAW, W.C. (1990). CENP-B: A major human centromere protein located beneath the kinetochore. *J. Cell. Biol.* 110, 1475- 1488.
- COOKE, H.J., BROWN, W. and RAPPOLD, G.A. (1985). Hypervariable telomeric sequences from human sex chromosomes are pseudoautosomal. *Nature* 317, 687-692.

- COOPER, D.N., TAGGART, M.H. and BIRD, A.P. (1983). Unmethylated domains in vertebrate DNA. Nucl. Acid Res. 11, 647-58.
- CORDINGLEY, M. and HAGER, G.L. (1988). Binding of multiple factors to the MMTV promoter in crude and fractionated nuclear extracts. Nuc Acid Res. 16, 609-28.
- CROSS, S.H., ALLSHIRE, R. C., MC KAY, S.J., MC GILL, N.I. and COOKE, H.J. . (1989). Cloning of human telomeres by complementation in yeast. Nature. 338, 771-4.
- CSORDAS, A. (1990). On the biological role of histone acetylation. Biochem. J. 265, 23-38.
- CUSICK, M.E., LEE, K-S, DE PAMPHALIS, M.L. and WASSERMAN, P.M. (1983). Structure of chromatin at deoxyribonucleic acid replication forks: nuclease hypersensitivity results from both prenucleosomal deoxyribonucleic acid and on an immature chromatin structure. Biochemistry 22, 3873- 3884.
- CUSICK, M.E., DE PAMPHALIS, M.L. and WASSERMAN, P.M (1984). Dispersive segregation of nucleosomes during replication of Simian Virus 40 chromosomes. J. Mol. Biol. 178, 249- 271.
- DE-LANGE, T., SHUIE, L., MYERS, R., COX, D.R., NAYLOR, S.L. , KILLERY, A.M. and VARMUS, H.E. (1990). Structure and variability of human chromosome ends. Mol. Cell Biol. 10, 518- 527.
- DE LANGE, T. (1992). Human telomeres are attached to the nuclear matrix. EMBO J. 11, 717- 724.
- DICKINSON, P., COOK, P. R. and JACKSON, D.A (1990). Active RNA polymerase I is fixed within the nucleus of HeLa cells. EMBO J. 9, 2207- 2214.
- DIJKWEL, P.A., MULLENDERS, L. H. F. and WANKA, F. (1979). Analysis of the attachment of replicating DNA to a nuclear matrix in mammalian interphase nuclei. Nucl. Acid Res. 6, 219- 230.
- DIJKWEL, P.A. and HAMLIN, J. L. (1988). Matrix attachment regions are positioned near replication initiation sites, genes and an interamplicon junction in amplified Dihydrofolate reductase domain of chinese hamster ovary cells. Mol. Cell. Biol. 8, 5398- 5409.
- DILWORTH, S.M., BLACK, S. J. and LASKEY, R.J. (1987). Two complexes that contain histones are required for nucleosome assembly in vitro: role of nucleoplasmin and NI in Xenopus egg extracts. Cell 51, 1009- 1018.

- DORN, R., HEYMANN, S., LINDIGKEIT, R. and REUTER, G. (1986). Suppressor mutation of position affect variegation in *Drosophila melanogaster* affecting chromatin properties. *Chromosoma* 93, 398-403.
- DURRIN, L.K., MANN, R. K., KAYNE, P.S. and GRUNSTEIN, M. (1991). Yeast histone H4 N terminal sequence is required for promoter activation in vivo. *Cell* 65, 1023- 1031.
- EARNSHAW, W.C., HALLIGAN, B., COOKE, C.A., HECK, M.M.S. and LIU, L.F. (1985). Topoisomerase II is a structural component of mitotic chromosome scaffolds. *J. Cell. Biol.* 100, 1706- 1715.
- EARNSHAW, W.C. and ROTHFIELD, N. (1985). Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma*. 91, 313- 321.
- ELGIN, S.C. (1988). The formation and functions of DNase I hypersensitive sites in the process of gene activation. *J. Biol. Chem.* 263, 19259-62.
- ELICEIRI, B., LABELLA, T., HAGINE, K., SRIVASTOVA, A., PILLA, G., PALIMIERI, G. and D'URSO, M. (1991). Stable integration and expression in mouse cells of yeast artificial chromosomes harboring human genes. *Proc. Natl. Acad. Sci. USA* 88, 2179-83.
- FAN, J-B., CHIKASHIGE, Y., SMITH, C.L., NIWA, O., YANAGIDA, M. and CANTOR, C.R. (1988). Construction of a Not I map of the fission yeast *Schizosaccharomyces pombe* genome. *Nuc. Acid Res.* 17, 2801- 2818.
- FARR, C., FANTES, J., GOODFELLOW, P. and COOKE, H.J. (1991). Functional reintroduction of human telomeres into mammalian cells. *Proc. Natl. Acad. Sci. USA* 88, 7006-10
- FEINBERG, A.P., VOGELSTEIN, B. (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137, 266-7.
- FINCH, J.T. and KLUG, A. (1976). Solenoid model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA* 73, 1897- 1901.
- FOTEDAR, R. and ROBERTS, J. M. (1989). Multistep pathway for replication- dependent nucleosome assembly. *Proc. Natl. Acad. Sci. USA* 86, 6459- 6463.
- FRANK, D., KESHET, I., SHANI, M., LEVINE, A., RAZIN, A. and CEDAR, H. (1991). Demethylation of CpG islands in embryonic cells. *Nature* 351, 239- 241.

- GARDINER-GARDEN, M. and FROMMER, M. (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.* 196, 261- 282.
- GASSER, S.M. and LAEMMLI, U. K. (1986). Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell* 46, 521- 530.
- GASSER, S.M., AMATI, B.B., CARDENAS, M.E. and HOFFMAN, J.F.X. (1989). Studies on scaffold attachment sites and their relation to genome function. *Ann. Rev. Cytology.* 57-96
- GNIRKE, A., BARNES, T. S., PATTERSON, D., SCHILD, D., FEATHERSTONE, T. and OLSON, M.V. (1991). Cloning and in vivo expression of the human GART gene using yeast artificial chromosomes. *EMBO. J.* 10, 1629- 1634.
- GNIRKE, A. and HUXLEY, C. (1991). Transfer of the human HPRT and GART genes from yeast to mammalian cells by microinjection of YAC DNA. *Somat. Cell. Mol. Gen.* 17, 573- 580.
- GOTTESFELD, J.M. (1980). Organisation of 5S genes in chromatin of *Xenopus Laevis*. *Nucl. Acid Res.* 8, 905- 922.
- GRANT, S.G. and CHAPMAN, V. M. (1988). Mechanisms of X- chromosome regulation. *Ann. Rev. Genet.* 22, 199- 233.
- GREIDER, C.W. and BLACKBURN, E. H. . (1989). A telomeric sequence in the RNA of tetrahymena telomerase required for telomere repeat synthesis. *Nature* 337, 331- 337.
- GREIDER, C.W. (1990). Telomeres, telomerase and senescence. *Bioessays* 12, 363- 369.
- GRUSS, C., GUTIERREZ, C., BURHUMS, W.C., DE PAMPHALIS, M.L. , KOLLER, T. and SOGO, J.M. (1990). Nucleosome assembly in mammalian cell extracts before and after DNA replication. *EMBO J.* 9, 2911-2922.
- GURLEY, L.R., D'ANNA, J. A., BARHAM, S.S., DEAVEN, L.L. and TOBEY, R.A. (1978). Histone phosphorylation and chromatin structure during mitosis in chinese hamster cells. *Eur. J. Biochem.* 84, 1- 15.
- HAAF, T. and SCHMID, M. (1989). 5- Azadeoxycytidine induced undercondensation in the giant X chromosomes of *Microtus agrestis*. *Chromosoma* 98, 93- 98.

- HAGLE, E. (1972). Strukturverändernde Wirkung von Fudr auf polytene Chromosomen und Beziehungen zwischen Replikationsdauer und Bruchhäufigkeit von Querscheiben. *Chromosoma* 33, 297-301.
- HARLEY, C.B., FUTCHER, A. B. and GREIDER, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature* 345, 458-460.
- HARRIS, H. (1970). *Cell fusion*. Clarendon press. Oxford.
- HARRISON, L.J., JACK, E.M., ALLEN, T.D. and HARRIS, R. (1983). The fragile X: A scanning electron microscope study. *J. Med. Genet.* 20, 280-5.
- HASTIE, N.D., DEMPSTER, M., DUNLOP, M.G., THOMPSON, D.K. and ALLSHIRE, R.C. (1990). Telomere reduction in colorectal carcinoma and in association with ageing in humans. *Nature* 346, 866-868.
- HEBBES, T.R., THORNE, A. W. and CRANE-ROBINSON, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* 7, 1395- 1402.
- HECK, M.M. and EARNSHAW, W.C. (1986). Topoisomerase II- a specific marker for cell proliferation. *J. Cell Biol.* 103, 2569-81.
- HILL, R.E., FAVOR, J., HOGAN, B.L., TON, C.C., SAUNDERS, G.F., HANSON, I.M., PROSSER, J., JORDAN, T. and VAN HEYNINGEN, V. (1991). Mouse small eye results from mutations in a paired like homeobox containing gene. *Nature* 354, 522-5.
- HIRANO, T., KINOSHITA, N., MORIKAWA, K. and YANAGIDA, M. (1990). Snap helix with knob and hole: essential repeats in *S. pombe* nuclear protein nuc 2⁺. *Cell* 60, 319- 328.
- HOFFMANN, J. F-X., LAROCHE, T., BRAND, A.H. and GASSER, S.M. (1989). Rap-I factor is necessary for loop formation in vitro at the silent mating type locus HML. *Cell* 57, 725- 737.
- HOFFMANN, J. F-X. and GASSER, S. M. (1991). Identification and purification of a protein that binds the yeast ARS consensus sequence. *Cell* 64, 951- 960.
- HOLLIDAY, R. (1987). The inheritance of epigenetic defects. *Science* 238, 163- 170.
- HOZAK, P., HASSAN, A.B., JACKSON, D.A. and COOK, P.R. (1993) Visualisation of replication factories attached to a nucleoskeleton. *Cell* 73, 361-74.

- HSIEH, C-L. and LIEBER, M. R. (1992). CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. *EMBO J.* 11, 315- 325.
- HUMPHRIES, S.E., YOUNG, D. and CARROLL, D. (1979). Chromatin structure of the 5S ribonucleic acid genes of *Xenopus Laevis*. *Biochemistry* 18, 3223- 3231.
- HUXLEY, C. and GNIRKE, A. (1991). Transfer of yeast artificial chromosomes from yeast to mammalian cells. *Bioessays* 13, 545550.
- HUXLEY, C., HAGINO, Y., SCHLESSINGER, D. and OLSEN, M.V. (1991). The human HPRT gene on a yeast artificial chromosome is functional when transferred to mouse cells by cell fusion. *Genomics* 9, 742- 750.
- INGLIS, J.D. and HILL, R. E. (1991). The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain. *EMBO J.* 10, 255- 61.
- IP, Y.T., JACKSON, V., MEIER, J. and CHALKLEY, R. (1988). The separation of transcriptionally engaged genes. *J. Biol. Chem.* 263, 14044-52.
- IZARRALDE, E., MIRKOVITCH, J. and LAEMMLI, U.K. (1988). Interaction of DNA with nuclear scaffolds in vivo. *J. Mol. Biol.* 200, 111- 125.
- JACKSON, D.A., MC CREADY, S. J. and COOK, P.R. (1981). RNA is synthesised at the nuclear cage. *Nature* 292, 552- 555.
- JACKSON, D.A., MC CREADY, S. J. and COOK, P.R. (1984). Replication and transcription depend on attachment of DNA to the nuclear cage. *J. Cell. Sci. Suppl.* 1, 59- 74.
- JACKSON, D.A. and COOK, P. R. (1985). Transcription occurs at a nucleoskeleton. *EMBO J.* 4, 919- 925.
- JACKSON, D.A. and COOK, P. R. (1986). A cell cycle dependent DNA polymerase activity that replicates intact DNA in chromatin. *J. Mol. Biol.* 192, 65- 76.
- JACKSON, D. A. and COOK, P. R. (1986b). Replication occurs at the nucleoskeleton. *EMBO J.* 5, 1403- 1410.
- JACKSON, D.A. and COOK, P. R. (1988). Visualisation of a filamentous nucleoskeleton with a 23nm axial repeat. *EMBO J.* 7, 3667-3677.

- JACKSON, D.A., DICKINSON, P. AND COOK, P.R. (1990). The size of chromatin loops in HeLa cells. *EMBO J.* 9, 567-71.
- JACKSON, V. (1990). In vivo studies on the dynamics of histone DNA interaction: evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. *Biochemistry* 29, 719- 731.
- JAKOBOVITS, A., MOORE, A. L., GREEN, L. L., VERGARA, G. J., MAYNARD-CURRIE, C.E., AUSTIN, H.A., KLAPHOLZ, S. (1993). Germ line transmission and expression of a human-derived yeast artificial chromosome. *Nature* 362, 255-258.
- JARMAN, A.P. and HIGGS, D.R. (1988). Nuclear scaffold attachment sites in the human globin gene complexes. *EMBO J.* 7, 3337-44.
- JOBLANKA, E., GOITEIN, R., SPERLING, K., CEDAR, H. and MOREAS, M. (1986). 5-aza-C-induced changes in the time of replication of the X-chromosomes of *Microtus agrestis* are followed by non random reversion to a late pattern of replication. *Chromosoma* 95, 81-8.
- JONES, R.H., MORENO, S., NURSE, P. and JONES, N.C. (1988). Expression of the SV40 promoter in fission yeast identification and characterisation of an AP-1 like factor. *Cell* 53, 659- 667.
- KAMAKAKA, R.T. and THOMAS, J. O. (1990). Chromatin structure of transcriptionally competent and repressed genes. *EMBO J.* 9, 3997- 4006.
- KAMPER, J., MEINHARDT, F., GUNGE, N. and ESSER, K. (1989). In vivo construction of linear vectors based on killer plasmids from *Kluyveromyces lactis*; selection of a nuclear gene results in attachment of telomeres. *Mol. Cell Biol.* 9, 3931-7.
- KAS, E. and LAEMMLI, U. K. (1992). In vivo topoisomerase II of the drosophila histone and satellite III repeats: DNA sequence and structural characteristics. *EMBO J.* 11, 705- 716.
- KAUFER, N.F., SIMANIS, V. and NURSE, P. (1985). Fission yeast *Schizosaccharomyces pombe* correctly excises a mammalian RNA transcript intervening sequence. *Nature* 318, 78-80.
- KAYNE, P.S., KIM, U-J., HAN, M., MULLEN, J.R., YOSHIZAKI, F. and GRUNSTEIN, M. (1988). Extremely conserved histone H4 N terminus is dispensible for growth but essential for repressing the silent mating loci in yeast. *Cell* 55, 27- 39.
- KEARSEY, S. (1984). Structural requirements for the function of a yeast chromosomal replicator. *Cell*.

- KELLUM, R. and SCHEDL, P. (1991). A position effect assay for boundaries of higher order chromosomal domains. *Cell* 64, 941-950.
- KELLUM, R. and SCHEDL, P. (1992). A group of scs elements function as domain boundaries in an enhancer- blocking assay. *Mol. Cell. Biol.* 12, 2424-2431.
- KIPLING, D. and COOKE, H.J. (1990). Hypervariable ultra-long telomeres in mice. *Nature* 347, 400-2.
- KLEINSCHMIDT, A.M. and MARTINSON, H.G. (1984). Role of histone tyrosines in nucleosome formation and histone-histone interactions. *J. Biol. Chem.* 259, 497-503.
- KLEINSCHMIDT, A.M., PEDERSON, T., TANI, T. and OHSHIMA, Y. (1990). An intron containing *Schizosaccharomyces pombe* u6 RNA gene can be transcribed by human RNA polymerase III. *J. Mol. Biol.* 211, 7-9.
- KLEMPNAUR, K-H., FANNING, E., OTTO, B. and KNIPPERS, R. (1980). Maturation of newly replicated chromatin of Simian virus 40 and its host cell. *J. Mol. Biol.* 136, 359- 374.
- KOCH, M.H.J., VEGA, M. C., SAYERS, Z. and MICHON, A.M. (1987). The superstructure of chromatin and its condensation mechanism III. *Eur. Biophys. J* 14, 307- 319.
- KOLSTO, A-B., KOLLIAS, G., GIGURERE, V., ISOBE, K-I., PRYDZ, H. and GROSVELD, F. (1986). The maintenance of methylation free islands in transgenic mice. *Nucl. Acid Res.* 14, 9667- 9677.
- KUHN, R.M., CLARKE, L. and CARBON, J. (1991). Clustered tRNA genes in *Schizosaccharomyces pombe* centromeric DNA sequence repeats. *Proc. Natl. Acad. Sci. USA* 88, 1306- 1310.
- KUCHERLAPATI, R. and SKOULTCHI, A. (1984). Introduction of purified genes into animal cells. *CRC. Crit. Rev. Biochem.* 16, 349-79.
- LAIRD, C., JAFFE, E., KARPEN, G., LAMB, M. and NELSON, R. (1987). Fragile sites in human chromosomes as regions of late replicating DNA. *Trends in Genet.* 3, 274- 281.
- LANGNER, K-L., WEYER, U. and DOERFLER, W. (1986). Trans effect of the E1 region of adenovirus on the expression of a prokaryotic gene in mammalian cells: resistance to 5' CCGG 3' methylation. *Proc. Natl. Acad. Sci. USA* 83, 1598- 1602.

- LARSON, D.D., SPANGLER, E.A. and BLACKBURN, E.H. (1987). Dynamics of telomere length variation in *Tetrahymena thermophila*. *Cell* 50, 477- 483.
- LEBKOWSKI, J.S. and LAEMMLI, U.K. (1982). Evidence for two levels of DNA folding in histone depleted hela interphase nuclei. *J. Mol. Biol.* 156, 309-24.
- LEDBETTER, D.H., AIRHART, S.D. and NUSSBAUM, R.L. (1986). Caffeine enhances fragile (X) expression in somatic cell hybrids. *Am. J. Med. Genet.* 23, 445-55.
- LEONHARDT, H., PAGE, A.W. WEIER, H., and BESTOR, T.H. (1992) A targetting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71, 865-73.
- LEVIS, R.W. (1989). Viable deletions of a telomere from a *Drosophila* chromosome. *Cell* 58, 791- 801.
- LEVY- WILSON, B. and FORTIER, C. (1989). The limits of the Dnase I sensitive domain of the human apolipoprotein B gene coincide with the locations of chromosomal anchorage loops and the defined 5' and 3' boundaries of the gene. *J. Biol. Chem* 264, 21196- 21204.
- LEWIS, J.D., MEEHAN, R.R., HENZEL, W.J., MAURER FOGY, I., JEPPESEN, P., KLEIN, F. and BIRD, A.P. (1992). Purification, sequence and cellular localisation of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905-14.
- LI, E., BESTOR, T.H. and JAENISCH, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915-26.
- LITTLEWOOD, T.D., HANCOCK, D. C. and EVAN, G. I. (1987). Characterisation of a heat shock-induced complex in the nuclei of cells. *J. Cell. Sci* 88, 65- 72.
- LOCK, L.F., MELTON, D. W., CASKEY, T. and MARTIN, G.R. (1986). Methylation of the mouse hprt gene differs on the active and inactive x chromosomes. *Mol. Cell. Biol.* 6, 914- 924.
- LOCK, L.F., TAKAGI, N. and MARTIN, G.R. (1987). Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell* 48, 39- 46.
- LASSON, R. and LACROUTE, F. (1983). Plasmids carrying the yeast OMP decarboxylase structural and regulatory genes: transcription in a foreign environment. *Cell* 32, 371-7.
- LOWARY, P.T. and WIDOM, J. (1989). Higher order structure of *Saccharomyces cerevisiae* chromatin. *Proc. Natl. Acad. Sci. USA* 86, 8266- 8270.

- LUNDBLAD, V. and BLACKBURN, E. H. (1990). RNA dependent polymerase motifs in Est-1: tentative identification of a protein component of an essential yeast telomerase. *Cell* 60, 529-530.
- MAHBUBANI, T., ELDER, J. K. and BLOW, J.J. (1992). DNA replication initiates at multiple sites on plasmid DNA in *Xenopus* egg extracts. *Nucl. Acid Res.* 20, 1457- 62.
- MANUELIDIS, L., LANGER-SAFER, P. R. and WARD, D.C. . (1982). High resolution mapping of satellite DNA using biotin labeled probes. *J. Cell. Biol.* 95, 619- 625.
- MARIN, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesises TTAGGG repeats. *Cell* 59, 521-29.
- MATSUMOTO, H., FUKUI, T. K., NIWA, O., SUGAWARA, N., SZOSTAK, J.W. and YANAGIDA, M. (1987). Identification of healed terminal DNA fragments in linear minichromosomes of *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* 7, 4424- 4430.
- MASUMOTO, H., MASOKATA, H., HURO, Y., NOZAKI, N. and OKAZAKI, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeres satellite. *J. Cell. Biol.* 109, 1963- 1973.
- MATTHEWS, H.R. and MORTON BRADBURY, E. (1978). The role of H1 histone phosphorylation in the cell cycle. *Ex. Cell. Res.* 111, 343-351.
- MAUNDRELL, K., HUTCHISON, A. and SHALL, S. (1988). Sequence analysis of ARS elements in fission yeast. *EMBO J.* 7, 2203- 2209.
- MC CREADY, S.J., GOODWIN, J., MASON, D.W., BRAZELL, I.A., and COOK, P.R. (1980). DNA is replicated at the nuclear cage. *J. Cell. Sci* 46, 365- 386.
- MC GHEE, J.D., NICKOL, J. M, FELSENFELD, G. and RAU, D.C. (1983). Higher order structure of chromatin: orientation of nucleosomes within the 30nm chromatin solenoid is independent of species and spacer length. *Cell* 33, 831- 841.
- MEEHAN, R.R., LEWIS, J. D., MC KAY, S., KLEINER, E.L. and BIRD, A.P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* 58, 499-- 507.
- MEEHAN, R., LEWIS, J. D., CROSS, S., NAN, X., JEPPESEN, P. and BIRD, A.P. (1993). Transcriptional repression by methylation of CpG. *J. Cell. Sci.* 16 (Suppl), 9-14.

- MILLS, A.D., BLOW, J.J., WHITE, J.G., AMOS, W. B., WILCOCK, D. and LASKEY, R.A. (1989). Replication occurs at discrete foci spaced throughout nuclei replicating in vitro. *J. Cell Sci.* 94, 471-478.
- MIRKOVITCH, J., MIRAULT, M-E. and LAEMMLI, U.K. (1984). Organisation of the higher order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* 39, 223- 232.
- MIRKOVITCH, J. (1986). Genes and loops in 320000 base pairs of *Drosophila melanogaster* chromosome. *J. Mol. Biol.* 190, 255258.
- MIRKOVITCH, J., GASSER, S. M. and LAEMMLI, U.K. (1988). Scaffold attachment of DNA loops in metaphase chromosomes. *J. Mol. Biol.* 200, 101- 109.
- MITCHELL, A.R. (1992). Hypomethylation of human heterochromatin detected by restriction enzyme nick translation. *Exp. Cell.Res.* 202, 203-206.
- MONAGHAN, A.P., DAVIDSON, D.R., SIME, C., GRAHAM, E., BALDOCK, R., BHATTACHARYA, S.S. and HILL, R.E. (1991). The msh- like homeobox genes define domains in the developing vertebrate eye. *Development* 112, 1053- 1061.
- MONK, M. (1986). Methylation and the X chromosome. *Bioessays* 4, 204- 208.
- MONK, M., BOUBELICK, M., and LEHNERT, S. (1987). Temporal and regional changes in DNA methylation in the embryonic extraembryonic and germ cell lineages during mouse embryo development. *Dev.* 99, 371- 382.
- MORIN, G. (1989). The human telomere terminal transferase is a ribonucleoprotein that synthesises TTAGGG repeats. *Cell* 59, 521- 529.
- MOTTUS, R., REEVES, R. and GRIGLIATTI, T.A. (1980). Butyrate suppression of position effect variegation in *Drosophila melanogaster*. *Mol. Gen. Genet* 178, 465- 469.
- MURAKAMI, S., MATSUMOTO, T., NIWA, O. and YANAGIDA, M. (1991). Structure of the fission yeast centromere cen3: direct analysis of the reiterated inverted region. *Chromosoma.* 101, 214- 221.
- NACHEVA, G.A., GUSHIN, D. Y., PREOBRAZHENSKAYA, O.V., KARPOV, V.L. and EBRAALIDSE, K.K., MIRZABEKOV, A.D. (1989). Change in the pattern of histone binding to DNA upon transcriptional activation. *Cell* 58, 27- 36.

- NAKAMURA, H., MORITA, T. and SATO, C. (1986). Structural organisation of replicon domains during DNA synthesis phase in mammalian nucleus. *Exp. Cell. Res* 165, 291-297.
- NAKASEKO, Y., ADACHI, Y., FUNAHASHI, S, NIWA, O. and YANAGIDA, M (1986). Chromosome walking shows that a highly homologous repetitive sequence present in all the centromere regions of fission yeast. *EMBO J* 5, 1011- 1021.
- NAKASEKO, Y., KIOSHITO, N. and YANAGIDA, M. (1987). A novel sequence common to the centromere regions of *S. pombe* chromosomes. *NUC. ACID RES.* 15, 4705- 15.
- NAKAYASU, H. and BEREZNEY, R (1989). Mapping replication sites in eukaryotic cell nucleus. *J. Cell Biol.* 108, 1-11.
- OHLENBUSCH, H. H., OLIVERA, B.M., TUAN, D. and DAVIDSON, N. (1967). Selective dissociation of histones from calf thymus nucleoprotein. *J. Mol. Biol.* 25, 299-315.
- OLINS, A.L. and OLINS, D.E. (1974). Spherical chromatin units. *Science* 183, 330-2.
- PACHNIS, V., PENNY, L., ROTHSTEIN, R. and CONSTANTINI, F. (1990). Transfer of a yeast artificial chromosome carrying human DNA from *Saccharomyces cerevisiae* into mammalian cells. *Proc. Natl. Acad. Sci. USA* 87, 5109- 5113.
- PALMER, D.K., O'DAY, K., WERNER, M.H. , ANDREWS, B.S. and MARGOLIS, R.L. (1987). A 17KDa centromere protein (CENP-A) copurifies with nucleosomes core particles and with histones. *J. Cell. Biol.* 104, 805- 814.
- PAROUSH, Z. , KESHET, I., YISRAELI, J. and CEDAR, H. (1990). Dynamics of demethylation and activation of the activation of a-actin gene in myoblasts. *Cell* 63, 1229- 1237.
- PAULSON, J.R. and LAEMMLI, U.K. (1977) The Structure of histone depleted metaphase chromosomes. *Cell* 12, 817-28.
- PAULSON, J.R. (1989). Scaffold morphology in histone depleted HeLa metaphase chromosomes. *Chromosoma* 97, 289- 295.
- PAVAN, W., HEITER, P. and REEVES, R.H. (1990). Modification and transfer into an embryonal carcinoma cell line of a 360 Kilobase human derived yeast artificial chromosome. *Mol. Cell. Biol.* 10, 4163- 4169.

- PHI-VAN, L., VON KRIES, J.P, OSTERTAG, W. and STRATLING, W.H. (1990). The chicken lysosome 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. *Mol. Cell. Biol.* 10, 2302-2307.
- PLUTA, A.F., SAITOH, N., GOLDBERG, I. and EARNSHAW, W.C. (1992). Identification of a subdomain of CENP-B that is necessary and sufficient for localisation to the human centromere. *J. Cell. Biol.* 116, 1081- 1093.
- POLIZZI, C. and CLARKE, L. (1991). The chromatin structure of centromeres from fission yeast: differentiation of the central core that correlates with function. *J. Cell. Biol.* 112, 191- 201.
- POLLACK, Y., STEIN, R., RAZIN, A. and CEDAR, H. (1980). Methylation of foreign DNA sequences in eukaryotic cells. *Proc. Natl. Acad. Sci. USA* 77, 6463-6467.
- PORTEOUS, D.J. . (1987). Chromosome mediated gene transfer: a functional assay for complex loci and an aid to human genome mapping. *Trends in Genet.* 3, 177- 182.
- PRUNELL, A. and KORNBERG, R. D (1978). Relation of nucleosomes to DNA sequences. *C.S.H.S.Q.B.* XLII, 103- 108.
- QUERY, C.C., BENTLEY, R.C. and KEENE, J.D. (1989). A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* 57, 89-101.
- RATTNER, J.B. and LIN, C. C. (1985). Radial loops and helical coils coexist in metaphase chromosomes. *Cell* 42, 291- 296.
- RATTNER, J.B. and LIN, C. C. (1987). The higher order structure of the centromere. *Genome* 29, 588- 593.
- REIGER, R., MICHAELIS, A. and GREEN, M.M. (1976). *Glossary of genetics and cytogenetics.* Springer-verlag. New York.
- REUTER, G., GIARRE, M., FARAH, J., GAUZE, J., SPEIRER, A. and SPEIRER, P. (1991). Dependence of position effect variegation in *Drosophila* on dose of a gene encoding an unusual zinc finger protein. *Nature* 344, 219-23.
- ROSE, D., THOMAS, W. and HOLM, C. (1990). Segregation of recombined chromosomes in meiosis I requires DNA topoisomerase II. *Cell* 60, 1009-17.

- RUSSELL, P.R. (1983). Evolutionary divergence of the mRNA transcription initiation mechanism in yeast. *Nature* 301, 167169.
- SAMBROOK, J., FRITSCH, E.F. and MANIATIS, T. (1989). *Molecular cloning: a laboratory manual*. Cold Spring Harbor Press.
- SEIDMAN, M.M. and LEVINE, A. J. and WEINTRAUB, H. (1979). The assymetric segregation of parental nucleosomes during chromosome replication. *Cell* 18, 439- 449.
- SELIG, S., OKAMURA, K., WARD, S. and CEDAR. H. (1992). Delineation of DNA replication time zones by fluorecence in situ hybridisation. *EMBO J.* 11, 1217-25.
- SHAMPAY, J., SZOSTAK, J. W. and BLACKBURN, E.H.. (1984). DNA sequences of telomeres maintained in yeast. *Nature* 310, 154-157.
- SHIPPEN-LENTZ, D. and BLACKBURN, E. H. (1990). Functional evidence for an RNA template in telomerase. *Science* 247, 546- 552.
- SIMMONS, T., LIPMAN, M. and HODGE, L.D. (1978). Uptake and early fate of metaphase chromosomes ingested by the WI-L2 human lymphoid cell line. *Somat. Cell Genet.* 4, 55-76.
- SMITH. M.F., ATHEY. B. D, WILLIAMS. S.P. and LANGMORE. J,P (1990). Radial density distribution of chromatin: evidence that chromatin fibers have solid centers. *J. Cell. Biol* 110, 245-254.
- SMITH, S. and STILLMAN, B. (1989). Purification and characterisation of CAF-I, a human cell factor required during DNA replication in vitro. *Cell* 58, 15- 25.
- SMITH, S. and STILLMAN, B. (1991). Stepwise assembly of chromatin during DNA replication in vitro. *EMBO J.* 10, 971- 980.
- SPERLING, K. and RAO, P. (1974). Mammalian cell fusion. replication behaviour of heterochromatin as observed by premature chromosome condensation. *Chromosoma* 45, 121-31.
- SOGO, J.M., STAHL, H., KOLLER, T. and KNIPPERS, R. (1986). Structure of replicating Simian Virus 40 minichromosomes. *J. Mol. Biol.* 189, 189- 204.
- SOUTHERN, E. (1975). Detection of specific sequences among DNA fragments seperated by gel electrophoresis. *J. Mol. Biol.* 98, 503.

- SOUTHERN, P. and BERG, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 promoter. *J. Mol. Appl. Genetics* 1, 327- 341.
- STANLEY BURNETT, T. and GALLIMORE, P. H. . (1985). Introduction of cloned human papillomavirus 1a DNA into rat fibroblasts: integration, de novo methylation and absence of cellular morphological transformation. *J. Gen. Virol.* 66, 1063- 1072.
- STARLING, J.A., MAULE, J. HASTIE, N.D and ALLSHIRE, R.C. (1990) Extensive telomere arrays in mouse are hypervariable. *Nuc. Acid Res.* 18, 6881-88.
- STEIF, A., WINTER, D. M., STRATLING, W.H. and SIPPEL, A.E. (1989). A nuclear DNA attachment element mediates elevated and position independent gene activity. *Nature* 341, 343- 345.
- STEIN, A. and MITCHELL, M. (1988). Generation of different nucleosome spacing periodicities in vitro. *J. Mol. Biol* 203, 1029- 1043.
- STRAKA, C. and HORZ.,W. (1991). A functional role for nucleosomes in the repression of a yeast promoter. *EMBO J.* 10, 361-368.
- STRATLING, W.H. and DOLLE, A. (1986). Chromatin structure of the chicken lysosyme gene domain as determined by chromatin fractionation and micrococcal nuclease digestion. *Biochemistry* 25, 495- 502.
- SUBIRANA, J.A., MUNOZ-GUERRA, S., RADERMACHER, M. and FRANK. J (1983). Three dimensional reconstruction of chromatin fibers. *J. Biol. Str. Dyn.* 1, 705- 714.
- SUBIRANA, J.A., MUNOZ-GUERRA, S., AYMAMI, J., RADERMACHER, M. and FRANK, J. (1985). The layered organisation of nucleosomes in 30 nm chromatin fibers. *Chromosoma* 91, 377- 390.
- SUMNER, A.T. (1990). In *Chromosome banding*, Unwin Hyman.
- SYBENGA, J. (1972). *General Cytogenetics*, Amsterdam: North Holland.
- SZOSTAK, J.W. and BLACKBURN, E. H. (1982). Cloning yeast telomeres on linear plasmids. *Cell* 29, 154- 157.
- SZYF, M., SCHIMMER, B.P. AND SEIDMAN, J.G. (1989). Nucleotide-sequence-specific *de novo* methylation in a somatic murine cell line. *Proc. Natl. Acad. Sci. USA* 86, 6853-57.

- TAKAHASHI, K., MURAKAMI, S., CHIKASHIGE, Y. , NIWA, O. and YANAGIDA, M. (1991). A large number of tRNA genes are symmetrically located in fission yeast centromeres. *J. Mol. Biol.* 218, 1317.
- TAKAHASHI, K., MURAKAMI, S., CHIKASHIGE, Y. , FUNABIKI, H., NIWA, O. and YANAGIDA, M. (1992). A low copy number central sequence with strict symmetry and unusual chromatin structure in fission yeast chromosomes. *Mol. Biol. Cell* 3, 819-835.
- TAZI, J. and BIRD, A. (1990). Alternative chromatin structure at CpG islands. *Cell.* 60, 909- 920.
- THOMA, F., KOLLER, Th. and KLUG, A. (1979). Involvement of histone H1 in the organisation of the nucleosome and of the salt dependent superstructures of chromatin. *J. Cell. Biol* 83, 403- 427.
- THOMA, F. and ZATCHEJ, M. (1988). Chromatin folding modulates nucleosome positioning in yeast minichromosomes. *Cell* 55, 945- 953.
- THORNE, A.W., KMICIEK, D., MITCHELSON, K., SAUTIERE, P. and CRANE ROBINSON, C. (1990). Patterns of histone acetylation. *Eur. J. Biochem.* 193, 701-13.
- TODA, T., NAKASEKO, Y., NIWA, O. and YANAGIDA, M. (1984). Mapping of rRNA genes by integration of hybrid plasmids in *Schizosaccharomyces pombe*. *Current Genetics.* 8, 93-97.
- TON.,C.C.T., HIRVONEN, H., MIWA, H., WEIL, M.M., MONAGHAN, P., JORDAN, T., VAN HEYNIGAN, V., HASTIE, N.D., MEIJERS-HEIJBOER, H., DRECHSLER, M., ROYER-POKORA, B., COLLINS, F., SWAROOP, A., STRONG, L.C. and SAUNDERS, G.F.(1991). Positional cloning and characterisation of a paired box and homeobox containing gene from the *Anaridia* region. *Cell* 67, 1059- 1074.
- TURNER, B. and FELLOWS, G. (1989). Specific antibodies reveal ordered and cell cycle related use of histone H4 acetylation sites in mammalian cells. *Eur. J. Biochem* 179, 134- 139.
- TURNER, B.M. (1991). Histone acetylation and control of gene expression. *J. Cell. Sci.* 99, 13-20.
- UDVARDY, A., MAINE, E. and SCHEDL, P. (1985). The 87A7 Chromomere: identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J. Mol. Biol.* 185, 341- 358.
- UEMERA, T., OHKURA, H., ADACHI, Y., MORINO, K., SHIOZAKI, K. and YANAGIA. M. (1987). DNA topoisomerase II is required for condensation and separation of mitotic chromosomes. *Cell* 50,

- UZAWA, S. and YANAGIDA, M. (1992). Visualisation of centromeric and nucleolar DNA in fission yeast by fluorescence in situ hybridisation. *J. Cell. Sci* 101, 267- 275.
- VAN HOLDE, K.E. (1989). *Chromatin*, New York: Springer Verlag.
- VASSILEV, L.T., BURHAMS, W. C. and DE PAMPHALIS, M.L. (1990). Mapping an origin of DNA replication at a single copy locus in exponentially proliferating mammalian cells. *Mol. Cell. Biol.* 10, 4865- 4689.
- VERKERK, A.J.M., PIERETTI, M., SUTCLIFFE, J.S., FU, Y., KUHL, D.P.A., PIZUTTI, A., REINER, O., REICHARDS, S., VICTORIA, M.F., ZHANG, F., EUSSEN, B.E., OMMEN, G.B., BLONDEL, L., RIGGINS, G.J., CHASTAIN, J.L., KUNST, C.B., GALSAARD, H., CASKEY, C., NELSON, D.L., OOSTRA, B.A., WARREN, S.T. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting variation in fragile X syndrome. *Cell* 65, 905-14.
- VON KRIES, J.P., BUHRMESTER, H. and STRATLING, W.H. (1991). A matrix/ scaffold attachment region binding protein: identification, purification and mode of binding. *Cell.* 64, 123- 135.
- WARD, M., SCOTT, R. J., DAVEY, M.R., CLOTHIER, R.H., COCKING, E.C. and BALLS, M. (1986). Transfer of antibiotic resistance genes between yeast and mammalian cells under conditions favoring cell fusion. *Somat. Cell Mol. Genet.* 12, 101- 109.
- WARD, W.S., PARTIN, A. W. and COFFEY, D.S. (1989). DNA loop domains in mammalian spermatozoa. *Chromosoma* 98, 153- 159.
- WEISBROD, S. and WEINTRAUB, H. (1981). Isolation of actively transcribed nucleosomes using immobilised HMG 14 and 17 and an analysis of α - globin chromatin. *Cell* 23, 391- 400.
- WIDOM, J. and KLUG, A. (1985). Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. *Cell* 43, 207- 213.
- WIGLER, M., LEVY, D. and PERUCHO, M. (1981). The somatic replication of DNA methylation. *Cell* 24, 33-40.
- WILKIE, A.O., LAMB, J., HARRIS, P.C., FINNEY, R.D. and HIGGS, D.R. (1990). A truncated human chromosome 16 associated with alpha thalassemia is stabilised by addition of telomeric repeats.

- b- globin gene expression. Proc. Natl. Acad. Sci. USA 85, 4638- 4642.
- YISRAELI, J., ADELSTEIN, R. S., MELLOUL, D., NUDEL, U., YAFFE, D. and CEDAR, H. (1986). Muscle- specific activation of a methylated chimeric actin gene. Cell 46, 409- 416.
- YU, G.L., BRADLEY, J.D., ATTARDI, L.D. and BLACKBURN, E.H. (1990). In vivo alteration of telomere sequences and senescence caused by mutated telomerase RNA's. Nature 344, 126- 132.
- YUNIS, J.J. (1981). Mid prophase chromosomes: The attainment of 2000 bands. Human Genetics 56 293- 8.
- ZHANG, X-Y., FITTLER, F. and HORZ, W. (1983). Eight different highly specific nucleosome phases on a- satellite DNA in the african green monkey. Nucl. Acid Res. 11, 4287- 4306.
- ZHANG, X-Y. and HORZ, W. (1984). Nucleosomes are positioned on mouse satellite DNA in multiple highly specific frames that are corelated with a diverged subrepeat of nine base pairs. J. Mol. Biol. 176, 105- 120.