The Amplification Of The Chinese Hamster Ovary Glutamine Synthetase Gene And Its Flanking DNA

A thesis submitted for the Degree of
Doctor of Philosophy at the
University of Glasgow

by

Bruce Hayward

Institute of Genetics
Church Street
Glasgow

October 1987

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

Acknowledgements

I thank my supervisor, Richard Wilson, for his help, crazy ideas and criticisms- experimental and literary- in the course of this work. My thanks too to the GS group for their encouragement and help, Pete Sanders and Jos Anderson especially. Thanks also to the 6th floor, Mary especially, for their friendly "nagging" and for their enzymes. This research would not have been possible without the Media Ladies who provided sterile labware and bacterial culture media. Thanks, and more, to Lesley for bearing with me.

My studies were financed by the Science and Engineering Research Council.

Contents

Abbreviations

Chemicals

DNA - deoxyribonucleic acid

RNA - ribonucleic acid

EDTA - ethylenediaminetetra-acetic acid (disodium salt)

EtBr - ethidium bromide

EtOH - ethanol

SDS - sodium dodecylsulphate

DTT - dithiothreitol

ATP - adenosine triphosphate

Tris - tris (hydroxymethyl) amino ethane

X-gal - 5-bromo-4chloro-3indolyl-Bgalactoside

Measurements

bp - base pairs

kbp - kilo base pairs (10^3 bp)

cpm - counts per minute

mA - milliamps (10^{-3} amps)

°C - degrees Centigrade

g - centrifugal force equal to gravitational acceleration

g - gramme

mg - milligramme (10^{-3} g)

ug - microgramme (10^{-6} g)

ng - nanogramme (10^{-9} g)

l - litre

ml - millilitre (10^{-3} l)

ul - microlitre (10^{-6} l)

M - molar (moles per litre)

mM - millimolar

pH - acidity (negative $\log_{10}$ [molar concentration H^+ ions])

V - volts

m - metre

cm - centimetre (10^{-2} m)

mm - millimetre (10^{-3} m)

uCi - microcuries (10^{-3} Curie)

Summary

By multistep selection of cells for increasing resistance to methionine sulphoximine, a specific inhibitor of glutamine synthetase, an amplified cell line was produced with approximately 1000 copies of the glutamine synthetase (GS) gene. The genomic GS gene was isolated from the DNA of this cell line and the cloned DNA used to probe the regions flanking the GS gene. By conventional agarose gel electrophoresis it was shown that the DNA containing the GS gene was homogeneous for at least 40 kbp in all copies of the GS amplification unit and that the hypomethylation patterns at the 5' of the GS gene were also homogeneous. An estimate of the size of the amplified unit suggested that it was 60 kbp in size. Pulse field gel electrophoresis was used to analysis the relationship of one unit to another and several possible models are discussed. PFGE suggested that a single unit was 75 kbp in size and analysis of other related cell lines suggested that this size is a reduction from the size of the unit in a less amplified line.

Karyotypic and *in situ* hybridization studies of the parental and highly amplified cell line revealed the presence of several rearranged marker chromosomes one of which was shown to be the site of the amplified DNA.

CHAPTER 1

Introduction

1.1 Introduction

The amplification of DNA sequences (ie. an increase in their copy number with respect to the rest of the genome) is a universal phenomenon. Its occurrence has been recognised for many years in prokaryotes, playing an important role in adapting to a changing environment without the loss of genetic function and with the potential to be readily reversible. In eukaryotic systems the importance of gene amplification has grown in the past few years from that of a developmental curiosity and cultured cell oddity to now touch areas as diverse as the tumour cell, biotechnology and the maintenance of the genome itself.

DNA amplification in eukaryotic systems is the subject of this chapter because of its relevance to the rest of this work and because DNA amplification in prokaryotes apparently proceeds via different mechanisms and therefore reveals little of DNA amplification in eukaryotes.

1.2 Programmed Amplification

1.2.1 Chorion genes in **Drosophila**

The amplification of the chorion gene clusters in the ovarian follicle cells of Drosophila is a developmentally programmed response to the necessity of synthesising large quantities of eggshell proteins within a small period of time (Spradling, 1981; Spradling & Mahowald, 1980). A different method of ensuring sufficient quantities of chorion proteins at the proper time is taken by the silk moth (Jones & Kafatos, 1980) which possesses multiple copies of these genes rather than amplifying single copy genes.

Drosophila possess six abundant chorion proteins the genes for which are grouped on two chromosomes. Chorion genes s38 and s36 are found on the X chromosome at position 7F1.2 together with 4 minor chorion genes while the genes for s18, s19, s16 and s15 are on chromosome 3 at 66D11.15. During oogenesis the chorion gene clusters of both pairs of chromosomes are amplified, the chromosome 3 clusters approximately 60-fold and the X chromosome clusters approximately 15-fold. The amplified region extends

approximately 50 kbp to either side of a central point with the sequences near the centre being five times more abundant than sequences near the end (Spradling, 1981). This amplification results from repeated cycles of DNA replication from a single origin as evidenced by complex multiforked structures originating from a single point seen with the electron microscope (Osheim & Miller, 1983), and also when an inversion in the central portion of the cluster changes the sequences that are amplified (Spradling & Mahowald, 1981). Orr-Weaver & Spradling (1986) cloned part of the chromosome 3 cluster and identified a 510 bp fragment which regulates the transcription of the s18 gene. This fragment is essential for the amplification of the chromosome 3 chorion gene cluster and its sequence contains partial homologies to a series of perfect and imperfect AATAC repeats, the presence of some of which are required for amplification of the X-linked chorion gene cluster (Spradling *et al*, 1987).

The detail of the expression patterns of these genes was investigated by Parks & Spradling (1987) who showed that the chromosome 3 cluster was expressed after the X chromosome cluster and that the expression of the minor X chromosome chorion genes was spatially regulated.

1.2.2 Other programmed amplifications

Similar developmentally programmed amplifications have been observed in the larvae of *Rhynchosciara* (Glover *et al*, 1982) and *Trichosia pubescens* (Amabis & Amabis, 1984) flies.

In *Xenopus laevis* germ cells a 10-40 fold amplification of rDNA genes occurs (Kalt & Gall, 1974) which is transient in male germ cells but becomes a 1000-fold amplification (Gall, 1968) via a rolling circle mechanism in oogonia (Rochaix *et al*, 1974) to produce extrachromosomal rings containig between 15 (Hourcade *et al*, 1973) and 100 (Thiebaud, 1979) copies of rDNA genes.

In the protozoa *Tetrahymena* the transcriptionally active macronucleus contains 1000 rDNA genes derived from the single rDNA gene present in the quiescent micronucleus. The amplified rDNA genes are present as linear extrachromosomal palindromes (Yao *et al*, 1978) which possess sequences capable of functioning

Figure 1.1

Drug	Target Protein	Reference
Methionine sulphoximine	Glutamine synthetase	Sanders & Wilson (1984)
Methotrexate	Dihydrofolate reductase (two alleles)	Kaufman et al (1981) Lewis et al (1982)
PALA	CAD	Wahl et al (1982)
Cadmium	Metallothionein I	Gick & McCarty (1982)
Pyrazofurin	UMP Synthetase	Suttle (1983)
Adenosine, alanosine and uridine	Adenosine deaminase	Yeung et al (1983)
"HAT" medium	Mutant HPRT	Zownir et al (1984)
Compactin	3-hydroxy-3-methylglutaryl CoA reductase	Luskey et al (1983)
Difluoromethylornithine	Ornithine decarboxylase	Leinonen et al (1987)
Hydroxyurea	M2 subunit of ribonucleotide reductase	Wright et al (1987)
Ouabain	Na, K-ATPase -subunit	Emanuel et al (1986)
Multidrug	170 kDa P-glycoprotein	de Bruijn et al (1986)
Albizzin	Asparagine synthetase	Andrulis et al (1983)

Facing page 3

as yeast telomeres (Szostak & Blackburn, 1982) and yeast replication origins (Kiss et al, 1981). Interestingly linear, 42 kbp palindromic extrachromosomal elements are responsible for resistance, by amplification of ADH4, to antimycin A in yeast (Walton et al, 1986)

1.3 Amplification by recombination

Gene amplification is a common occurrence in prokaryotes occurring spontaneously and at high frequency (10^{-1} to 10^{-5}). It has been reviewed by Anderson & Roth (1977). The process requires a system for homologous recombination (Straus & Straus, 1976) and the end points of the amplified region are frequently ribosomal RNA genes or short regions (12 bp) of homology (Edlund & Normark, 1981) which suggested that unequal exchange between homologous regions generated the amplification. Whoriskey et al (1987) showed that short repeated sequences of up to 15 bp, sometimes involving broken homology, occurred at the junctions of their amplified sequences. Unequal exchange has also been implicated in the state of flux of the tandemly repeated rDNA cluster of yeast where up to 7 rDNA units can be duplicated together (Klein & Petes, 1981; Petes, 1980).

1.4 Amplification in cultured cells

Overproduction of a protein which is the target of a specific inhibitor is a common mechanism of resistance to that inhibitor and an increase in numbers of the target gene is a common way of achieving that overproduction. Resistance to inhibitors can be elevated eventually several thousand-fold and the target protein can constitute up to 30% of total protein (Sanders & Wilson, 1984). Fig 1.1 provides a list of genes which have been amplified in the course of selection for resistance to specific inhibitors. With the exception of resistance to multiple drugs (eg. vincristine, adriamycin, colchicine, actinomycin D, cytochalasin B) where the amplification is believed to involve a transport protein the amplifications involve the protein targeted by the inhibitor in question.

Resistance to inhibitors is achieved by step-wise selection

of cells capable of growth in gradually increasing concentrations of the inhibitor and the resistant mutants constitute between 10^{-4} and 10^{-7} of the cell population. In the case of PALA resistance, for example, initial selection was for resistance to PALA levels of 15 to 60 uM (Giulotto et al, 1986) and after five rounds of selection resistance to 25 mM PALA was achieved (Ardeshir et al, 1983).

In general the frequency of resistance caused by gene amplification is so high that resistance due to regulatory alterations or structural mutation is not observed but examples involving ribonucleotide reductase (Wright et al, 1987) and ornithine decarboxylase (McConlogue et al, 1986) have been documented. While it is true to say that the level of gene amplification is approximately reflected by the increase in protein level it is not equivalent to the possession of multiple copies of the original gene. de Bruijn et al (1986) show that in one of their cell lines the levels of P-glycoprotein were greater than expected from the amplification of the DNA sequences alone while another fortuitously amplified gene was proportionately underexpressed. The expression of the target gene can be reduced relative to the copy number as shown by some PALA-resistant cell lines where enzyme activity was less than half that expected from the CAD gene copy number (Zieg et al, 1983).

Mayo & Palmiter (1982) showed that glucocorticoid regulation of the metallothionein I gene was lost but heavy metal regulation was retained after it was amplified by selection for cadmium resistance. Luskey et al (1983) showed that the amplification of the gene for 3-hydroxy-3-methylglutaryl coenzyme A reductase did not affect its regulation by low density lipoproteins.

Gene amplification can also occur to cause a phenotypic reversion in circumstances where the gene involved has become less active; Zownir et al (1984) showed phenotypic reversion was due to the amplification of the temperature sensitive HPRT gene rather than the appearance of another mutant form and spontaneous revertants of the $HPRT^-$ cell line of Fenwick et al (1984) were also shown to be amplificants of the HPRT gene. Similarly reversion of a viral thymidine kinase gene with an inactive

promoter was via gene amplification (Roberts & Axel, 1982).

1.5 Amplification in whole organisms

Examples of amplifications which confer resistance to drugs are not solely confined to cell lines. Two species of *Leishmania*, *L.mexicana amazonensis* and *L.tropica*, have been used to produce variants resistant to tunicamycin (Kink & Chang, 1987) and methotrexate (Coderre *et al*, 1983) respectively. In the field strains of the *Culex pipiens* mosquito and *Musca domestica* housefly have been described (Hyrien & Buttin, 1986) which, by virtue of an amplified detoxifying esterase, are resistant to an organophosphate insecticide. An amplified region of DNA of unknown function has also been reported in some wild strains of the mouse *Mus musculus domestica* (Weith *et al*, 1987).

1.6 Amplification and oncogenesis

The altered expression of oncogenes has been suggested as a major contributory factor in tumour progression and several have been shown to be amplified in some tumours. The *myc* oncogene family has been shown to be amplified in many tumours- N-*myc* in neuroblastomas (Emanuel *et al*, 1985; Shiloh *et al*, 1986); c-*myc* in glioblastoma (Trent *et al*, 1986), COLO 320 (derived from colonic neuroendocrine cells) (Schwab *et al*, 1986) and cell lines derived from a variety of organs (Kinzler *et al*, 1986). Johnson *et al* (1987) have shown that relapse patients with amplified *myc* oncogenes in small cell lung cancers survive a shorter period of time than relapse patients without amplified *myc* oncogenes.

Other oncogenes which have been shown to be amplified include K-*ras* in an ovarian carcinoma (Filmus *et al*, 1986) and in adrenal tumour cells (George *et al*, 1984) and c-*erb*B2 in mammary tumours (van der Vijver, 1987).

Winter & Perucho (1986) have shown that malignant transformation of NIH 3T3 cells transfected with human *ras* oncogenes is a gradual and reversible process dependent upon the abundance of *ras* sequences and transcripts. *Met* in spontaneously transformed NIH 3T3 fibroblasts was shown to be amplified (Cooper *et al*, 1986).

In a number of cases the oncogene has not only been amplified but also rearranged. Schwab *et al* (1986) showed an abnormal c-*myc* transcript was preferentially produced in COLO 320 cells containing 25 copies each of the normal and 5'-rearranged c-*myc* genes. Gilman (1987) reports that exon 1 of c-*myc* appears to contain a block to transcription elongation and that exon 1 is regularly mutated in Burkitt tumours. c-*abl* in a chronic myelogenous leukaemia (Collins & Groudine, 1987) has been rearranged and amplified. The amplification and rearrangement of oncogenes is often associated with chromosomal translocations and chromosomal abnormalities are frequently associated with tumour cells. Hunts *et al* (1985) showed that the EGF receptor gene in A431 cells was amplified, rearranged and at the translocation point of a rearranged chromosome 7. A c-*myc* sequence was translocated to within 1 kbp of an Igh heavy chain switch region in a Burkitt's lymphoma (Battey *et al*, 1983) while Collins (1986) showed that a c-*abl* (from chromosome 9) had been amplified and become associated with a Philadelphia chromosome (9:22 translocation). Selden *et al* (1983) showed that a c-*abl* had become associated with a C immunoglobulin gene (from chromosome 22) and the two had coamplified and become associated with an abnormal chromosome.

1.7 Amplification Of Transfected DNA

The transfection of selectable genes into recipient cell lines has shown them to be capable of amplification whether present as cDNAs carried on an expression vector plasmid (Fendrock *et al*, 1986; Kaufman *et al*, 1986) or a bacterial gene (Cartier *et al*, 1987; Chapman *et al*, 1983) or a mammalian gene (Gasser *et al*, 1982) present with its flanking regions on a cosmid (Milbrandt *et al*, 1983) or when present on an entire chromosome (Meinkoth *et al*, 1987). Because transfected DNA sequences are ligated together by the recipient cell (Perucho *et al*, 1980) forming tandem arrays which readily amplify, the co-amplification of an unselectable, but commercially valuable, gene (eg. B-interferon) with that of a selectable gene (dihydrofolate reductase) offers a way of increasing the yield of commercially

valuable proteins (McCormick et al, 1984).

Amplification of transfected sequences has also been used to attempt to define the role of flanking sequences in the amplification process. The results of transfection studies show that the site of chromosomal integration is apparently random and can take place at telomeres (Kaufman et al, 1983), centromeres and the main body of the chromosome (Wahl et al, 1984). Genomic environment has been shown to be important with Carroll et al (1987) revealing that a CAD cosmid transformant line which was capable of amplification at frequencies at least 100-fold greater than those of its cotransformed lines had integrated its CAD sequences in close proximity to a replication origin. Meinkoth et al (1987) showed that chromosome transfer from a species noted for unstable gene amplification to one associated with stable amplification resulted in stable gene amplification. However the effect of introducing large amounts of recombinogenic DNA into cells is an unknown and it could affect subsequent amplification processes. In some cases (Bostock & Clark, 1983) the unit amplified in the transfected line is indistinguishable from the donor unit while Kaufman et al (1983) using a cDNA found amplification units ranging from tens of kilobases to 1000 kbp. Milbrandt et al (1983) used a 45 kbp DHFR-containing cosmid to transform a $DHFR^-$ variant of the cell line from which the cosmid was derived and found that the size of the resulting amplification unit was close to that of the original unit (135 kbp) (Milbrandt et al, 1981).

Roberts et al (1983) used cotransfection of two plasmid sets, one bearing a functional APRT gene and a truncated thymidine kinase gene and the other bearing random human sequences, to study gene amplification by selecting for thymidine kinase activity. They found a gradient of amplification reminiscent of that found in Drosophila chorion gene amplification and suggested that multiple initiations from one replication origin followed by multiple recombinations were responsible for the observed structure. This result is the simplest structural analysis for an entire population of amplified units and lends credence to one model of the

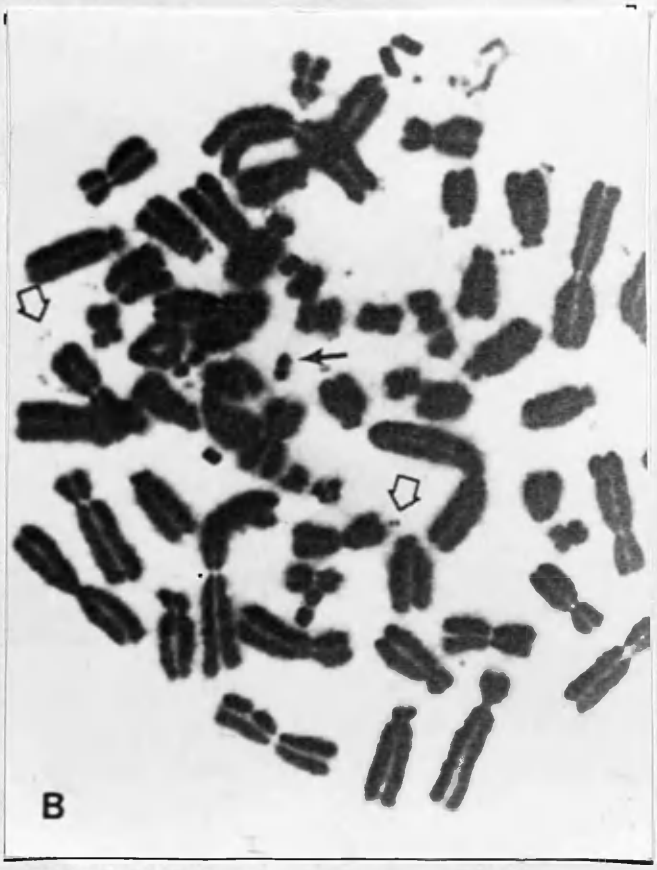


Fig 1.2 An example of a double minute containing cell line (Kuo et al, 1982). The open arrows indicate small double minutes and the closed arrows large double minutes.

amplification process but the recombination observed between the commonly repeated plasmid vector sequences is unlikely to represent the state of affairs in a normal genome.

1.8 Stability of amplification and chromosomal structure

In mammalian cells gene amplification is usually associated with the appearance of one of two types of unusual chromosome structure. The degree of stability evinced by the amplification is connected to these structures and reflects the tendency, or lack of it, of the species towards a stable karyotype (Schimke, 1984). The difference in stability of the amplification in the absence of selection is marked with a stable amplification declining slowly over 6-12 months and an unstable amplification losing 50% of its amplified genes in 2 weeks (Schimke, 1984).

1.8.1 Unstable amplifications

Unstable amplification is associated with small extrachromosomal elements called double minutes (Brown *et al*, 1981) which were originally observed in a variety of human and animal tumours. They take the form of small paired structures and were shown by Barker & Hsu (1978) to be acentromeric and to segregate randomly during metaphase apparently by association *en masse* with the ends or sides of true chromosomes such that cells in the same population can contain from zero to several thousand double minutes (Barker & Stubblefield, 1979). Electron microscopy shows them to be composed of two spherical objects (see fig 1.2), each approximately 0.5 uM in diameter, containing approximately 1000 kbp of DNA and joined together by chromatin fibres (Barker & Stubblefield, 1979). It is estimated that a double minute can contain between one and twenty genes (Schilling *et al*, 1982) depending upon its size. Recently pulse field gel electrophoresis was used to show that a 650 kbp extrachromosomal element, presumably related to the larger double minutes, behaved as a linear, rather than circular, structure (Maurer *et al*, 1987) Because they segregate randomly at metaphase the numbers of double minutes in each daughter cell is unlikely to be equal. In the absence of selection for amplified genes the metabolic load imposed by high numbers of double minutes causes the cells

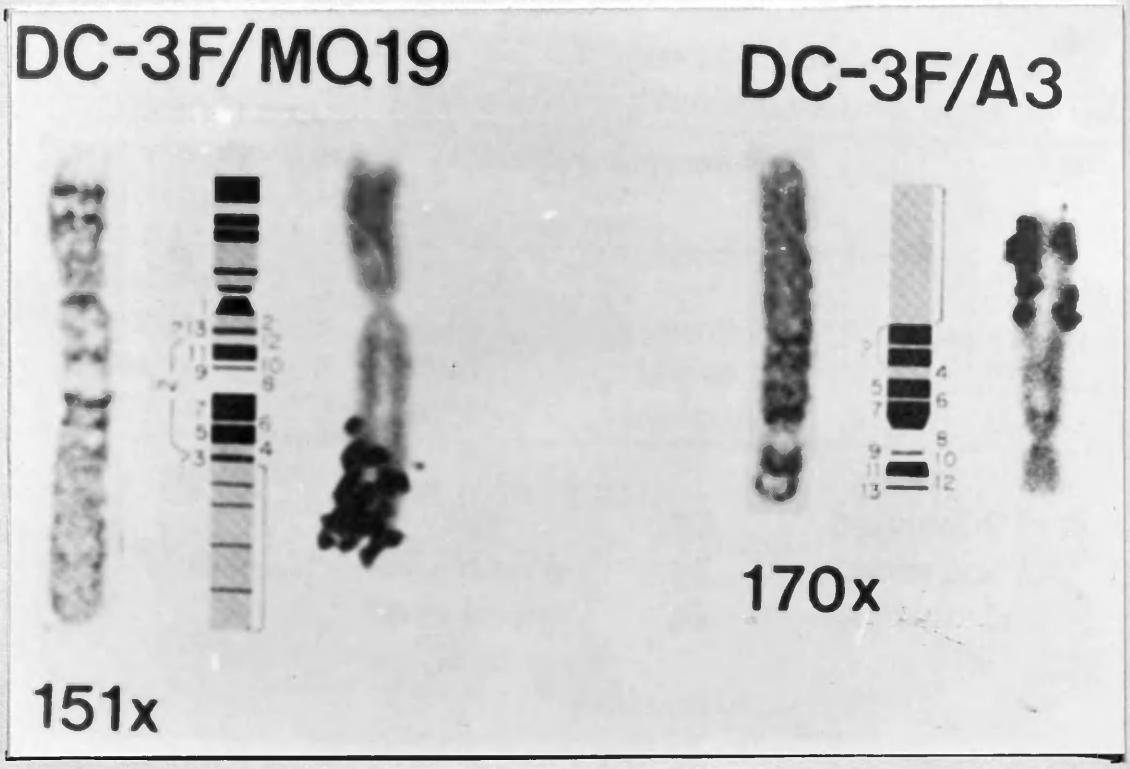


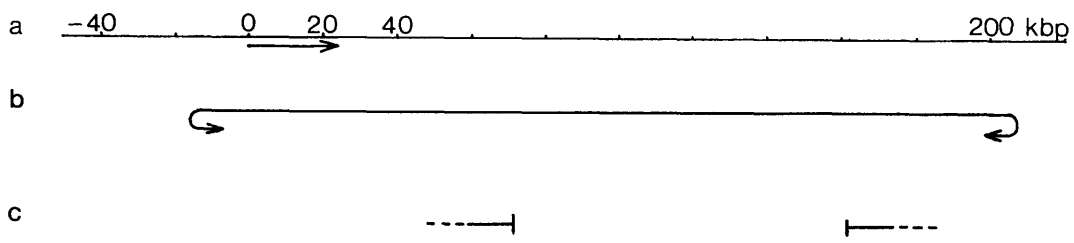
Fig 1.3 Two examples of homogeneously staining regions from methotrexate-resistant cell lines (Biedler, 1982). In both cases an HSR-containing trypsin-Giemsa banded chromosome is on the left and a diagrammatic representation in the centre with the HSR indicated by a bracket. The right hand chromosome shows the localization of the amplified dihydrofolate reductase genes to the HSR by <u>in situ</u> hybridization. The numbers indicate the degree of amplification.

containing them to grow more slowly than those cells with fewer double minutes and the cells containing the fewest double minutes readily multiply to dominate the population (Kaufman *et al*, 1981). Their small size (approximately 3% of the size of human G-group chromosomes) enables them to be readily separated from chromosomes during DNA preparation and thus affords an extra stage of purification to those wishing to clone amplified sequences (George & Powers, 1981). Double minutes are commonly found in mouse cell lines (Brown *et al*, 1981) and human cell lines (Trent *et al*, 1986) with amplified gene sequences and only rarely in hamster cell lines (Kuo *et al*, 1982) and then in conjunction with chromosome abnormalities and in a minority of the cell population (<20%).

1.8.2 Stable amplifications

After prolonged periods of selection in culture a previously unstable amplification may become stable (Kaufman *et al*, 1981) with the integration of the amplified genes into the chromosomes. The metaphase chromosomes of any species possess a characteristic pattern of dark/light banding after trypsin-Giemsa treatment (Seabright, 1971) and a disruption of this pattern provides a means of identifying and mapping chromosome abnormalities. The abnormality characteristic of stable gene amplificationis the homogeneously staining region (HSR) an expanded region of chromosome lacking a banding pattern and staining with intermediate intensity throughout its length (see fig 1.3). Numerous *in situ* hybridization studies have shown the HSR to be the location of the amplified genes (Wahl *et al*, 1982; Kellems *et al*, 1982). HSRs can be located close to the site of the original unamplified gene (Wahl *et al*, 1982), although the scale of banding patterns and *in situ* hybridization is too large to show whether the HSR is merely close to or actually coincident with the site of the original gene, or at some other chromosomal location (Alitalo *et al*, 1983). The production of an HSR may involve some degree of chromosomal rearrangement as evinced by the correspondence between the site of an HSR and a translocation in many cases. Biedler (1982) showed that the HSR was the site of translocation in a number of Chinese hamster DHFR amplificants

1.4a



1.4b

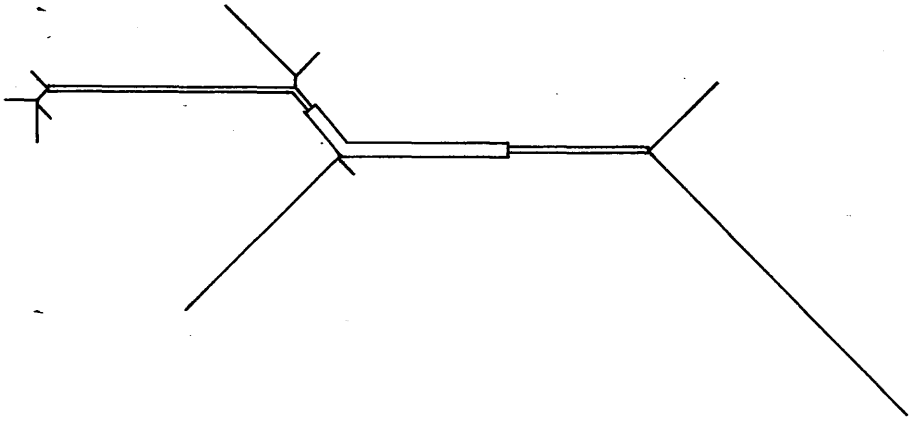


Fig 1.4a The homogeneous amplification units cloned by Looney & Hamlin (1987). Line (a) shows the extent of the units found in a directly repeated form with the numbering beginning at the DHFR gene transcription start. Line (b) indicates those sequences contained in the units found in an inverted repeat structure and line (c) shows the region deleted (lying between the vertical bars) in 10% of the units.

Fig 1.4b The non-homogeneous amplification units cloned by Federspiel et al (1984). The central box represents the 32 kbp DHFR gene (5' to 3', left to right). The flanking boxes/lines represent alternative structures with their width representing the frequency of occurrence; their junctions represent the position at which the alternative maps diverge.

and that the locations of these HSRs were in different regions of the q arm of chromosome 2 close to the site of the original gene. Another study (Flintoff et al, 1984) showed that an HSR could be near the site of the original gene while more telomeric chromosomal regions had been translocated elsewhere.

Both the size and copy number of the amplification unit contribute to the visibility of the chromosomal location of the units. Kaufman et al (1983) using transfected DNA sequences produced a 200-fold amplification of the DHFR gene which produced a visible HSR in two cell clones containing approximately 2 x 10^5 kbp (= 1000 kbp/unit) and another clone where the amplification site was not visible arguing for a very small (50 kbp) amplified unit.

1.9 Amplification at the DNA level

At the DNA level amplification becomes a confusing phenomenon capable of producing structures of remarkable simplicity or unbelievable complexity (see fig 1.4). This plethora of variation is most easily reconciled with a single procedural theme when amplification is viewed as a multistep process beginning with the initial confusion of the first amplification event and leading from a collection of coamplified fragments, through subsequent events, to an amplified unit which imposes a minimum load upon its host cell.

Initial attempts to quantify an amplification unit made use of highly amplified lines where a set of restriction fragments could be seen in ethidium bromide stained gels and the size deduced by simple addition. Sanders & Wilson (1984) estimated the size of the glutamine synthetase amplified unit to be approximately 50 kbp while Milbrandt et al (1981) found their DHFR unit to be approximately 135 kbp in size. That these represent only minimum sizes was shown later by Looney & Hamlin (1987) who succeeded in cloning the units of Milbrandt et al (1981) and revealing that they were 220 and 260 kbp in size. Other workers (de Bruijn et al, 1986) have reported a homogeneous structure at least 1100 kbp in size containig 5 different transcriptional units in a multidrug-resistant cell line and that

other cell lines similarly selected produce different structures. Wahl *et al* (1982) estimated that 500 kbp coamplified with each CAD gene by dividing the total size of the amplified region, deduced from metaphase spreads, by the number of CAD genes present. A later study (Ardeshir *et al*, 1983) supported this size estimate but showed that different fragments were amplified to different degrees in the same cell line and that in 12 different amplified lines only the 44 kbp immediately surrounding the CAD gene was consistently amplified. Federspiel *et al* (1984) investigated the DNA structure around amplified DHFR genes on double minutes and showed there was a multiplicity of different restriction maps and rearrangements within a single cell line some of which originated within the third intron of the DHFR gene (see fig 1.4). Meinkoth *et al* (1986) showed that a rearrangement near amplified CAD genes resulted in the production of two CAD mRNA species. The studies of Ardeshir *et al* (1983) and Federspiel *et al* (1984) show that:

(i) amplifed regions are not necessarily composed of identical units.

(ii) different sequences are coamplified in different independent events.

(iii) novel joints (ie. contiguous sequences derived from noncontiguous sequences in the unamplified parental line) are unique to each cell line and 'hot spots' for rearrangements are not evident.

Giulotto *et al* (1986) analysed the first CAD gene amplification event and discovered that as much as 10000 kbp could be coamplified along with each CAD gene and that subsequent amplification events reduce this to around 1000 kbp per CAD gene.

In situ hybridization studies have demonstrated the clustering of amplified genes which necessarily means that each gene must have some structural relationship to the genes flanking it. The examples cited above have been unable to ascertain this relationship because of the large expanses of flanking DNA and its evident inhomogeneity.

Ford *et al* (1985) transfected Rat-1 cells with a portion of the polyoma virus genome in a search for cellular enhancer

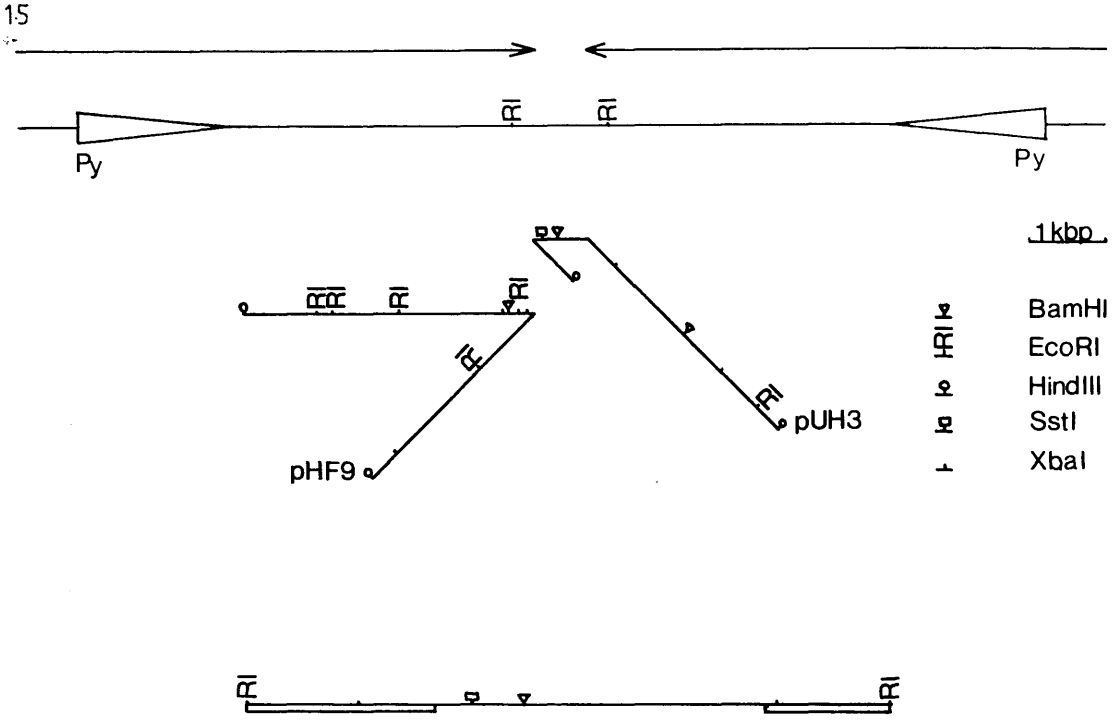


Fig 1.5 The inversion point of Ford et al (1985) cloned by Passananti et al (1987). The arrows represent the extent of the inverted repeat with the region between them being non-inverted. Plasmids pHF9 and pUH3 contain the unrearranged parental sequences which have become associated with the inverted repeat structure and their horizontal portions represent the regions involved in the inversion point. Below is a close-up of the 1.3 kbp EcoRI fragment containing the inversion point, the boxes represent pHF9-derived sequences.

sequences. One of the resultant cell lines was shown to contain 20-40 copies of Py DNA present in inverted repeat structures at least 22 kbp long (see fig 1.5). A reannealing-S1 nuclease digestion procedure was used to specifically identify inverted repeats. Using this procedure Ford & Fried (1986) showed that the c-*myc* genes amplified in four human tumour cell lines and the CAD genes amplifed in three BHK cell lines were arranged in the inverted repeat configuration. While this technique cannot reveal how far from the inversion point the genes are it does say that the sequences between the genes and the inversion point are homogeneous. Saito & Stark (1986) succeeded in cloning the inversion point from another CAD-amplified line. Looney & Hamlin (1987) have cloned the whole of a DHFR amplified unit from a single cell line and have found that there are two types (see fig 1.4). One unit type constituted 10-15% of the units, was 260 kbp long and arranged in a direct repeat structure while the other unit type was 220 kbp long arranged in an inverted repeat structure and its sequence was entirely contained within that of the 260 kbp unit. A variant which had deleted some 90 kbp accounted for 10% of the population. The homogeneity of these units was also shown by Burhans *et al* (1986) who showed that each unit possessed the same replication origin and that replication initiation was synchronous and occurred early in S phase.

1.10 Dynamics of amplification

Three factors affect the frequency at which amplifications occur:

(i) the relative growth rates of wild-type and amplified cells.

(ii) the rate at which amplifications are lost.

(iii) the rate at which they are produced.

It is a common observation that cells containig amplified DNA grow more slowly in non-selective media than cells lacking amplified DNA. Grund *et al* (1983) found that the doubling time of their multidrug-resistant cell lines had increased to between 18 and 23 hours from the 14 hours of the parental line. Kaufman *et al* (1981) showed that the growth rate of cells was inversely

related to the degree of gene amplification.

The loss rate of amplifications is difficult to quantify. In highly amplified lines (>50 gene copies) the rate can be dramatic as shown by Kaufman et al (1979) whose cell line took 120 doublings to lose the majority of its amplified DHFR copies present on double minutes but retained a stable7-fold amplification. The stability of amplification in HSR-containing cells is certainly greater than that of double minute-containing cells and HSR-containing cells can maintain amplifications stably for periods in the order of a year when subcloned. Fendrock et al (1986) suggest that an HSR is capable of changing at a rate of 1% per cell division though how this process would occur is unclear since unequal sister chromatid exchange, which potentially could alter the size of an HSR, is not increased in frequency within an HSR compared to the rest of the chromosome complement (Morgan & Fero, 1987).

Various models have been proposed to account for the mechanism of DNA amplification and the dynamics of the acquisition of amplified DNA has been extensively studied to shed light upon the viability of the models. To address the problem correctly only cells which have recently undergone a primary amplification event can be studied because other factors could affect secondary events and confuse the issue.

Using a procedure which ensured that the amplification event observed occurred in the last cell divisions before selection was applied Zieg et al (1983) found that 11 independent CAD-amplificants had amplified one CAD gene between 12 and 20 times and that this conferred resistance to levels of PALA, the selective agent, up to 10 times that used for the initial selection. An analysis of the rate of amplification (the number of resistant colonies formed per cell per generation) showed that it was relatively independent of the PALA concentrations used for selection which means that a standard amplification event produces a 6 to 10-fold increase in CAD gene copy number. The proportion of cells in the population resistant to a given level of PALA was dependent on the level of PALA.

Kaufman et al (1979) showed that the flourescence emission

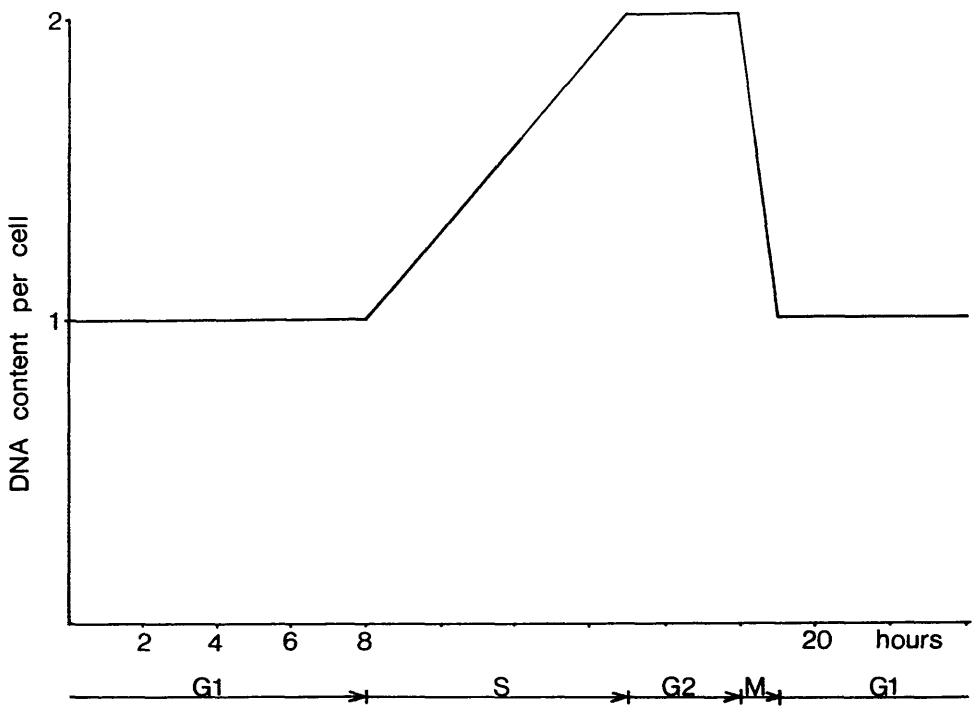


Fig 1.6 An idealised cell cycle showing the approximate duration of each stage of the cycle and its relationship to the DNA content of the cell (measured in arbitary units).

of cells saturated with a flourescent derivative of methotrexate (the inhibitor of DHFR) was proportional to the dihydrofolate reductase enzyme levels and this in turn was proportional to the degree of methotrexate resistance and the number of amplified DHFR genes. By coupling this knowledge to a flourescence-activated cell sorter (FACS) Johnston et al (1983) were able to analyse a population of cells that were not under any selective pressure. This technique revealed that the population was heterogeneous with respect to DHFR gene copy number and that 3% of the population had a flourescence level at least twice that of average. By selecting this 3% as the basis for a new population and then reselecting the cells with the highest flourescence in subsequent populations it was possible to produce a population with an average 40-fold increase in DHFR copy number after only 10 sortings. They concluded that amplification events occurred at a rate of approximately 1×10^{-3} per cell division in previously unamplified cells and that cells already possessing amplifications underwent further events at rates as high as 3×10^{-2} per cell division. A subset of cells therefore contains a DHFR copy number greater than the average in any population.

Mariani & Schimke (1984) suggested that this amplification phenomenon could come about by rereplication of DNA during S-phase. An idealised cell cycle is depicted in fig 1.6. By using a synchronised cell culture they were able to show that a transient blockage of replication by hydroxyurea after the initiation of S phase increased the frequency of DHFR amplification 10-100 fold. This blockage caused virtually all the DNA replicated prior to the blockage, which could be up to 10% of the genome, to be rereplicated after its removal. Under the same experimental conditions d'Anna et al (1986) found that DNA undergoing replication prior to the blockage could become extended to lengths greater than average replicon size. However Morgan et al (1986) using cytogenetic techniques found no evidence of rereplication under these conditions and suggested abnormal segregation of chromosomal fragments induced by the hydroxyurea followed by rereplication in a second cell cycle was responsible for the amplification. Hill & Schimke (1985) used FACS to show

that after hydroxyurea treatment the subpopulation with increased DHFR gene content showed an increased incidence of chromosomal abnormalities such as polyploidy, endoreduplication, chromosome fragmentation and extrachromosomal DNA.

Other work has confirmed the importance of S phase in the amplification process. Kleinberger et al (1986) used carcinogens to enhance methotrexate resistance and showed that the carcinogens used induced a transient enhancement of methotrexate resistance due to DHFR gene amplification. This enhancement increased for 72 to 96 hours after treatment and declined thereafter and the enhancement was maximal when the cells were treated in early S phase. Johnston et al (1986) showed that a blockage of the cell cycle at the G_1-S phase boundary by the inhibition of DNA synthesis resulted in an overproduction of DHFR enzyme (whose production is cell-cycle regulated and maximal in early S phase) and this was due to an alteration in the expression patterns of the parental DHFR genes. The levels of other proteins were also increased and the authors suggest that the overabundance of these proteins, including DHFR, presumed to be involved in DNA synthesis, allows the observed overreplication of DNA to occur when the blockage is removed and DNA synthesis resumes. Rice et al (1986) used a combination of FACS and staining with Hoechst 33342, a viable DNA dye, to analyse an asynchronous cell population. They found that hypoxia, associated with portions of solid tumours, caused an inhibition of DNA synthesis and that release from hypoxia caused overreplication of DNA. FACS also showed that the subpopulation containing overreplicated DNA was more resistant to methotrexate than was the subpopulation containing normal quantities of DNA. By incorporating a pulse of BrdUrd into an asynchronous population and measuring total DNA content Rice et al (1986) were able to distinguish between the G_1, S and G_2/M phases of the cell cycle. This enabled them to show that the cells in S phase when hypoxia was applied were the ones which underwent DNA overreplication when hypoxia was removed and that more than 60% of S phase cells (34% of total cells) did so. Furthermore cells were maximally resistant to methotrexate 12 hours after release from hypoxia.

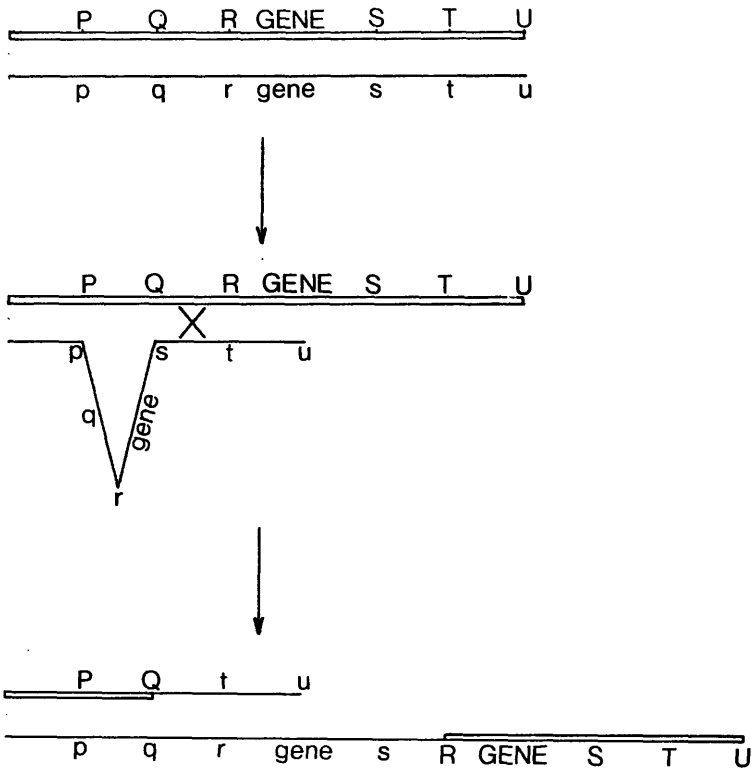


Fig 1.7 Unequal sister chromatid exchange. The first line shows the two chromatids correctly aligned, these then become misaligned (second line) and a recombination event occurs which results in a 2:0 distribution of the gene copies between the two sister chromatids (third line).

Other factors can contribute to the frequency of amplification. That the position in the genome of the gene under selection alters the frequency of amplification was shown by Wahl et al (1984) using a transfected CAD-containing cosmid. One transformant had an amplification frequency at least 100-fold higher than that of other transformants. By using a subclone of this line which carried the amplified CAD genes on 250 kbp episomes, Carroll et al (1987) were able to show that the CAD genes had acquired a functional origin of replication, most probably by integrating close to it in the original transformant. Interestingly the episomes showed virtually identical replication patterns to those of ribosomal RNA genes which are known to replicate in early S phase. Lavi (1981) showed that amplification of an integrated SV40 genome requires a functional origin of replication.

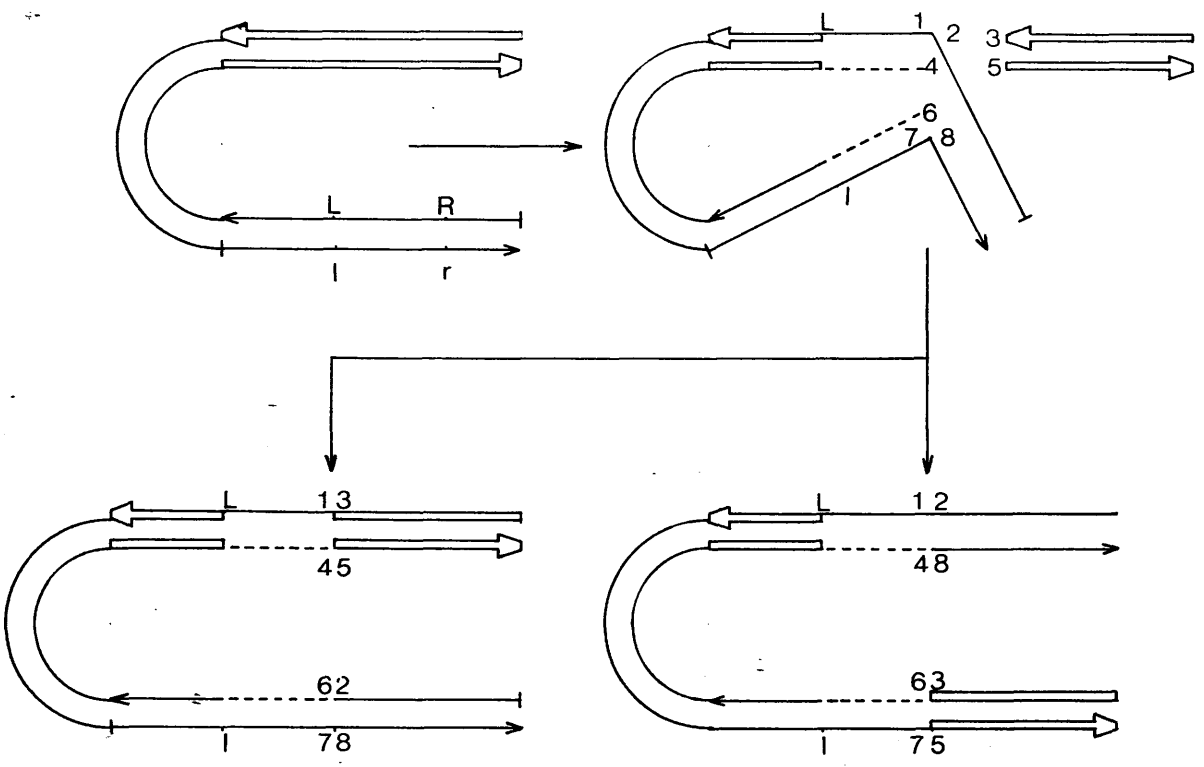
Another factor is the genetic background. By selecting simultaneously for resistance to methotrexate and PALA, Giulotto et al (1987) were able to isolate cells which showed an enhanced tendency toward DNA amplification as shown by an increase in the frequency of resistance to other selective agents. The frequency of doubly-resistant cells in the original screening was approximately 1×10^{-6}. Because the frequency of classical mutation is of this order or lower and since the observed doubly-resistant frequency is the multiplicant of three factors (proportion of cells with "amplificator" phenotype in the parental population; the proportion of "amplificator" cells possessing amplified DHFR genes; the proportion of "amplificator" cells possessing amplified CAD genes) the authors suggest that the "amplificator" phenotype may itself be due to a gene amplification phenomenon.

1.11 Models of amplification

Gene amplification is a multifaceted problem and the models available are content to address themselves to the question of how the extra DNA is generated and to let the details take care of themselves.

One possible model is unequal sister chromatid exchange

1·8a



1·8b

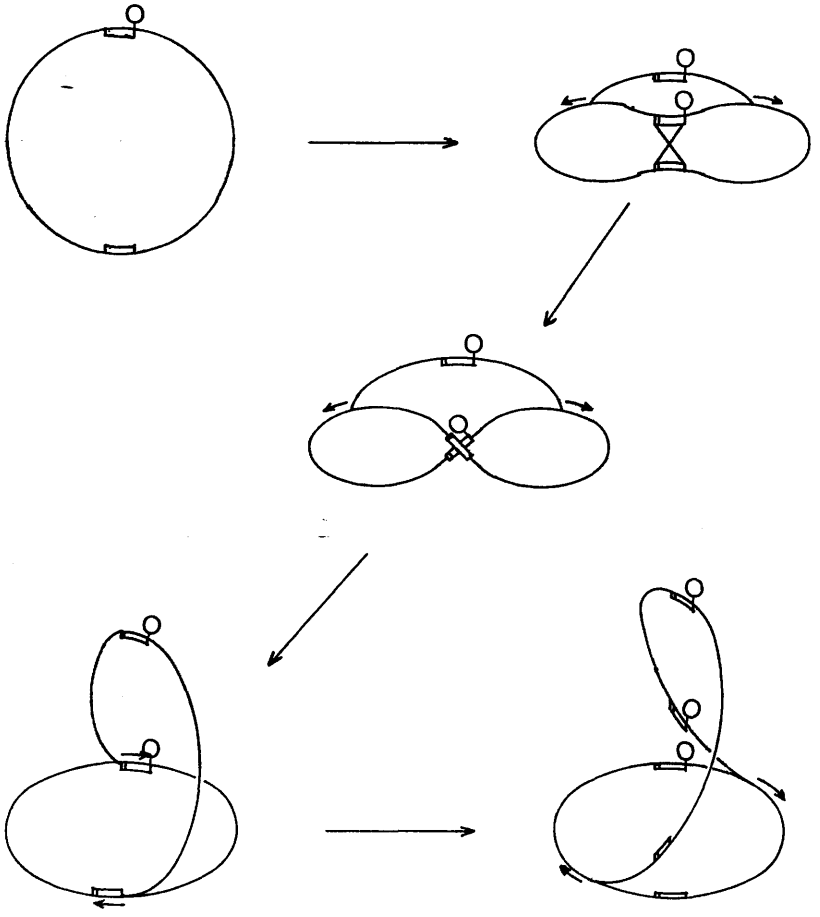


Fig 1.8a The effects of transposition in linear chromosomes (Stark & Wahl, 1984), if the curved lines are ignored the event is interchromosomal rather than the intrachromosomal event depicted. The transposing sequence is defined by LR, lr in the first diagram. It transposes into a new site and DNA replication (dotted lines) occurs (second diagram). The potential recombination points are numbered and depending on their association a simple insertion at a new site (bottom left) or an inversion (intrachromosomal) or a translocation (interchromsomal) can occur (bottom right).

Fig 1.8b The proposed overreplication model for yeast 2uM plasmids (Futcher, 1986). DNA replication begins bidirectionally from "O", the origin. Soon after it commences one of the replicated inverted repeats (open boxes) recombines with the unreplicated inverted repeat. This flips the direction of one replication fork so that both move in the same direction to produce a loop of DNA containing many copies of the same sequence.

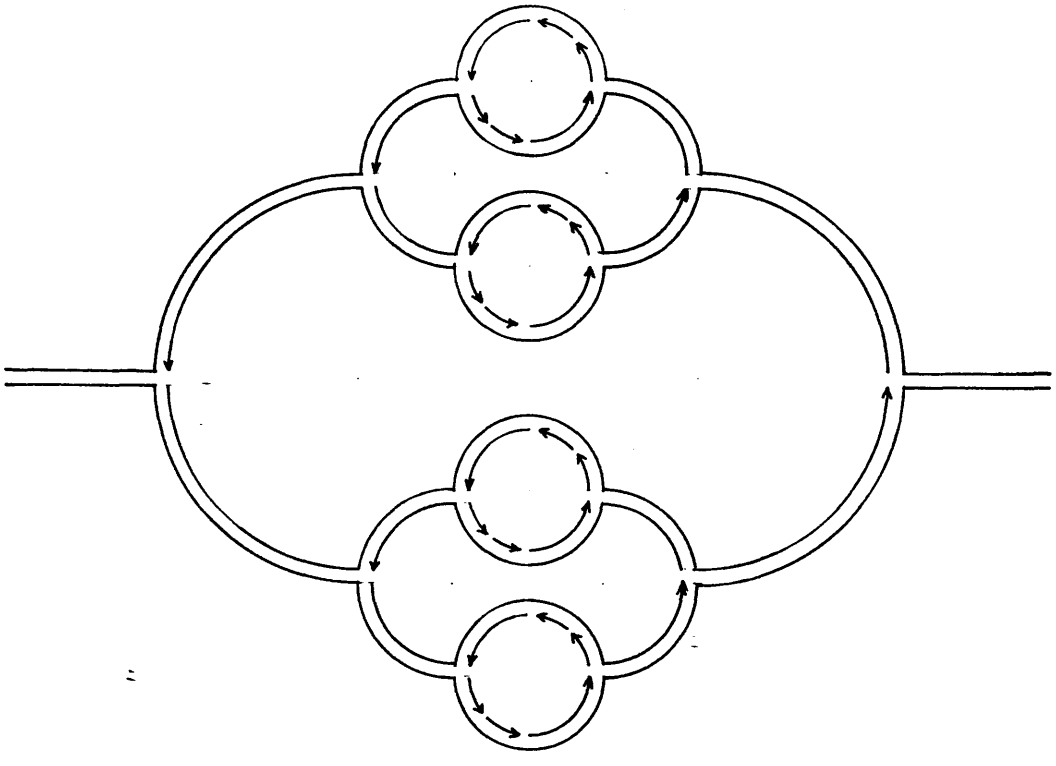


Fig 1.9 The "onionskin"- a model of multiple initiations of DNA replication from a single replication origin producing a nested series of replication forks. Note that only the outermost strands are connected to the chromosome and that all other strands possess free, and therefore recombinogenic, ends.

where two chromatids become misaligned at mitosis and reciprocal recombination generates a duplication and a deletion (see fig 1.7). The weight of evidence seems firmly against this model. Unequal exchange can generate only an increase of two-fold in the copy number in any one exchange event and multiple events must be postulated to generate the 6-12 fold increase observed. Unequal exchanges are equally likely to generate a loss of copy number as a gain and this would contradict the rapid generation and relative stability shown by amplified chromosomal structures. Unequal exchange does not explain the observations of structures containing DNA from various parts of the genome or chromosomal rearrangements frequently associated with amplification. Unequal exchange does not predict the importance of the replication process in the primary events of amplification.

Another possible model suggests that transposition, akin to that seen in bacteria, may be the prime mover in amplification (see fig 1.8). Transposition in bacteria can lead to reproduction of the transposing sequence at its insertion (acceptor) site or to the formation of a cointegrate intermediate on which both old and new transposing sequences are present. Cointegrate structures form when either the donor or acceptor structure is circular and since eukaryotic chromosomes are linear the results of transposition are capable of producing a reciprocal translocation between chromosomes or an inversion of sequences between the old and new transposing sequences on one chromosome.

The final general model is that of multiple initiations of replication from one replication origin. In its tidiest form multiple initiations produce the "onionskin" model- (see fig 1.9) a nested series of replication forks which produce a gradient of decreasing amplification from the centre outwards and which is responsible for the programmed amplification of the Drosophila chorion genes.

Both of these models are one-way processes in that the same mechanism which produces the amplification cannot be invoked to remove it which is in keeping with the stability of amplifications and both require that replication takes place.

Transposition has the advantage of providing a direct link

between the movement of DNA and its amplification which multiple initiations do not. A single transposition does not however explainthe multiple copies produced in one amplification event and transposition must therefore be postulated to induce multiple initiations at its insertion site. One potential consequence of bacterial transposition is the production of multiple copies by rolling circle replication where only one initiation of replication is required to give rise to multiple linear direct repeats; it does however require a circular starting element the generation of which would remove the direct association between movement and amplification in mammalian transposition. Transposition must be considered as an unusual event in mammalian cells and it would be required to occur many times in one cell cycle to satisfy the observations of multiple copies and the quantity of the genome which is overreplicated.

The common occurrence of inverted repeat structures in amplified DNA and their formation early inthe amplification process led Passananti et al (1987) to propose that the initial event in amplification is the formation of a circular extrachromosomal element containing two genes in inverted orientation and one replication origin. By analogy to the model proposed by Futcher (1986) for the overreplication of the yeast 2 uM plasmid they propose that after replication has initiated, a homologous recombination event flips the direction of one of the replication forks and a double rolling circle is produced which produces a loop of inverted repeats (see fig 1.8). This model requires inverted repeats to function which means it cannot directly account for direct repeats. It also requires the production of a specific extrachromosomal intermediate and it will only produce homogeneous units.

Multiple initiations from a single origin of replication as a model for gene amplification has several attractions, not least because it generates a great deal of unattached, highly recombinogenic DNA which can be held to produce almost any structure given an appropriately modelled recombination event. Another point in its favour is that it has been shown that multiple initiations produce the Drosophila chorion gene

amplification.

Schimke et al (1986) have taken the consequences of overreplication and loss of replication control well beyond the boundaries of the DNA amplification phenomenon. Their model suggests that inhibition of normal DNA synthesis results in the temporary decoupling of protein production and DNA synthesis with the result that an overaccumulation of proteins involved in DNA synthesis occurs. When DNA synthesis resumes these proteins are capable of multiple reinitiations of synthesis with the resulting overreplication of DNA. DNA replication occurs in the nucleoskeleton (Razin et al, 1986; Jackson & Cook, 1986), as does transcription (Jackson & Cook, 1985) which has been shown to be involved in chorion gene amplification, and the chromosomal DNA exists as loops anchored at the nucleoskeleton (Vogelstein et al, 1980). Schimke et al (1986) observe that this structure will produce free ends in close contact, if multiple initiations occur, a situation where recombination is likely to occur (Anderson & Eliason, 1986). The potential outcomes of a modelled recombination in this situation include circular or linear extrachromosomal elements, chromatid breakage, chromatid exchange, inversion or a breakage-bridge fusion chromosome all of which have been observed by Hill & Schimke (1985). Because of the nucleoskeleton localization of the replication machinery it is likely that in addition to bringing free ends derived from a single origin into close proximity free ends from several origins will be in close proximity which could recombine to produce the DNA rearrangements seen. Roth & Wilson (1986) showed that complementarity of between one and six bases at the ends of DNA duplexes was sufficient to induce end-joining of otherwise nonhomologous sequences.

Schimke et al (1986) propose that overreplication and recombination may have profound implications. Examples of amplified oncogenes from tumour lines are now common in the literature and Schimke et al (1986) suggest that this reflects a loss of replication control which could be a primary event in the creation of a tumourigenic cell. Aging at the cellular level is characterised by an increase in chromosomal aberrations and

overreplication-recombination is one way to create them.

1.12 Amplification and evolution

Evolution too may be affected by overreplication-recombination since it provides a means of creating multigene families, changing genome size, altering sequence copy number, rearranging chromosomes or changing ploidy. Schimke et al (1986) note that major genomic change is often associated with population divergence in plants and that 90-98% of mammalian germ cells degenerate, both of which could be a consequence of the loss of replication control. Bernstein (1977) has suggested that meiosis may function to screen chromosomes and genomes for aberrations by chromosomes pairing which can act to repair aberrant sequences or to prevent the primordia from becoming gametes.

Multigene families display features reminiscent of amplified genes. Numerous studies have revealed genes within families to be arranged in direct repeats and clustered, calliphorin genes (Schenkel et al, 1985), kallikrein genes (Mason et al, 1983); or clustered and in inverted repeats, apolipoproteins (Karathanasis, 1985), mouse major urinary proteins (Bishop et al, 1985). The mouse major urinary proteins are arranged in inverted pairs where one of the pair is functional and the other is inactivated by a nonsense mutation in codon 7 (Ghazal et al, 1985). Units which contain multiple transcription modules are known, for example units containing all five types of histone gene are arranged in tandem repeats in Xenopus (Perry et al, 1985).

In addition to producing multigene families DNA amplification may also play a role in maintaining their homogeneity and hence function. Weiner & Denison (1982) suggest that a functional member of a multigene family may be periodically amplified and overwhelm the accumulated sequence divergence of the rest of the family. This would lead to the accumulation of divergent non-functional genes and the authors note that there are 10 times more human U1 small nuclear RNA pseudogenes than there are functional genes.

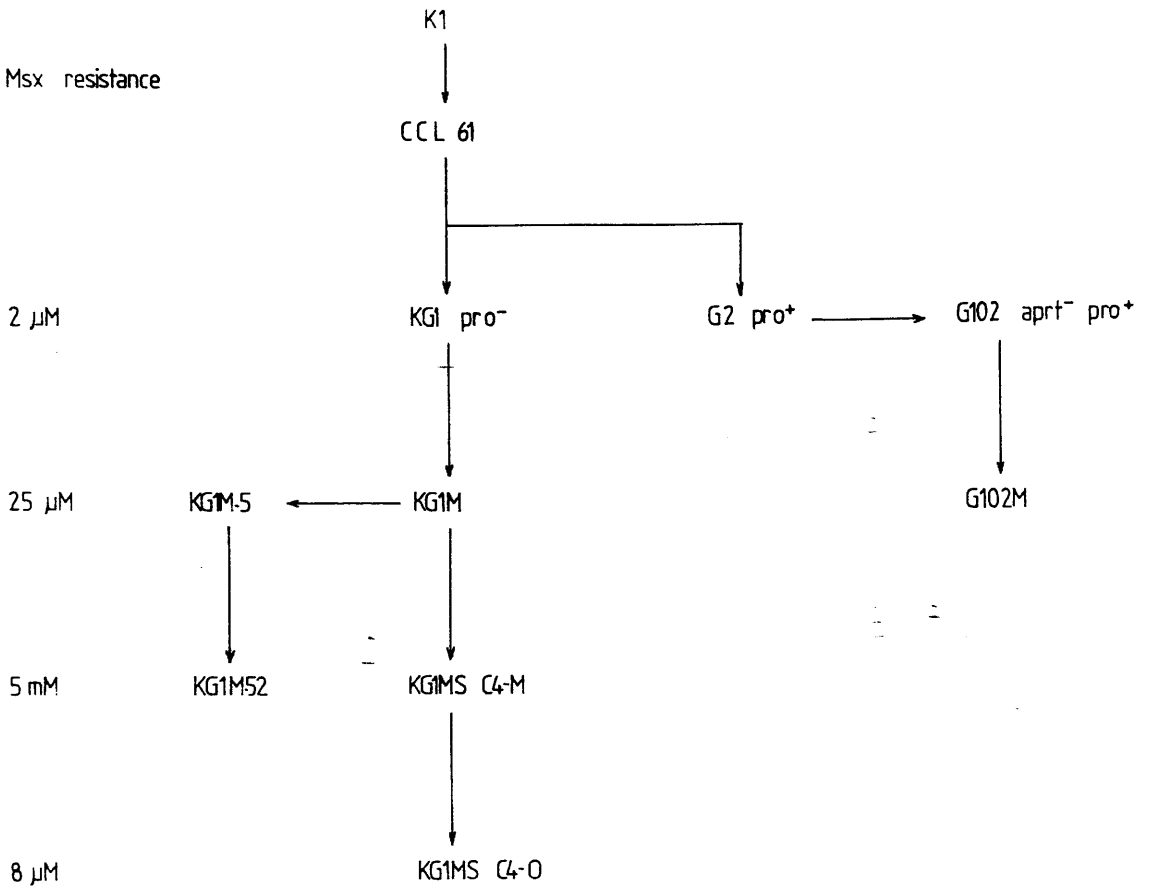


Fig 1.10 A diagram shows the derivation of the various Chinese hamster ovary K1 cell lines used in the study of the glutamine synthetase gene system. The level of methiomine sulphoximine to which each cell line is resistant is listed on the left-hand side. The KG1M-5 line is a subclone of the KG1M line and KG1M-52 is an uncloned descendant line of KG1M-5. C4M is a subclone of the KG1MS line subcloned after KG1MS achieved resistance to levels of 5mM Max.

Facing page 22

1.13 Glutamine synthetase- an introduction to the system

1.13.1 Glutamine synthetase- the protein

The glutamine synthatase enzyme is a universal housekeeping enzyme found in organisms as diverse as bacteria, plants, chickens and Chinese hamsters. Converting, as it does, glutamate and ammonia to glutamine its action links the TCA cycle with the metabolism of the amino acids and hence the enzyme represents an important regulatory point in cellular metabolism.

The level of GS protein varies from vertebrate tissue to tissue and can constitute between 0.01 - 0.1% of soluble protein in most tissues and >1% in certain hepatocytes, adipocytes and glial cells. GS activity is regulated at the transcriptional level by the action of glucocorticoid steroids and cAMP whilst external glutamine appears to regulate GS levels post-transcriptionally via ADP-ribosylation. In Chinese hamsters GS is a 42 kDa protein whose predicted amino acid sequence shows strong homologies to bovine and other GS sequences and regions of homology to the GS of the cyanobacterium Anabaena and plants (Hayward et al, 1986).

The creation of glutamine from glutamate and ammonia is a complex reaction. The hydrolysis of ATP leads to the formation of an enzyme-gamma glutamyl-phosphate complex which is attacked by ammonia to yield glutamine, phosphate and ADP. To date no regions of the GS protein have been identified as being potential substrate binding sites by virtue of their homology to other known binding domains.

1.13.2 The CHO cell lines- a history

Methionine sulphoximine (Msx) acts as a transition state analogue for all known glutamine synthetase proteins and it is an irreversible inhibitor. Studies of the glutamine synthetase in CHO cells were initiated by Dr R.H.Wilson because of the useful properties of the CHO cells- good somatic genetics, the ability to grow in defined media, the existence of auxotrophs allowing the use of complementation genetics.

The CHO-K1 line was obtained from the American Type Culture Collection as CCL 61. Analysis of these cells showed that 0.1%

had reverted to a *pro*$^+$ phenotype and this subpopulation was designated G2 (see fig 1.10). From G2 cells was isolated the G102 line which has the *aprt*– phenotype making it an otherwise wild-type cell line capable of complementation with the KG1 cell line for fusion/complementation studies. The remainder of CCL 61 had retained the *pro*– phenotype and were designated KG1 (for K1-Glasgow).

Methionine sulphoximine is toxic to both KG1 and G102 cells at levels exceeding 2 uM. Stepwise selection for KG1 or G102 cells with resistance to concentrations of Msx above 2 uM produced cell lines resistant to levels of 25 uM (KG1M and G102M). By complementation/cell fusion studies invoving the Msx-sensitive parental lines it was shown that the cause of Msx-resistance acted as a semi-dominant genetic marker. The site of action of Msx was shown to be specific to the generation of glutamine when 20 uM glutamine was shown to relieve the toxic effect of 20 mM Msx when both were applied to Msx-sensitive cells.

Further stepwise selection for Msx resistance of the KG1M line produced the KG1MS line, resistant to 5 mM Msx, of which KG1MSC4-M (C4M) is a clonal derivative. This line shows an increase in GS protein level which can be seen on SDS-PAGE gels. Fron the C4M line was derived the revertant KG1MSC4-O (C4O) line which is resistant to only 8 uM Msx. Dr R.H.Wilson (pers. comm.) reports that when Msx is not present in the C4M growth medium, where its presence is essential to maintain the selection for the 1000 active GS genes, the cells appear to fail to grow or divide for a few days before growth initiaites at a few foci, and from these foci came C4O.

1.13.3 Glutamine synthetase- the amplified DNA

Parallel restriction digests of KG1, C4M and C4O DNA, when gel electrophoresed showed that discrete restriction fragments could be seen in C4M DNA lanes but not in KG1 or C4O DNA lanes. By probing these gels with a probe made from total poly-A$^+$ mRNA, radiolabelled using reverse transcriptase, a subset of these C4M-specific fragments was identifed as containing the amplified glutamine synthetase gene. Of these fragments an 8.2 kbp BglII

fragment was selected as being large enough to contain the majority of the GS DNA sequences and small enough to be unambiguously identifed on an agarose gel. This fragment was isolated from a LMP-agarose gel and, after purification, cloned into the BamHI site of the pUC-9 vector plasmid. After screening the inserts with radiolabelled total C4M poly-A^+ mRNA two plasmid species were isolated and designated pGS-1 and pGS-2. pGS-2 was later shown to contain the pGS-1 insert in reverse orientation.

Using a series of serial dilutions of pGS-1 and C4M DNA digested with BglII which were gel electrophoresed, Southern blotted and probed with radiolabelled pGS-1 the copy number of GS genes in C4M cells was shown to be approximately 1000.

A lambda library of C4M DNA was created by ligating size selected C4M DNA partially digested with Sau3A to lambda vector L47.1 DNA digested with BamHI. Probing this library with pGS-1 produced two lambda clones which were used to extend the cloned region 3.4 kbp (lambda 86/2) and 5.4 kbp (lambda 11/3) to either side of pGS-1 (Sanders & Wilson, 1984). These lambda clones were subcloned into pUC-9 plasmids and a description of these and their later history is to be found in chapter 3.

Having located the exons of the GS gene within this cloned region, initially by S1 nuclease protection experiments and later by heteroduplex mapping, the GS cDNA was cloned at Celltech for sound commercial reasons and to allow a comparison of the CHO GS cDNA sequence with the available genomic sequence and with the sequence of other proteins. The cDNA was cloned from C4M mRNA and the procedure involved a classical first strand synthesis using oligo-dT primers and a second strand synthesis using RNase H enzyme which is capable of priming directly off RNA. A presumptive full length cDNA clone was isolated from a plasmid library created by the ligation of C-tailed cDNA to pBR322 plasmid G-tailed at the PstI site. Selection of the clone (pGSC-45) was acheived by probing initially for 5' sequences and then rescreening for 3' sequences (Hayward *et al*, 1986). pGSC-45 was shown to be 2.8 kbp long, the expected size of the larger of the two species of mRNA which code for GS protein (Sanders & Wilson, 1984).

CHAPTER 2

Materials and Methods

Chemicals and Biochemicals

General chemicals and organic compounds	B.D.H., Hopkins and Williams Koch Light Laboratories, May and Baker
Media	Difco, Oxoid
General biochemicals	Sigma
Agarose, X-gal	BRL
Antibiotics	Sigma
Restriction enzymes	BRL, Boehringer Mannheim New England Biolabs
^{35}S-dNTP, ^{32}P-dNTP	NEN
^{3}H-dCTP	Amersham

General Solutions

TE: 10mM Tris.Cl, 1mM EDTA, pH 7.5

PBSA: 10g NaCl, 0.25g KCl, 1.15g Na_2HPO_4, 0.212g KH_2PO_4, pH 7.2 in 1l dH_2O

20 x SSC: 175.3g NaCl, 88.2g NaCitrate, pH 7 in 1l dH_2O

2.1.1 Cell culture

Cells in routine culture were maintained in 25cm^2 plastic culture vessels under 10ml of appropriate media at 37^{o}C and were passaged when confluent (approximately once a week).

The cells were passaged by the removal of the old media by aspiration followed by a wash with 5ml of versene, also removed, followed by treatment with 0.5ml trypsin-versene to detach the cell sheet. The loose cells were squirted through a pipette several times after the addition of 9.5ml of fresh media and, if necessary, their number assessed using a haemocytometer. Fresh cultures were then initiated with 1/4 to 1/120th volume of the detached cells made up to 10ml with appropriate media. The cultures were then gassed to the correct pH (as judged by the media colour) using 5% CO_2.

When experiments involving genomic DNA or large numbers of cells were planned cells were cultured first in 125cm^2 vessels and when the cultures were confluent transferred to burrlers to increase cell number to 1-3 x 10^8.

Media solution

300 ml water

36 ml BHK 21 (Glasgow Dulbecco's Minimal Eagle's M x 10)

4 ml Non-essential amino acids solution

4 ml 50mM proline

4 ml nucleotides solution

4 ml 100mM sodium pyruvate

13.4 ml 7.5% sodium bicarbonate (pH indicator)

and

40 ml of foetal calf serum or dialysed foetal calf serum

Non-essential amino acids were alanine, aspartate, glycine, serine at 100 uM and asparagine, glutamate and proline at 500 uM. Nucleotide solution was adenosine, guanosine, cytidine, uridine at 30 uM and thymidine at 10 uM.

Versene: 8g NaCl, 0.2g KCl, 1.15g Na_2HPO_4, 0.2g KH_2PO_4, 0.2g EDTA, 1.5ml 1% Phenol Red, pH 7.1-7.3 in 1l. Prefilter and autoclave before use.

2.1.2 Freezing cells

Cells were grown to confluency in 125cm^2 vessels and the cell sheet detached as for passaging (except that 10ml versene, 2ml trypsin-versene and 18ml of media were used). The cell number was determined using a haemocytometer and the cells then pelleted at 1500 rpm for 5 minutes in a bench centrifuge. The cells were then resuspended in freezing medium at 5-10 x 10^5 cells/ml and frozen at -70^{o}C in 1ml aliquots.

Cells were recovered from storage by thawing an aliquot at room temperature and adding it to 9ml of appropriate media in a 25cm^2 vessel and incubating it at 37^{o}C.

Freezing solution

8 volumes of the growth medium

1 volume of foetal calf serum

1 volume sterile dimethyl sulphoxide

2.1.3 Selection for glutamine synthetase gene amplification

This was carried out using subcloned cell lines in media containing 10% dialysed foetal calf serum and appropriate concentrations of methionine sulphoximine. See chapter 5 for

details.

2.1.4 Cell lines

Two stock cell lines were routinely cultured- both Chinese hamster ovary cell lines. The KG1 line, obtained as CCL 61 from the American Type Culture Collection, Rockville MD., was cultured in media containing 10% foetal calf serum and the KG1MSC4-M line in media containing 10% dialysed foetal calf serum and 5mM methionine sulphoximine.

2.1.5 Dialysis of foetal calf serum

Foetal calf serum contains glutamine which is removed by dialysis. 500ml of serum was dialysed at 4°C against 54g NaCl in 2 litres PBSA and 6 litres of distilled water for 12 hours twice, and then against 72g NaCl in 8 litres of dH_2O for a further 12 hours.

The dialysed serum was then Millipore filtered and 1, 0.3, and 0.1ml of the filtrate were incubated at 37°C for 2 days in 20ml of bacterial growth medium (L-broth) to check for bacterial contamination. Uncontaminated serum was stored at -20°C while contaminated serum was refiltered and retested.

2.1.6 Subcloning

Cells were grown in a petri dish under 2-3ml of media which was changed daily in a CO_2 incubator. When cell colonies were visible a cloning ring was used to isolate a colony which was then washed with versene and removed into a Pasteur pipette after trypsin-versene treatment. Each colony so isolated was used as the basis for a fresh subclone line.

2.2.1 Bacteria

The Escherichia coli strain used in this study was ΔM15 and intended for use with the blue-white X-gal selection of the pUC8,9,18 and 19 series of plasmids developed by Vieira and Messing (1982) and Yanish-Perron et al (1985). It had the following genotype- Δ[lac pro] Ø80dlacZΔM15

It was grown in liquid culture in L-broth (plus 50ug/ml ampicillin when containing pUC plasmids) at 37°C and was aerated

by shaking.

When grown on solid support it was plated onto 25ml L-agar plates (containing 50µg/ml ampicillin and 20µg/ml X-gal when appropriate) and incubated overnight inverted at 37°C.

For extended storage stabs in L-agar were made and stored, sealed, at room temperature or 1ml of a log phase culture was mixed with 1ml of a 20% glycerol, 2% peptone solution and stored at either -20 or -70°C.

L-Broth was 10g tryptone, 5g yeast extract, 5g NaCl, 1g glucose, 20mg thymine, made up to 1 litre in distilled water and adjusted to pH 7 with NaOH.

L-Agar was L-broth plus 15g/l agar and lacking glucose.

They were sterilised by autoclaving at 120°C for 15 minutes.

Ampicillin and **X-gal** (5-Bromo-4chloro-3-indolyl-β-galactoside) were stored solid at -20°C and made up as required. The X-gal was dissolved in dimethyl sulphoxide at 10mg/ml.

2.2.2 Transformation

A 2.5ml starter culture was grown overnight and 1ml of this transferred to 20ml of fresh L-broth and this was grown, shaking, at 37°C for 3 hours. The cells were pelleted (2 minutes at 2.7 Krpm) and resuspended in 10ml of ice-cold 50mM calcium chloride. After 10 minutes on ice the cells were repelleted as before and resuspended in 2ml of the calcium chloride and kept on ice. 200ul aliquots were transformed by the addition of the appropriate DNA. After 15 minutes on ice the cells were heat-shocked (2.5 minutes at 42°C or 5 minutes at 37°C) and returned to the ice for a further 10 minutes. 800µl of L-broth was then added to each aliquot and the result incubated at 37°C for 60 minutes. The culture was then plated out in 200µl aliquots and the plates incubated overnight at 37°C.

2.2.3 Selection for plasmid inserts

ΔM15 lacks a functional β-galactosidase gene but produces a truncated form of the enzyme lacking part of the alpha domain which can be complemented by the alpha polypeptide encoded by the pUC plasmid region encompassing the polylinker region. Acting together these polypeptides are capable of converting X-gal to a

blue metabolised form. An insertion into the polylinker region will prevent this complementation and leave the X-gal in its white, unmetabolised, form. A simple visual inspection of the blue or white colonies will reveal those harbouring plasmids containing inserts which can then be further analysed.

Colonies selected for further analysis were patched out onto plates in 1 cm^2 and the resultant smear scraped off using a sterile toothpick and placed in 200µl single colony final sample buffer for 15 minutes. This was spun in an Eppendorf tube (12000 rpm, 15 minutes) and 30-40µl of each sample loaded onto an agarose gel and electrophoresed to determine the relative sizes of the inserts.

Colonies selected from this second round of selection were grown up in 20 ml L-broth and plasmid DNA prepared using a modified version of the technique of Birnboim and Doly (Maniatis, 1982). The plasmid DNA was then analysed by restriction endonuclease digestion and agarose and/or acrylamide gel electrophoresis.

Single Colony Gel Loading Buffer: 2% Ficoll, 1% SDS, 0.01% Bromophenol Blue, 0.01% Orange G in TEA

2.3 DNA Preparation and Manipulation

2.3.1 Ethanol precipitation of DNA

The DNA solution was made 0.3M NaOAc with the addition of 3M NaOAc, pH5.2, and two volumes of absolute ethanol were added. The solutions were mixed and placed at -20°C for at least an hour to precipitate the DNA before centrifugation at 10 Krpm or more for at least 10 minutes.

2.3.2 Plasmid preparation-this was an appropriately scaled version of the alkaline lysis technique of Birnboim and Doly described in Maniatis *et al* (1982) with various final purification methods.

Solutions

I 50mM glucose
10mM EDTA
25mM Tris.Cl pH 8
4 mg/ml lysozyme

II 0.2 M NaOH
1% SDS

III 3 M Potassium Acetate pH 4.8
(100ml solution contains 60ml 5 M potassium acetate, 11.5ml glacial acetic acid, 28.5ml water)

For a 200 ml bacterial culture the following procedure was employed: The bacteria were pelleted by centrifugation at 10 Krpm for 10 minutes and resuspended in 10 ml of solution I. After 5 minutes at room temperature 20 ml of solution II was added and the mixture placed on ice for 5 minutes before 15 ml of solution III was added. After a further 5 minutes on ice the preparation was centrifuged at 10 Krpm for 10 minutes at 4°C. The supernatant was transferred to 250 ml pots through medical gauze to remove any particulate matter and 90 ml of ethanol added to precipitate the nucleic acids. After storage at -20°C for 60 minutes it was centrifuged at 10 Krpm for 10 minutes.

For 200ml preparations the nucleic acid precipitate was resuspended in 6ml of TE and 6.6g of caesium chloride and 1.5mg ethidium bromide added. This was centrifuged at 49 Krpm, 20°C for 16 hours in a Beckman Ti 70 rotor in an ultracentrifuge. After centrifugation the plasmid band (either visible or flourescent under ultra-violet illumination) was removed. Two volumes of TE were added and the ethidium bromide removed by partitioning into several changes of propan-2-ol. The volume of the aqueous phase was maintained at its original level by further additions of TE during the partitioning process. The DNA was then ethanol precipitated, washed in 70% ethanol, dried and resuspended in TE at 0.1-1mg/ml. Its purity and intactness were assayed by gel electrophoresis.

For 20 ml and 1.5 ml preparations the precipitate was resuspended in water and extracted twice against an equal volume of phenol:chloroform (2:1) and once against an equal volume of

chloroform before being reprecipitated.

2.3.3 Genomic DNA preparation

The cell sheet is removed using trypsin-versene and the burrlers washed with PBSA solution at 4°C. The cells are pelleted and resuspended in ice-cold TE at approximately 10^8 cells/ml. 10 volumes of 0.5M EDTA (pH 8), 100 µg/ml proteinase K, 0.5% Sarcosyl is added. The solution is left at 50°C fpr three hours and swirled occasionally. The DNA is gently extracted against phenol 3 times (the phenol forms the upper phase) and is then dialysed against 4l of 50mM Tris.Cl (pH 8), 10mM EDTA, 10mM NaCl with several changes of the solution. Treat the DNA with 100 µg/ml DNAase-free RNAase at 37°C for 3 hours and extract twice against phenol/chloroform. Dialyse extensively against TE. (Maniatis et al, 1982).

2.3.4 Restriction digests

Four digestion buffers were routinely used containing various concentrations of NaCl. These were stored as x10 concentrates a at -20°C prior to use. The digestion conditions used were derived from the New England Biolabs enzyme catalogue (see p 7-29 and p 110). Digestion volumes were generally as small as possible while taking into account the requirements to have a final glycerol concentration of <5% and to dilute the DNA solution at least fourfold. The amount of restriction enzyme used to ensure complete digestion was 2-3 fold in excess for plasmid DNA and 5-10 fold excess for mammalian DNA (based on 1 enzyme unit per ug of DNA). Digestion periods of over one hour were routinely used and for large quantities of mammalian DNA (>20µg) a second aliquot of enzyme was frequently added two hours after the start of incubation. The reaction was stopped by the addition of EDTA to 10 mM final concentration.

Restriction enzyme buffers x 10

Low salt : 100mM Tris.Cl pH 7.5
100mM $MgCl_2$

Medium salt : 100mM Tris.Cl pH 7.5
100mM $MgCl_2$
500mM NaCl

High salt : 500mM Tris.Cl pH 7.5
100mM $MgCl_2$
1000mM NaCl

150mM salt : 500mM Tris.Cl pH7.5
100mM $MgCl_2$
1500mM NaCl

2.3.5 Ligations

These were generally performed in a final volume of 20µl with insert sequences in a 3-4 fold excess over vector sequences, 1-2 U of DNA ligase and 2µl of a 10 mM ATP solution. Incubation was at 16°C overnight or, occassionally, at room temperature for 3 hours. The reaction was stopped by freezing at -20°C or immediate use in the transformation of E.coli. Generally 5 µl of the reaction mixture was added to 200µl of the cells in calcium chloride. Vector alone was ligated and transformed as a control.

Ligation buffer x 10 500mM Tris.Cl pH 7.4
100mM $MgCl_2$
100mM dithiothreitol
10mM ATP
1 mg/ml BSA

2.3.6 DNA preparation for pulse field gel electrophoresis-

The cell sheet was detached from the burrler by trypsin-versene treatment and broken up by passage through a pipette after the addition of cold PBSA. A haemocytometer was used to discover the number of cells present and these were pelleted using a bench centrifuge. They were resuspended in PBSA which was prewarmed to 39°C. Molten 2.5% low melting point agarose (which had previously been treated with DEAE-cellulose) at 39°C was then added and 100ul aliquots put into the slots of the block former which was then placed at 4°C until the agarose had set. The PBSA:agarose ratio was 2:1 and the final volume gave a cell concentration of 10^6 cells/100µl. Using an alcohol sterilised spatula end the cell blocks were pushed into the lysis mixture (0.5ml/block) and incubated at 50°C for 16-48 hours after which the blocks and mixture were stored at 4°C.

Lysis solution 0.5 M EDTA pH 9.5
1% sarcosyl
0.25 mg/ml proteinase K

2.3.7 Restriction endonuclease digestion for PFGE

The DNA handling techniques described are modifications of those of Bernards et al (1986).

The blocks to be digested were immersed overnight in 10mM Tris.Cl pH 7.5, 10mM EDTA, 0.1mM PMSF (10 ml per block) at room temperature and then transferred to the appropriate restriction buffer (lacking dithiothreitol) plus 0.1mM PMSF and 0.2% gelatin (10 ml per block) for 1 hour. Digestion took place in 50µl restriction buffer containing the restriction enzyme(s) per half block digested. 30 units of enzyme were used for each digestion which lasted for 6 hours at the appropriate incubation temperature unless otherwise stated. The reaction was stopped by the addition of EDTA to 50mM final concentration.

PMSF was phenylmethylsulphonyl flouride (BDH) and stored solid at room temperature. It was dissolved in acetone prior to use.

2.4 Pulsed field gel electrophoresis

See fig 4.2. The system used was that described by Carle and Olson (1984). A 10 x 10 cm agarose gel was placed between two sets of platinum electrodes placed at the diagonals of the gel. Alternation of the current between these sets of electrodes provides the directionally varying electric field used to separate molecules >20kbp in length. The length of time each electrode pair was activated was the same for both pairs and varied from 5 seconds to 40 seconds depending on the resolution of sizes required. The voltage supplied was chosen according to the required run length and was supplied by a Vokam 2197 unit capable of producing 0-2500 V. It was alternated between the electrode pairs by the use of an RS 348-403 switch coupled to a RS 345-375 timing mechanism. Standard agarose in 0.25 x TBE was used to form the gel and the gel slots were able to accomodate half an agarose block. The blocks were placed in the gel slots using an alcohol sterilised spatula and sealed in place using

molten 1% low melting point agarose in PBSA. The 0.25 x TBE buffer was recirculated through a heat exchanger during operation to maintain a temperature of 14°C. As a precaution to prevent buffer movement washing the gel from its horizontal glass plate a narrow strip of double-sided sellotape was placed on the glass plate at its top and bottom and the molten agarose poured onto it. After gel electrophoresis was complete the gel was removed, stained with ethidium bromide and treated thereafter as a standard agarose gel. Fresh buffer was added prior to each run (approximately 1500ml) until the gel was 5-10 mm submerged to ensure that the buffer did not become exhausted and the same amount of buffer was removed at the end of each run. It was impractical to completely drain the system owing to the amount of peripheral tubing involved.

Size markers for the system were lambda multimers. Lambda DNA produced within the department was mixed with PBSA and DEAE-cellulose treated 2.5% low melting point agarose made in PBSA (2:1 ratio) to give a final DNA concentration of 1ug/ml. The mixture was allowed to set in the block former and kept in PBSA solution for at least two weeks at room temperature to allow the cohesive ends to come together and produce multimers. The blocks were stored at 4°C in PBSA.

2.5.1 Metaphase spreads for karyotypic analysis

An actively dividing sub-confluent flask of cells was selected and 100 ul of a 100ug/ml solution of colchicine added to it to arrest the cells in metaphase. After 90 minutes of incubation at 37°C the medium was carefully drawn off and the cells gently washed with 5ml of versene which was also carefully drawn off. 0.5ml of trypsin-versene was then added and the culture vessel gently tapped until the majority of the cells had detached (visualised using x100 inverting microscope). The cells were collected by the gentle addition of 5ml of the media/versene mix and the return of 5.5ml to the mixture. This was then centrifuged at 1500 rpm for 5 minutes in a bench centrifuge. The majority of the liquid was aspirated off and the cells resuspended in the remainder using gentle passage up and down a Pasteur pipette. 10ml of 75 mM KCl was then added and the mixture

left to stand at room temperature for 20 minutes (KG1) or 25 minutes (C4M). The cells were then repelleted as before and again most of the liquid aspirated off and the cells gently resuspended in the remainder. The suspension was slowly added to 10ml of ice-cold fresh fixative (Analar grade methanol: glacial acetic acid, 3:1) and left on ice for 10 minutes. After centrifugation and resuspension as before (the pellet is usually invisible at this stage) a further 10ml of ice-cold fixative is added and left on ice for a further 10 minutes. After another round of centrifugation and resuspension a drop was dropped onto a clean slide and examined with a x100 inverting microscope. If a reasonable number of well-spread, but not dispersed, chromosome spreads were apparent the remainder of the suspension was dropped onto slides from approximately 30 cms and allowed to air dry.

2.5.2 Trypsin-Giemsa banding

After drying for 3 days the spreads are ready for use. Placed horizontal a slide is covered by a trypsin-versene solution for 30 to 90 seconds (depending on how readily the chromosomes are digested) and this is washed off with PBSA. A 10% solution of Giemsa stain (Gurr R66) in a phosphate buffer (2mM Na_2HPO_4, 2mM KH_2PO_4, pH7) covers the slide for 2 to 5 minutes (depending on the staining efficiency of the Giemsa batch) and this is washed off using distilled water and the slide placed near vertically to air-dry. The slide was then examined at x40 and x250 magnification and if successfully banded more slides were so treated and subsequently photographed using a Leitz Vario-Orthomat automatic photomicrographic system mounted on a Leitz Orthoplan universal largefield microscope with Kodak Panatomic-X 400 ASA black and white film. The magnification was approximately x1000.

Slides were cleaned by immersion in 99% ethanol 1% HCl.

2.5.3 In Situ hybridization

Metaphase spreads were prepared as for trypsin-Giemsa banding except they were used within 2 days of preparation. The method was that of Pardue (1985). Slides were placed in an upright rack which was transferred from one solution to another.

1 Digest with RNase A (boiled for 10 minutes previously) at 100 µg/ml in 2 x SSC for 60 minutes at 37°C.

2 Wash in 2 x SSC (3 x 10 minutes).

3 Dehydrate with 70% ethanol (2 x 10 minutes) followed by 95% ethanol (1 x 5 minutes) and allow to air dry.

4 Denature DNA with 70% formamide in 2 x SSC for 2 minutes at 70°C.

5 Dehydrate as in 3.

6 50 µl of the probe solution is placed on each slide. (See note).

7 Seal under a coverslip using cowgum paste.

8 Incubate at 37°C for >11 hours.

9 Remove the coverslip and wash in 50% formamide in 2 x SSC (2 x 10 minutes) at 39°C.

10 Wash in 2 x SSC for 10 minutes at 39°C.

11 Dehydrate as in 3.

12 Cover with autoradiographic emulsion and expose in the dark at 4°C for at least 3 days.

13 Develop for 5 minutes in Kodak D19 developer, gently rinse in tap water, and fix for 4 minutes in Ilford Hypam and rinse gently again. The temperature should be 18-20°C.

14 While wet stain for 1 minute in 5% Giemsa stain in water, rinse gently in tap water and then in distilled water.

15 Dry the slide completely and mount a coverslip using D.P.X. (BDH)

16 Examine and photograph under a microscope as for trypsin-Giemsa banding.

Notes

The formamide is deionized prior to use by stirring with a deionizing resin (Amberlite MB-1 from Sigma) for 30 minutes and filtering to remove the resin.

The probe is prepared by random primer synthesis using 50 uCi of ^{3}H-dNTP instead of ^{32}P-dNTP. The probe solution consists of 50% formamide, 0.3M NaCl, 30mM trisodium citrate, 40mM sodium phosphate pH 6, radiolabelled DNA at 0.05-0.2 µg/ml and a 1000-fold excess of sheared salmon sperm DNA. The DNAs are denatured by boiling for 10 minutes and are added to the remainder of the

components of the probe solution at 70°C just prior to its application.

Autoradiographic emulsion is Ilford K5 emulsion diluted 1:2 emulsion:dH_2O which is stored in 15ml aliquots in the dark at 4°C and thawed to 47°C just prior to use.

20 x SSC is 175.3g NaCl, 88.2g NaCitrate in 1l, pH 7.

Coverslips were boiled in 1M HCl and rinsed in 10mM sodium phosphate pH7 prior to use to neutralise any soluble alkali they contained.

2.6 DNA radiolabelling

Nick translation was used initially to radiolabel DNA probes but latterly the technique of random synthesis off hexanucleotide primers has been employed which is consistently more efficient and reproducible than nick translation.

2.6.1 Nick translation- the final volume is 50 µl, 9 µl of which comprise labelling solutions, leaving 41 µl to be apportioned between the probe DNA solution, the radiolabel solution and water. 5 µl of 10 x nick-translation buffer is mixed with 1 ul of a 1mM solution of each of the three cold deoxynucleotides (=3 µl) the probe DNA solution (0.1-1 µg DNA) and the radiolabel solution (30-50 µCi of either ^{35}S-dNTP or ^{32}P-dNTP) and water. 0.5 µl of a 0.1 µg/ml DNAase I solution (prepared by a 10^4 dilution of a 1 mg/ml solution into nick translation buffer in 50% glycerol which can be stored at -20°C) is added on ice and left for 2 minutes. 5 units of E.coli DNA polymerase I is added and incubated for 1 hour at 16°C. The reaction is stopped by 2 µl 0.5 M EDTA.

Nick translation buffer x 10

0.5M Tris.Cl pH 7.2
0.1M $MgSO_4$
1mM DTT
500 ug/ml BSA (fraction V)

2.6.2 Random primer synthesis- the final volume is 50 µl, 13 µl of which comprise the labelling solutions, leaving 37 µl to be apportioned between the probe DNA solution, distilled water and radiolabel. 100 to 400 ng of DNA are used and 30 to 50 µCi of a

^{32}P radiolabel. The probe DNA solution plus the water are boiled for 10 minutes and rapidly placed on ice. 10µl of the appropriate labelling solution is added (these are made up lacking one of the four deoxynucleotides) followed by 2 µl of BSA (10 mg/ml, fraction V) and the radiolabel. 5 units of Klenow fragment are then added and an overnight room temperature incubation begun. The reaction is stopped either by separation directly through a Sephadex column (G50-150 in TE) or by adding 50 µl of stop solution (which permits a 'cold' chase synthesis) waiting 30 minutes and then separating DNA from unincorporated nucleotides using a Sephadex column.

Random primer synthesis solutions

The final labelling solution was a composite of several subsidiary solutions and these are given here.

Solution I : 1.25M Tris.Cl pH 8
0.125M $MgCl_2$

Individual dNTP solutions were made 100mM in TE

Solution A : 1 ml of solution I
18µl B-mercaptoethanol
3 of the dNTP solutions (5µl each)

Solution B : 2M HEPES made to pH 6.6 with 4M NaOH

Solution C : Hexadeoxyribonucleotides (Pharmacia, PL No. 2166) in TE at 90 O.D units/ml

Labelling solution is A:B:C,100:250:150

Stop solution : 20mM NaCl
20mM Tris.Cl pH 7.5
2mM EDTA
0.25% SDS
1mM of each dNTP

The labelling solution was stored at -20°C and the stop solution at 4°C.

2.7 Southern blotting

Two versions of Southern blotting (Southern, 1975) were employed depending on the membrane type onto which the DNA was transferred. For most purposes the nylon-based Biodyne A (Pall Ultrafine Filtration Corp.) membrane was used according to the manufacturers instructions while for probing single copy

sequences nitrocellulose (Schleicher & Schuell) membranes were used.

In both cases the agarose gel was gently shaken in 0.2 M HCl at room temperature for 15 minutes followed by 30 minutes gentle shaking in 1.5M NaCl, 0.5M NaOH.

For DNA transfer to Pall membrane the gel was immersed in 3M sodium acetate pH 5.5 for 30 minutes while for nitrocellulose 0.5M Tris.Cl pH 7.5, 1.5M NaCl was used. DNA transfer was accomplished by the capillary action of the migration of 20 x SSC from the reservoir through the gel and membrane and subsequent tissue layers.

After overnight transfer the Pall membrane was immediately baked at 80°C for 1-2 hours or irradiated with ultraviolet light for 2 minutes to fix the DNA on the membrane. The nitrocellulose was briefly soaked in 3 x SSC and allowed to air-dry before baking at 80°C for 1-2 hours to fix the DNA.

2.7.1 Probing DNA bound to Pall membranes

The membrane was prehybridized for 1-3 hours at the hybridization temperature (68°C) in 4 ml hybridization solution per 100 cm^2 membrane plus 100µg/ml nonhomologous DNA. This solution was removed and an appropriate volume replaced containing the radiolabelled probe (2ml solution/100 cm^2) and incubated for >16 hours. After incubation was complete the blot was washed 3 times for 30 minutes in wash buffer at room temperature and put to autoradiograph thereafter.

Pall solutions

Hybridization solution: 5 x Denhardt's buffer
5 x SSPE
0.2% SDS
100µg/ml nonhomologous DNA

Wash buffer: 5mM sodium phosphate pH 7
1mM EDTA
0.2% SDS

100 x Denhardt's buffer: 2% Ficoll
(filter sterilised) 2% Polyvinylpyrolidone
2% BSA

20 x SSPE: 3.6M NaCl
0.2M sodium phosphate pH 7.7
20mM EDTA

For both Pall and nitrocellulose membranes the nonhomologous DNA and the probe DNA were denatured by boiling for 10 minutes followed by their addition to hybridization solution prewarmed to the incubation temperature. The nonhomologous DNA used was salmon sperm DNA made up at 10mg/ml and sheared by repeated passage through an 18 guage needle.

At least 5 x 10^6 cpm (^{32}P) or (^{35}S) were used to probe a blot. Autoradiography was performed at -70°C and with intensifying screens for ^{32}P-labelled probes.

2.7.2 Probing DNA bound to nitrocellulose membranes

The membrane was prewetted using 2 x SSC prior to sealing into the hybridization bag (for ease of handling). The membrane was prehybridized for 1-2 hours at 68°C in the hybridization solution plus 100µg/ml nonhomologous DNA before the radiolabelled probe DNA was added and incubation continued for a further 16 hours. After incubation the membrane was washed progressively in 5 x, 1 x, and 0.2 x SET with 0.1% sodium pyrophosphate, 0.1% SDS added.

Hybridization solution: 4 x SET
0.1% sodium pyrophosphate
5 x Denhardt's buffer

20 x SET: 3M NaCl
0.4M Tris.Cl pH 7.8
20mM EDTA

2.8 Conventional agarose gel electrophoresis

Various gel sizes and buffers were used in the course of this investigation. The horizontal gel sizes ranged from minigels of 5 x 7.5 cms holding 25 mls of agarose to large gels of 16.5 x 20 cms holding 200 mls of agarose. Ethidium bromide was added while the agarose was molten at 0.5µg/ml to eliminate the necessity to stain the gel after electrophoresis. Electrophoresis buffers were 1 x TEA, 1 x TBE, and 0.5 x TBE. The size of DNA

restriction fragments (both visible and as revealed by radiolabelled probing) was estimated from graphs of the $\log_{10}$ molecular size plotted against the distance migrated according to the formula: logM = A - B x D (Helling et al, 1974)

M = molecular size in bp

D = distance migrated

B and A = arbitrary constants

Molecular size standards were obtained by the use of 123 bp or 1 kbp ladders from BRL (520-56135A and 520-56155A respectively) or the restriction of lambda cI857Sam7 DNA (Phillippsen et al, 1978).

2.8.1 Electrophoresis buffers

10 x TBE: 108g Tris, 55g Boric Acid, 9.3g EDTA made up to 1 litre with distilled water; pH is 8.3

10 x TEA: 48.4g Tris, 16.4g NaOAc, 3.6g EDTA made up to 1 litre with distilled water, pH adjusted to 8.2 with acetic acid.

2.8.2 Photographing of gels: Ethidium bromide stained gels were viewed on a 254nm uv transilluminator and photographed using a Polaroid camera loaded with Polaroid 4 x 5 Land film (no 57) or a Pentax 35mm SLR loaded with Ilford HP5 film. Both cameras were fitted with a Kodak Wratten filter No.9.

2.9 Extraction of DNA from agarose

Two procedures were employed to remove a desired band of DNA from agarose after gel electrophoresis and ethidium bromide staining. Both are described in Maniatis et al (1982).

If the agarose was low melting point the DNA band was excised using a scalpel while under long wave ultraviolet illumination. Five volumes of TE pH 8 were added to the agarose block and it was then heated at 65°C for 10 minutes to melt the agarose. After the solution had cooled to room temperature it was extracted four times against phenol and once against phenol:chloroform before the DNA was precipitated using ethanol. The recovery of the DNA was checked by gel electrophoresis after the DNA was resuspended in TE.

If the agarose was standard type the DNA band was visualised

using long wave ultrviolet illumination and a slit made in the gel in front of the DNA band. Into this slit was placed a piece of Whatman 3MM papaer backed by dialysis tubing. Electrophoresis was continued until the DNA had migrated into the Whatman paper. The paper and the tubing were then removed into a 400µl Eppendorf tube which had had a needle hole made in its bottom. This tube was placed inside a 1.5ml Eppendorf tube and the two centrifuged for 15 seconds. The eluate (in the 1.5ml tube) was removed and 100ul of elution buffer added to the paper and the tubes recentrifuged to recover the eluate. This was repeated twice more. The combined eluates were then extracted twice against an equal volume of phenol:chloroform and the DNA was thenethanol precipitated. Recovery of the DNA was checked using gel elctrophoresis.

Elution solution: 0.2M NaCl
50mM Tris.Cl pH 7.6
1mM EDTA
0.1% SDS

2.10 Descriptions of the Southern blots found in the text

All gels are orientated to show the wells at the top of the page and the direction of DNA migration is down the page. Unless otherwise stated all the gels were treated with HCl during the Southern transfer procedure and the autoradiography occurred at -70°C in the presence of intensifying screens.

2.10.1 Conventional agarose gels

All of the gels were run in 1 x TEA buffer, transferred to Pall membrane and probed with ^{35}S-labelled complete plasmids labelled using the nick translation procedure.

The specific details for each figure are given as follows: %agarose: probe used, incorporation as cpm x 10^8/ug DNA, cpm x 10^6 used: ug DNA/lane: autoradiograph incubation time

Fig 3.4a: 1%: pGS-216, 0.6, 0.2: 0.75: 1 month: No HCl
3.4b: 1%: pGS-268, 0.7, 6: 0.75: 2 months : No HCl
3.6 : 1%: pGS-268, 0.7, 6: 1: 4 days
5.4a: 1%: pGS-1, 1, 5: 1: 10 days
5.4b: 0.8% : pGS-1, 5: 0.75: 8 days

2.10.2 Pulse field gels

Each track was loaded with an agarose plug containing 5 x 10^5 cells and all probings were done using ^{32}P-labelled DNA fragments isolated from agarose. All probes were made using the random primer synthesis method which gave approximately 65% incorporation in all cases, equivalent to an incorporation of 4 x 10^8 cpm/ug of probe. Unless otherwise stated the gels were composed of 1% agarose in 0.25 x TBE. The specific details of each figure are given as follows: membrane type: cpm x 10^6 applied, fragment type: electrophoresis conditions- run time in hours, run voltage, switching frequency in seconds- for both first (i) and second (ii) regimes where appropriate: autoradiograph incubation time.

Fig 4.5: Pall: 4, I: 8, 300, 5: 2 days
4.7: Pall: 4.8, I: 7, 300, 5: 24 hours
4.9: Pall: 7, II: (i) 6, 300, 20, (ii) 16, 110, 5: 5 hrs
4.10b: Pall: 7, II: (i) 16, 105, 41, (ii) 6, 290, 5: 17 hrs
4.11a: Pall: 8, III: (i) 6, 300, 20, (ii) 16, 110, 5: 26 hrs

4.11b: Pall: 2, IV: (i) 16, 105, 41, (ii) 6, 290, 5: 17 hrs

4.13: Pall: 8, pGS-113: 15, 300, 50: 24 hrs: Gel was 1.5%

5.1: Nitro: 8, pGS-113: (i) 14.5, 117, 41, (ii) 6, 290, 5: 1 hr at room temperature

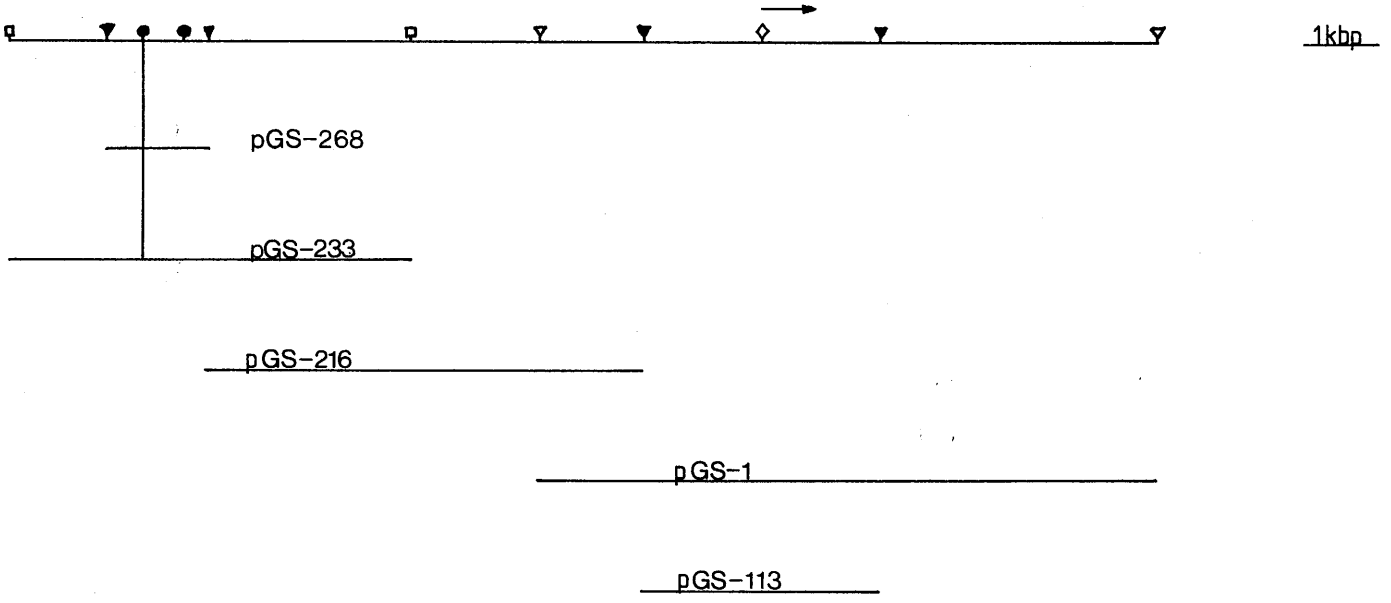
5.2: Nitro: 8, pGS-113: (i)16, 105, 41, (ii)6, 290, 5: 17 hrs

5.5: as fig 4.9

5.6: Nitro: 8, pGS-113: (i)16, 105, 41, (ii) 6, 290, 5: 1 wk

CHAPTER 3

Initial Experiments to Analyse the Glutamine Synthetase Gene.



Facing page 45

Fig 3.1 A diagram showing the cloned regions of the glutamine synthetase gene. The arrow indicates the location of the exon defined by S1 nuclease protection experiments. The lines below the restriction map define the genomic inserts in various plasmids and the sequences to the left of the vertical line are lambda sequences. The symbols are ● BamHI, ▽ BglII, ▽ EcoRI, ▼ HindIII and ◇ KpnI.

3.1 Isolation of the genomic glutamine synthetase gene

When high molecular weight genomic DNA from C4M is digested with restriction enzymes and size fractionated on agarose gels discrete bands are visible which are not present in similar restriction digests of KG1 DNA .By summing the sizes of these bands an estimate of the size of the amplified unit was made as 60 kbp. On the assumption that an increase in the gene copy number of GS and an increase in the level of GS protein would be matched by an increase in the level of GS mRNA reverse transcriptase was used to produce a radiolabelled probe complementary to total C4M mRNA. When this was used to probe size-fractionated restriction-digested C4M DNA an 8.2 kbp BglII fragment (among others) was identified. This 8.2 kbp band was purified from LMP agarose by Dr. P.G.Sanders and cloned into the BamHI site of the pUC-9 vector plasmid (Vieira and Messing 1982) to produce the recombinant plasmid pGS-1.

Dr. R.H.Wilson created a lambda library of C4M by cloning Sau3A-partially digested C4M DNA into the BamHI site of the lambda cloning vector L47.1 (Loenen & Brammar, 1980). Plasmid pGS-1 was then used as a probe by Dr. R.H.Wilson to isolate two overlapping lambda clones (86/2 and 11/3) which extended the cloned GS DNA to approximately 15 kbp.

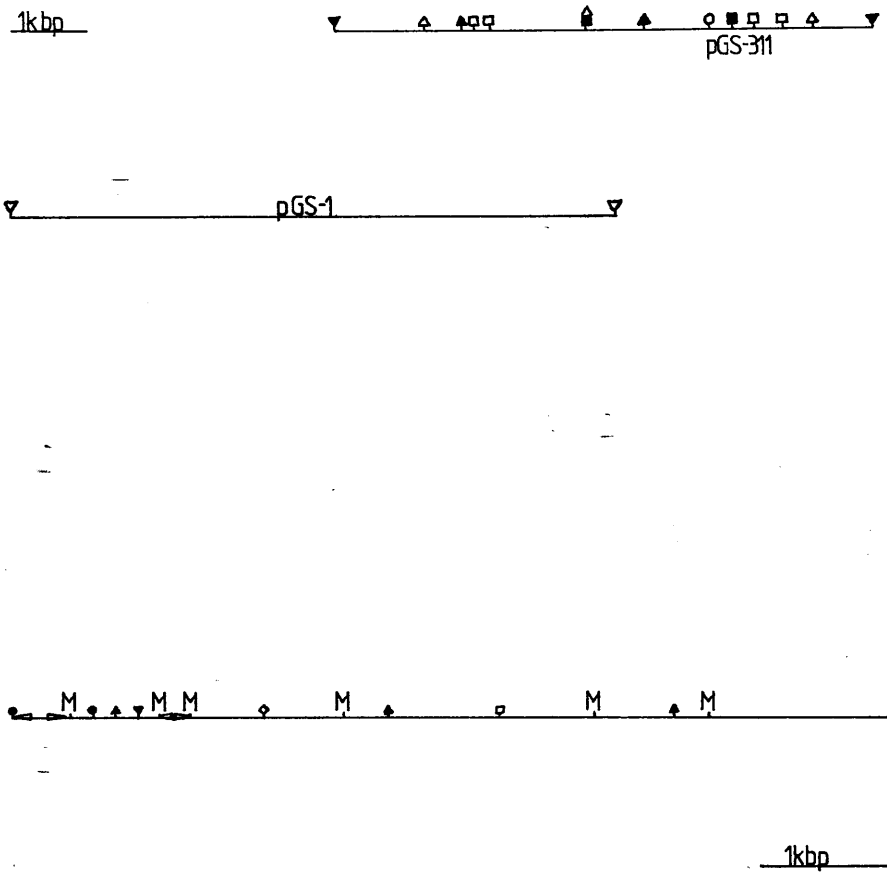
One major exon had been located (see fig 3.1) using S1 nuclease protection experiments.

3.2 Initial experiments to characterise the cloned region

Given a unit size of at least 60 kbp and a partially characterized region of 15 kbp as its only known region it was desirable both to further characterize the known region and to use probes from it to further extend the known region. Since it was likely that the DNA containing the GS gene would be relatively stable because of the selection for functional GS genes and hence DNA fidelity any information relating to the amplification process would only be obtained through study of the DNA at some distance from the GS gene.

3.2.1 pGS-311

Towards this end the genomic DNA within lambda 86/2 was subcloned by me into the pUC-9 vector. A HindIII digestion of



Figures 3.2 and 3.3

Fig 3.2 The restriction map of pGS-311 showing its overlap to the 3' end of pGS-1.

Fig 3.3 The restrction map of the inserts of pGS-216 and pGS-233 showing the location of the MspI sites. The map runs from the BamHI-lambda junction in pGS-233 to the 3' end of pGS-216 (left to right). The arrows indicate the extent of areas which could contain further MspI sites. The symbols are ▼ BamHI, ▼ BglII, ▼ EcoRI, ▼ HindIII, ◇ KpnI, M MspI, ▲ PstI, ▲ SstI, ▼ XbaI.

86/2 DNA was performed and the fragments produced cloned into HindIII cut pUC-9 and used to transform E.coli M15. One plasmid containing a 6.8 kbp insert represented the most 3' section of the cloned region (3' being defined with respect to the direction of transcription of the GS gene) and was designated pGS-311.

Further restriction analysis using a variety of single and double restriction enzyme regimes produced the map shown in fig 3.2 which did not differ from the map of the genomic DNA obtained by probing restriction digests with pGS-1. Having mapped pGS-311 it was screened for regions which were both 3' to the region contained in pGS-1 and free from repetitive DNA sequences. A Southern blot was made of various restriction digests of pGS-311 and this was probed with radiolabelled total KG1 genomic DNA. Any hybridization detected would be due to highly repeated sequences in the probe DNA annealing to similar sequences in the plasmid DNA. The result showed that the 3' region of pGS-311 contained extensive repetitive sequences and it could not therefore be used as a reliable probe to further extend our knowledge of the amplification unit. Further experiments involving the 3'-side of the GS gene were not therefore attempted.

3.2.2 Locating 5' HpaII sites using cloned DNA

Several plasmid subclones of the other, 5', lambda 11/3 had been made by Dr. P.G.Sanders and undergraduates A.Chudleigh and P.Bell. Two of these were the 5.5 kbp EcoRI fragment which included 1.9 kbp of lambda sequence contained in pGS-233 and the 5.9 kbp HindIII fragment contained in pGS-216.

On the assumption that the most 5' GS exon lay within the cloned region the CCGG sites within pGS-216 and pGS-233 were located by HpaII digestion. The commonest form of base alteration in eukaryotic DNA is the methylation of the cytosine of a CpG pair. The precise role of this modification, or lack of it, in gene regulation is still a matter of debate but CG-rich "islands" and hypomethylation are often associated with the promoter regions of eukaryotic genes. The restriction map of pGS-216 and pGS-233 is shown in fig 3.3. It can be seen that a cluster of HpaII sites exists at the most 5' area of the cloned C4M DNA- a result indicative of a CG "island" and by inference the start of

the GS gene (see Bird, 1986). Because the DNA was cloned, rather than native, digestion of it could not reveal the methylation status of the CCGG sites mapped.

3.2.3 5'-hypomethylation studies

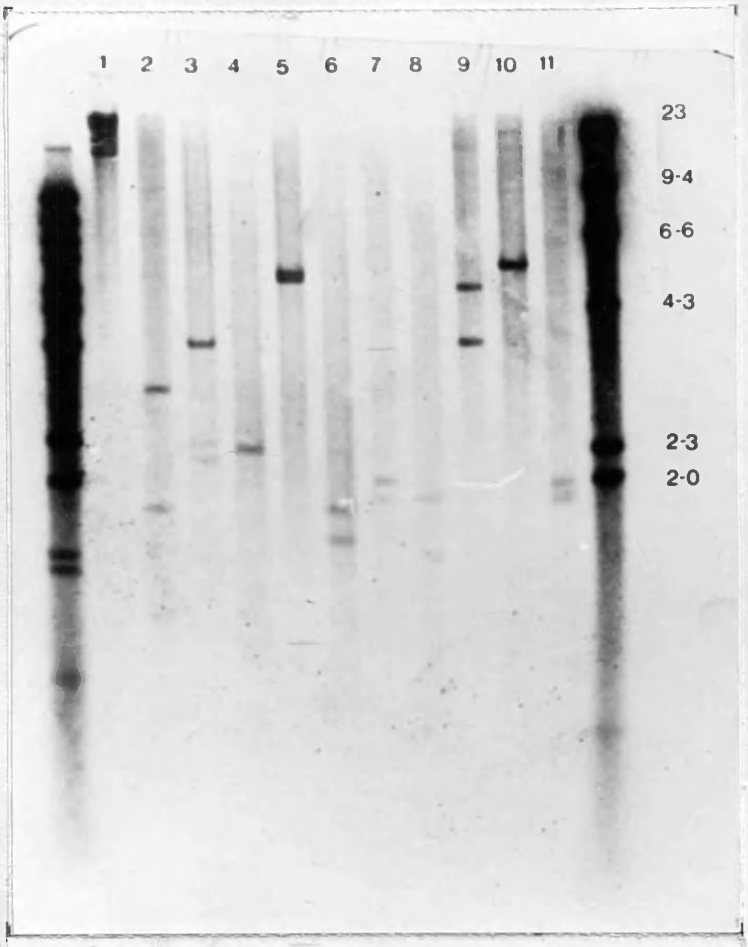
Two isoschizomeric enzymes- HpaII and MspI- exist for the CCGG sequence but differ in their sensitivity to methylation status. MspI is unaffected by methylation status and will cleave either CCGG or $CC^{Me}GG$ whereas HpaII will only cleave CCGG. To identify the methylated CCGG sites a series of parallel double digests utilising HpaII or MspI with a variety of other restriction enzymes were performed on genomic C4M DNA. The results were split, gel electrophoresed and transferred onto two identical Southern blots. One blot was probed with radiolabelled pGS-216 and the other with pGS-268. The results are shown in fig 3.4. Given the map of CCGG sites presented previously (fig 3.3) for the cloned region of the amplification unit an equivalent map can be constructed for the genomic data.

The pattern generated by MspI digestion of C4M DNA is, within the bounds of experimental variation, the same as that pattern found in the cloned DNA and this is to be expected as MspI is methylation insensitive. This identity of pattern between C4M DNA (the population of amplified units) and the cloned DNA (an individual of that population) indicates that lambda 11/3 did not, by chance, contain an atypical unit.

It can be seen from the two autoradiographs that (a) presents a simpler situation than (b) judged by the numbers of subsidiary bands present in (a) and (b). The plasmid mapping revealed regions containing clusters of CCGG sites whose individual sites lie too close together to be mapped, or numbered, using 1 % agarose gels. It should be noted that although MspI digestion will reveal only the positions of those CCGG sites at the extremities of the cluster these sites may be resistant to HpaII digestion while other, more central, CCGG sites are not.

Fig 3.5 shows the regions which would hybridise to the probes pGS-216 and pGS-268. Fig 3.4a shows that only the two largest of the several potential MspI fragments hybridised strongly with pGS-216 (these are marked by open boxes in fig

3.4a



3.4b

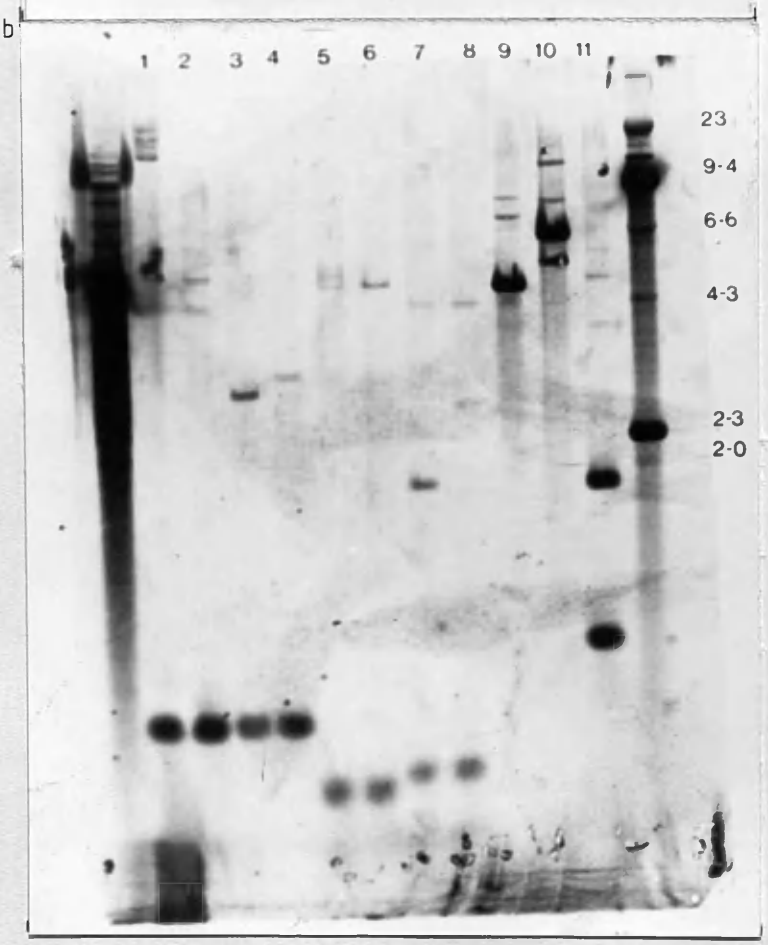


Fig 3.4 The mapping of the hypomethylated region towards the 5'-end of the GS gene using C4M DNA. 3.4a was probed with pGS-216 and 3.4b with pGS-268. In both figures:
Lane 1: C4M DNA digested with HpaII.
2: MspI
3: HpaII + EcoRI
4: MspI + EcoRI
5: HpaII + HindIII
6: MspI + HindIII
7: HpaII + PstI
8: MspI + PstI
9: EcoRI
10: HindIII
11: PstI
The size markers are in kbp. For details see p.43

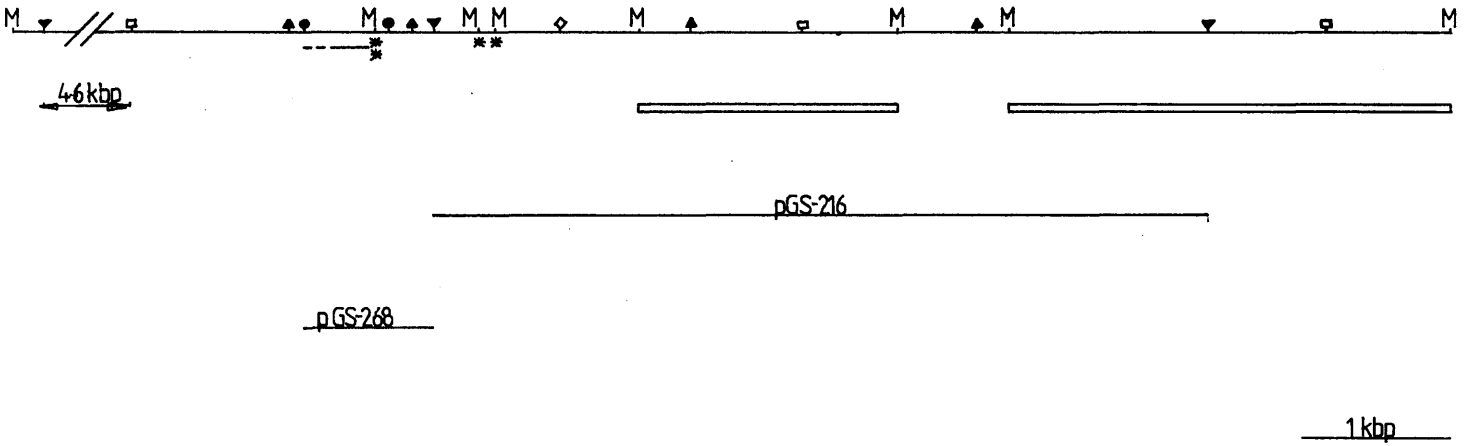


Fig 3.5 The restriction map of the 5' region of the genomic glutamine synthetase gene derived from analysis of figures 3.3 and 3.4. The open boxes show the MspI fragments which hybridised successfully to pGS-216 and below them are the regions covered by the pGS-216 and pGS-268 probes. The stars on the map show CCGG sites which are always unmethylated (two stars) or unmethylated in different units (one star). The broken line indicates a region which must contain more unmethylated CCGG sites. See fig 3.3 for the restriction symbols.

3.5). The failure of the smaller (<1 kbp) to hybridise is presumably due to poor transfer during the blotting procedure, a problem which did not arise in the gel which is represented in fig 3.4b. All fragments produced by EcoRI, HindIII, or PstI digestion alone hybridised successfully with pGS-216. Double digests with MspI and EcoRI, HindIII, or PstI produce the expected results. (EcoRI + MspI) digestion leaves only the truncated 3'-MspI fragment sufficiently large to hybridise successfully; (HindIII + MspI) digestion leaves the 5'-MspI fragment unaltered while truncating the 3'-MspI fragment; (PstI + MspI) digestion truncates the 3'-PstI to a size indistinguishable from its originally smaller more-5' fragments while further shortening the central PstI fragment and rendering the remains of the 5'-PstI fragment incapable of hybridizing successfully.

The results obtained with HpaII replacing MspI indicate that the majority of CCGG sites exist in the $CC^{Me}GG$ form and are resistant to HpaII digestion. HpaII digestion alone produces several fragments with sizes in excess of 10 kbp which indicates that (i) there are no unmethylated CCGG sites in the vicinity of the GS gene, excluding those marked in fig 3.5 and (ii) not all sites which can exist in the unmethylated form do so in every amplified unit.

The precise location of the HpaII sites (ie unmethylated CCGG sequences) can be derived from analysis of the double digests. (HpaII + EcoRI) digestion cleaves only the larger of the two EcoRI fragments to produce two fragments of approximately 2.2 and 2.1 kbp which indicates the presence of at least one HpaII site close to the 5' terminus of the HindIII fragment contained in the pGS-216 probe. (HpaII + HindIII) digestion produces a doublet of 5.4 and 5.2 kbp suggesting two HpaII sites are present at the 5' terminus. (HpaII + PstI) digestion patterns suggest that the lower of the PstI bands is in fact composed of two populations of DNA fragments only one of which is cleaved by HpaII digestion and that the fragments so produced are too small to be detected. The final result indicates the presence of two HpaII-sensitive sites separated by as little as 200 bp where (i) the more 5' of which is always unmethylated and the more 3' of which is only unmethylated in a proportion of the GS genes or

3.6

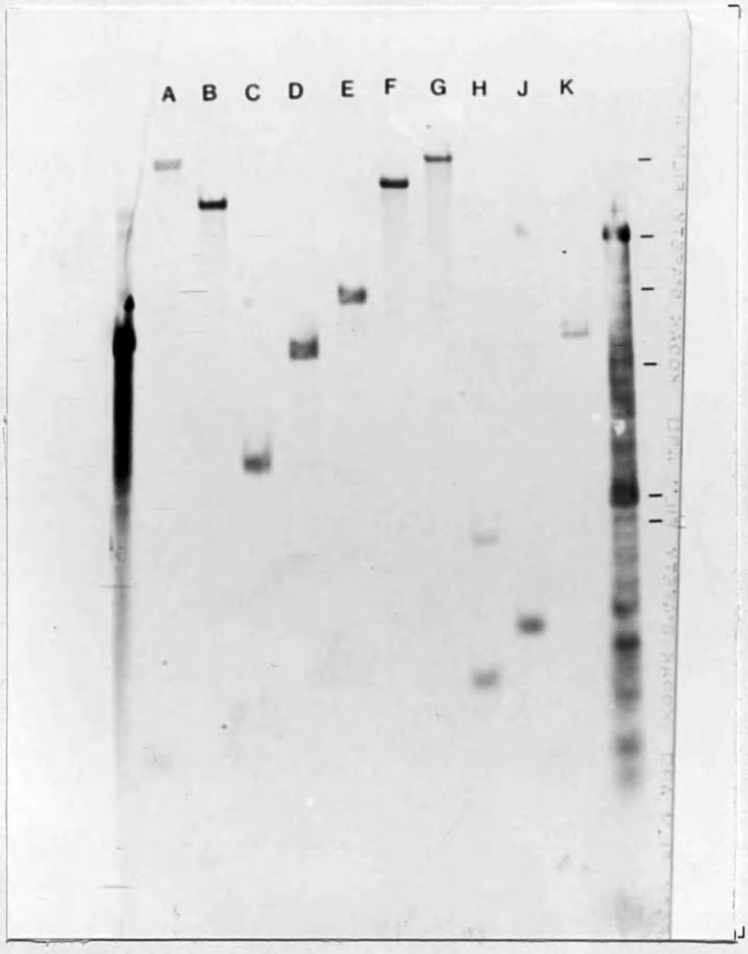


Fig 3.6 Upstream probing with pGS-268 of C4M DNA showing the homogeneity of the GS amplification unit. The size markers are lambda HindIII- 23, 9.4, 6.6, 4.3, 2.3 and 2.0 kbp. Lane A, C4M DNA digested with BamHI; B, BglI; C, BglII; D, EcoRI; E, HindIII; F, HpaI; G, KpnI; H, PstI; J, PvuII; K, XbaI.**For details see p.43**

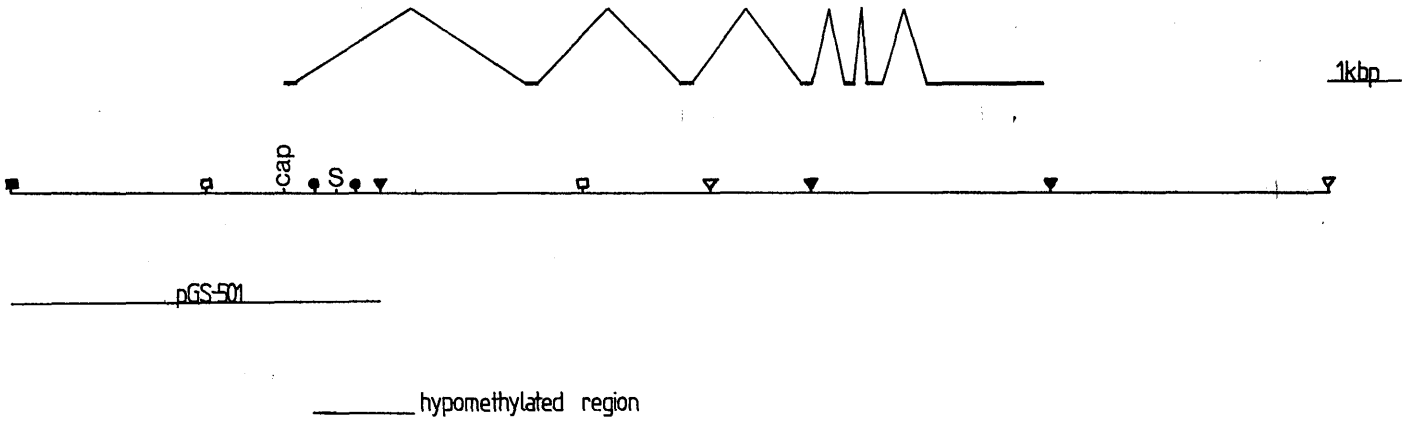
(ii) at least one or other site is unmethylated in every GS gene.

The data present in autoradiograph (b) confirms and extends the result derived from autoradiograph (a) while adding complexity in the form of many subsidiary bands. Considering only the major bands it can be seen that the major species is a very small fragment which is trimmed by both HindIII and PstI digestion. The location of the probe DNA is shown in fig 3.5 and it can be seen that a HpaII/MspI fragment whose 3' terminus was the 5' HpaII site in (a) would indeed be appropriately trimmed by further digestion with HindIII or PstI. Since the probe size is greater than the size of this sole fragment (870 bp and 600 bp respectively) it follows that since no other fragments are revealed after (HpaII/MspI + HindIII) digestion more unmethylated CCGG sites must exist beyond the 5' terminus of this fragment which are sufficiently close to one another to generate fragments too small to be detected. The appearance of a single HpaII fragment also indicates that the 3' unmethylated CCGG (equivalent to the 5' CCGG of (a)) must always be unmethylated.

The data shows that of the CCGG sites in the environs of the GS gene only a subset are usually found in the unmethylated state and that this unmethylated subset exists as a tight cluster of sites towards the 5' end of the GS gene and suggest the existence of a CG-rich 'island'. In this cluster one or other of the two most 3' sites are unmethylated in the GS gene sequences while the next site is always unmethylated and those in the 5' tight cluster are probably only unmethylated in a proportion of the GS gene sequences.

A complete analysis of the minor bands of autoradiograph (b) is not possible due to the plethora of bands. Two points are worthy of note however. The next 5' CCGG sequence would appear to lie close to the 3' side of the next HindIII site and would also appear to mark the beginning of another cluster of CCGG sites which show variation in their methylation status. The minor band beneath the major HindIII band suggests that a small proportion of the amplified units have undergone a rearrangement or deletion which has reduced the size of this HindIII fragment.

3.3 Extending the restriction map of the GS unit



Facing page 51

Fig 3.7 Up-to-date map of the glutamine synthetase gene. Above the restriction map is the exon/intron structure of the GS gene with the heavy lines (exons) connected by light lines (spanning the introns). The positions of the cap site (cap) and the SstII site (S) are marked and the regions covered by the insert in pGS-501 and the hypometylated areas are indicated.

To map the DNA beyond the cloned region C4M genomic DNA was cut with a variety of restriction enzymes, gel electrophoresed and Southern blots made. These blots were probed with various radiolabelled fragments derived from the cloned region and the results obtained are shown in fig 3.6. The results show that the structure of the amplified units is homogeneous for a distance of approximately 20 kbp upstream of the cloned region.

3.4 Summary

An 8.2 kbp BglII band was cloned into pUC-9 after its identification by differential hybridization of radiolabelled total C4M and KG1 cDNAs. The cloned region was extended further with two overlapping fragments derived from a C4M lambda library. Radiolabelled cloned probes were used to map the proximate regions of the amplification unit and to identify a hypomethylated CCGG cluster towards the 5' end of the glutamine synthetase gene.

3.5 Further developments

Heteroduplex mapping by Dr. L.Coggins (Beatson Institute, Glasgow) using cloned DNA and C4M mRNA revealed the intron-exon structure of the GS gene and I participated in the cloning of the GS cDNA (Hayward et al, 1986). Sequence analysis of the cDNA and corresponding genomic regions (A.Hussain, this laboratory) showed the presence of an untranslated 5' exon not contained in the previously cloned DNA. By cloning a 5.0 kbp XbaI-HindIII fragment from low melting point agarose C.Gibbs (this laboratory) isolated pGS-501 which contained the 5'-untranslated exon. This data is shown in fig 3.7.

It can be seen that the location of the hypomethylated sites does correspond to the 5' region of the GS gene and sequence analysis and a more detailed restriction analysis by other members of the group have shown that this region is rich in CG sequences.

CHAPTER 4

The Amplification Unit of KG1MSC4-M Cells

Figure 3.6 showed that the restriction patterns of the DNA surrounding the GS gene was the same for a distance of approximately 40 kbp taking into account the cloned region as well. As far as could be investigated, therefore, with the available probes the DNA was homogeneous and the relationship of one GS gene to another was still obscure as was the form of the DNA outwith the probed region.

To try to gain some measure of any variation in the structures of the amplified units an attempt was made to produce a probe which contained all amplified DNA. This entailed a procedure modified from that detailed by Caizzi and Bostock (1982). C4M genomic DNA was sheared and radiolabelled. After heating to 100 ^{o}C to completely denature it it was allowed to reanneal to $C_{o}t$ 10^{-1} at which point only sequences present in copy numbers in excess of 10000 would be double stranded. Hydroxyapatite was used to remove these double stranded sequences and the remainder of the DNA was split into two portions. To one portion nothing was added and to the other a 1000-fold excess of cold KG1 DNA. This competitor DNA would have the effect of diluting 1000-fold the specific activity of any non-amplified sequences while only decreasing 2-fold the specific activityof sequences amplified 1000 times. The probes created both with and without competitor DNA were used to probe identical genomic DNA Southern blots to see if a subset of bands which hybridized to probe minus competitor would hybridize to probe plus competitor and would thus reveal a collection of amplified DNA fragments. In the event no hybridization was detected and no further attempts were made to pursue this line of enquiry.

4.1 The Plan to map the amplification units

Analysis of the region of the amplification unit accessible by the probes available showed that the data obtained could not shed light upon the rest of the amplification unit nor reveal the existence and nature of any DNA interspersed between or flanking the units.

Three approaches were planned to investigate the amplification phenomenon on a larger scale than could be attempted using conventional agarose gel electrophoresis. The

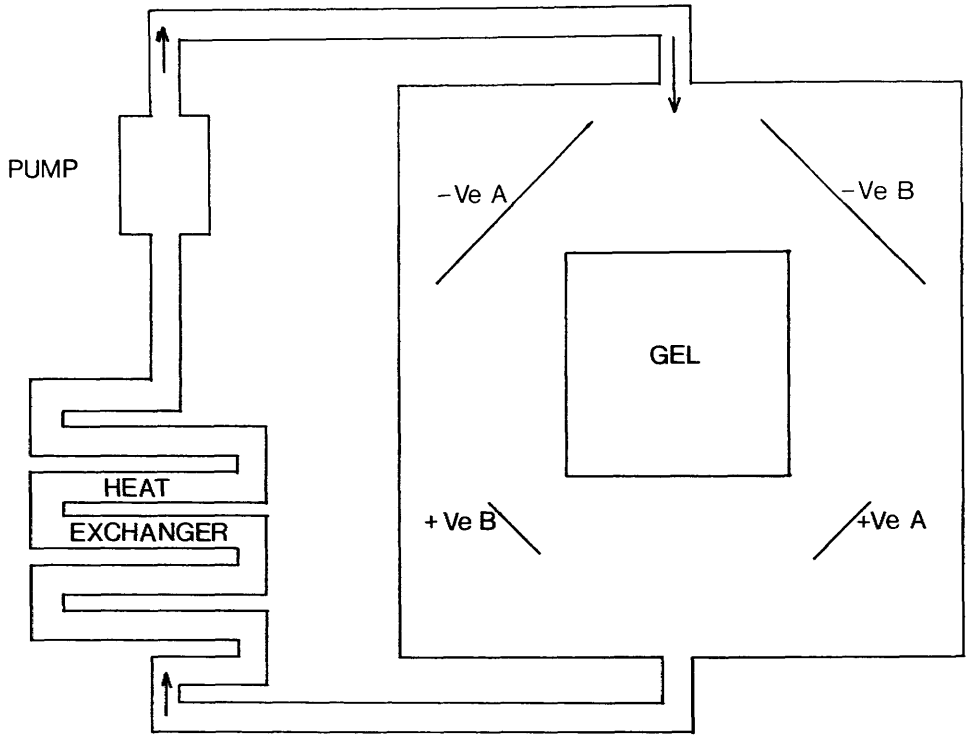


Fig 4.1 A diagrammatic representation (not to scale) of the pulse field gel electrophoresis equipment. It shows the components of the peripheral system, the direction of buffer recirculation and the positioning and relative sizes of the electrodes.

Facing page 53

first of these approaches is reported in chapter 6. It involved a karyological examination of the parental and amplified cell lines and aimed to discover the rearrangements at the chromosomal level associated with the amplification of the glutamine synthetase gene. An *in situ* hybridization to locate the amplified GS genes was performed to discover whether the amplification units were in a single cluster, several or scattered throughout the genome which result would bear upon the question of structural variation within the units and the nature of any flanking DNA.

The second approach was to attempt to isolate several examples of the uncloned regions of the amplification unit. An analysis of the restriction patterns of these clones would have shown any variation in the structure of the amplified units and whether the amplified units were separated from each other by stretches of "random" DNA. Numerous attempts were made to construct a C4M cosmid library by ligating Sau3A-partially digested C4M DNA to BamHI-digested cosmid vector pJB8 and to package the result using an *in vitro* packaging system but no attempt was successful.

The third approach was to use pulse field gel electrophoresis to separate large DNA restriction fragments to attempt to define an entire amplification unit.

4.2 Pulse Field Gel Electrophoresis

This technique was developed by Schwartz & Cantor (Schwartz *et al*, 1982; Schwartz & Cantor, 1984) to separate very large DNA molecules. It has been used to visualise yeast (Carle & Olson, 1985) and trypanosome (van der Ploeg *et al*, 1984a & b; Johnson & Borst, 1986) chromosomes with important consequences for karyological analysis and gene expression studies. With the advent of restriction enzyme digestion techniques allowing the digestion of DNA embedded in agarose (Bernards *et al*, 1986) several groups have applied PFGE to mammalian genetics. Brown & Bird (1986) investigated C+G 'islands', Hardy *et al* (1986) the human major histocompatability complex, Graham *et al* (1987) the chorionic gonadotropin B-subunit gene and Kenwrick *et al* (1987) the Duchenne muscular dystrophy region.

4.2

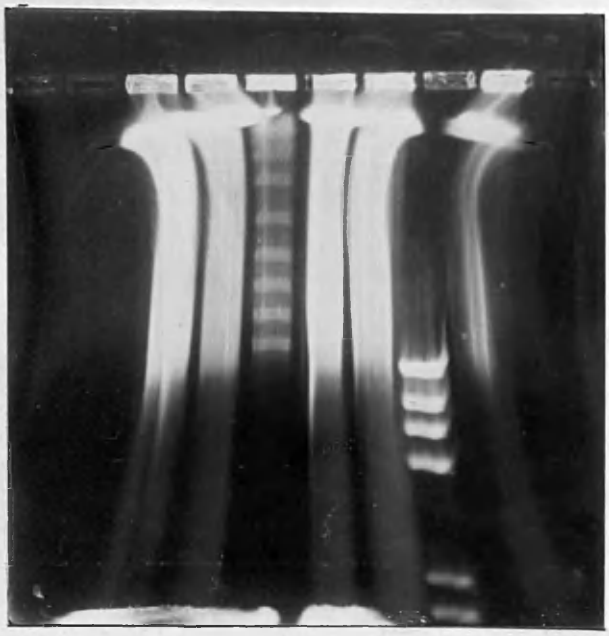


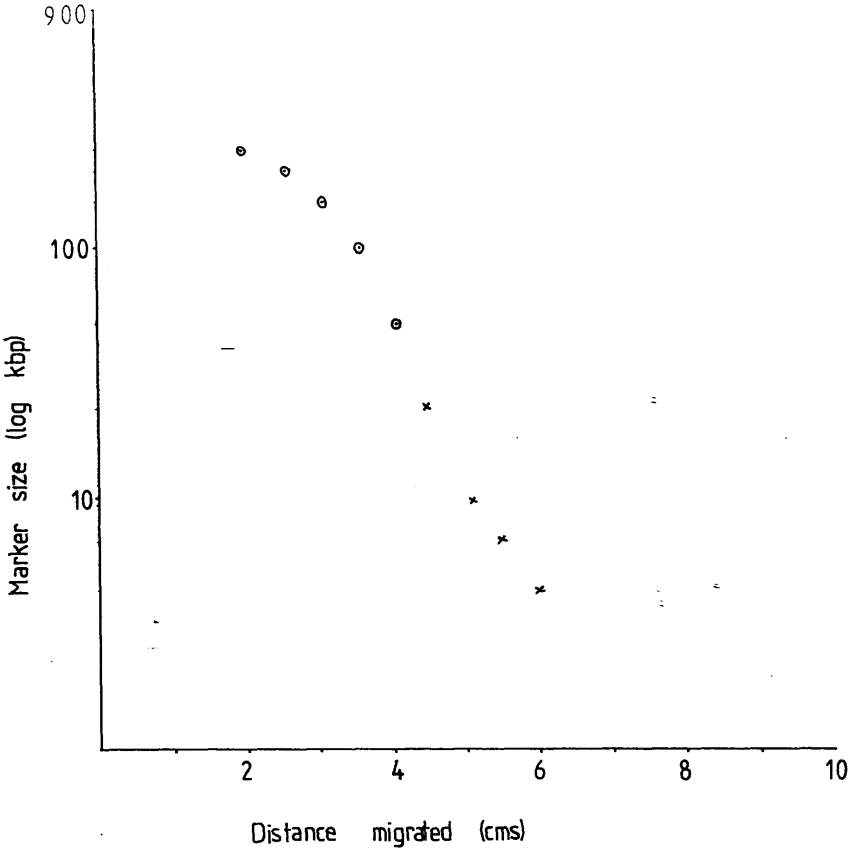
Fig 4.2 The compression zone of a pulse field gel electrophoresis run (indicated by the black lines).

Several variations in design now exist (Anand, 1986; McPeek et al, 1986; Carle et al, 1986; Cardiner et al, 1986). The apparatus used in my experiments was built in the Genetics Dept. workshop of Edinburgh University to a design of Carle & Olson (1984) and is diagrammed in fig 4.1.

In conventional gel electrophoresis a single continuous electric field is applied parallel to the gel. The DNA, being negatively charged, migrates towards the positive electrode at a rate related to its length. In this system DNA fragments can be envisaged as rigid structures which migrate through the matrix formed by agarose molecules. If the DNA fragments are above a certain length they are uniformly retarded in the matrix and are not resolved. Pulse field gel electrophoresis operates by exposing the DNA to a variable electric field supplied, in this case, by the alternate use of two pairs of electrodes. Because the direction of the field changes as the electrode pair activated is changed the DNA fragments are required to alter their orientation within the gel matrix. The fragments are not rigid while this process occurs and can therefore migrate through the gel matrix. The ease with which this reorientation can be performed decreases with increasing fragment length and therefore migration distance through the gel becomes, as with conventional gel electrophoresis, a measure of the size of the DNA fragment.

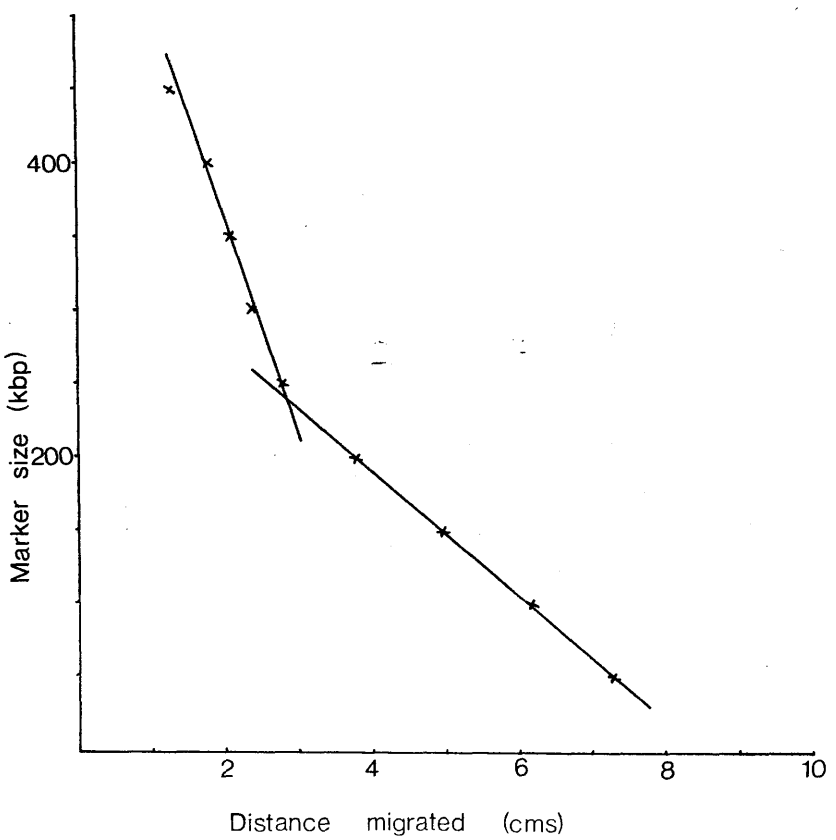
The time taken by the DNA fragment to completly reorientate to the new direction of the electric field is a function of its length. By choosing one or more switching frequencies the maximum separation of the size range of interest can be achieved. To separate fragments of up to 400 kbp I used two switching frequencies- 40 seconds or 20 seconds in the first part of the separation and 5 seconds in the second part. The penalty for achieving this resolution was the presence of a severe compression zone where the majority of fragments longer than 400 kbp accumulated (see fig 4.2). Separation of molecules of up to 2000 kbp (eg yeast chromosomes) requires a longer switching frequency (50 seconds), a higher voltage (300 V) and a longer run time (18 hours).

Because of the size of the fragments which can be resolved using pulse field gel electrophoresis is so large the handling of



Figures 4.3 and 4.4 Graphs showing the migration characteristics of DNA fragments of various lengths in the PFGE apparatus.

4.4



the DNA requires that opportunity for mechanical shearing should be kept to a minimum. This requirement rules out the use of DNA in solution and techniques for embedding living cells (or organisms) in agarose and then making the DNA they contain accessible have been developed. The technique I used to encapsulate, lyse and restrict DNA was a version of that of Bernards et al (1986) with PBSA solution, routinely used in cell culture, replacing the (PSG) solution used for trypanosome handling. The size markers used were concatamers of lambda DNA which were formed by the annealing of cos sites of lambda suspended in agarose.

The system comprised a gel tank in a closed circuit of tubing through which the electrophoresis buffer was recirculated and which allowed the system to be maintained at 10-15°C by passage through a heat exchange system. The volume of the system was 2.5-3 litres and of this volume 1.5 litres was replaced with fresh buffer prior to each run. Although the voltage required to operate the system was relatively modest, at most 300 V, the stress caused by momentary breaks in the circuit during the mechanical switching from one electrode pair to the other required a robust power supply. The power supply used initially a Chandos E.37 was unable to meet these requirements of the system and was substituted with another sequencing powerpack, an LKB 2197.

The buffer originally used was 1 x TBE but it was found that the ionic strength of this buffer was too great and caused the current to increase as the electrophoresis continued eventually causing the voltage to drop as the powerpack reached the limits of its output. This problem was remedied by the introduction of 0.25 x TBE as the electrophoresis buffer.

4.3 PFGE Experiments

It will be seen from the various photographs and autoradiographs presented in this section that DNA fractionated by pulse field gel electrophoresis adopts a curved track through the gel. A graphical representation of the straight line length of migration, that is one taking no account of the lateral migration of the DNA, reveals that there are two types of

4.5

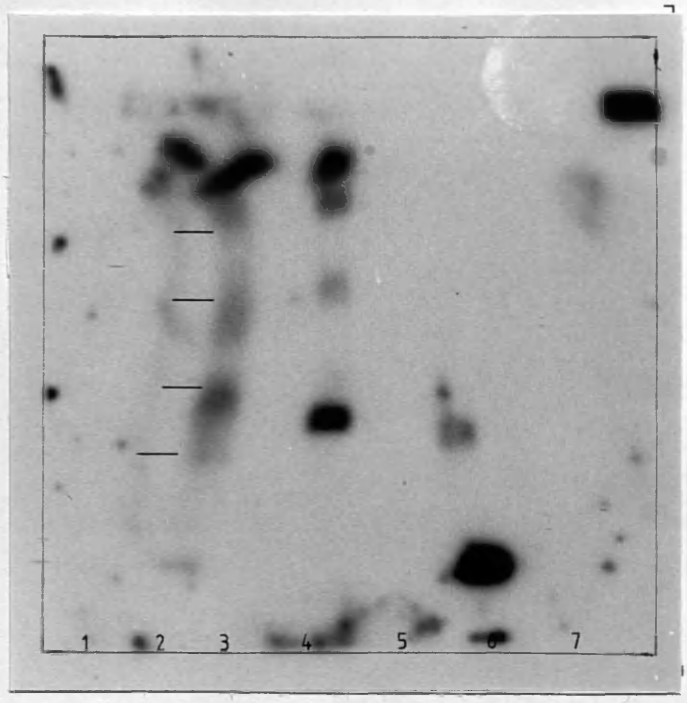


Fig 4.5 Southern blot of pulse field gel (PFG) probed with fragment I (see fig 4.6). Lane I is lambda multimer markers- 50, 100, 150 and 200 kbp; lane 2 C4M, lane 3 KG1 DNA digested with BamHI; 4 (C4M), 5 (KG1) with SstII; and 6 (C4M), 7 (KG1) with XhoI. **For details see p.43**

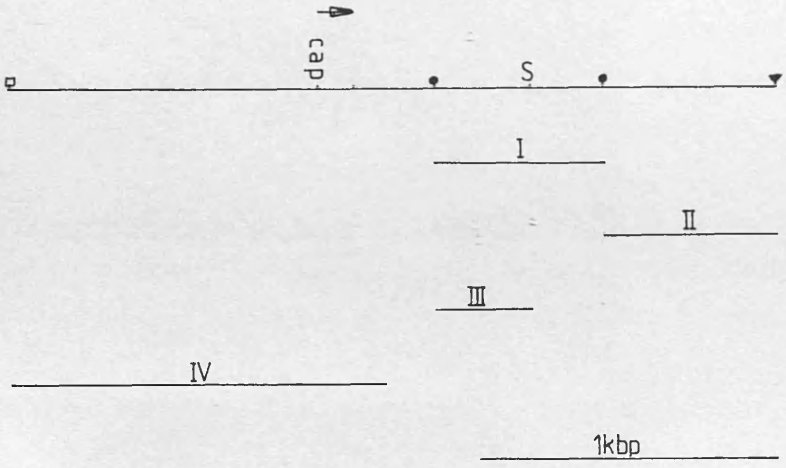
separation (fig 4.3 and 4.4) Those fragments which would normally be resolved on a conventional 1% agarose gel (<30kbp) migrate a distance proportional to the log of their size while larger fragments migrate a distance proportional to their linear length. The migration distance of these larger fragments also appears to be related to the pulse frequencies used in that a graphical representation will show two gradients, one created by the movement of very large fragments (>200kbp) and one by the lesser fragments (see fig 4.4) which could be generated by the greater ease of migration of smaller fragments at faster pulse frequencies.

Unfortunately the distortion due to curvature cannot be discounted when attempting to compare fragments from different lanes and this "curvature factor" means that the sizes of fragments can only be estimated using the marker lane as a reference.

The first successful PFGE experiments were performed using a 5 second switch frequency alone which resolved up to 200 kbp (see fig 4.5) and utilised restriction enzymes which were known to produce large fragments in the region of the GS gene. The PFGE gel was Southern blotted and probed with the fragmentI shown in fig 4.6. The autoradiograph shows that (i) DNA fragments specific to the amplified line are being detected and (ii) the digestion procedure itself does not significantly degrade the DNA and (iii) the detection of more than one fragment in each of the C4M digestion lanes indicates a complex situation exists. Lacking markers for sizes below 50 kbp the size of the major XhoI fragment cannot be estimated but, based on conventional gel electrophoresis sizing, it should be 18 kbp. The secondary XhoI is approximately 70 kbp in size. The three SstII fragments have sizes of 80, 160 and 210 kbp. The BamHI fragments are poorly resolved but a similar situation of multiple bands is suggested in the autoradiograph. Two explanations could account for this multitude of bands. Firstly they could represent partial digestion products of a homogeneous and ordered array as the presence of discrete bands larger than the first fragment (which must contain the known homogeneous region of the GS gene) can only be accounted for if all GS genes possess the same outlying

4.6

Fig 4.6 Diagram of fragments used as probes in the PFGE experiments. The restriction map is a detail of the region around the start of the GS gene (see fig 3.7), the arrow is the first exon and the direction of transcription.



4.7

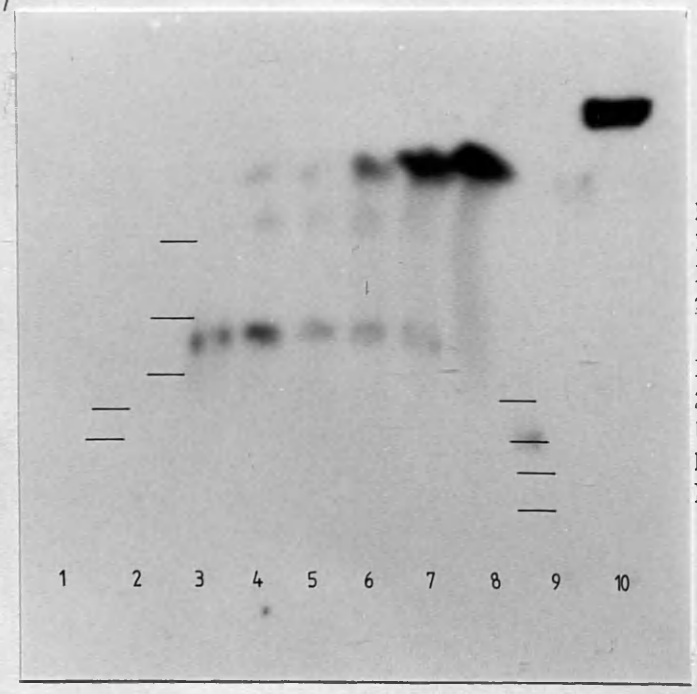


Fig 4.7 Southern blot of PFG probed with fragment I. Lane 1, λ XhoI markers; 2, lambda multimers; 9, λ HindIII. Lanes 3-7 C4M DNA digested with SstII: 25, 6.4, 1.6, 0.4 and 0.1 U. Lane 8, with ClaI. Lane 10, undigested.
For details see p.43

restriction patterns. Alternatively the bands could represent several families of restriction polymorphisms, each one homogeneous but differing in the position of the SstII site distant from the GS gene.

To distinguish between these two possibilites a series of digests were performed using different amounts of SstII enzyme from 0.1 U to 25 U (see fig 4.7) and the result shows that the multiple bands are due to partial digestion of an ordered array as the polymorphic population explanation would predict that the relative amounts of DNA in each fragment would be independent of enzyme concentration. The sizes of the two SstII fragments resolved are 75 and 150 kbp.

4.4 The initial model of the amplification units

The fact that a summation of the sizes of C4M-specific restriction fragments gives approximately 60 kbp of highly amplified DNA when coupled with the data from SstII digestions suggests that the 75 kbp SstII fragment could represent a complete amplification unit and that these units could be arranged in arrays such that the 150 kbp fragment could represent two complete amplification units and so on (see fig 4.8).

This idea predicts that any other restriction enzymewhich cuts the amplification unit once will also yield fragments the same size as those generated by SstII cleavage. The recognition sequence of SstII is CCGCGG and contains two CpG sequences which are rare in eukaryotic DNA. In an attempt to find other equally rare sites several restriction enzymes were employed- MluI (ACGCGT), NotI (GCGGCCGC), SalI (GTCGAC) and SfiI ($GGCCN_5GGCC$)- all of which would be expected to cut rarely. Digestion with MluI produced a fragment which migrated no further than the compression zone indicating a size in excess of 400 kbp and the NotI digestion either failed to work or produced a fragment so large it was unable to migrate from the well. Both the MluI and NotI lanes showed DNA had been digested when compared to the DNA of the control, sham, digest lane; data not shown.

Fig 4.9 shows the results of SalI, SfiI and SstII digestions of C4M DNA when Southern blotted and probed with radiolabelled fragmentII (see fig 4.6). The sizes of the SstII and major SalI

4.8

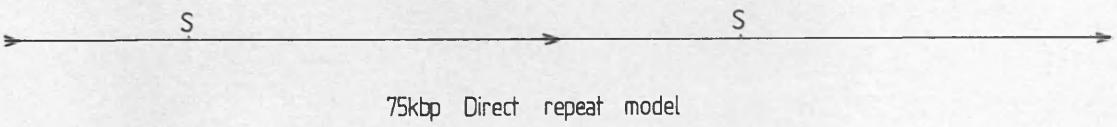


Fig 4.8 Initial model of the amplification units. The arrowheads define the end of a unit and its orientation. S is SstII.

4.9

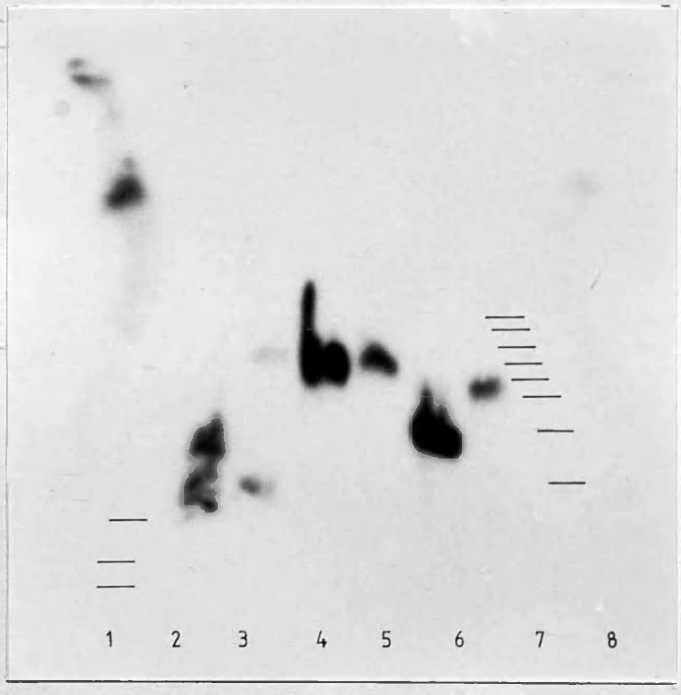


Fig 4.9 Southern blot of PFG probed with fragment II. Lane 1, λHindIII markers (23, 9.4 and 6.6 kbp); lane 8 lambda multimers. Lane 2, C4M DNA digested with SalI; lane 4, SfiI; lane 6, SstII.
For details see p.43

fragments lie between 60 and 90 kbp while the size of the SfiI fragment lies between 180 and 290 kbp. The subsidiary SalI fragments have sizes of 30 and 40 kbp. The apparent identity in size between the SstII fragment and the major SalI fragment lends credence to the concept of an amplification unit approximately 75 kbp in size which contains a series of SalI sites 3' of the GS gene. The size of the SfiI fragment (the true size of this fragment is almost certainly less than the figure stated because of the "curvature factor") is however greater than that of the SstII fragment and cannot be accounted for within a head-to-tail arrangement of homogeneous 75 kbp units. One possible explanation could be the occurrence of a SfiI recognition sequence in every other unit in the array though the existence of this degree of order in unit arrays is less easy to credit than a simple array.

An alternative is that the SfiI fragment represents the true unit size and that it is fortuitous that the two SstII fragments contained within it are the same size and hence give a misleading partial digest gel; but further increasing the size of the amplification unit conflicts with the length of the amplified DNA derived from conventional gel electrophoresis.

4.5 Later models of the amplification units

A series of SfiI or SstII partial digests of C4M DNA were made and size separated by PFGE, Southern blotted and probed with fragment II in order to ascertain the nature of the amplified units. The result is shown in fig 4.10. The sizes of the fragments are listed in the legend and were estimated taking account of the "curvature factor". Assuming that the first three SstII fragments (numbered s= 1-3 up the side of fig 4.10b) are those fragments visible in the photograph of the gel (4.10a) their size can be taken directly from the photograph. It is apparent that the sizes of the SstII and SfiI fragments preclude the simple direct repeat structure proposed in fig 4.8 and that previous estimates of the size of the first SstII fragment were in error.

Fig 4.11 shows the hybridization pattern found when the Southern blots of fig 4.9 and 4.10b were stripped of their first probe and reprobed with fragments III and IV respectively. Fig 4.11a shows that the size of the C4M SalI fragments revealed has

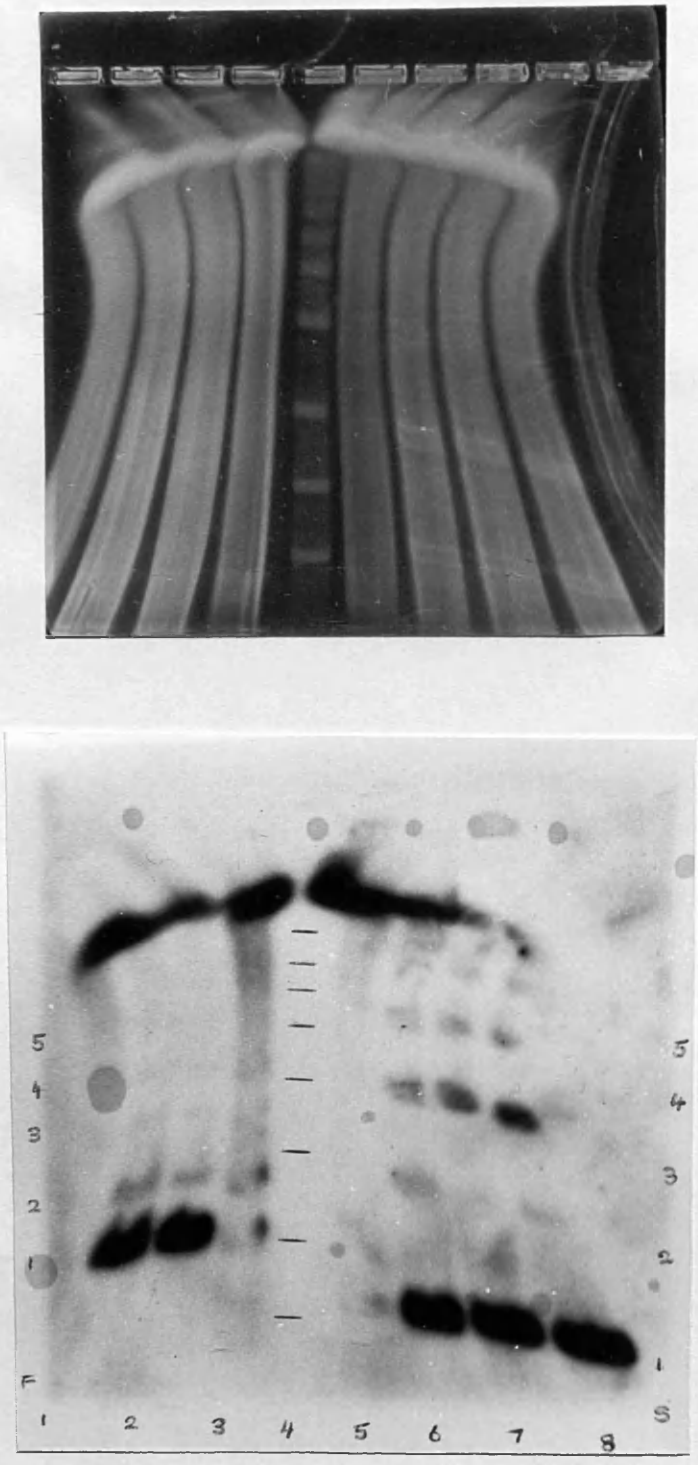


Fig 4.10 PFG of C4M DNA digested with various amounts of SfiI (lanes 1-3) or SstII (lanes 5-8). 4.10a shows the ethidium bromide stained gel showing the SstII doublet structure while 4.10b shows the equivalent Southern blot probed with fragment II. For details see p 43. The five SstII bands (s=1-5) have sizes of 55, 140, 195, 250-275 and 350-400 kbp while the five SfiI bands (f=1-5) have sizes of 105, 140, 185, 210 and 250-275 kbp.

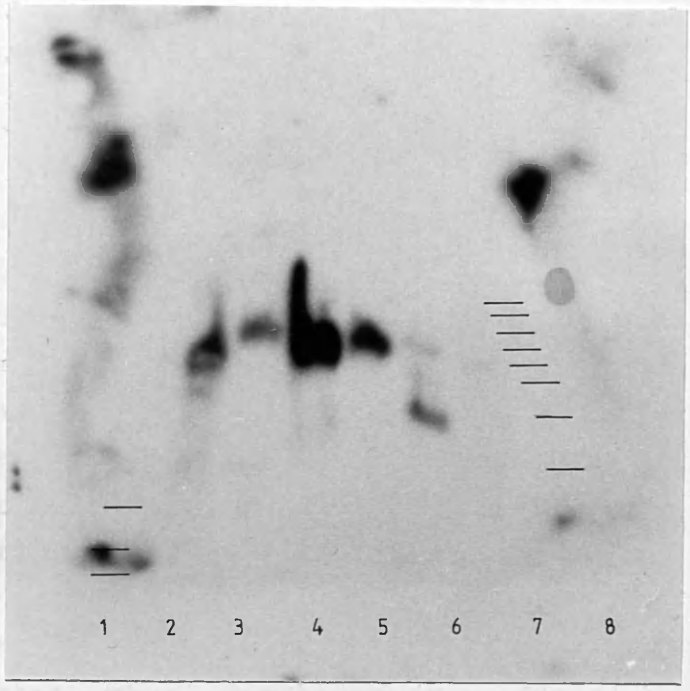
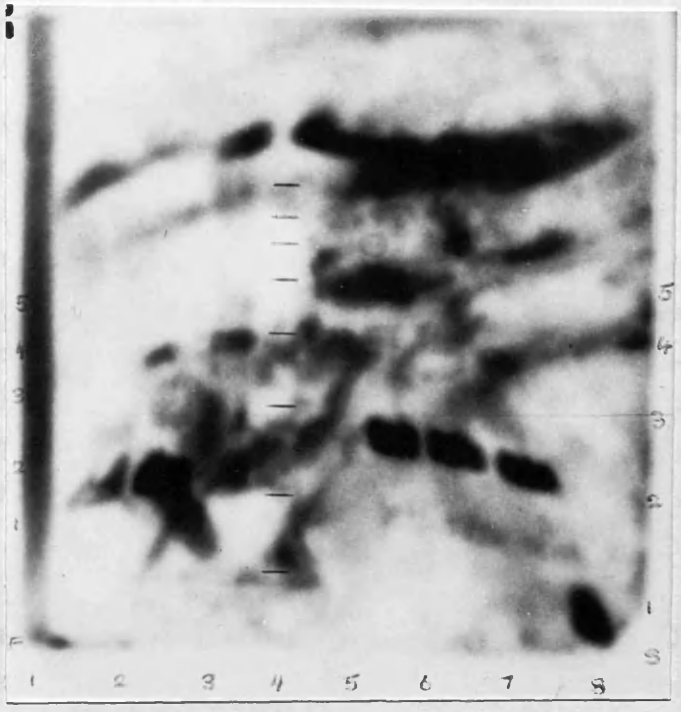


Fig 4.11 Reprobings of the Southern blots of fig 4.9 (4.11a) and 4.10b (4.11b) with fragments III and IV respectively to show the differences in hybridization patterns between fragment II and fragments III and IV. See p 43 for details.



increased above the size of the first SstII fragment and that the intensity of signal from that SstII fragment has dropped dramatically compared to fig 4.9. Fig 4.11b shows that, when probed with fragment IV, the first SstII fragment is not revealed at all while the first SfiI fragment is still shown.

Consideration of the data contained in figs 4.10b and 4.11b allows certain proposals to be made about the structure of the DNA containing the amplified glutamine synthetase gene. The s=2 SstII fragment is revealed relatively poorly by fragment II probing, that it is revealed at all might in fact be an artefact due to the fact the fragment II probe was prepared from a plasmid (pGS-233) which contains sequences 5' of the closer SstII site whose size is sufficiently close to that of fragment II to allow for some contamination during fragment preparation.

It follows from the hybridization patterns of the s=1 and s=2 SstII fragments that they represent two distinct SstII fragments, one 55 kbp long (s=1) containing the bulk of the GS gene and one 140 kbp long (s=2) containing the promoter region of the GS gene. Since both contain portions of the GS gene they must be linked in a consistent fashion throughout the amplified unit population. Assuming that an ordered structure does not cease beyond the SstII sites at the ends of the s=1 and s=2 fragments or that the next SstII fragments are large enough to migrate in the compression zone it follows that the s=3, 4 and 5 bands are associated with the s=1 and 2 fragments and that these bands are also consistent in size throughout the amplified unit population.

Fig 4.10b shows no hybridization in the SstII lanes around 110 kbp. This lack means that two s=1 fragments cannot be placed together in the unit structure. The s=3 band is 195 kbp in size and contains a fragment II sequence and therefore a s=1 fragment: s=3 is composed of two SstII fragments, one 55 kbp and one 140 kbp- which is remarkably close to the size of s=2. Since the s=1 and s=2 fragments must adjoin the observation that fragment IV does not conclusively reveal the s=3 band must be artefactual in nature. The s=4 band has a size between 250 and 275 kbp and contains both fragment II and fragment IV sequences which means it cannot be constructed of two s=2 fragments alone (=280 kbp) and therefore it is most likely to consist of two s=1 fragments

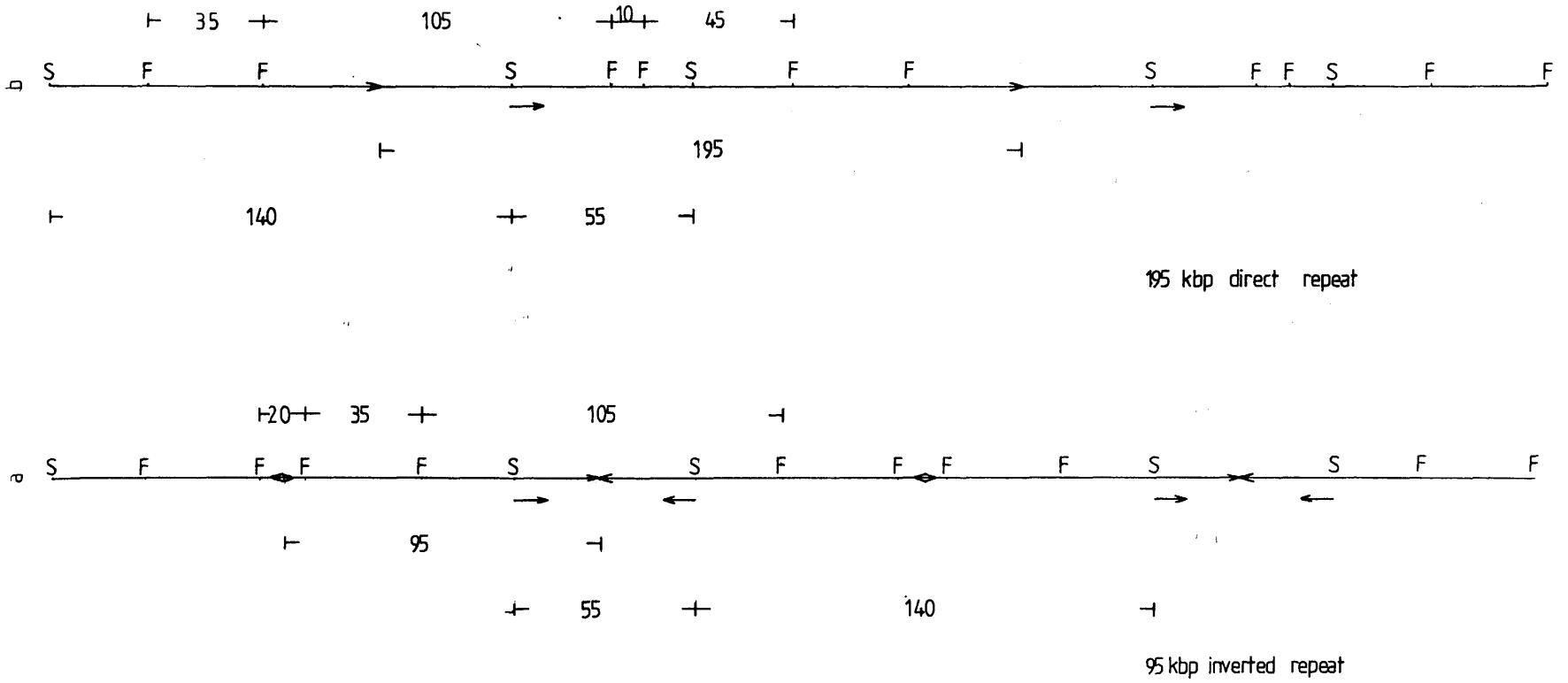


Fig 4.12 Later models of the amplification units - 4.12a shows the inverted repeat model and 4.12b the direct repeat model. In 4.12a the ⟶✕⟵ and ⟶◇⟵ represent the inversion points while in 4.12b ⟶ represents a nominal end of the unit. In both diagrams S represents a SstII site, F a SfiI site and → the glutamine synthetase gene. The various distances are in kbp.

Facing page 61

and one s=2 fragment (=255 kbp). The s=5 band has a size between 350 and 400 kbp and two s=1 and two s=2 fragments together will give 395 kbp.

Having established that neither the s=1 nor the s=2 fragments appear to join directly to another of their own fragment type and since the s=3, 4 and 5 bands do not require the existence of further SstII fragments to account for their sizes the proposed map of the SstII fragments is simple. It consists of alternating s=1 and s=2 fragments.

A similar analysis can be done on the SfiI fragments. The f=1 SfiI fragment is 105 kbp in size and is revealed by both fragment II and IV which means the f=1 fragment must overlap the s=1 and s=2 SstII fragments. A variety of possible overlaps can be considered. The placings of the SstII fragments does not impose an orientation in that they could be in a direct repeat structure or an inverted repeat structure with the inversion points situated at the centres of each SstII fragment making them 95 kbp apart (see fig 4.12a). Assuming an inverted repeat structure two possible placings are available for the f=1 SfiI fragment. If one of the f=1 SfiI sites is placed within the s=1 fragment it is clear that the other f=1 site will map beyond the inversion point in the s=2 fragment which would violate both the symmetry and homogeneity of the units, this mapping possibility is, therefore, an unlikely one. The other possibility (see fig 4.12a) in the inverted repeat structure is for the f=1 fragment to completely contain the s=1 fragment and, since inverted repeats require symmetry this would place the f=1 SfiI sites 25 kbp beyond the SstII sites. The distance between these SfiI sites across the s=2 inversion point is then 90 kbp. The f=2 SfiI band is 140 kbp in size which implies that a distance of 35 kbp (140-105) lies between the f=1 site and the f=2 site. The symmetry of inverted repeats requires that the f=2 site is added on to both sides of the s=2 inversion point and the distance between the f=2 sites is then 20 kbp. The f=3 band is 45 kbp larger than the f=2 band and this size cannot be accounted for in this model. Two models which incorporate the simple alternation of the s=1 and s=2 fragments are diagrammed in fig 4.12.

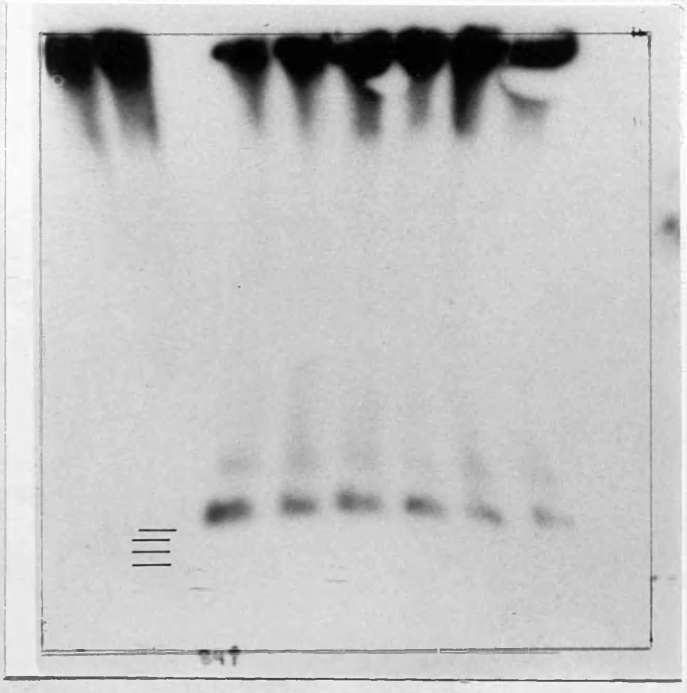


Fig 4.13 Southern blot of PFG run under condition capable of resolving up to 2000 kbp. C4M DNA was digested with various levels of SstII and probed with pGS-113. The lines are lambda multimers. See p 43 for details.

Facing page 64

contradictory. If the doublet is a running artefact it is both odd and repeatable, if it is real it should further complicate the SfiI restriction pattern and appear on other Southern blots- it could be present in fig 4.9 and it is not present in fig 5.1.

4.8 Probing beyond s=5

In an attempt to discover how many SstII fragments are linked together pulse field gels were run which were capable of resolving up to 2000 kbp. C4M DNA was digested with various amounts of SstII and electrophoresed. A typical result is shown in fig 4.13. It can be seen that the same two fragments are revealed in all lanes. The photograph of the ethidium bromide stained gel (not shown) shows that the C4M DNA front did not migrate as far as the largest lambda multimer band and the SstII fragments seen probably represent the fragments previously discussed which have migrated anomalously under this regime.

CHAPTER 5

The Amplification Units of other KG1-derived Cell Lines

5.1

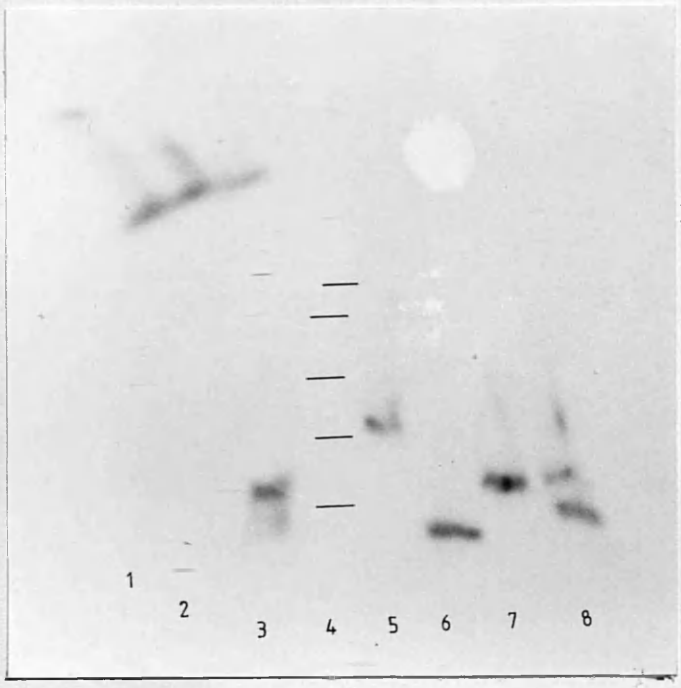


Fig 5.1 Nitrocellulose Southern blot of PFG probed with pGS-113 (see fig 3.1). Lane 4, lambda multimers. Lane 1, C4M DNA digested with ClaI; 2, MluI; 3, SalI; 5, SfiI; 6, SstI; 7, SstII; 8, XhoI.
For details see p.43

5.2

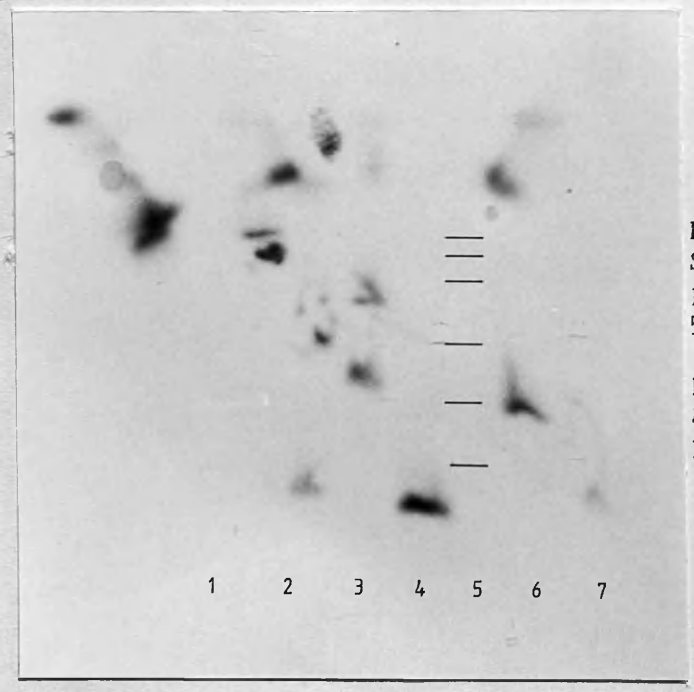


Fig 5.2 Nitrocellulose Southern blot of PFG probed with pGS-113. Lane 5, lambda multimers. Lane 1, KG1M DNA digested with MluI; 2, SalI; 3, SfiI; 4, SstI; 6, SstII; 7, XhoI.**For details see p.43**

5.1 Introduction

Fig 5.1 shows the basic restriction pattern of the C4M amplified unit. This structure is the end result of a selection process begun when the parental KG1 cells were first challenged with methionine sulphoximine. Since that initial event the GS gene population has been increased approximately 1000-fold and has been maintained at that level for many cell generations. In that time it is very probable that the amplification unit will have evolved into a form different from the DNA surrounding the original GS gene. Work by Dr. A.P. MacCabe (this laboratory) has shown that the C4M GS gene is identical to that of KG1 to some 20 kbp upstream of the CAP site differing only by an insertion of 0.5 kbp several kilobases upstream of the CAP site. 20 kbp downstream of the CAP site the restriction maps of the C4M and KG1 diverge. I have shown that the KG1 GS gene is contained within a 190 kbp SfiI fragment which constitutes a significant difference from the 110 kbp SfiI fragment of C4M and implies that significant structural alterations have occurred during the development of the amplification unit.

5.2 The amplification unit of the KG1M cell line

A cell line exists which is an intermediate in the evolution of KG1 to C4M. KG1M is resistant to levels of 25 uM Msx and is capable of readily amplifying further the size of its GS population. Fig 5.2 shows a pulse field gel restriction mapping of KG1M using the same enzymes as fig 5.1. A comparison of the two restriction patterns reveals several things. The size of the Sst I fragment (approx. 18 kbp by conventional mapping), which contains the majority of the GS gene coding sequences, has remained constant as has the Xho I fragment. The Sal I fragment of the KG1M unit is approximately the same size as the Xho I fragment and appears to be a single fragment, the Sal I fragment of C4M has altered in size. The C4M Sal I fragment exhibits complex behaviour, in this autoradiograph it shows a primary band equivalent in size to the C4M Sst II fragment with a smear below it while in fig 5.5 multiple C4M Sal I fragments are visible the largest of which is the size of the C4M Sst II fragment. This C4M Sal I complexity was also observed by Dr. A.P.MacCabe using

5.3

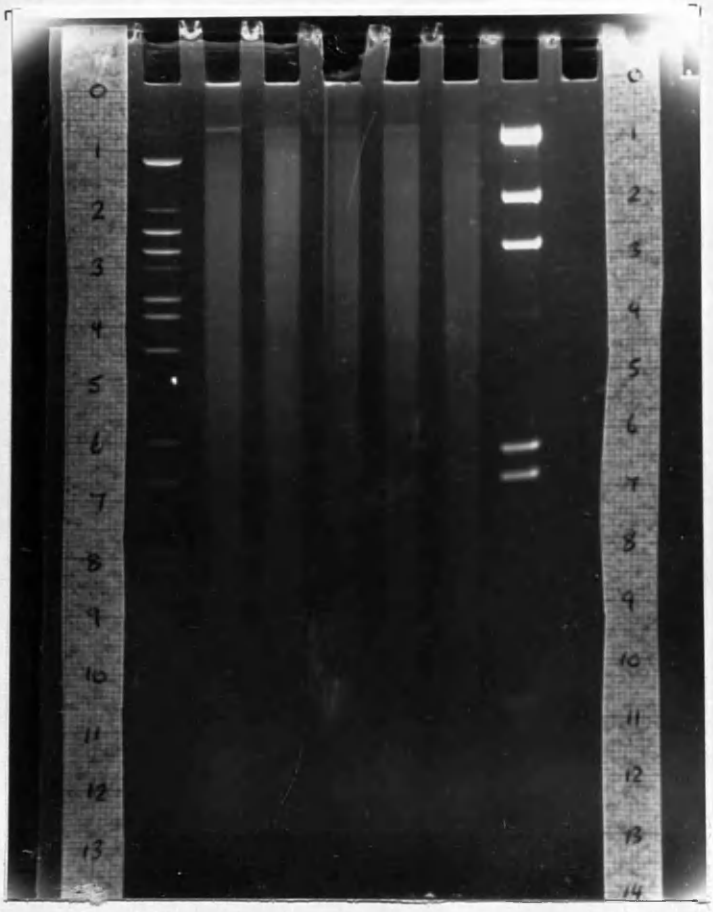


Fig 5.3 Photograph of ethidium bromide stained gel of KG1M.52 DNA restricted with BglII, EcoRI, HindIII, PstI, PvuII (left to right) to show the visible discrete bands of amplified fragments. The markers are λBstEII (left) and λHindIII (right).

conventional gel electrophoresis.

While the C4M SalI fragment size has increased, that of the SstII fragment has decreased, relative to that of the KG1M Sst II, from 90 kbp to 75 kbp, and the size of the Sfi I fragments are the same at 110 kbp.

The differences in restriction pattern indicates that some rearrangement in the structure of the amplified unit has occurred in the progression from KG1M to C4M. Since the GS gene of C4M cells and the parental cells is the same and since this region includes the mapped Sst II site and the probe used lies 3' to this site any rearrangement must have occurred downstream of the GS gene. The size of the Sfi I fragment does not change between KG1M and C4M units while the size of the Sst II fragment does, therefore the Sst II fragment cannot be contained within the Sfi I fragment. This means that one SfiI site must lie within the 55 kbp SstII fragment which is incompatible in the inverted repeat structures discussed in chapter 4 but easily accomodated in the direct repeat structure. Fig 5.2 shows a second SfiI band whose size is larger than the f=2 C4M SfiIU band (see fig 4.10b) and this is consistent with a direct repeat model.

Although reports in the literature do indicate that the amount of DNA coamplified with each gene does decrease as the copy number of the gene increases (Giulotto *et al*, 1986), presumably for reasons of the metabolic load of large quantities of excess DNA or for the greater ease of secondarily amplifying a smaller amplification unit, it is unlikely that the reduction in unit size from KG1M to C4M falls into this category. Giulotto *et al* (1986) observed a decrease from 10^4 kbp DNA amplified per gene to 10^3 kbp per gene making a decrease of 45 kbp fairly insignificant. Furthermore the postulated size of the KG1M unit (approx. 250 kbp) makes it one of the smallest "natural" amplification units, comparable to the 220/260 kbp units of Looney & Hamlin (1987) which have a copy number of at least 500 in CHO cells. It is more probable that the present C4M units represent a stochastic event which has come to dominate the population and are not the end result of a necessary decrease in size.

5.4a

Fig 5.4 Southern blots of DNAs of cell lines derived from KG1M, both have λ HindIII size markers and were probed with pGS-1. **Fig 5.4a**: Lane A, KG1M.52 DNA digested with HindIII; B, C4M with HindIII; C, KG1M.51 with HindIII; D, C4M with BglII; E, KG1M.51 with BglII; F, KG1M.52 with BglII.

5.4b

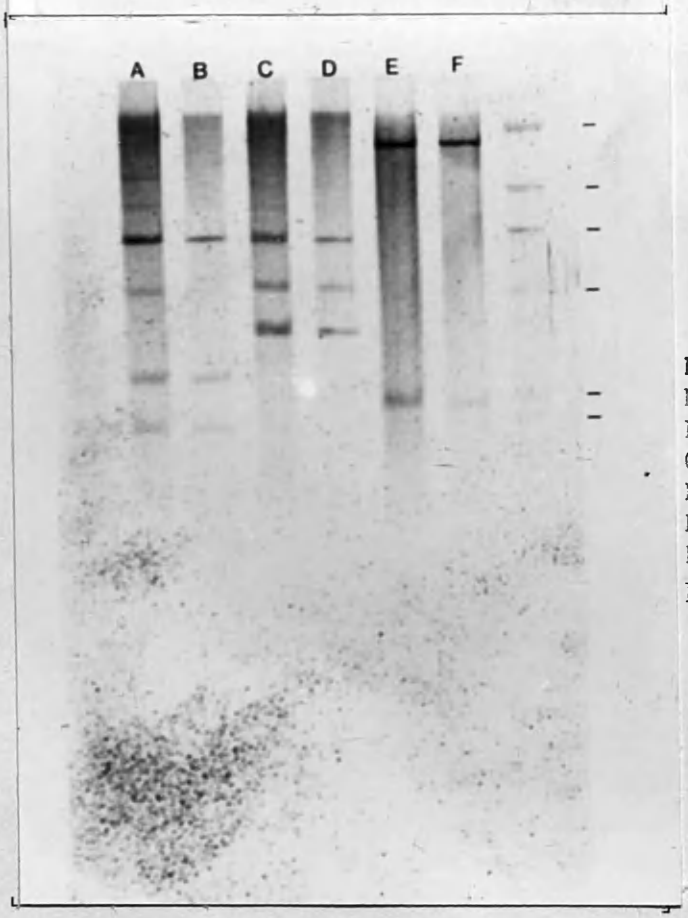


Fig 5.4b: Lane A, KG1M.51 DNA digested with HpaI; B, KG1M.52 with HpaI; C, KG1M.51 with KpnI; D, KG1M.52 with KpnI; E, KG1M.51 with SstI; F, KG1M.51 with SstI.

For details see p.43

size.

5.3 The amplification unit of a cell line derived from KG1M

In order to attempt to analyse the variability of the amplification process clones were isolated from each of three cell lines, KG1, KG1M and KG1MS-C40, and used as the starting points for rounds of stepwise selection for increasing resistance to Msx. Those lines with an initial resistance close to wild type, ie. KG1 and C40 derived lines, failed to achieve any consistent increase in resistance to Msx after 7 passages while the line derived from KG1M had, in the same number of passages, become resistant to 5 mM Msx, a 200-fold increase in resistance levels. Genomic DNA was isolated from these KG1M-derived cell lines within 3 passages of their becoming resistant to 5 mM Msx which reduces the possibility of alterations to the amplified unit as a result of extended culture. The DNA was restricted and gel electrophoresed (see fig 5.3). As is the case for C4M DNA discrete bands can be seen, a summation of the sizes of these bands produces 110-120 kbp indicating that the amplification unit in these cell lines is somewhat larger than that of C4M. Numerous Southern blots of these DNAs electrophoresed beside equivalent digests of C4M DNA were made and probed (see fig 5.4a and b) and the results showed that the DNA structure of all these cell lines appeared identical in the immediate vicinity of the GS gene. KG1M.52 DNA was prepared for PFGE, again derived from cells within 3 passages of becoming resistant to 5 mM Msx. A PFGE of equivalent digests of C4M and KG1M.52 was performed and Southern blotted (see fig 5.5). Probing this shows that the two amplification units are not the same. Because of the compression of the size markers and the intensity of the bands accurate sizing of the larger bands is not possible. It can be seen that the KG1M.52 unit structure is simpler than that of the C4M unit judging by the difference in Sal I restriction patterns. It can also be seen that the size of the KG1M.52 Sal I fragment lies between 23 and 50 kbp, the Sfi I fragment is approximately the same size as that of C4M and the Sst II fragment is larger then that of C4M. A comparison of these relative sizes to those of fig. 5.5 suggests that the amplified unit in KG1M.52 cells may

5.5

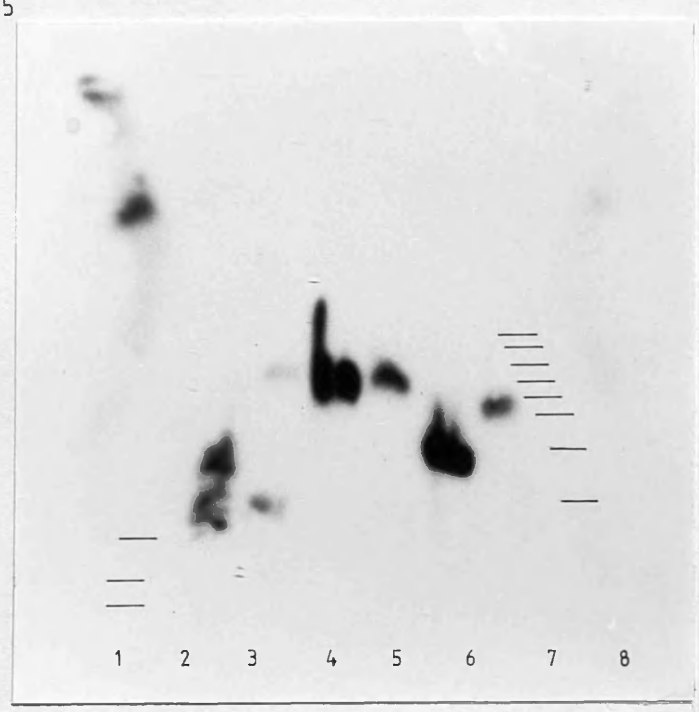


Fig 5.5 Southern blot of PFG probed with fragment II. Lane 1, λ HindIII size markers; lane 8 lambda multimers. Lane 2, C4M DNA digested with SalI; 4, SfiI; 6, SstII. Lane 3, KG1M.52 DNA digested with SalI; 5, SfiI; 7, SstII. For details see p.43

5.6

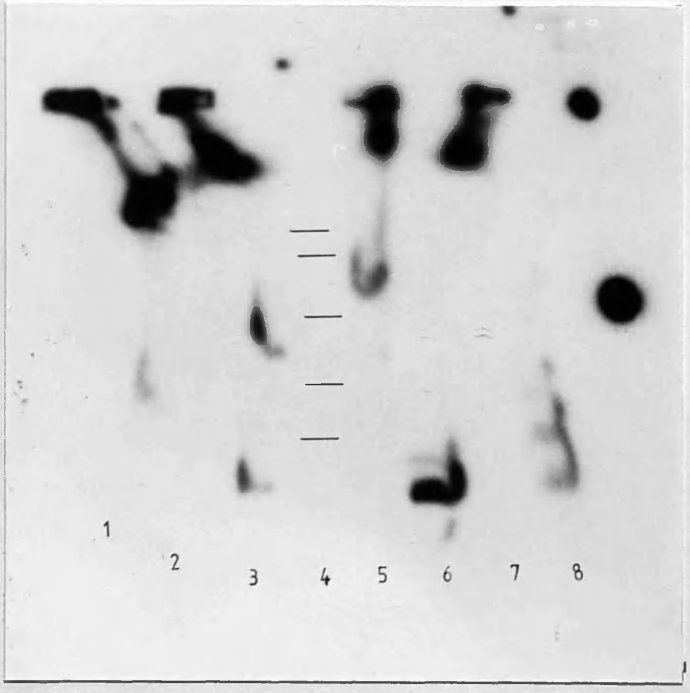


Fig 5.6 Nitrocellulose Southern blot of PFG probed with pGS-113. Lane 4, lambda multimers. Lane 1, KG1MSC4-0 DNA digested with ClaI; 2, MluI; 3, SalI; 5, SfiI; 6, SstI; 7, SstII; 8, XhoI.
For details see p.43

well be the same as that of KG1M cells. If this is the case it reveals that the secondary amplification process does not require a DNA rearrangement and that the alteration in the C4M unit may therefore be a consequence of extended culture rather than amplification.

5.4 The amplification unit of the revertant cell line, C4O

Fig 5.6 shows the restriction pattern of C4O cells equivalent to figs 5.1 and 5.2. Considering that these experiments were designed to be equivalent to each other it is possible to say, very approximately, from the incubation times of the various autoradiographs that the number of hybridising sequences in the C4O cells has been reduced at least 10-fold relative to the KG1M cells and at least 200-fold relative to the C4M cells. Based on their relative positions and the lack of complexity in the Sal I track, the Sal I, Sst I and Xho I bands suggest that the DNA structure around the GS gene resembles that of KG1M rather than that of C4M. Since C4O is descended from C4M rather than KG1M it is possible to account for the local structure by supposing that either the C4M rearrangement event occurred after the C4O lineage was begun or that the reduced numbers of GS genes in C4O have been derived from one of the several alternatives present in C4M cells and it happens to be the one which resenbles the KG1M GS gene. An alternative explanation is to suggest that the C4O GS genes are derived from a population of genes which were separated from the C4M GS genes and present in small numbers, possibly a relic of early amplification events or direct descendants of the other ancestral GS gene. The lack of a Sst II band is, presumably, a result of nondigestion rather than very major rearrangement although the Sfi I band is approx. 175 kbp in size which means some change has occurred to enlarge the amplification unit.

CHAPTER 6

The Chromosomes of KG1 and KG1/MSC4-M Cells and the Localization of the Amplified Glutamine Synthetase Genes by *In Situ* Hybridization

Facing page 69

6.1 Introduction

The DNA of a cell is packaged within the nucleus in the form of chromosomes. At metaphase in the cell cycle chromosomes condense into a state where they can be visualised using a microscope. Numerous methods have been devised for investigating metaphase chromosomes. Several dyes exist which bind to chromatin or specific subsets thereof in a consistent fashion giving the opportunity to map chromosomes. Tritiated DNA probes may also be used, in a fashion analogous to probing DNA immobilised on membranes, to localise sequences of interest to specific regions of a chromosome.

The technique of trypsin-Giemsa banding was developed by Seabright (1971) and involves a partial digestion of metaphase chromosomes followed by staining with the Giemsa dye. The result is a consistent series of dark and light bands which identifies each chromosomal region and permits detailed analysis of rearrangements, additions or deletions of the chromosome.

An analysis of the banding patterns of the parental line and its glutamine synthetase-amplified daughter line was undertaken to provide data on the degree of alteration between them and between the parental line and other sublines of the Chinese hamster ovary cell line.

6.2 CHO in the literature

The Chinese hamster, *Criticelus griselus*, has a diploid

Fig 6.1: A diagrammatic representation of the chromosomes of Chinese hamster ovary cells. Line (a) shows the karyotype of the female Chinese hamster depisting at least one example of each chromosome pair. Chromosome pairs 1, 6, 9 and 11 are only represented once as neither of the pair is rearranged. The standard karyotype contains 22 chromosomes. Line (b) shows the CHO karyotype of Worton *et al* (1977) and the brackets on the right hand sides of the chromosomes indicate their relationship to the wild-type chromosomes. The standard karyotype contains 21 chromosomes due to the loss of an 11. Line (c) shows the further rearrangements common to the K1 lineage cells analysed by Worton (1978). Brackets to the left hand sides of the chromosomes indicate their relationship to the CHO chromosomes.

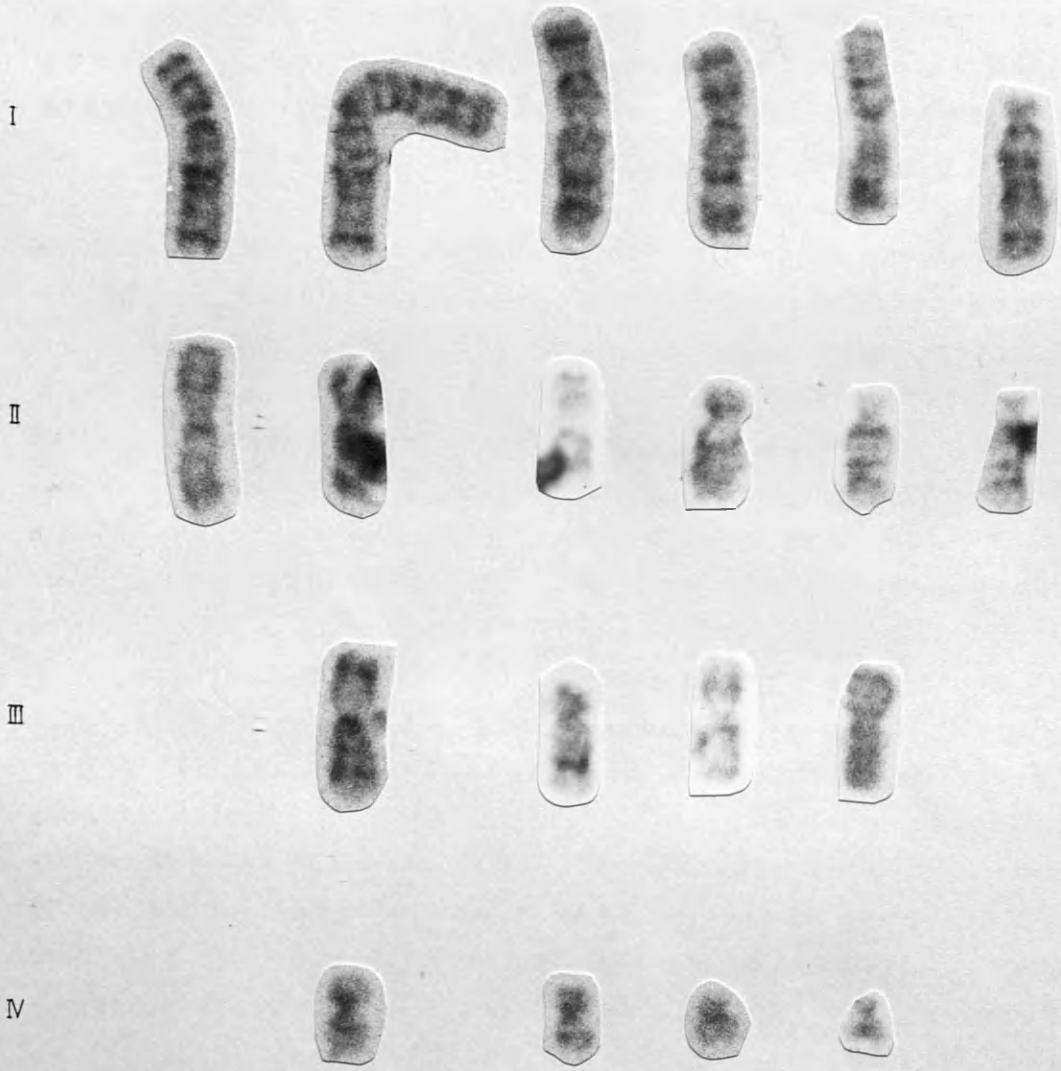
state of 22 chromosomes, 10 autosomal chromosome pairs and the X/Y pair. In 1957 Puck et al (1958) began the culture of the Chinese hamster ovary (CHO) cell line and several karyotypic investigations have been undertaken using various sublines since that time. Kao and Puck (1969) observed a quasidiploid karyotype of 21 chromosomes containing 9 altered chromosomes (which were designated Z-1 to Z-9) and 3% less total chromatin than the normal chromosome complement.

A more detailed investigation by Deaven and Petersen (1973) revealed a quasidiploid karyotype of modal number 21 which contained 13 rearranged chromosomes (Z chromosomes). Using these results as a reference point Worton and his co-workers investigated the variations in karyotype between sublines of the CHO line (Worton et al, 1977; Worton, 1978). They showed that although there were many reaarangements common to all sublines each subline possessed one or more rearrangements unique to itself.

Two CHO sublines were analysed by Worton (1978), both of the K1 lineage from which my parental line- KG1- is derived, and compared to his reference CHO karyotype. Both lines had 3 new rearranged chromosomes in common derived from the breakage and fusion of one normal chromosome and two Z chromosomes though in one line one of these new chromosomes was further added to. A diagrammatic representation of the derivation of all the chromosomes common to these two sublines is presented in fig 6.1 (see fig 1s of Worton et al, 1977; Worton. 1978).

6.3 KG1 chromosomes in comparison to literature chromosomes

56 metaphase spreads of KG1 were photographed after trypsin-Giemsa banding and 36 quasidiploid and one quasitetraploid karyotypes were analysed. The most common karyotype contained 20 chromosomes. Seventeen chromosomes were seen which were identifiable as chromosomes described in the literature and which included the three K1 marker chromosomes. Of the remaining three chromosomes one was too small to be unambiguously identified and the other two were marker chromosomes unique to KG1 designated K and A. The standard karyotype is presented in fig 6.2 and contains the K1 marker chromosome (Z-3q) in the $(Z\text{-}3q)p^{+}$ form. In



KG1 karyotype including deletion variant of chromosome 1, Δ1.
Left to right: I : Δ1,1,2,Z-2,t(8q,Xq),Z-4
II : t(Xp;Z-3p),(Z-3q)p+,Z-8,Z-9,6,6
III: Z-7,K,Z-5,A
IV: 9,10,11/Z-13,Z-12

Facing page 71

addition to marker chromosomes K and A a small proportion of the karyotypes contained another, additional, marker chromosome designated G.

6.4 The karyotypic variation of KG1 chromosomes

Karyotypic variation can manifest itself in several ways. Cells can become quasitetraploid in nature, that is the quasidploid complement can be approximately doubled within the cell, or chromosomes can be lost or further rearranged while the quasidiploid state is maintained.

Qualitative

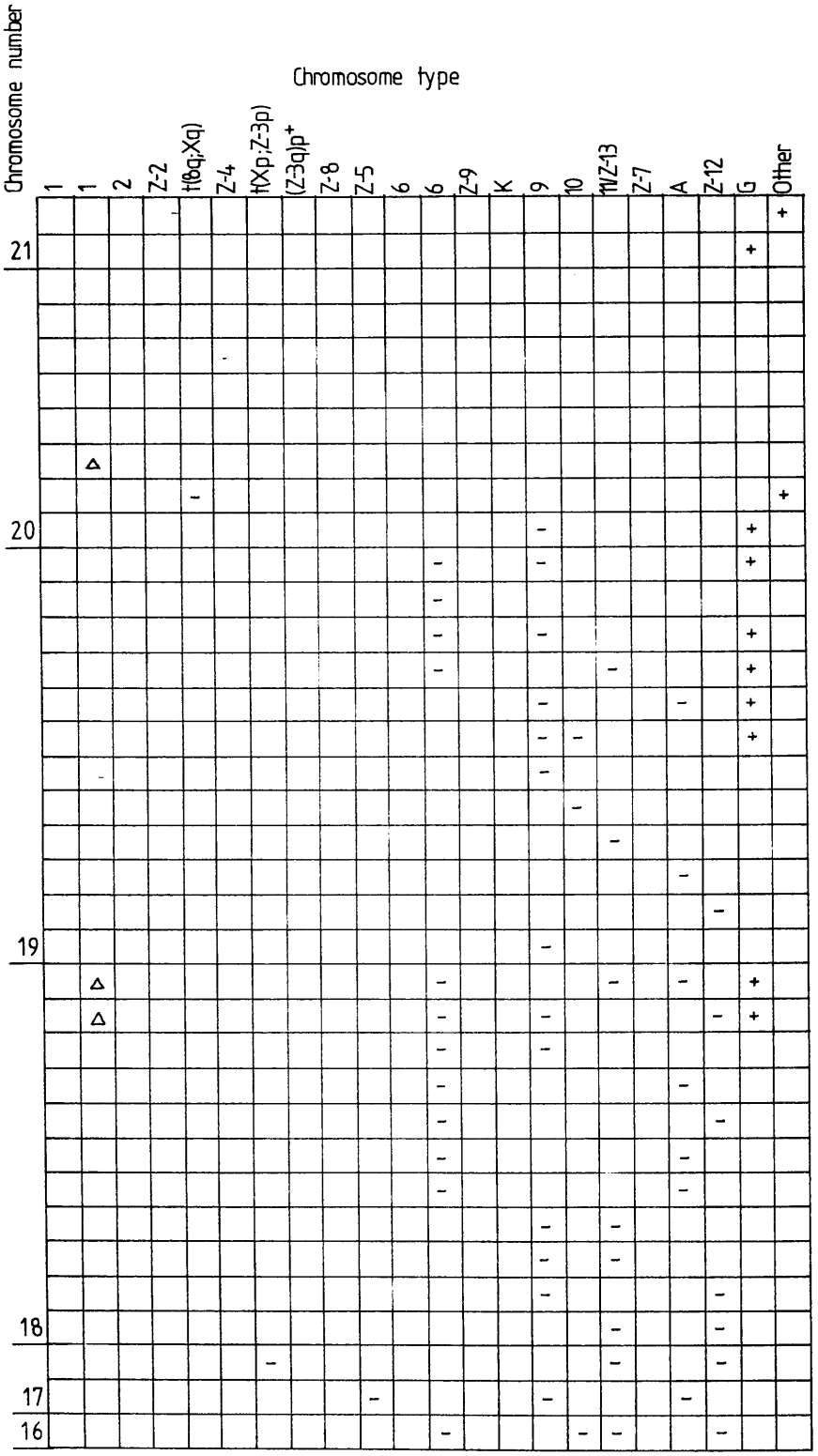
The majority of the variation seen was due to the complete absence of one or more of the smaller chromosomes. Three partial deletions of one p-arm of a chromosome 1 were seen, as was a probable breakage due to the fixation process in the same place on the p-arm. A loss of the t(8q;Xp) was seen coupled with the appearance of a new chromosome resembling the $(Z\text{-}3q)p^+$.

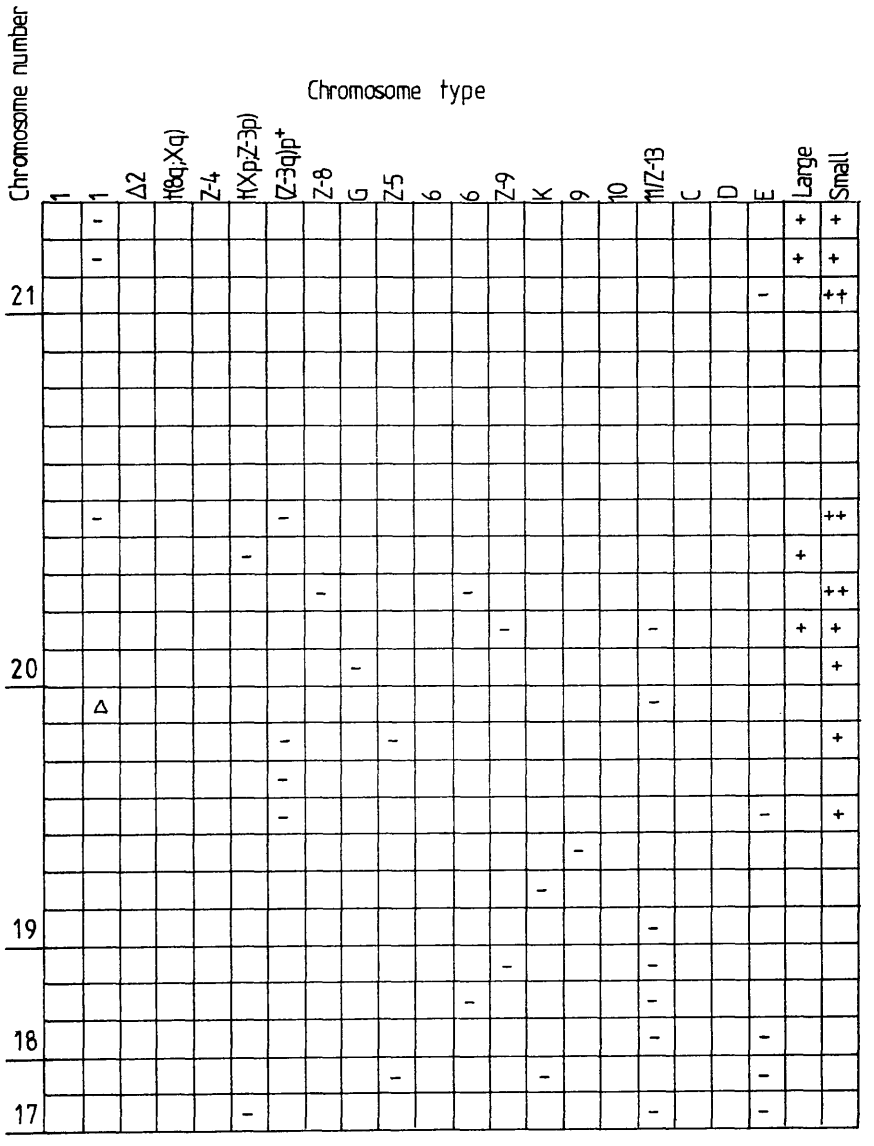
Quantitative

The karyotypic variation due to chromosome loss was very considerable in this analysis. 60% of the 110 metaphase spreads of the standard CHO line examined by Worton et al (1977) possessed one of two karyotypes each containing 21 chromosomes and differing solely by the addition of an extra band to the q-arm of one of the 6 chromosome pair. Other CHO lines investigated in the same study showed a similarly homogeneous population. Both Deaven and Petersen (1973) and Worton et al (1977) show that the vast majority of their cells (38 of 50 and >60% respectively) possess the modal number of chromosomes and that this is numerically coincident with the most common karyotype. This was not the case in this analysis.

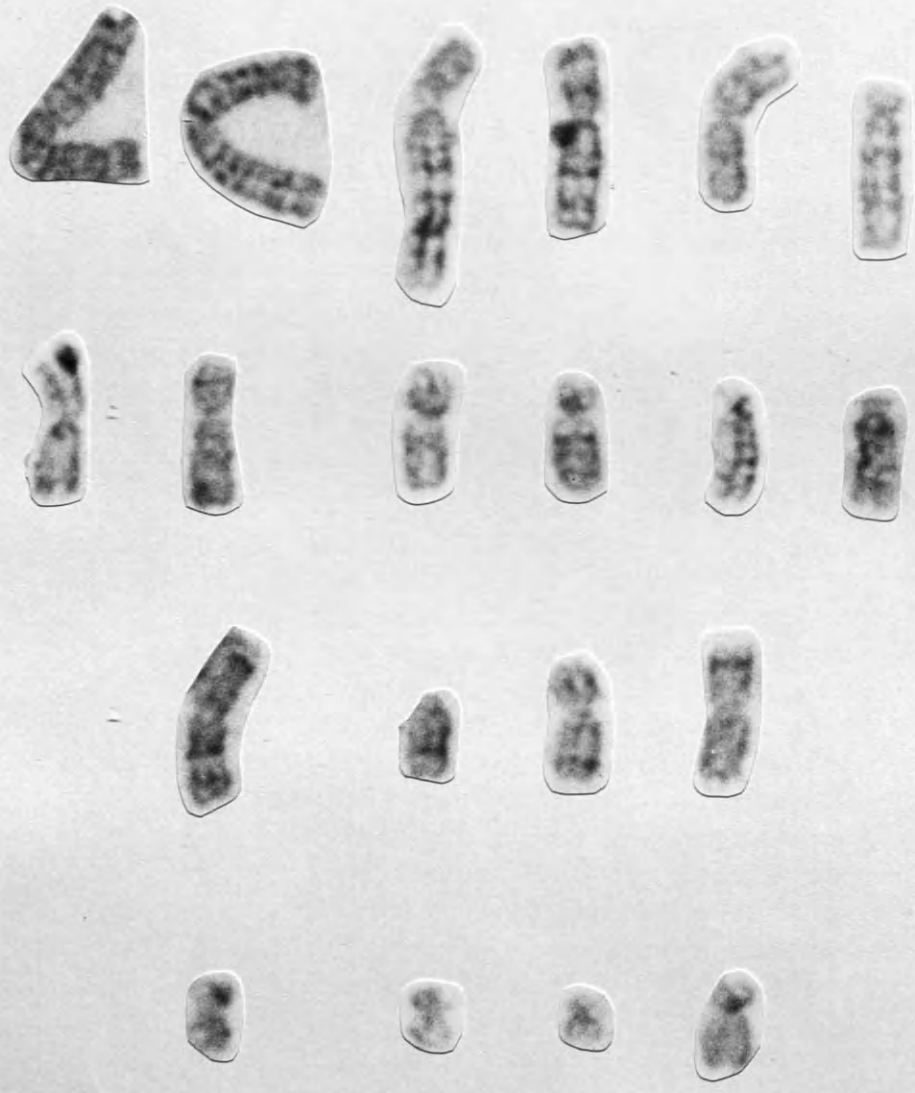
The most common karyotype occurred 6 times (8 including two with additional chromosomes) and possessed 20 chromosomes. While 21 chromosomes is the modal number usually reported in the literature 20 chromosomes was the modal number of the subline possessing $(Z\text{-}3q)p^+$ (Worton, 1978) and this was due to the fusion of the Z-10 chromosome with one of the 6 chromosome pair. This

6-3
KG1


Chromosome number
Chromosome type
1
1
2
Z-2
t(6q;Xq)
Z-4
t(Xp;Z-3p)
(Z-3q)p+
Z-8
Z-5
6
6
Z-9
K
9
10
mZ-13
Z-7
A
Z-12
G
Other
21
20
19
18
17
16



Figures 6.3 and 6.4 Tables of the karyotypic variation of the 36 KG1 (6.3) and 25 C4M (6.4) karyotypes analysed. The "-" sign shows a designated chromosome is missing while a "+" sign indicates the presence of a new chromosome. In the KG1 table the new chromosomes are the G subpopulation marker or "other" while in the C4M table they are "large" or "small". The Δ indicates the deletion of a telomeric portion of the p-arm of a chromosome 1, an example of which is shown in fig 6.2



Standard C4M karyotype:
left to right: I: 1,1,C,Δ2,t(8q;Xq),Z-4
II: t(Xp;Z-3p),(Z-3q)p+,Z-8,Z-9,6,6
III: D,K,Z-5,G
IV: 9,10,11/Z-13,E

Facing page 73

fusion has not occurred in KG1. KG1 lacks a Z-10, a 9 and either an 11 or a Z-13 chromosome. It is most probable that a fusion between the 9 and the 11/Z-13 produced the A chromosome and that the addition of material to Z-10 (a telomeric chromosome) produced the acrocentric K chromosome. Twelve karyotypes possessed 19 chromosomes and a further eleven possessed 18. Thus the commonest karyotype does not occur most of the time nor is there a modal number of chromosomes. See fig 6.3 for a detailed description of the karyotypic variation.

6.5 The C4M cell line

C4M cells possess approximately 1000 copies of the GS gene which, assuming a unit size of 75 kbp (see chapter 4), would comprise at least 2% of the haploid genome (based on a genome size of 3×10^6 kbp). An addition to the genome of this size should produce some karyotypic effects.

Forty-five C4M metaphase spreads were photographed after trypsin-Giemsa banding and 25 quasidiploid and 2 quasitetraploid spreads were analysed. Sixteen chromosomes were seen which closely resembled chromosomes present in KG1; four C4M specific chromosomes were seen three of which were the products of rearrangements and one whose origin is uncertain but was shown to contain the amplified GS DNA. See fig 6.5. Other rearranged chromosomes were observed occassionally.

6.6 Karyotypic variation of the C4M cell line

Qualitative

Like the KG1 karyotypes the majority of the variation was due to the loss of complete chromosomes. Unlike KG1 karyotypes this loss involved both larger and smaller chromosomes and the generation of novel chromosomes. C4M cells also showed a higher frequency of quasitetraploidization than KG1 cells.

Quantitative

In the quasidiploid spreads analysed twelve species of chromosome were lost one or several times. The commonest karyotype possessed 20 chromosomes (n=5) and 20 was also the

modal number of chromosomes (n=10). Five quasitetraploid spreads were seen in the 45 spreads photographed. See fig 6.4 for a detailed description of the karyotypic variation.

Consensus

The most common karyotype contains 20 chromosomes. Sixteen of these are KG1 chromosomes but their number includes the G chromosome which is present only in a subpopulation of KG1. The KG1 chromosomes lost in C4M are the 2, Z-2, Z-7, Z-12 and the A which have been replaced by the Δ2, C, D and E chromosomes; those retained are the 1 pair, t(8q;Xq), Z-4, t(Xp;Z-3p), (Z-3q)p^+, Z-8, Z-9, Z-5, the 6 pair, K, 9, 10, 11/Z-13 and the G chromosome. See fig 6.5

6.7 Analysis of the karyotypic variation of KG1 and C4M

Numerical variation

The numerical variation in the karyotypes of both KG1 and C4M is more pronounced than any in the literature, especially as the frequency of chromosomal rearrangements does not show a corresponding increase. Several possibilities suggest themselves to account for this variation:-

(i) random chromosome loss incurred during the fixation of the metaphase spreads.

(ii) a genetic tendency of the KG1 line (and consequently its daughter line C4M) to lose chromosomes.

(iii) a lessening of the selective pressure on the genome due to the culture conditions allowing viability to be maintained with chromosomes lost.

(iv) the occurrence of stressful culture conditions just prior to the harvesting of the spreads causing karyotypic aberrations.

None of these explanations are adequate in themselves to account for all the numerical variation. The possibility that stressful culture conditions can alter karyotype was addressed by Rice *et al* (1986) and was found to be the case. However far from causing chromosome loss stress causes overreplication of DNA and

the consequent generation of rearranged chromosomes and amplified DNA. It is therefore unlikely that this possibility affects the numerical variation observed.

The culture conditions under which KG1 and C4M are grown, though different, would seem to be more stringent than those employed by Worton et al (1977) who were using a medium which contained a greater variety of amino acids than does our media and also several other additives not used in our media (see Stanners, 1971, for their media recipe). It is therefore unlikely that the variation in chromosome number is due to a lessening of selective pressure due to the culture conditions of KG1 and C4M.

A genetic tendency to lose chromosomes would have a serious negative effect on the viability of those cells possessing it unless it coexisted with appropriately beneficial culture conditions. The endpoint of such a happy conjunction would be the definition of a new minimum karyotype which should be apparent from this analysis unless those spreads were taken from a culture just starting down that path. Worton et al (1977) found that while some lines showed a greater tendency to produce quasitetraploids (up to 15% of cells analysed in one line) these lines still showed a strong mononumerical bias in their quasidiploid cells. A genetic tendency to randomly lose chromosomes is therefore unlikely both because it would adversely affect cell viability and because the observed tendency is to do precisely the opposite, ie. increase chromosome number. Furthermore the C4M line shows a marked increase in the tendency to produce quasitetraploids (5 of 45, >10%, as against 1 of 56, <2%) which seems at odds with a tendency to lose chromosomes.

Loss of chromosomes during fixation is certain to happen, especially in the KG1 line as I observed KG1 cells were very liable to disrupt if not handled gently. Furthermore in the KG1 line most of the chromosomes lost were the smaller ones which would most easily have been borne away by the fixative after the nucleus ruptures . In the C4M line all sizes of chromosome were lost although smaller ones were lost somewhat more frequently than larger ones. Thus the pattern of loss varies between the two lines and unless chromosome loss during preparation affected them differently it cannot therefore account for all the chromosome

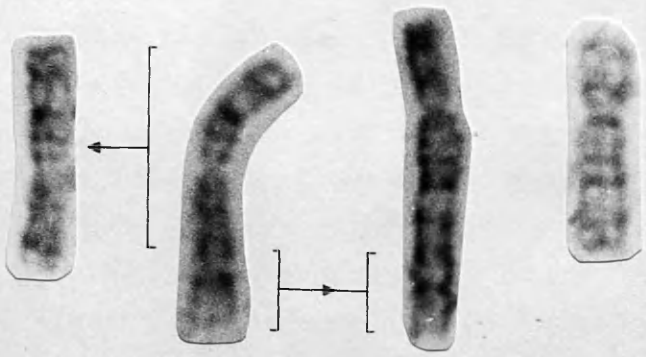


Fig 6.6 The derivation of the Δ2 and C chromosomes of C4M. From left to right, the Δ2 chromosome and the 2 chromosome from KG1 (both upside down with respect to figs 6.5 and 6.2 respectively). Next is the C chromosome and the KG1 Z-2 chromosome all aligned to show homologous bands. The brackets indicate the redisribution of the 2 chromosome to the Δ2 and C chromosomes.

loss. Since chromosome loss due solely to the fixation process cannot account for the different patterns of loss observed in the KC1 and C4M lines other influences must act upon these cell lines. The loss of many chromosomes in the C4M line is illusory since for each characteristic chromosome lost a new one frequently appears and this is especially true of the larger chromosomes apparently lost (see fig 6.4) leaving only the loss of smaller chromosomes to be explained by the vicissitudes of the fixation process. KG1 cells have lost few large chromosomes but many smaller ones arguing for losses during fixation. Atop this loss however the frequent loss of a chromosome 6 and/or 9 suggests that these chromosomes can be lost without significantly harming cell viability.

Both cell lines show an ability to rearrange one of the chromosome 1 pair. Three KG1 cells had a deletion of part of the p-arm and in addition to this alteration three C4M cells either lost the whole chromosome completely or rearranged it beyond recognition.

6.8 The origins of the C4M marker chromosomes

Two types of karyotypic abnormality have been found to be associated with gene amplification- homogeneously staining regions and double minutes (see Schimke, 1982, references therein). Double minutes are small extrachromosomal bodies associated with unstable amplification events and no evidence of their existence was seen in this investigation.

As previously stated the majority of the C4M chromosomes appear to be the same as the KG1 chromosomes they originated from. The retention of the majority of chromosomes in an unaltered form has been observed in several other CHO gene amplification events- dihydrofolate reductase (Flintoff *et al*, 1984; Milbrandt *et al*, 1981), multidrug resistance (Grund *et al*, 1983), asparagine synthetase (Andrulis and Siminovitch, 1982) and vincristine resistance (Teeter *et al*, 1986; Kuo *et al*, 1982). The last example also involved the creation of a quasitetraploid state.

The C chromosome and the Δ2 chromosome

The Δ2 chromosome lacks the telomeric portion of its p-arm

and this has been translocated to the end of the q-arm of the Z-2 chromosome which is designated C. This rearrangement is shown in figure 6.6.

The D chromosome

This chromosome appears to be the result of addition of material to Z-7 on the p-arm and, possibly, on the q-arm.

The E chromosome

The E marker chromosome is of unknown origin being a small undefined acrocentric chromosome.

The K marker chromosome, unique to KG1, is also present indicating the relationship between KG1 and C4M.

6.9 Is C4M derived from a distinct subpopulation of KG1?

The G chromosome is present in almost all the C4M karyotypes but only appearsin a sub-population of KG1. This fact suggests that the C4M line was derived from the G-containg KG1 subpopulation which could either be due to a fortuitous initial selection or a reflection of a greater tendency towards genomic variability in the G subpopulation.

The C4M karyotype is more unstable than that of KG1. This is shown in two ways, firstly C4M cells have a greater tendency to produce a quasitetraploid state (5 of 45) than KG1 cells (1 of 56) and secondly C4M cells have a greater tendency to undergo chromosomal rearrangements than KG1 cells. Only 2 of 36 KG1 spreads (5.5 %) showed evidence of chromosomal alteration while 10 of 25 (40 %) C4M spreads possessed rearranged chromosomes. This difference in chromosome stability could be due to the selection of an unstable KG1 subpopulation as the progenitors of the C4M line. This follows from the fact that to survive the selection procedure required an overproduction of glutamine synthetase protein and this was achieved by an overproduction of DNA containing the glutamine synthetase gene. This overreplication of DNA must have caused some chromosomal alteration. Cells which were more prone to chromosomal instability (including overreplication of DNA) would therefore be

6.7

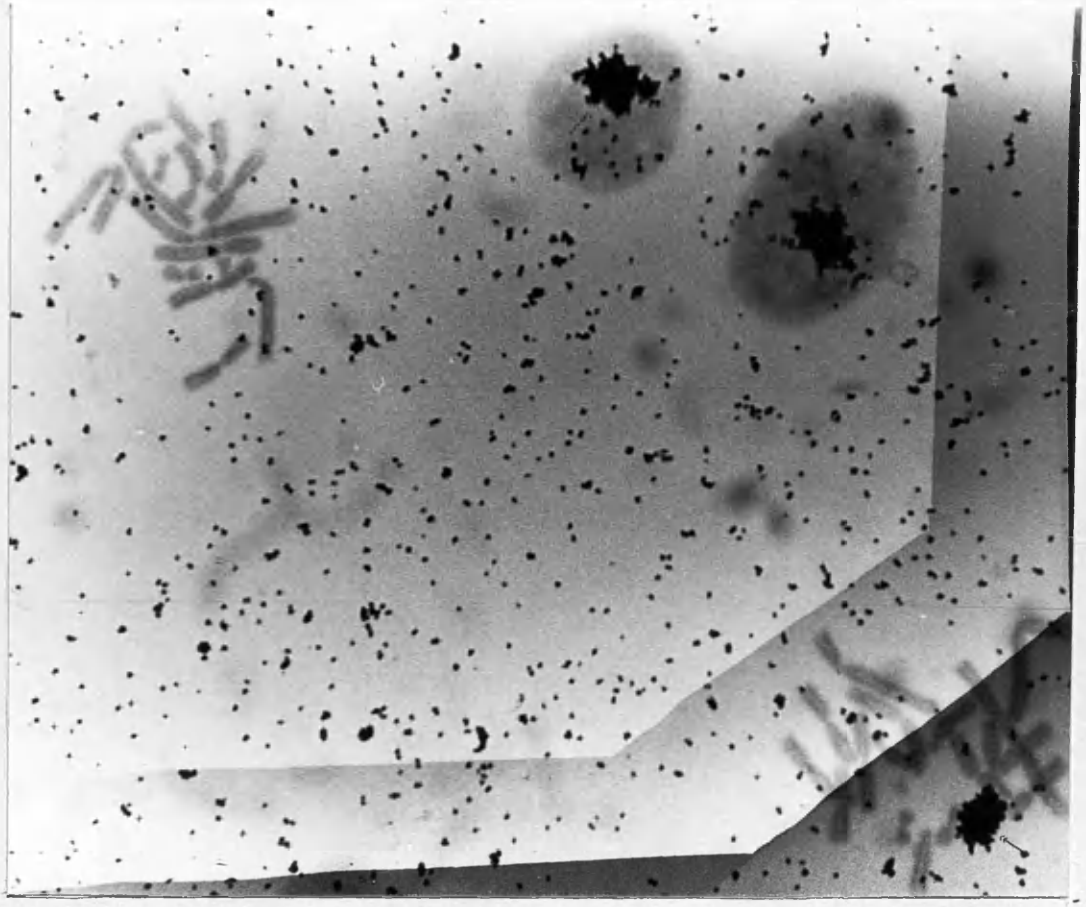


Fig 6.7 Photograph showing the localization of the amplfied GS sequences (arrowed) by *in situ* hybridization to metaphase spreads and interphase nuclei. The metaphase spread in the left hand corner is an example of a non-hybridising spread.

at a selective advantage and would tend to compose the cell population which survived the selection procedure and would become C4M.

An alternative explanation for the difference is that C4M culture conditions may be more stressful than those for KG1 and this may cause a greater degree of chromosomal instability (Rice et al, 1986). To survive in 5 mM methionine sulphoximine a huge quantity of glutamine synthetase protein is required to "soak up" the inhibitor to ensure some active GS remains. Sanders and Wilson (1984) estimate approximately 30% of the cellular protein may be GS in C4M cells and the effect of this excess protein on the osmotic balance and metabolic processes could be stressful in some way.

6.10 In Situ hybridization to localise the amplified region

Many studies (Wahl et al, 1982; Kellems et al, 1982) have used in situ hybridization techniques to localise the site(s) of the DNA amplified in their mutant cell lines. It has been found almost invariably that the site of the amplified DNA corresponds to the position of the homogeneosly staining region in the karyotype and that this region is located on a marker chromosome associated with the mutant cell line.

To identify the locality of the amplified GS DNA in the C4M line an in situ hybridization experiment was performed. Using ^{3}H-labelled dCTP in a random primer synthesis reaction (replacing the equivalent ^{32}P-labelled dCTP) the genomic fragment of the plasmid pGS-113 was radiolabelled and then purified on a Sephadex G50-50 column.

Separate sets of slides bearing metaphase spreads and cells of both C4M and KG1 cell lines were prepared in the same way as for karyotyping purposes. These slides were then processed, incubated with the radiolabelled probe, washed and coated with autoradiographic emulsion (Pardue, 1985).

Upon development and staining 35 metaphase spreads of C4M were photographed, 24 were quasidiploid and 11 were quasitetraploid. Due to the extended exposure the background, as represented by the random occurrence of silver grains, was relatively high (see fig 6.7). No photographs of KG1 metaphase

spreads were made as only background grains were present.

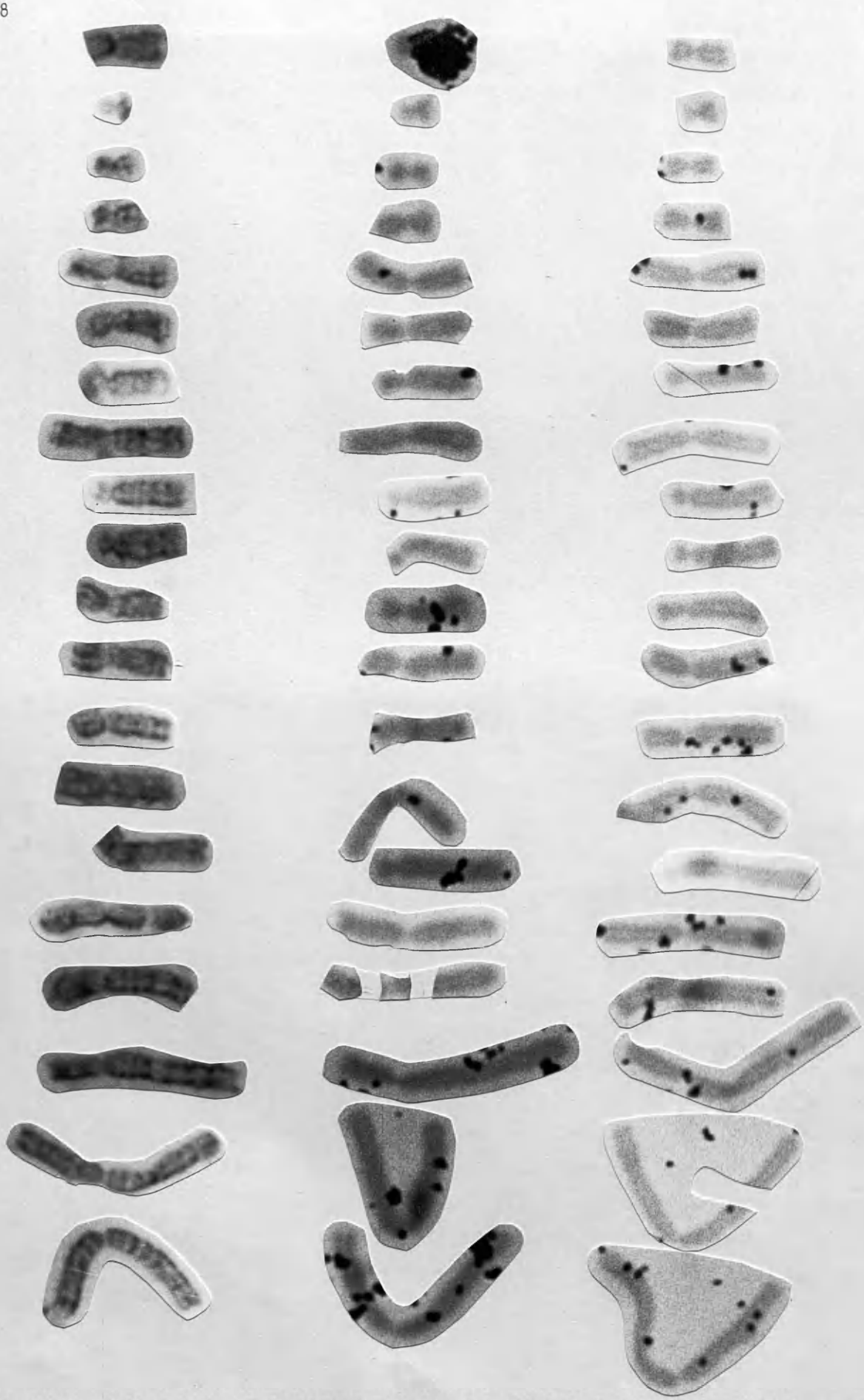
Eleven C4M quasidiploid metaphase spreads were analysed and the results revealed a karyotype like that described previously. A detailed analysis was not possible because the chromosomes were only Giemsa stained instead of being trypsin-Giemsa banded, however the modal number of chromosomes was twenty and an appropriate number of large, medium and small chromosomes were present.

As can be seen from fig 6.7 a very specific region has been located by the tritiated-GS probe (arrows) Furthermore this region of probe complementarity can be detected in the nuclei of the cells as well as on the chromosome spreads. The unlabelled metaphase spread indicates that not all cells possess this hybridising region. A sample of 721 cells were scored for the presence or absence of this region and 89 were found to lack it. This result implies that just over 12% of C4M cells lack all, or the vast majority, of the amplified GS DNA.

A number of conclusions can be drawn from this result. All bar two of the spreads containing this region showed an approximately comparable area covered by grains (the other two had larger areas covered). Assuming that this area equivalence is not an artifact caused by total overexposure of various different degrees of amplification it implies that the chromosome harbouring this hybridizing region contains a standard number of copies of GS DNA and C4M cells either have this region in its entirety or they do not.

Lack of this region implies that cells are unable to make sufficent GS protein to be resistant to 5 mM Msx. The existence of these cells suggests therefore that cells containing the region can export GS protein in sufficent quantities to enable both themselves and the GS-reduced cells to survive or that the amplification of GS DNA is so unstable as to be lost in 12% of cells every generation.

Resistance to Msx is an unstable phenomenon in the absence of Msx. Dr R.H.Wilson (pers comm) suggests that the presence of 1000 active copies of the GS gene may, by virtue of the huge quantities of GS protein they generate, have a deleterious effect upon cellular metabolism. When C4M cells are initially cultured



Facing page 80

in media free from Msx they remain quiescent for several days, neither growing nor dividing, until growth is observed at scattered foci. These foci represented the founder cells of the cell line KG1MSC4-0 which is resistant to 8uM Msx.

It has been observed that the promoter elements of genes are closely associated with DNA sequences which have the capacity to bind to the nucleoskeleton (Gasser & Laemmli, 1986). It is postulated that this cohabitation brings the active promoter elements close together in a region enriched with enzymes involved in transcription. It is interesting to note that the hybridization over interphase nuclei is clustered rather than dispersed, see fig 6.7.

Fig 6.8 shows representative quasidiploid metaphase spreads of C4M cells. It can be seen from spread (b) that the amplifed region is not located on the large marker chromosomes but is instead located on the q-arm of the small E marker chromosome. Spread (c) is an example of the non-hybridizing karyotype yet it possesses the modal number of 20 chromosomes which are akin to those if spread (b). An analysis of the arm length ratios and chromosome size with respect to that of Z-4 established that the dimensions of the E chromosome were demonstrably different from those of the three smaller chromosomes, 9,10 and 11/Z-13 and three similarly sized chromosomes, 6,Z-9 and K. A similar analysis of the chromosomes visible in the <u>in situ</u> karyotypes showed that they did not represent the E chromosome which implies that the E chromosome possesses the GS-hybridising region.

That the amplified GS DNA should not be located on a large

Fig 6.8 The localization of the amplified GS sequences to the q-arm of the E chromosome. Top line is a representative, trypsin-Giemsa banded, C4M karyotype ordered (left to right) 1, 1, C, .2. t(8q;Xq), Z-4, t(Xp;Z-3p), (Z-3q)p^+, Z-8, Z-9, 6, 6, D, K, Z-5, G, 9, 10, 11/Z-13 and E. The middle line is similarly ordered and shows the localization of silver grains over the E chromosome and the lower line shows a non-hybridising spread possessing a smaller chromosome in place of the E.

marker chromosome is unusual. Whether the correlation between amplified DNA and large chromosomes represents merely the greater statistical likelihood that the amplified DNA will be located in the majority portion of the genome contained within the larger chromosomes or does indeed represent some undefined requirement for the survival and progression of the amplifed DNA is open to question. Conversely the evident instability of the GS amplified DNA could be due to its location on the small E chromosome or to some inherent instability which would manifest itself regardless of chromosomal location. The mechanism of the loss of the amplified GS sequences is unclear, of the six karyotypes lacking the E chromosome three had additional small chromosomes which could have been the remnants of the E chromosome and three had no such chromosomes so either deletion of the amplified locus or complete chromosome loss can appear to account for the loss of some amplified sequences.

An alternative answer to the question of the instability of the amplified DNA is to suggest that the DNA resides, in part, on extrachromosomal double minutes. This is an unlikely explanation for several reasons. Firstly Chinese hamster cell lines are only rarely reported to contain double minutes. Secondly if double minutes were present they would have to coexist alongside the E chromosome, unless precisely segregated, within the same cell and the rapid emergence of deamplified cells would still require the loss of the amplified sequences on the E chromosome so the known instability of double minute amplifications would not wholly explain the very rapid loss of Msx resistance. Thirdly the use of the fluorescent stain DAPI, which is capable of revealing DNA of the order of 150 kbp, shows no extrachromosomal DNA is present in C4M cells (Dr. R.H.Wilson, pers comm). Fourthly the migration characteristics of undigested C4M DNA in conventional and pulse field gels renders it unlikely that extrachromosomal molecules smaller than 150 kbp exist.

Whether the rearrangement of the 2, Z-2 and Z-7 chromosomes played some part in the amplification process or merely occurred serendipitously in parallel to the amplification of the GS genes is unknown. Since the chromosomal location of the parental GS gene is unknown it is possible that it lies in one of the

rearranged chromosomes and that the initial amplification event caused a rearrangement at that site.

Assuming that the amount of chromatin on a chromosome is directly proportional to that chromosome's length, or part thereof, measurements indicate that the q-arm of the E marker constitutes between 2 and 3% of the genome. Given an amplified unit repeated 1000 times with a minimum size of 70 kbp within a genome of 3×10^6 kb the amplified GS DNA will constitute at least 2% of the genome - a result which suggests that (i) the vast majority of the amplified GS DNA lies on the q-arm of the E marker chromosome and (ii) the bulk of the DNA present on that q-arm is amplified in nature.

CHAPTER 7

Overall Discussion and Proposals for Future Experiments

7.1 Overall Discussion

A cell line containing approximately 1000 copies of the glutamine synthetase gene has been produced by multistep selection for increasing resistance to methionine sulphoximine. I have shown that the karyotype of the C4M cell line contains several new chromosomes - the products of rearrangements of the parental KG1 chromosomes. The C4M line exhibits an increase in karyotypic variability relative to the KG1 line, possessing aberrant chromosomes and quasitetraploid cells with greater frequency. By *in situ* hybridization I showed that the GS genes were confined to the long arm of a small chromosome. A rough estimate of the size of the q-arm and the number of GS units present suggests that the DNA comprising the q-arm may be almost entirely amplified units.

Conventional restriction mapping and gel electrophoresis established that the DNA surrounding and including the GS gene was homogeneous over some 40 kbp and that there was a consistent hypomethylation pattern in the 5' region of the GS gene.

Pulse field gel electrophoresis showed that the DNA surrounding any one GS gene was homogeneous for distances of at least 300 kbp. It further showed that other cell lines showed different, but homogeneous, restriction patterns and that the restriction patterns of the C4M line differed from those of the KG1 and KG1M lines from which it was derived.

Several groups have analysed the structure of the amplified DNA from various cell lines. Their efforts to construct a consistent amplified unit have met with various degrees of success. The presence of inverted repeat structures was shown to be widespread by Ford & Fried (1986); occurring in every one of seven cell lines tested although there is no way of knowing if the inverted repeat structure is the sole arrangement. Cloning of the reflection point (Passananti *et al*, 1987; Saito & Stark, 1986; Looney & Hamlin, 1987) showed that the symmetry was not perfect but interrupted by a 1-2 kbp unique fragment at the centre. To date only Looney & Hamlin (1987) have succeeded in cloning an entire amplified unit and although they succeeded in defining the ends of the unit they were unable to deduce the relationship of one unit to another. PFGE offered an alternative

to cloning and defining an average unit and the results from it allows the consideration of various models of the structure of the amplicon cluster.

Fig 7.1 shows some of the possible models all of which have the following in common:

(a) there are no "random" ,ie. multiple unique, sequences present.

(b) the units are arranged tandemly with no interspersed sequences present.

(c) the structure of a unit is simple, ie no multiple internal repeats.

(d) each unit contains a GS gene.

Various observations must be incorporated into any successful model:

(i) conventional restriction analysis gives a unit size of 50-60 kbp.

(ii) the PFGE restriction pattern is homogeneous.

(iii) the presence of two discrete SstII fragments containing the GS gene.

(iv) the regular pattern of partial digests of Sst II or Sfi I.

(v) the differences and similarites in the restriction patterns of C4M and KG1M.

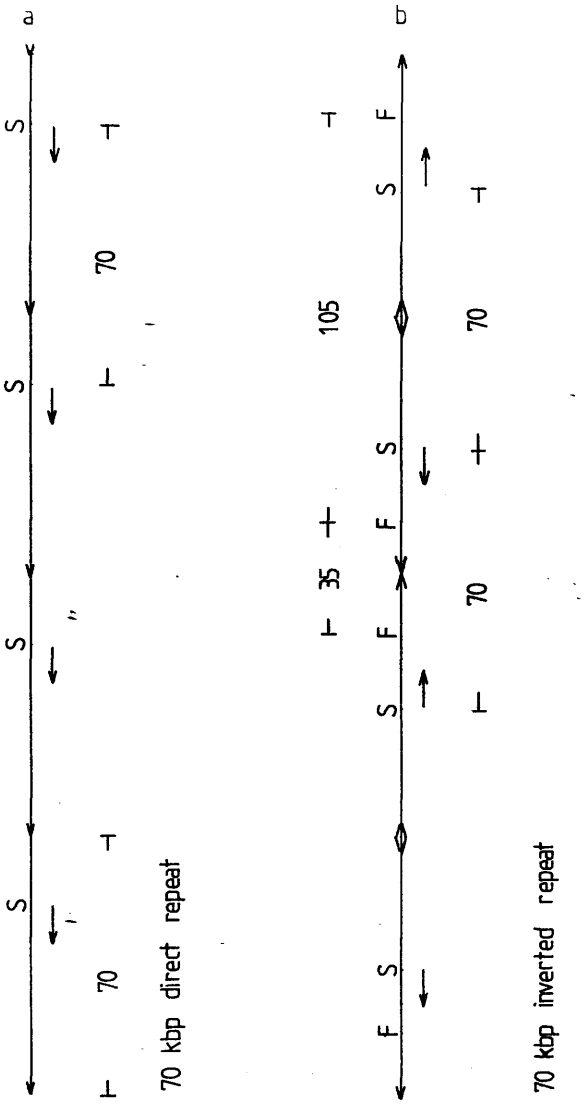
(vi) the correspondence in size between the Sst II fragment and the partial digest fragments of Sal I and Xho I.

(vii) the relative underhybridization of the second Sst II fragment.

(viii) the difference in hybridization patterns of gels probed with fragments to the 5' and 3' of the SstII sites.

Figure 7.1 shows some of the possible models of the amplified unit. The 70 kbp direct repeat unit (7.1a) model is highly implausible since neither the SstII nor the SfiI data are compatible with a tandem array composed of a single 70 kbp unit.

The 70 kbp inverted repeat unit (7.1b) model predicts that the SstII fragments revealed by fragments II or IV should be the same size (approx 75 kbp) and this is not borne out by the hybridization pattern of fig 4.10b and 4.11b. If the SstII



→ Glutamine synthetase gene

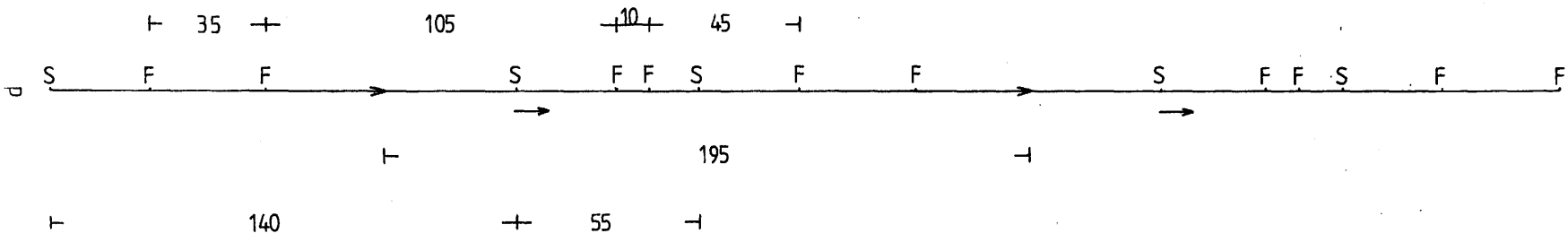
⊢ ⊣ Sizes in kbp

S Sst II site

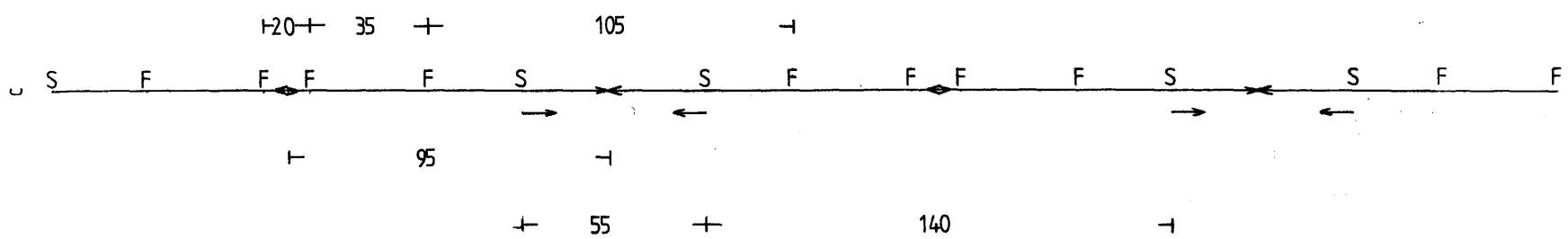
F Sfi I site

Ends of inverted repeat

End of direct repeat



195 kbp direct repeat



95 kbp inverted repeat

site is placed asymmetrically in order to produce a fragment of 55 kbp in size for fragment II to reveal it predicts that the other SstII fragment which would be revealed by fragment IV would be 85 kbp in size and again this is not borne out by the hybridization pattern of fig 4.11b.

The 95 kbp inverted repeat unit (7.1c) and 195 kbp direct repeat (7.1d) share similar flaws. Both predict that the SfiI restriction pattern should have a 195 kbp band, equivalent to two units or one unit respectively, and the presence of a band this size is definitely questionable in fig 4.10b. Similarly both models predict the presence of SfiI bands of lesser size which also may not be present, in the case of the inverted repeat SfiI bands of 160 and 175 kbp are expected and in the case of the direct repeat SfiI bands of 115 and 160 kbp. In addition the inverted repeat model would have difficulty in accommodating the reduction in size of the smaller SstII fragment from 90 kbp in KG1M to 55 kbp in C4M, while the 105 kbp SfiI fragment, which contains completely the smaller SstII fragment, remains with its size unaltered.

Both of the larger size models (7.1c and 7.1d) explain easily the difference in hybridization patterns generated by fragments II and IV, an observation which must be passed off as an hybridization artifact in the smaller size models.

Models which incorporate a third SstII fragment into their structure are of too great a size to be usefully speculated upon given the current data.

7.2 Future experiments

Several suggestions for localised restriction mapping are included in chapter 4 but the possibility that the amplified unit could be larger than 195 kbp will not be resolved by such procedures.

The only way to satisfactorily resolve these modelling problems is to create a cosmid library of C4M DNA and use that to map, in detail, an amplified unit. Because one cosmid cannot accommodate even the minimum size of a unit the resultant structure will be an average unit composed of pieces of several units. It will however be capable of defining the size of the unit and the presence or absence of inverted repeats.

Although a considerable amount of work has been devoted to the primary amplification event little has been devoted to later events.

Questions such as - is a secondary event just a 10-fold amplification of a single unit like the primary event?

Does a secondary event involve more than one unit?

Is the increase in copy number arithmetic or geometric?

How does the structure of the amplified unit alter as its copy number increases?

Is the size of the unit fixed during amplification?

Could it change during amplification or as a function of continued culture?

In the KG1M line is an ideal starting point to answer these questions. It readily undergoes further amplification and possesses a homogeneous amplified unit. Cloning of this unit would furnish a variety of probes, both near and distant to the GS gene which could be used as monitors for structural changes during amplification and in continued culture while PFGE would allow the mapping of these alterations at the multiunit level.

At the chromosomal level questions of whether secondary amplifications result in major chromosomal changes or whether the amplified units multiply in their original, primary, position could also be addressed using the KG1M cell line.

On a different tack the existence of 1000 copies of a housekeeping gene in a single cell provides an excellent vehicle for studies of the genes' chromatin structure. C4M cells remain responsive to glucocorticoids, which increase the level of GS protein, and studies of the DNAase I hypersensitivity patterns should reveal the DNA area affected by glucocorticoid stimulation. The high copy number should allow the relatively easy application of techniques which allow the probing of the fine structure of hypersensitive sites to reveal the interaction of protein(s) and DNA within them (Jackson & Felsenfeld, 1985) and to later characterise and isolate the proteins involved (Emerson et al, 1985).

Investigation has shown that replication and transcription occur at the nucleoskeleton (Jackson & Cook, 1985b and 1986) and

Gasser & Laemmli (1986) have succeeded in defining DNA regions to the 5' and 3' of several genes which bind to the nucleoskeleton. Similar experiments using C4M cells could define equivalent attachment regions whose position and sequence could be ascertained and related to the transcriptional control of the GS gene. Using a synchronised C4M cell culture it should be possible to define the replication origin and plasmid clones of this region could then be used to analyse the interaction between the origin, its chromatin structure and the nucleoskeleton.

Bibliography

Alitalo K, Schwab M, Lin CC, Varmus HE, Bishop JM (1983) Homogeneously staining chromosomal regions contain amplified copies of an abundantly expressed cellular oncogene (c*myc*) in malignant neuroendocrine cells from a human colon carcinoma.
Proc. Natl. Acad. Sci. USA **80**, 1707-1711

Amabis DC, Amabis JM (1984) Effects of ecdysterone in polytene chromosomes of *Trichosia pubescens*.
Devl. Biol. **102**, 1-9

Anand R (1986) Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules.
Trends In Genet. **2**, 278-283

Anderson RA, Eliasson SL (1986) Recombination of homologous DNA fragments transfected into mammalian cells occurs predominantly by terminal pairing.
Mol. Cell. Biol. **6**, 3246-3252

Anderson RP, Roth JR (1977) Tandem genetic duplications in phage and bacteria.
Ann. Rev. Microbiol. **31**, 473-505

Andrulis IL, Duff C, Evans-Blackler S, Worton R, Siminovitch L (1983) Chromosomal alterations associated with overproduction of asparagine synthetase in albizzin-resistant Chinese hamster ovary cells.
Mol. Cell. Biol. **3**, 391-398

Andrulis IL, Siminovitch L (1982) Amplification of the gene for asparagine synthetase.
pp 75-80 Gene Amplification (Cold Spring Harbour)

Ardeshir F, Giulotto E, Zieg J, Brison O, Liao WSL, Stark GR (1983) Structure of amplified DNA in different Syrian hamster cell lines resistant to **N**-(Phosphonacetyl)-**L**-aspartate.
Mol. Cell. Biol. **3**, 2076-2088

Barker PE, Hsu TC (1978) Are double minutes chromosomes?
Exp. Cell Res. **113**, 457-458

Barker PE, Stubblefield E (1979) Ultrastructure of double minutes from a human tumour cell line.
J. Cell Biol. **83**, 663-666

Battey J, Moulding C, Taub R, Murphy W, Stewart T, Potter H, Lenoir G, Leder P (1983) The human c-*myc* oncogene: structural consequences of translocations into the IgH locus of Burkitt lymphoma.
Cell **34**, 779-787

Bernards A, Kooter KM, Michels PAM, Hoberts RMP, Borst P (1986) Pulsed field gel electrophoresis of DNA digested in agarose allows the sizing of the large duplication unit of a surface antigen gene in trypanosomes.
Gene **42**, 313-322

Bernstein H (1977) Germ line recombination may be primarily a manifestation of DNA repair processes.
J. Theor. Biol. **69**, 371-380

Biedler J (1982) Evidence for transient or prolonged existence of amplified DNA sequences in antifolate-resistant, vincristine-resistant and human neuroblastoma cells.
pp 39-46 Gene Amplification (Cold Spring Harbour)

Bird AP (1986) CpG-rich islands and the function of DNA methylation.
Nature **321**, 209-213

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA.
Nucl. Acids Res. **7**, 1513-

Bishop JO, Selman GG, Hickman J, Black L, Saunders RDP, Clark AJ (1985) The 45-kb unit of major urinary protein gene organization is a gigantic imperfect palindrome.
Mol. Cell. Biol. 5, 1591-1600

Bostock CJ, Clark EM (1983) Gene amplification in methotrexate-resistant mouse cells.**V** Intact amplification units can be transferredto and amplified in methotrexate-senisitve mouse L cells.
Chromosoma **88**, 31-41

Brown PC, Beverley SM, Schimke RT (1981) Relationship of amplified dihydrofolate reductase genes to double minute chromosomes in unstably resistant mouse fibroblast cell lines.
Mol. Cell. Biol. **1**, 1077-1083

Brown WRA, Bird AP (1986) Long-range restriction site mapping of mammalian genomic DNA.
Nature **322**, 477-481

Burhans WC, Selegue JE, Heintz NH (1986) Isolation of the origin of replication associated with the amplified Chinese hamster dihydrofolate reductase domain.
Proc. Natl. Acad. Sci. USA **83**, 7790-7794

Caizzi R, Bostock CJ (1982) Gene amplification in methotrexate resistant cells: IV Different DNA sequences are amplified in different resistant cell lines.
Nucl. Acids Res. **10**, 6597-6618

Carle GF, Frank M, Olson MV (1986) Electrophoretic separations of large DNA molecules by periodic inversion of the electric field.
Sci. **232**, 65-68

Carle GF, Olson MV (1984) Separation of chromosomal DNA molecules from yeast by orthogonal field alternation gel electrophoresis.
Nucl. Acids Res. **12**, 5647-5664

Carle GF, Olson MV (1985) An electrophoretic karyotype for yeast.
Proc. Natl. Acad. Sci. USA **82**, 3756-3760

Carroll SM, Gaudrey P, de Rose ML, Emery JF, Mienkoth JL, Nakkim E, Subler M, von Hoff DD, Wahl GM (1987) Characterization of an episome produced in hamster cells that amplify a transfected CAD gene at high frequency : functional evidence for a mammalian replication origin.
Mol. Cell. Biol. **7**, 1740-1750

Cartier M, Chang MW-M, Stanners CP (1987) Use of the Escherichia coli gene for asparagine synthetase as a selective marker in a shuttle vector capable of dominant transfection and amplification in animal cells.
Mol. Cell. Biol. **7**, 1623-1628

Chapman AB, Costello MA, Lee F, Ringold GM (1983) Amplification and hormone-regulated expression of a mouse mammary tumour virus-Ecogpt fusion plasmid in mouse 3T6 cells.
Mol. Cell. Biol. **3**, 1421-1429

Coderre JA, Beverley SM, Schimke RT, Santi DV (1983) Overproduction of a bifunctional thymidylate synthetase

dihydrofolate reductase and DNA amplification in methotrexate resistant Lieshmania tropica.
Proc. Natl. Acad. Sci. USA **80**, 2132-2136

Collins SJ (1986) Breakpoints on chromosome 9 and 22 in Philadelphia chromosome-positive chronic myelogenous leukaemia (CML) - amplification of rearranged c-abl oncogenes in CML blast crisis.
J. Clin. Invest. **78**, 1392-1396

Collins SJ, Groudine MT (1987) Chronic myelogenous leukaemia: amplification of a rearranged c-abl oncogene in both chronic phase and blast crisis.
Blood **69**, 893-898

Cooper CS, Tempest PR, Beckman MP, Heldin C-H, Brookes P (1986) Amplification and overexpression of the met gene in spontaneously transformed NIH 3T3 mouse fibroblasts.
EMBO J. **5**, 2623-2628

d'Anna JA, Church VL, Tobey RA (1986) Changes in H1 content, nucleosome repeat lengths and DNA elongation under conditions of hydroxyurea treatment that reportedly facilitate gene amplification.
Biochim. Biophys. Acta. **868**, 226-237

Deaven LL, Petersen DF (1973) The chromosomes of CHO, an anueploid Chinese hamster cell line: G-band, C-band and autoradiographic analyses.
Chromosoma **41**, 129-144

de Bruijn MHL, van der Bliek AM, Biedler JL, Borst P (1986) Differential amplification and disproportionate expression of five genes in three multidrug-resistant Chinese hamster cell lines.
Mol. Cell. Biol. **6**, 4717-4722

Edlund T, Normark S (1981) Recombination between short DNA homologies causes tandem duplications.
Nature **292**, 269-271

Emanuel BS, Balaban G, Boyd JP, Grossman A, Negishi M, Parmiter A, Glick MC (1985) N-myc amplification in multiple homogeneously staining regions in two human neuroblastomas.
Proc. Natl. Acad. Sci. USA **82**, 3736-3740

Emanuel JR, Garetz S, Schnieder J, Ash JF, Benz EJ (Jnr), Levenson R (1986) Amplification of DNA sequences coding for the Na,K-ATPase -subunit in ouabain-resistant C^+ cells.
Mol. Cell. Biol. **6**, 2476-2481

Emerson BM, Lewis CD, Felsenfeld G (1985) Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult B-globin gene: nature of the binding domain.
Cell **41**, 21-30

Federspiel NA, Beverley SM, Schilling JW, Schimke RT (1984) Novel DNA rearrangements are associated with dihydrofolate reductase gene amplification.
J. Biol. Chem. **259**, 9127-9140

Fendrock B, Destrempes M, Kaufman RJ, Latt SA (1986) Cytological, flow cytometric and molecular analysis of the rapid evolution of mammalian chromosomes containing highly amplified DNA sequences.
Histochem. **84**, 121-130

Fenwick RG (Jnr), Fuscoe JC, Caskey CT (1984) Amplification versus mutation as a mechanism for reversion of an HGPRT mutation.
Som. Cell Mol. Genet. **10**, 71-84

Filmus J, Trent JM, Pullano R, Buick RN (1986) A cell line from a human ovarian carcinoma with amplification of the K-*ras* gene.
Cancer Res. **46**, 5179-5182

Flintoff WF, Livingston E, Duff C, Worton RG (1984) Moderate level gene amplification in methotrexate-resistant Chinese hamster ovary cells is accompanied by chromosomal translocations at or near the site of the amplified DHFR gene.
Mol. Cell Biol. **4**, 69-76

Ford M, Davies B, Griffiths M, Wilson J, Fried M (1985) Isolation of a gene enhancer within an amplified inverted duplication after "expression selection".
Proc. Natl. Acad. Sci. USA **82**, 3370-3374

Ford M, Fried M (1986) Large inverted duplications are associated with gene amplification.

Cell **45**, 425-430

Futcher AB (1986) Copy number amplification of the 2 uM circle plasmid of Saccharomyces cerevisiae.
J. Theor. Biol. **119**, 197-204

Gall JG (1968) Differential synthesis of the genes for ribosomal RNA during amphibian oogenesis.
Proc. Natl. Acad. Sci. USA **60**, 553-560

Gardiner K, Laas W, Patterson D (1986) Fractionation of large mammalian DNA restriction fragments using vertical pulse-field gradient gel electrophoresis.
Som. Cell Mol. Genet. **12**, 185-195

Gasser C, Simonsen CC, Schilling JW, Schimke RT (1982) Expression of abbreviated mouse dihydrofolate reductase genes in cultured hamster cells.
Proc. Natl. Acad. Sci. USA **79**, 6522-6526

Gasser SM, Laemmli UK (1986) Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D.melanogaster.
Cell **46**, 521-530

George DL, Powers VE (1981) Cloning of DNA from double minutes of Y1 mouse adrenocortical tumour cells: evidence for gene amplification.
Cell **24**, 117-123

Ghazal P, Clark AJ, Bishop JO (1985) Evolutionary amplification of a pseudogene.
Proc. Natl. Acad. Sci. USA **82**, 4182-4185

Gick CC, McCarty KS (Snr) (1982) Amplification of the metallothionein-I gene in cadmium and zinc resistant Chinese hamster ovary cells.
J. Biol. Chem. **257**, 9049-9053

Gilman M (1987) Nuclear oncogenes- a meeting review.
Genes & Devel. **1**, 213-215

Giulotto E, Knights C, Stark GR (1987) Hamster cells with increased rates of DNA amplification, a new phenotype.
Cell **48**, 837-845

Giulotto E, Saito E, Stark GR (1986) Structure of DNA formed in the first step of CAD gene amplification.
EMBO J. **5**, 2115-2121

Glover DM, Zaha A, Stocker AJ, Santelli RV, Pueyo MT, de Toledo SM, Lara FJS (1982) Gene amplification in Rhynchosciara salivary gland chromosomes.
Proc. Natl. Acad. Sci. USA **79**, 2947-2951

Graham MY, Otani T, Boime I, Olson MV, Carle GF, Chaplin DD (1987) Cosmid mapping of the human chorionic gonadotropin B-subunit genes by field-inversion gel electrophoresis.
Nucl. Acids Res. **15**, 4437-4448

Grund SH, Patil SR, Shal HO, Pauur PC, Stadler JK (1983) Correlation of unstable multidrug cross-resistance in Chinese hamster ovary cells with a homogeneously staining region on chromosome 1.
Mol. Cell Biol. **3**, 1634-1647

Hardy DA, Bell JI, Long EO, Lindsten T, McDevitt HO (1986) Mapping of the class II region of the human major histocompatability complex by pulsed field gel electrophoresis.
Nature **323**, 453-455

Hayward BE, Hussain A, Wilson RH, Lyons A, Woodcock V, McIntosh B, Harris TJR (1986) The cloning and nucleotide sequence of cDNA for an amplified glutamine synthetase gene from the Chinese hamster.
Nucl. Acids Res. **14**, 999-1008

Helling RB, Goodman HM, Boyer HW (1974) Analysis of the R.EcoRI fragments of DNA from lambdoid bacteriophage and other viruses by agarose gel electrophoresis.
J. Virol. **14**, 1235-1244

Hill AB, Schimke RT (1985) Increased gene amplification in L5178Y mouse lymphoma cells with hydroxyurea-induced chromosomal aberrations.
Cancer Res. **45**, 5050-5057

Hourcade D, Dressler D, Wolfson J (1973) The nucleolus and the rolling circle.
Cold Spring Harbour Symp. Quant. Biol. **38**, 537-550

Hunts JH, Shimizu N, Yamamoto T, Toyoshima K, Merlino GT, Xu Y-H, Pastan I (1985) Translocation chromosome 7 of A431 cells contains amplification and rearrangement of EGF receptor gene responsible for production of variant mRNA.

Som. Cell Mol. Genet. **11**, 477-484

Hyrien O, Buttin G (1986) Gene amplification in pesticide-resistant insects.
Trends In Genet. **2**, 275-276

Jackson DA, Cook PR (1985a) A general method for preparing chromatin containig intact DNA.
EMBO J. **4**, 913-918

Jackson DA, Cook PR (1985b) Transcription occurs at a nucleoskeleton.
EMBO J. **4**, 919-925

Jackson DA, Cook PR (1986) Replication occurs at a nucleoskeleton.
EMBO J. **5**, 1403-1410

Jackson PD, Felsenfeld G (1985) A method for mapping intranuclear protein-DNA interactions and its application to a nuclease hypersensitive site.
Proc. Natl. Acad. Sci. USA **82**, 2296-2300

Johnston BE, Ihde DC, Makuch RW, Gazdar AF, Carney DN, Ole H, Russell E, Nan MH, Minna JD (1987) Myc family oncogene amplification in tumour cell lines established from small cell lung cancer patients and its relationship to clinical status and course.
J. Clin. Invest. **79**, 1629-1634

Johnston PJ, Borst P (1986) Mapping of VSG genes on large expression-site chromosomes of Trypanosoma brucei separated by pulsed field gradient electrophoresis.
Gene **43**, 213-220

Johnston RN, Beverley SM, Schimke RT (1983) Rapid spontaneous dihydrofolate reductase gene amplification shown by flouresence-activated cell sorting.
Proc. Natl. Acad. Sci. USA **80**, 3711-3715

Johnston RN, Feder J, Hill AB, Sherwood SW, Schimke RT (1986) Transient inhibition of DNA synthesis results in increased dihydrofolate reductase synthesis and subsequent increased DNA content per cell.
Mol. Cell. Biol. **6**, 3373-3381

Jones CW, Kafatos FC (1980) Structure, organization and evolution of developmentally regulated chorion genes in a silkmoth.

Cell **22**, 855-867

Kalt MR, Gall JG (1974) Observations on early gene development and premeiotic ribosomal DNA amplification in Xenopus laevis.
J. Cell Biol. **62**, 460-472

Kao F, Puck TT (1969) Genetics of somatic mammalian cells: IX Quantitation of mutagenesis by physical and chemical agents.
J. Cell. Physiol. **74**, 245-257

Karathansis SK (1985) Apolipoprotein multigene family: tandem organization of human apolipoprotein AI, CIII and AIV genes.
Proc. Natl. Acad. Sci. USA **82**, 6374-6378

Kaufman RJ, Brown PC, Schimke RT (1979) Amplified dihydrofolate reductase genes in unstably methotrexate-resistant cells are associated with double minute chromosomes.
Proc. Natl. Acad. Sci. USA **76**, 5669-5673

Kaufman RJ, Brown PC, Schimke RT (1981) Loss and stabilization of amplified dihydrofolate reductase genes in mouse sarcoma S-180 cell lines.
Mol. Cell. Biol. **1**, 1084-1093

Kaufman RJ, Murtha P, Ingolia DE, Yeung C-Y, Kellems RE (1986) Selection and amplification of heterologous genes encoding adenosine deaminase in mammalian cells.
Proc. Natl. Acad. Sci. USA **83**, 3136-3140

Kaufman RJ, Sharp PA, Latt SA (1983) Evolution of chromosomal regions containing transfected and amplified dihydrofolate reductase sequences.
Mol. Cell. Biol. **3**, 699-711

Kellems RE, Harper ME, Smith LM (1982) Amplified dihydrofolate reductase genes are located in chromosome regions containing DNA that replicates during the first half of S-phase.
J. Cell Biol. **92**, 531-539

Kenwrick S, Patterson M, Speer A, Fischbeck K, Davies K (1987) Molecular analysis of the Duchenne muscular dystrophy region using pulsed field gel electrophoresis.
Cell **48**, 351-357

Kink JA, Chang K-P (1987) Tunicamycin-resistant Leishmania mexicana amazonensis: expression of virulence associated with an increased activity of N-acetylglucosaminyl-

transferase and amplification of its presumptive gene.
Proc. Natl. Acad. Sci. USA **84**, 1253-1257

Kinzler KW, Zehnbauer BA, Brodeur GM, Seeger RC, Trent JM, Meltzer PS, Vogelstein B (1986) Amplification units containing human N-myc and c-myc genes.
Proc. Natl. Acad. Sci. USA **83**, 1031-1035

Kiss GB, Amin AA, Pearlman RE (1981) Two separate regions of the extrachromosomal ribosomal deoxyribonucleic acid of Tetrahymena thermophila enable autonomous replication of plasmids in Saccharomyces cerevisiae.
Mol. Cell. Biol. **1**, 535-543

Klein HL, Petes TD (1981) Intrachromosomal gene conversion in yeast.
Nature **289**, 144-148

Kleinberger T, Etkin S, Lavi S (1986) Carcinogen-mediated methotrexate resistance and dihydrofolate reductase amplification in Chinese hamster cells.
Mol. Cell. Biol. **6**, 1958-1964

Kuo T, Pathak S, Ramagli L, Rodiguez L, Hsu TC (1982) Vincristine-resistant Chinese hamster ovary cells.
pp 53-58 Gene Amplification (Cold Spring Harbour)

Lavi S (1981) Carcinogen-mediated amplification of viral DNA sequences in simian virus 40-transformed Chinese hamster embryo cells.
Proc. Natl. Acad. Sci. USA **78**, 6144-6148

Leinonen P, Alhonen-Hongisto L, Laine R, Janne OA, Janne J (1987) Human myeloma cells acquire resistance to difluoromethyl-ornithine by amplification of the ornithine decarboxylase gene.
Biochem J. **242**, 199-203

Lewis JA, Davide JP, Melera PW (1982) Selective amplification of polymorphic dihydrofolate reductase gene loci in Chinese hamster lung cells.
Proc. Natl. Acad. Sci. USA **79**, 6961-6965

Looney JE, Hamlin JL (1987) Isolation of the amplified dihydrofolate reductase domain from methotrexate-resistant Chinese hamster ovary cells.
Mol. Cell. Biol. **7**, 569-577

Luskey KL, Faust JR, Chin DJ, Brown MS, Goldstein JL (1983) Amplification of the gene for 3-hydroxy-3-methylglutaryl coenzyme A reductase but not for the 53 kDa protein in UT-1 cells.
J. Biol. Chem. **258**, 8462-8469

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory approach.
Cold Spring Harbour Laboratory Press

Mariani BD, Schimke RT (1984) Gene amplification in a single cell cycle in Chinese hamster ovary cells.
J. Biol. Chem. **259**, 1901-1910

Mason AF, Evans BA, Cox DR, Shine J, Richards RI (1983) Structure of mouse kallikrein gene family suggests a role in specific processing of biologically active peptides.
Nature **303**, 300-307

Maurer BJ, Lai E, Hankalo BA, Hood L, Attardi G (1987) Novel submicroscopic extrachromosomal elements containing amplified genes in human cells.
Nature **327**, 434-437

Mayo KE, Palmiter RD (1982) Glucocorticoid regulation of the mouse metallothionein I gene is selectively lost following amplification of the gene.
J. Biol. Chem. **257**, 3061-3067

McConlogue L, Dana SL, Coffino P (1986) Multiple mechanisms are responsible for altered expression of ornithine decarboxylase in overproducing variant cells.
Mol. Cell. Biol. **6**, 2865-2871

McCormick F, Trahey M, Innis M, Dieckmann B, Ringold G (1984) Inducible expression of amplified human beta interferon genes in CHO cells.
Mol. Cell. Biol. **4**, 166-172

McPeek FD (Jnr), Coyle-Morris JF, Gemmill RM (1986) Separation of large DNA molecules by modified pulsed field gradient gel electrophoresis.
Anal. Bioch. **156**, 274-285

Meinkoth J, Killary AM, Fournier REK, Wahl GM (1987) Unstable and stable CAD gene amplification: importance of flanking sequences and nuclear environment in gene amplification.

Mol. Cell. Biol. **7**, 1415-1424

Heinkoth JL, Legouy E, Brison O, Wahl GM (1986) New RNA species is produced by alternate polyadenylation following rearrangement associated with CAD gene amplification.
Som. Cell Mol. Genet. **12**, 389-350

Milbrandt JD, Azizkhan JC, Hamlin JL (1983) Amplification of a cloned Chinese hamster dihydrofolate reductase gene after transfer into a dihydrofolate reductase-deficient cell line.
Mol. Cell. Biol. **3**, 1274-1282

Milbrandt JD, Heintz NH, White WC, Rothman SM, Hamlin JL (1981) Methotrexate-resistant Chinese hamster ovary cells have amplified a 135-kilobase pair region that includes the dihydrofolate reductase gene.
Proc. Natl. Acad. Sci. USA **78**, 6043-6047

Morgan WF, Bodycote J, Fero ML, Hahn PJ, Kapp LN, Pantelias GE, Painter RB (1986) A cytogenetic investigation of DNA rereplication after hydroxyurea treatment: implications for gene amplification.
Chromosoma **93**, 191-196

Morgan WF, Fero ML (1987) Induced sister chromatid exchange is not increased in homogeneously staining regions that contain amplified genes.
Cancer Genet. Cytogenet. **26**, 245-251

Orr-Weaver TL, Spradling AC (1986) <u>Drosophila</u> chorion gene amplification requires an upstream region regulating s18 transcription.
Mol. Cell. Biol. **6**, 4624-4633

Osheim YN, Miller OL (Jnr) (1983) Novel amplification and transcriptional activity of chorion genes in <u>Drosophila melanogaster</u> follicle cells.
Cell **33**, 543-553

Pardue M (1985) in Nucleic Acid Hybridization: a practical approach. Ed. Hames BD, Higgins SJ (IRL Press).
pp 179-202

Parks S, Spradling A (1987) Spatially regulated expression of chorion genes during <u>Drososphila</u> oogenesis.
Genes & Devel. **1**, 497-509

Passananti C, Davies B, Ford M, Fried M (1987) Structure ofan inverted duplication formed as a first step in a gene amplification event: implications for a model of gene amplification.
EMBO J. **6**, 1697-1703

Perry M, Thomsen GH, Roeder RG (1985) Genomic organization and nucleotide sequence of two distinct histone gene clusters from Xenopus laevis.
J. Mol. Biol. **185**, 479-499

Perucho M, Hanahan D, Wigler M (1980) Genetic and physical linkage of exogenous sequences in transformed cells.
Cell **22**, 309-317

Petes TD (1980) Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes.
Cell **19**, 765-774

Philippsen P, Kramer RA, Davies RW (1978) Cloning of yeast ribosomal DNA repeat unit in SstI and HindIII lambda vectors using genetic and physical size selections.
J. Mol. Biol. **123**, 371-386

Puck TT, Cieciura SJ, Robinson A (1958) Genetics of somatic cell mammalian cells.**III** Long-term cultivation of euploid cells from human and animal subjects.
J. Exp. Med. **108**,945-956

Razin SV, Kekelidge MG, Lukanidin EM, Scherrer K, Georgiev GP (1986) Replication origins are attached to the nuclear skeleton.
Nucl. Acids Res. **14**, 8189-8207

Rice GC, Hoy C, Schimke RT (1986) Transient hypoxia enhances the frequency of dihydrofolate reductase gene amplification in Chinese hamster ovary cells.
Proc. Natl. Acad. Sci. USA **83**, 5978-5982

Roberts JM, Axel R (1982) Amplification and correction of transformed genes.
pp 251-255 Gene Amplification (Cold Spring Harbour)

Roberts JM, Buck LB, Axel R (1983) A structure for amplified DNA.
Cell **33**, 53-63

Rochaix J-D, Bird A, Bakken A (1974) Ribosomal RNA gene amplification by rolling circles.

J. Mol. Biol. **87**, 473-487

Roth DB, Wilson JH (1986) Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction.
Mol. Cell. Biol. **6**, 4295-4304

Saito I, Stark GR (1986) Charomids: cosmid vectors for efficient cloning and mapping of large or small restriction fragments.
Proc. Natl. Acad. Sci. USA **83**, 8664-8668

Sanders PG, Wilson RH (1984) Amplification and cloning of the Chinese hamster glutamine synthetase gene.
EMBO J. **3**, 65-71

Schenkel H, Kejzlarova-Lepesant J, Berreur P, Moreau J, Scheller K, Bregegere F, Lepesant J-A (1985) Identification and molecular analysis of a multigene family encoding calliphorin, the major larval serum protein of Calliphora vicina.
EMBO J. **4**, 2983-2990

Schilling J, Beverley S, Simonsen C, Crouse C, Setzer D, Feagin J, McGrogan M, Kohlmiller N, Schimke RT (1982) Dihydrofolate reductase gene structure and variable structure of amplified DNA structure in mouse cell lines.
pp149-154 Gene Amplification (Cold Spring Harbour)

Schimke RT (1984) Gene amplification in cultured animal cells.
Cell **37**, 705-713

Schimke RT, Sherwood SW, Hill AB, Johnston RN (1986) Overreplication and recombination of DNA in higher eukaryotes: potential consequences and biological implications.
Proc. Natl. Acad. Sci. USA **83**, 2157-2161

Schwab M, Klempnauer K-H, Alitalo K, Varmus H, Bishop M (1986) Rearrangement at the 5' end of amplified c-myc in human COLO 320 cells is associated with abnormal transcription.
Mol. Cell. Biol. **6**, 2752-2755

Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis.
Cell **37**, 67-75

Schwartz DC, Saffron W, Welsh J, Haas R, Goldenberg M, Cantor CR (1982) New techniques for purifying large DNAs and studying

their properties and packaging.
Cold Spring Harbour Symp. Quant. Biol. **47**, 189-195

Seabright M (1971) A rapid banding technique for human chromosomes.
Lancet **11**, 971-972

Shiloh Y, Korf B, Kohl NE, Sakai K, Brodeur GM, Harris P, Kanda N, Seeger RC, Alt F, Latt SA (1986) Amplification and rearrangement of DNA sequences from the chromosomal region 2p24 in human neuroblastomas.
Cancer Res. **46**, 5297-5301

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis.
J. Mol. Biol. **98**, 503-533

Spradling AC (1981) The organization and amplification of two chromosomal domains containing Drosophila chorion genes.
Cell **27**, 203-209

Spradling AC, de Cicco DV, Wakimoto BT, Levine JF, Kalfayan LJ, Cooley L (1987) Amplification of the X-linked Drosophila chorion gene cluster requires a region upstream from the s38 chorion gene.
EMBO J. **6**, 1045-1053

Spradling AC, Mahowland AP (1980) Amplification of genes for chorion proteins during oogenesis in Drosophila melanogaster.
Proc. Natl. Acad. Sci. USA **77**, 1096-1100

Spradling AC, Mahowland AP (1981) A chromosome inversion alters the pattern of specific DNA replication in Drosophila follicle cells.
Cell **27**, 203-209

Stanners CP (1971) Two types of ribosome in mouse-hamster hybrid cells.
Nature (New Biol.) **230**, 52-54

Stark GR, Wahl GM (1984) Gene amplification.
Ann. Rev. Biochem. **53**, 447-491

Straus DS, Straus LD (1976) Large overlapping tandem genetic duplications in Salmonella typhimurium.
J. Mol. Biol. **103**, 143-153

Suttle DP (1983) Increased levels of UMP synthase protein and mRNA in pyrazofurin-resistant rat hepatoma cells.
J. Biol. Chem. **258**, 7707-7713

Szostak JW, Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors.
Cell **29**, 245-255

Teeter LD, Atsumi S, Sen S, Kuo T (1986) DNA amplification in multidrug, cross-resistant Chinese hamster ovary cells: Molecular characterization and cytogenetic localization of the amplified DNA.
J. Cell Biol. **103**, 1159-1166

Thiebaud CH (1979) Quantitative determination of amplified rDNA and its distribution during oogenesis in Xenopus laevis.
Chromosoma **73**, 37-44

Trent J, Meltzer P, Rosenblum M, Harsh G, Kinzler K, Mashal R, Feinberg A, Vogelstein B (1986) Evidence for rearrangement, amplification and expression of c-myc in a human glioblastoma.
Proc. Natl. Acad. Sci. USA **83**, 470-473

Van der Ploeg LHT, Cornelissen AWCA, Michels PAM, Borst P (1984a) Chromosome rearrangements in Trypanosoma brucei.
Cell **39**, 213-221

Van der Ploeg LHT, Schwartz DC, Cantor CR, Borst P (1984b) Antigenic variation in Trypanosoma brucei analysed by electrophoretic separation of chromosome sized DNA molecules
Cell **37**, 77-84

Vieira J, Messing J (1982) The pUC plasmids and the M13mp7 derived system for insertion mutagenesis with synthetic universal primers.
Gene **19**, 259-268

Vogelstein B, Pardoll DM, Coffey DS (1980) Supercoiled loops and eukaryotic DNA replication.
Cell **22**, 79-85

Wahl GM, de Saint Vincent RB, de Rose ML (1984) Effect of chromosomal position on amplification of transfected genes in animal cells.
Nature **307**, 516-520

Wahl GM, Vitto L, Padgett RA, Stark GR (1982) Single copy and amplified CAD genes in Syrian hamster chromosomes localised by a highly sensitive method for *in situ* hybridization.
Mol. Cell. Biol. 2, 308-319

Walton JD, Paquin CE, Kaneko K, Williamson VM (1986) Resistance to antimycin A in yeast by amplification of ADH4 on a linear, 42 kbp palindromic plasmid.
Cell **46**, 857-863

Weiner AM, Denison RA (1982) Either gene amplification or gene conversion may maintain the homogeneity of the multigene family encoding human U1 small nuclear RNA.
Cold Spring Harbour Symp. Quant. Biol. **47**, 1141-1149

Weith A, Winking H, Brackmann B, Boldyreff B, Traut W (1987) Microclones from a mouse germ cell line HSR detect amplification and complex rearrangements of DNA sequences.
EMBO J. **6**, 1295-1300

Whoriskey SK, Nghiem V-H, Leong P-M, Masson J-M, Miller JH (1987) Genetic rearrangements and gene amplification in *Escherichia coli*: DNA sequences at the juncture of amplified gene fusions.
Genes & Devel. **1**, 227-237

Worton RG (1978) Karyotypic heterogeneity in CHO cell lines.
Cytogenet. Cell Genet. **21**, 105-110

Worton RG, Ho CC, Duff C (1977) Chromosome stability in CHO cells.
Som. Cell Genet. **3**, 27-45

Wright JA, Alam TG, McClarty GA, Tagger AY, Thelander L (1987) Altered expression of ribonucleotide reductase and role of M2 gene amplification in hydroxyurea-resistant hamster, mouse, rat and human cell lines.
Som. Cell Mol. Genet. **13**, 155-165

Yanish-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors.
Gene **33**, 103-119

Yao M-C, Blackburn EH, Gall JG (1978) Amplification of the rDNA genes in *Tetrahymena*.
Cold Spring Harbour Symp. Quant. Biol. **43**, 1293-1296

Yeung C-Y, Frayne LC, Al-Ubaidi MR, Hook AG, Wright DA, Kellems RE (1983) Amplification and molecular cloning of murine adenosine deaminase gene sequences.
J. Biol. Chem. **258**, 15179-15185

Zieg J, Clayton CC, Ardeshir F, Giulotto E, Swyryd EA, Stark GR (1983) Properties of single step mutants of Syrian hamster cell lines resistant to N-(phosphonacetyl)-L-aspartate.
Mol. Cell. Biol. 3, 2089-2098

Zownir O, Fuscoe JC, Fenwick R, Morrow J (1984) Gene amplification as a mechanism of reversion at the HPRT locus of V79 Chinese hamster cells.
J. Cell. Physiol. **119**, 341-348