DNA AMPLIFICATION IN MAMMALIAN CELLS

Richard P. GROVES

Thesis presented in partial fulfilment of the degree of Doctor of Philosophy to the University of London

February 1992

Department of Molecular Biology

Imperial Cancer Research Fund

Lincoln's Inn Fields

London WC2A 3PX.

Department of Biochemistry

University College London

Gower Street

London WC1 6BT.

ABSTRACT

DNA Amplification in Mammalian Cells

DNA amplification has been observed in mammalian cells derived from tumours and also in tissue culture cells. The initial amplification event is thought to be random and in some cases the amplified DNA contains a gene whose consequent over-expression can confer a growth advantage or reduced sensitivity to a cytotoxic compound. In tissue culture systems, specific drugs can be used to isolate cells which have previously amplified a specific gene. The mechanisms responsible for these DNA abnormalities have been difficult to study as the initial events are rare and the amplification products could previously only be visualised many cell generations after their formation. During this time secondary mechanisms may alter the structure and appearance of the amplified arrays making an evolutionary interpretation of the amplification difficult. Until recently, only indirect methods were available to study the structure of amplified arrays but advances in fluorescent *in situ* hybridisation of metaphase chromosomes have allowed direct visualisation of very early gene amplification events.

Part of this thesis describes an attempt to construct a model cell line in which a single predetermined locus could be amplified simultaneously throughout a cell population in a controlled manner. Cosmids covering the monkey CAD locus were available and one was used to construct a homologous replacement fragment containing an SV40 replication origin and a dominant selectable marker. The fragment was transfected into ts-COS cells and if a cell line could have been isolated in which the fragment had integrated by homologous recombination, it may have been possible to over-replicate the locus by a shift to the permissive temperature to mimic the 'onionskin' model of gene amplification.

Other chapters of this thesis are concerned with the evolution and stability of amplified DNA and also the analysis of amplified DNA by *in situ* hybridisation on human metaphase chromosomes. Amplified DNA is lost very rapidly from PALA-resistant mutants selected in a single selection step but in mutants selected in several steps of increasing drug concentration the amplified genes are lost more slowly, if at all. PALA resistance in human cells has not been extensively investigated and, following recent advances in the understanding of early events in gene amplification provided by *in situ* hybridisation analysis of PALA-resistant BHK cells, a human system would provide a more accessible genome for further experiments.

To Christine,

my long suffering supporter

Je me presse de rire de tout, de peur d'être obligé d'en pleurer.

Pierre-Augustin de Beaumarchaise (1732-1799)

I am grateful to everyone who has been supportive to me in my time at ICRF, in particular Katy Smith, Andrew Ackrill, Catherine Mao and last but by no means least, Jim Wright. I am particularly in debt to Katy Smith without whose help the *in situ* hybridisation projects could not have progressed and also to Mary Stark for assistance along the way. I finally wish to extend my thanks to George Stark for the opportunity to find out what science is really like. As James Cabell once wrote: 'I am willing to taste any drink once'... it was an interesting experience.

As a post-script, I owe several large favours to Tim Copsey who risked life and lungs motorcycling around London collecting and delivering copies of this thesis in itsrevised form, whilst I was otherwise engaged in France.

Contents

Figures

Tables

Abbreviations

Abbreviations used throughout are those suggested by the Biochemical Journal, *273*, pp 1-19 (1991). Additional abbreviations used are listed below.

BHK	Baby hamster kidney cells
CAD	A multifunctional enzyme encoding carbamoyl phosphate synthetase, aspartate transcarbamylase and dihydroorotase activties
CHO	Chinese hamster ovary cells
ts-COS	CV1, origin defective, SV40 containing monkey cells
DHFR	Dihydrofolate reductase
DMSO	Dimethyl sulphoxide
DTT	DL-Dithiothreitol
FITC	Fluorescein isothiocyanate
MTX	Methotrexate
PALA	*N*-(phosphonacetyl)-L-aspartate
PEG	Polyethylene glycol
SV40	Simian Virus 40
T-ag	Large T-antigen of SV40

Suppliers

Chemicals were purchased from Sigma Chemical Co. Ltd. or BDH, Ltd. with the following exceptions:

Acrylamide and bis-acrylamide (premixed)	National Diagnostics, USA
Agarose (all types)	BRL, USA
Agar, yeast extract, casamino acids and tryptone	Difco, USA
Caesium chloride (CsCl)	Koch Light, UK
Dulbecco's modified Eagle's medium (DMEM)	Flow Laboratories, UK
PALA	manufactured and provided by the National Cancer Institute, USA
Restriction and other enzymes	New England Biolabs, USA

Commonly Used Solutions

50x Denhardt's solution (aqueous)	1% Ficoll 1% polyvinylpyrrolidone 1% BSA
2.5x HTM (aqueous)	0.5 M Hepes (pH 6.6) 0.125 M Tris (pH 8) 12.5 mM $MgCl_2$ 20 mM β–mercaptoethanol
PBSA	1% NaCl 0.025% KCl 0.14% Na_2HPO_4 0.025% KH_2PO_4

SM	0.58% NaCl 0.2% $MgSO_4.7\ H_2O$ 50 mM Tris (pH 7.5) 0.01% gelatin
1x SSC (pH 7.5)	150 mM NaCl 150 mM sodium citrate
1x SSPE (pH 7.4)	150 mM NaCl 10 mM NaH_2PO_4 1 mM EDTA (pH 8)
1x TAE (pH 8)	40 mM Tris-acetate 1 mM EDTA
1x TBE	89 mM Tris-borate 89 mM boric acid 2 mM EDTA
TE	10 mM Tris (pH 7.5) 1 mM EDTA (pH 8)

Media Compositions

Luria-Bertani (LB) broth	1% Bacto-tryptone 1% NaCl 0.5% yeast extract
NZCYM broth	1% NZ-amine 0.5% NaCl 0.5% yeast extract 0.1% casamino acids 0.2% $MgSO_4.7\ H_2O$
Psi (φ) broth	2% tryptone 0.5% yeast extract 20 mM $MgSO_4.7\ H_2O$ 10 mM NaCl 5 mM KCl

Chapter I

Aspects of Gene Amplification

1.1

General

1.1.1

Introduction

Gene amplification - more accurately defined as DNA amplification - is the increase in copy number of part of the genome over the genome as a whole; it does not include phenomena such as polyploidy and chromosome duplication. Products encoded by genes present in amplified DNA are often expressed at elevated levels with consequent advantages and disadvantages (for reviews on DNA amplification see Stark and Wahl, 1984; Hamlin *et al.*, 1984; Schimke, 1985; Delidakis *et al.*, 1989; and Stark *et al.*, 1989).

DNA amplification has been observed in prokaryotes and also in eukaryotes, either as part of a developmentally regulated programme (Drosophila and Xenopus, for example) where gene products are required at rates beyond those which can be attained by the transcriptional and translational rates of a single gene, or as an aberrant event where abnormal cellular mechanisms give rise to amplified DNA. The latter phenomenon has been observed in higher eukaryotic tumours and cultured cell lines, providing such cells with a growth advantage or providing a survival advantage in the presence of toxic compounds. Amplification has not been seen in normal cells despite several spirited attempts (Wright *et al.*, 1990 and Tlsty, 1990), and it is possible that the aberrant cellular mechanisms responsible for DNA amplification are linked to those responsible for other forms of genomic instability such as chromosomal inversion, translocation and perhaps chromosome loss (Stark *et al.*, 1989).

This introduction will deal first with regulated (programmed) DNA amplification and then with the cytology of amplification as observed in tumours and drug resistant cells. These systems are then dealt with in more detail; in particular the CAD amplification system which is the focus of this thesis. Models of amplification are described as they are currently understood, but models in vogue as the main project

was embarked upon are included with the evidence that existed for and against them at that time. The main body of work for this thesis comprises the attempted construction of an artificial amplification system to mimic a spontaneous event. The model was based on Simian Virus 40 (SV40) replication and therefore a section of relevant background information about SV40 is included. The aims and theory of the proposed model system conclude the chapter.

1.2

Developmental Gene Amplification

1.2.1

Introduction

During the development of certain organisms large amounts of specific gene products are required over a short period of time. The amounts needed may be in excess of those attainable using transcriptional and translational regulatory mechanisms. An increase in the copy number of relevant genes to produce additional transcription templates has been employed in several organisms. Two well documented situations are the amplification of rDNA genes of Xenopus and the chorion genes of Drosophila (reviewed by Kafatos *et al.*, 1985 and Delidakis *et al.*, 1989).

1.2.2

Amplification in Xenopus

A mature Xenopus oocyte contains about 10^{12} ribosomes, and using known rates of transcriptional and translational efficiencies it is estimated that it would take 2 months for a primary oocyte to produce the necessary ribosomal proteins from a haploid gene. It would take 120 years for the repetitive rDNA genes to produce an equivalent number of rRNA molecules (Kafatos *et al.*, 1985). This demonstrates that the primary oocyte cannot produce sufficient amounts of the necessary gene products during oogenesis using transcriptional and translational mechanisms. In

both male and female frogs the rDNA genes are preferentially amplified about 10- to 40-fold in the gonial cells. At the beginning of meiosis the spermatocyte loses its extra copies but those in the oocyte undergo a second round of amplification to give a final 2500-fold increase in the copy number. At this point, approximately 70% of the nuclear DNA of the oocyte encodes rRNA. The mechanism for this amplification is thought to be via a rolling circle mechanism (see later) which replicates an extrachromosomal DNA fragment produced by an unknown mechanism.

1.2.3

Amplification in Drosophila

The eggshell (chorion) produced around each Drosophila melanogaster oocyte comprises about 20 proteins; six are major proteins, each present at about 1-4 x 10^{12} molecules per chorion and encoded by a single copy gene. During choriogenesis, major chorion protein mRNAs accumulate in about one hour and most of the protein is produced in 2 hours. To synthesize such large amounts of protein quickly the single copy chorion genes are preferentially amplified by 20- to 80-fold , depending on the chromosomal location of the particular gene. This system differs from the amplification of rDNA genes in Xenopus in one important aspect: in Xenopus the genes are amplified transiently - extra copies are produced for a short developmental period, and when they are no longer required they are degraded and lost from the cells as they continue their development; chorion cells, on the other hand, are terminally differentiated and the amplified genes are maintained in the cells throughout their existence. In chorion cells, the chorion genes are amplified by multiple rounds of local replication initiated at replication origins within or flanking each cluster of chorion genes (Osheim, Miller and Beyer, 1988). The amplified arrays exist as 80-100 kb gradients of amplified genes and are retained by the chorion cells throughout their existence, producing the proteins required in choriogenesis.

1.2.4

Conclusion

The cases described above are only two of several examples of regulated gene amplification, and all are important in developmental studies and as models of eukaryotic DNA replication. They are deliberate, tightly regulated events, and although they are included as examples of gene amplification, they do not represent another class of amplification where the event is the result of cellular mechanisms behaving in an abnormal manner.

1.3

Cytology of Amplified DNA

1.3.1

Visualisation of Amplified DNA

To date, amplified DNA has been observed in three forms. These are referred to as: extended chromosomal regions (ECRs; previously called homogeneously staining regions or HSRs); double minutes (DMs); and other episomal forms which are sub-microscopic and therefore differ from DMs at least in size. Where known, the DNA in these structures has been observed in head-to-head, tail-to-tail or head-to-tail tandem arrays, and the arrangement probably depends on the mechanism responsible for its generation and the topology of sequences prior to the amplification event (Wahl, 1989).

1.3.2

Extended Chromosomal Regions

ECRs (or HSRs) have been observed in G-banded chromosome spreads from tumour cells of hamster, mouse, rat and human origin, and also in drug resistant tissue cultured cells. They are seen as chromosomal tracts that do not exhibit the usual pattern of irregular stained bands but are stained either uniformly dark or light,

or show very faint regular bands (Cowell, 1982). These regions vary in length between tumours and cell lines, from hardly detectable to approaching the length of chromosome arms (250 kb to 'tens of megabases'; Smith *et al.*,1990). In a stable cell line the ECR is usually located at the same position for all cells in a population. ECRs can be situated on the chromosome where the original gene resided, or on other identifiable chromosomes, usually at chromosome ends. They can also be present on unidentifiable chromosomes formed by recombination and rearrangement of the ECR and other chromosomal elements.

Molecular analyses of DNA isolated from cell lines containing ECRs have shown the extent of DNA rearrangement and the co-amplification of sequences flanking the amplified gene. In one particular example, BHK cells selected with PALA showed amplification of the CAD gene and large amounts of flanking DNA (Ardeshir *et al.*, 1983 and Zieg *et al.*, 1983). It was also demonstrated that cell lines isolated in the same PALA selection experiment amplified the CAD gene and different flanking sequences to varying degrees, and even within cell lines the level of amplification of these sequences was variable. Although ECRs were observed to be stable between chromosomes and cell lines, the DNA within these regions could be relatively unstable.

1.3.3

Double Minutes

DMs are seen under a light microscope as paired, acentric chromatin bodies that stain poorly in metaphase spreads. They have been observed in tumour cells and in cell lines derived from mice, rats, hamsters and humans coincident with MTX, vincristine and metallothionein resistance, respectively (referenced in Hamlin *et al.*, 1984). DMs are unequally segregated at mitosis, and in resistant cells they are rapidly lost in the absence of the selective regime. Their number and size can vary widely between independent cell lines and also between individual cells of those lines; this is probably related to recombination and random segregation at mitosis. They replicate once per cell cycle and their accumulation in a cell is not thought to arise from

disproportionate replication (Cowell, 1982). DMs stain poorly with most cytological stains, including those that identify AT- and GC-rich regions, but in many cases the nature of the DNA contained in DMs is uncertain. However, DMs present in MTX-resistant cells have been shown to contain amplified DHFR sequences (Kaufman *et al.*, 1979).

1.3.4

Episomal Forms

Submicroscopic episomal elements were first described in amplified cells by Carroll *et al.* (1987) in a Chinese hamster cell system where an introduced CAD gene was amplified. In one particular clone the integrated gene gave rise to a circular DNA element. This episome was 250-300 kb in size, contained approximately 5 copies of the transfected sequence, and was semiconservatively replicated. An episome of ~650 kb containing amplified DHFR genes was described by Maurer *et al.* (1987) in human MTX-resistant cells containing a 250- to 300-fold DHFR amplification, and Von Hoff *et al.* (1988) have described a submicroscopic episome containing amplified MYC oncogenes in 2 tumour-derived cell lines. Episomes have also been implicated in the amplification of the multidrug resistance gene *mdr* 1 (Ruiz *et al.*, 1989). It has been suggested that submicroscopic episomes are the precursors of DMs on the basis of their replication and structural similarities. Their size (120-750 kb) has tempted some researchers to link episome size with the size of replication loop domains (replicons) and suggest a relationship between replication and episome formation (Wahl, 1989). Carroll *et al.* (1988) demonstrated that episomes in a PALA-resistant CHO cell system were the product of a chromosomal deletion event. They showed these episomes to be DM precursors, and the DMs were also shown to reintegrate into the genome at apparently random locations. However, the cell system used for the investigation was artificially supplied with the selectable gene by transfection, and the deletion/production of episomes may be artifactual in this particular case. It is possible that episomes containing endogenous amplified genes are not produced in the same manner and do not have the same properties as those

described by Carroll *et al.* in their artificial system, but these episomal forms containing endogenous genes may still be intermediates in certain mechanisms of DNA amplification.

1.3.5

Inter-relationship Between Amplified DNA Forms

The relationship between DMs, episomes and ECRs is unclear. ECRs could be the integrated products of DMs which may in turn be derived from episomes. Carroll *et al.* (1988) showed that DMs can integrate into chromosomes, apparently at random locations. Trask and Hamlin (1989) and Smith *et al.* (1990) demonstrated, using fluorescent *in situ* hybridisation techniques, that ECRs containing amplified genes are usually located at the locus of the unaltered endogenous gene. This is not consistent with a deletion/reintegration model as, following a deletion, no homology would remain in the chromosome for the extrachromosomal element to reintegrate at its original locus by homologous recombination. An episomal element could integrate into its homologous sister chromatid during metaphase by homologous recombination, leaving one daughter cell with a deletion and the other with a duplicated chromosomal arm (Smith *et al.*, 1990). In an alternative argument, excision events involving ECRs could give rise to episomes which in turn could give rise to DMs and so on. There is no conclusive evidence for either cycle of events, but recent developments (reviewed by Stark *et al.*, 1989) favour ECRs as the primary arrangement of amplified DNA in most cultured cell lines.

The environment and type of cell could be important factors in determining the nature of amplified DNA (Wahl *et al.*, 1984; Meinkoth *et al.*, 1987). In MTX-resistant mouse cells the amplified DHFR cells are nearly always maintained on DMs; in PALA-resistant BHK cells the amplified CAD genes have nearly always been located in ECRs, even at very early stages of amplification (Smith *et al.*, 1990). It is possible that ECRs and DMs represent the outcomes of different mechanisms of DNA amplification

1.4

Mammalian DNA Amplification

1.4.1

Amplification in Normal Cells

Very few reports exist of amplified genes in primary somatic or germline cells. Prody *et al.* (1989) reported a case where a human gene was amplified 100-fold in a farmer who had been exposed to agricultural organo-phosphorous pesticides over a long period of time. Individuals such as the farmer who have a 'silent' (defective) version of the gene are vulnerable to such pesticides. A son and a grandson of the farmer also carried the amplification but grandparents did not. It was suggested that the gene was amplified in the early stages of embryogenesis, spermatogenesis or oogenesis by a mechanism consistent with the 'onionskin' model. Wright *et al.* (1990) exposed several primary cell preparations (mammary epithelial cells, human keratinocytes and human diploid fibroblasts) to different cytotoxic drugs previously shown to select cells which had amplified genes encoding the drugs' targets, and failed to establish any resistant mutants from 5×10^8 cells. Tlsty (1990) performed a similar study on rodent and human primary cells with the conclusion that these cells lack a detectable frequency of gene amplification. Therefore, the ability to amplify DNA and/or maintain it is presumed to be extremely rare in normal cells. It is not known whether the absence of detectable amplified DNA reflects a cellular failure to amplify DNA sequences or whether it reflects the non-survival or even active destruction of a cell harbouring amplified DNA.

It is evident that during evolution genes have been duplicated and amplified to create large gene families of similar but different genes that have a common progenitor (Matera, Weiner and Schmid, 1990; Ohta, 1988). Examples of this include the families of globin, tRNA, histone, immunoglobulin, actin, myosin, collagen, histocompatibility antigen, ovalbumin and interferon genes (Hamlin *et al.*, 1984). It is therefore possible that, during the evolution of higher eukaryotes a form of amplification has occurred and has been tolerated, but the structure of these

multigene families suggests that mechanisms different to those in drug-resistant cells may have been responsible. One major difference is the juxtaposition of duplicated genes without intervening non-gene sequences (Hamlin *et al.*, 1984); in the light of the following sections it will be apparent that the mechanisms proposed for amplification are at a loss to produce such an arrangement.

1.4.2

Amplification in Tumours

Primary specimens from a number of tumours have been shown to contain amplified DNA. Figure 1.1 lists a number of oncogenes that have been found to be amplified in tumours. The myc gene family is often involved (for an example see Zehnbauer *et al.*, 1988) and increasing myc copy number correlates with increasing metastatic potential and poor prognosis. Myc amplification is also correlated with increased cloning efficiency, shortened doubling times, decreased expression of differentiation markers and decreased adhesiveness in cell cultures (Klein and Klein, 1985). Amplified arrays in the form of DMs or ECRs (see below) have been observed in tumours from many animals including man, but the nature of the amplified sequences has not been established. There are also strong correlations between species and tumour type, and the manifestation of the amplified DNA array. Amplification of oncogenes or other genes may help the progression of a tumour once transformation has occurred, or alternatively, it may be a contributory tumourigenic factor. An increase in an oncogenic gene product through gene amplification may provide a tumour cell with a growth advantage, particularly in the development of growth factor and substrate independence.

Figure 1.1

Amplified Cellular Genes in Tumours and Tumour-Derived Cell Lines

These genes have been found to be amplified to some degree in the tumours shown. The list represents a small fraction of the many tumours and cell lines that have been found to contain amplified DNA. Amplification has not been shown to be the cause of any tumour in which it is present but it may reflect the general chromosomal instability and abnormallity in such cells (Hamlin *et al.*, 1984).

Oncogene	Tumour(s), Tumour Cell Lines.
c-*myc*	Lung, breast, Burkitt's lymphoma, myeloid leukemia cell line HL-60
N-*myc*	Neuroblastoma, retinoblastoma
L-*myc*	Lung
Ki-*ras*	Lung, ovary
c-*abl*	leukemia cell line K-562
Epidermal growth factor receptor	Glioma, squamous cell carcinoma

1.5

Amplification in Cultured Cells

1.5.1

Introduction

Amplification has been observed in many cultured cell lines derived both from tumours and transformed, immortalized cells (for extensive examples see reviews listed in 1.1). DNA is thought to be amplified in a random fashion through out the genome as cells divide, but the amplified DNA is probably unstable and not maintained in the absence of selection pressures for its maintenance. Cells with amplified sequences may be lost from the population either because of retarded growth or from the lethal or deleterious effects of genes over-expressed due to amplification.

Cytotoxic inhibitors can be used to isolate resistant cells from a sensitive cell population and the mechanism of resistance depends on the specific inhibitor and in some cases, the amount of inhibitor used. A common form of resistance is over-production of the target protein so that sufficient molecules remain uninhibited to continue their cellular function. This can be achieved by an increase in copy number of the gene encoding the target protein. Figure 1.2 lists some of the inhibitors to which cells can become resistant through amplification of a particular gene. In most of these cases amplification is the most frequent mode of resistance, so that other mechanisms such as regulatory mutations, altered inhibitor binding or transport and metabolic mutations are rarely observed. These mechanisms can account for some MTX-resistant mutants where resistance can also be acquired by altered inhibitor transport or by reduced affinity of DHFR for MTX (Schimke, 1984). Occasionally genes encoding non-target proteins can be amplified; for example, cells selected with one specific drug can become cross-resistant to a variety of other drugs which are often functionally unrelated (see figure 1.2). Such cross-resistant cells sometimes have amplified MDR genes (Biedler *et al.*, 1988; Roninson *et al.*, 1984) encoding a 170 kDa P-glycoprotein which is thought to actively transport certain

Figure 1.2

Gene Amplification in Drug-Resistant cells

Cell survival in the listed drugs or selective agents usually requires amplification of the appropriate gene. In the case of 'multiple drugs' the protein listed confers Multiple Drug Resistance (MDR) syndrome where cell resistant to one drug are often resistant to others. 'Multiple drugs' include: Vincristine, maytansine, adriamycin, Baker's antifolate, colchesine, actinomycin D, cytochalasin B, emetine, puromycin and melphalan.

Selective Agent.	Gene/Binding Protein.
•Methotrexate (MTX)	Dihydrofolate reductase (DHFR)
•PALA	CAD
•Ouabain	Na^+K^+ ATPase
•Cadmium	Metallothionein (I)
•6-Azauridine or pyrazofurin	UMP synthetase
•Adenosine, alanosine and deoxycoformycin	Adenosine deaminase
•Methionine sulphoximine	Glutamine synthetase
•5-Fluorodeoxyuridine	Thymidylate synthetase
•α-Methyl- or α-difluoro-methyl-ornithine	Ornithine decarboxylase
•Hydroxyurea or deoxynucleotides	Ribonucleotide reductase
•Aphidicolin	Ribonucleotide reductase
•Adenine and coformycin	AMP deaminase
•Albizziin or β-aspartyl hydroxamate	Asparagine synthetase
•Mycophenolic acid	IMP-5'-dehydrogenase
•Tunicamycin	N-Acetylglycoaminyl transferase
•Borrelidin	Threonyl tRNA synthetase
•Multiple drugs (see legend)	170 kDa P-glycoprotein

classes of compounds from the cell and reduce the cellular concentration and effect of the inhibitor.

The MTX and PALA systems are the best characterised, and most work to date in understanding gene amplification in Dr. George Stark's laboratory has been done using the PALA system which is described in detail below.

1.5.2

CAD and PALA

CAD is the acronym for an enzyme with three functional domains which is involved in the UMP biosynthetic pathway (figure 1.3a): Carbamoyl phosphate synthetase, Aspartate transcarbamoylase and Dihydroorotase (Stryer, 1981). Carbamoyl phosphate synthetase catalyses the formation of carbamoyl phosphate from glutamine and ATP. Aspartate transcarbamoylase converts carbamoyl phosphate and aspartate into *N*-carbamoyl aspartate which is cyclised into a pyrimidine ring by the dihydroorotase activity to form dihydroorotate. The pathway continues through the activity of other enzymes to produce uridylate (UMP) which is a major pyrimidine nucleotide. In mammals, CAD is a single polypeptide of about 200 kD and is encoded by a single gene. In BHK cells the gene covers 25 kb with about 37 intervening sequences (Padgett *et al.*, 1982) and produces a 7.9 kb mRNA (Padgett *et al.*, 1979), of which about 6.5 kb has been cloned as a cDNA which lacks about 1.5 kb of 5' sequence (Shigesada *et al.*, 1985). This sequence was used as a probe for the CAD gene in the experiments described later.

N-(phosphonacetyl)-L-aspartate (PALA; figure 1.3b) was designed as a transition state analogue to E. coli aspartate transcarbamoylase and binds reversibly to the aspartate transcarbamoylase domain of CAD (Collins and Stark, 1971). It can enter cells, inhibit UMP production and, in the absence of uridine, cause cell death by pyrimidine nucleotide starvation. Cells can also use uridine as a precursor for UMP production through an alternative biochemical pathway and therefore, cultured cells are selected with PALA in medium containing serum from which uridine has been removed by dialysis. There is a danger in these selections that the cells are further

Figure 1.3

a) Biosynthetic Pathway of UMP Synthesis

CAD catalyses the first three steps of the pathway to convert glutamine to dihydroorotate which includes the formation of a pyrimidine ring. Dihydroorotate is further converted to UMP by subsequent enzymes and is a major pyrimidine nucleotide in the cell.

b) *N*-(phosphonacetyl)-L-aspartate (PALA)

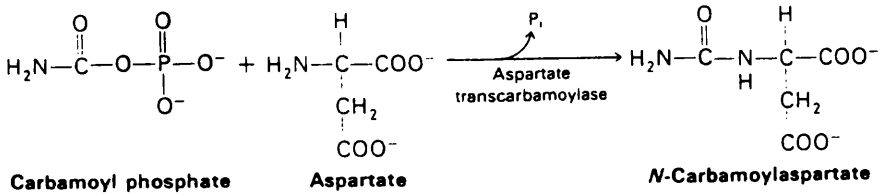
This transition state analogue binds tightly but reversibly to the aspartate transcarbamoylase activity of CAD. In cells deprived of uridine its presence inhibits proliferation and causes death by pyrimidine starvation.

a.

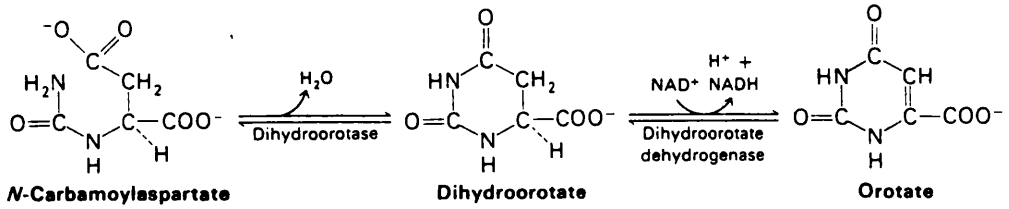
i) The carbamoyl phosphate synthetase domain of CAD catalyses the conversion of glutamine to carbamoyl phosphate and glutamate:

Glutamine + 2 ATP + HCO_3^- ⟶ carbamoyl phosphate + 2 ADP + glutamate

ii) Aspartate transcarbamylase activity of CAD catalyses the formation of N-carbamoylaspartate from carbamoyl phosphate and aspartate:

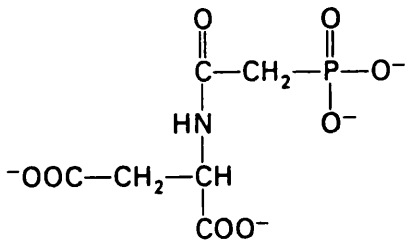


iii) The pyrimidine ring is formed by the cyclizing of N-carbamoylaspartate, catalysed by dihydroorotase domain of CAD:



b.

N-(phosphonacetyl)-L-aspartate (PALA)



compromised by the unintentional removal of small molecules such as growth factors along with uridine in the dialysis. There is therefore a risk that the selection is not only for PALA resistance but also for a decrease or loss of dependence on these lost molecules. This risk can be controlled for by growing cells in medium supplemented with dialysed serum in the absence of any drug, and all cell lines tested to date appear to grow as they do in normal serum.

Swyryd and co-workers (1974) showed that PALA blocked proliferation of cultured mammalian cells but that some cells were unaffected and showed large increases in aspartate transcarbamoylase activity. Wahl *et al.* (1979) went on to show that this increased activity was due to amplification of the CAD gene. Up to the present time, PALA resistance has never been shown to be achieved by any other mechanism. It is important to stress that PALA has not been shown to cause CAD gene amplification and its presence in culture media allows survival of cells which have already amplified the DNA that includes the CAD gene.

1.5.3

Amplification of CAD in BHK Cells

Amplified DNA containing the CAD gene has always been observed as ECRs in BHK cells resistant to PALA, and in most cases it has the same chromosomal location (the short arm of chromosome B9) as the original locus (Wahl *et al.*, 1982; Smith *et al.*, 1990). The amplified CAD genes of mutant cell lines selected in increasing concentrations of PALA are very stable and the extra copies of CAD are not lost when the cells are grown out of selection over long periods of time. In contrast, cells selected in a single step at low PALA concentrations lose their extra CAD genes within 30 days of growth out of selection. This time increases to approximately 90 days for cell lines selected in two selection steps (Saito *et al.*, 1989). Smith and colleagues (1990) demonstrated that single-step mutants analysed very early after selection had extended short arms of chromosome B9 and more 'mature' single-step mutants had condensed chromosomal extensions of the same chromosomal region. This was in agreement with the observations of Saito *et al.*

(1989) where the amount of DNA co-amplified with CAD in the first step of selection was reduced in later rounds of amplification. Estimates of the amount of DNA amplified in the initial amplification event range from 500 kb per extra copy of CAD (Wahl *et al.*, 1982) to 10, 000 kb (Giulotto, Saito and Stark, 1986) or 'tens of megabases' (Smith *et al.*, 1990) and it is likely that the earlier one looks, the larger and more accurate the estimate will be.

1.5.4

Positional Effects

The frequency of amplification can be greatly affected by chromosomal position. Wahl , Robert de Saint Vincent and Derose (1984) showed that transfected Syrian hamster CAD genes were amplified at different frequencies in different independent CHO transformants. This difference was ascribed to the chromosomal location of the integrated DNA and the effect of sequences flanking the integration site. Transformants containing integrations near the centromere retained the donated genes at the original position; amplified genes were located in other chromosomes or in extra-chromosomal elements. Transformants which had insertions in the middle of a chromosomal arm showed no obvious cytogenetic rearrangements after amplification.

Amplification control elements (ACEs) that act in *cis* during Drosophila chorion gene amplification have been described (Spradling *et al.*, 1987). They have been found in each gene cluster and mutations in these elements render the replicon unable to amplify. There are several other amplification-enhancing regions (AERs) in the area, and their presence is also required for normal amplification of the chorion genes (Delidakis and Kafatos, 1989). In higher eukaryotes where gene amplification has not been found to play a part in development, *cis*-acting elements have also been detected. McArthur and Stanners (1991) and Beitel *et al.* (1991) have described elements which can stimulate amplification of a DHFR-containing plasmid in a variety of cell lines.

1.6

Models of DNA Amplification

1.6.1

Introduction

Amplified DNA is found in different forms, influenced by factors such as species, locus, inhibitor dose and maturity of the amplified DNA. The diversity and inter-relationships between the amplified structures may reflect different mechanisms at work. A common feature of many models are inverted duplications. These structures may have special properties and effects that allow DNA surrounding them to be further amplified.

1.6.2

Model Requirements

Any model that attempts to explain how DNA becomes amplified must address the following observations:

a) the amount of DNA amplified can be very large, up to 10, 000 kb per gene (Giulotto *et al.*, 1986; Smith *et al.*, 1990; Trask and Hamlin, 1989; Windle *et al.*, 1991);

b) amplification can manifest itself as linear intrachromosomal or circular extrachromosomal arrays;

c) novel joints (Ardeshir *et al.*, 1983; Ford and Fried, 1986; Saito *et al.*, 1989) are found in the amplified DNA but recombination hot-spots, on the whole, do not seem to exist (for an example of such a recombinational hot spot, see Hyrien *et al.* [1987]);

d) high levels of drug resistance are achieved by successive selection in increasing amounts of drug, but for some inhibitors, eg. PALA, resistance to a low level of drug in a single step, mediated by amplification, produces mutants that are resistant to much higher drug concentrations without any apparent further amplification (Zieg *et al.*, 1983);

e) intrachromosomal arrays appear to reside on the same chromosome as the original endogenous gene which is present and apparently unaltered at its wild type chromosomal location (Trask and Hamlin, 1989; Smith *et al.*, 1990);

f) the spontaneous rate of amplification can be 10^{-6} to 10^{-4} events/cell/division;

g) agents that block DNA replication (eg. hydroxyurea , MTX) or damage DNA (eg. UV light, MTX, 5-aza-deoxycytidine) can stimulate amplification of some genes (Sharma and Schimke, 1989; Sherwood, Schumacher and Schimke, 1988).

The models described below each satisfy most of the points above but it is unlikely that a single mechanism is responsible for all instances of DNA amplification. As stated before, several known factors can influence the outcome of an amplification event and so while a particular model may have merit in one situation, it may have none in another. Four major models are schematically shown in figure 1.4.

1.6.3

The Disproportionate Replication or 'Onionskin' Model

Figure 1.4 shows a schematic diagram of this model. A bidirectional origin initiates replication more than the statutory once per cell cycle. This unscheduled replication produces a structure which must be resolved if the chromosome is to segregate normally at mitosis. The figure shows three ways in which this structure could be resolved. Recombination between different duplexes would generate an intrachromosomal array, while recombination within a duplex would generate circular extrachromosomal elements. Two replication forks in very close proximity could produce linear extrachromosomal elements whose fate would be determined by subsequent recombination events. The model is derived from the observations of Bullock and Botchan (1982) and Roberts *et al.* (1983) in an attempt to explain the excision mechanism of integrated SV40 sequences. How the cellular replication origin can escape normal control is not known. Mariani and Schimke (1984) showed that DHFR genes were amplified in a single cell cycle following inhibition of DNA synthesis with hydroxyurea after the DHFR region had been replicated. When the block was removed all the DNA , including that previously replicated, was re-replicated

Figure 1.4

The Four Main Models of Gene Amplification

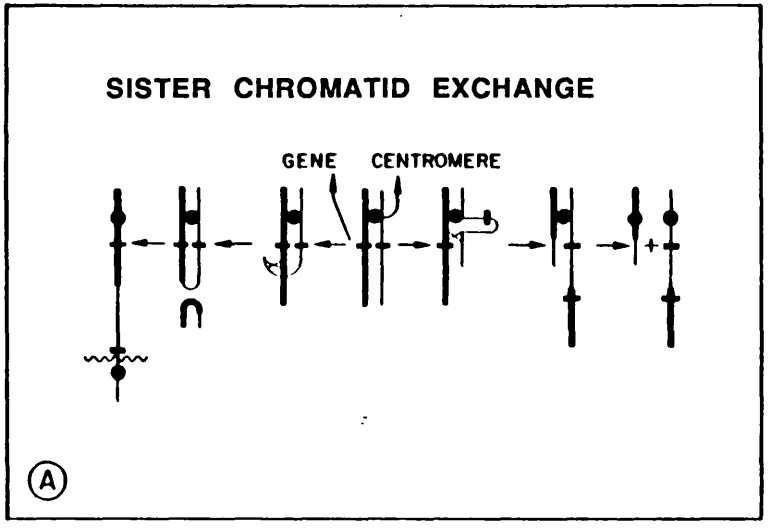
A) *Unequal sister chromatid exchange*: In the initial event a large DNA fragment is transfered from one chromatid to the other. The result is a duplication (one extra copy) on one chromatid and a deletion on the other.

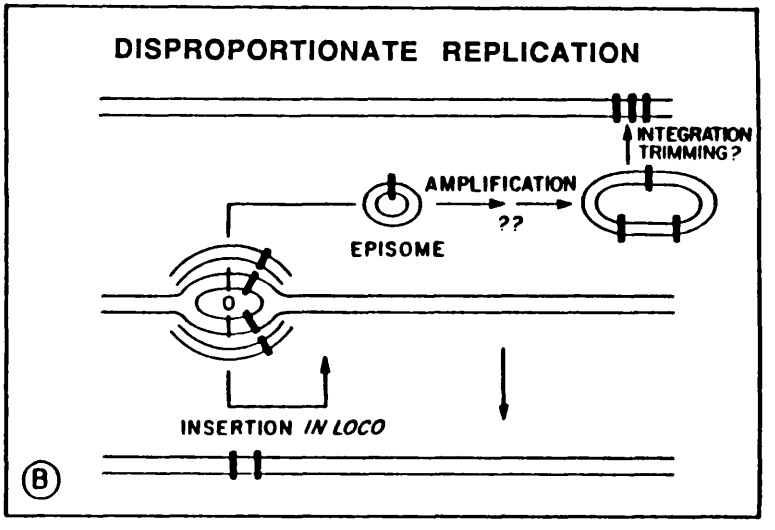
B) *Disproportionate Replication*: This is the 'onionskin' model; a replication origin misfires and initiates replication more than once in a single cell cycle. Recombination is required to resolve the mitotically unstable structure and this recombination determines the DNA arrangement.

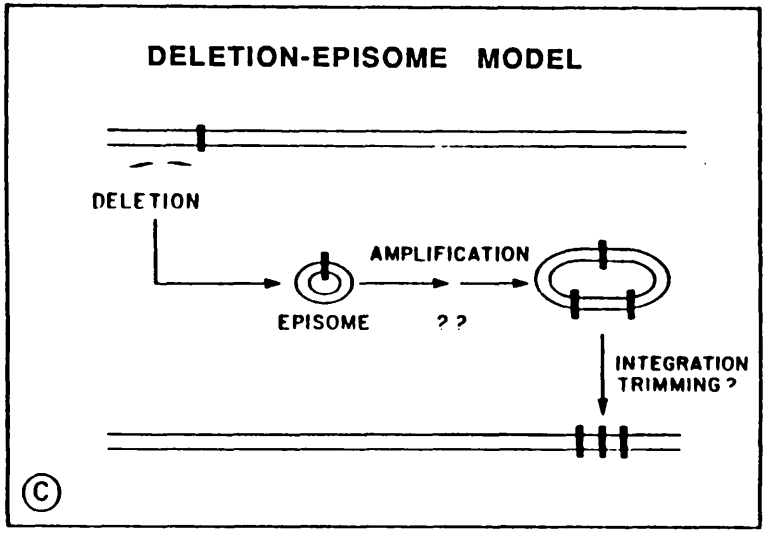
C) *The Deletion/Episome Model*: A chromosomal deletion is produced by recombination and this molecule is randomly segregated and amplified. Further recombination can increase the episome size, possibly to the size of DMs, or the episome may reintegrate into the genome.

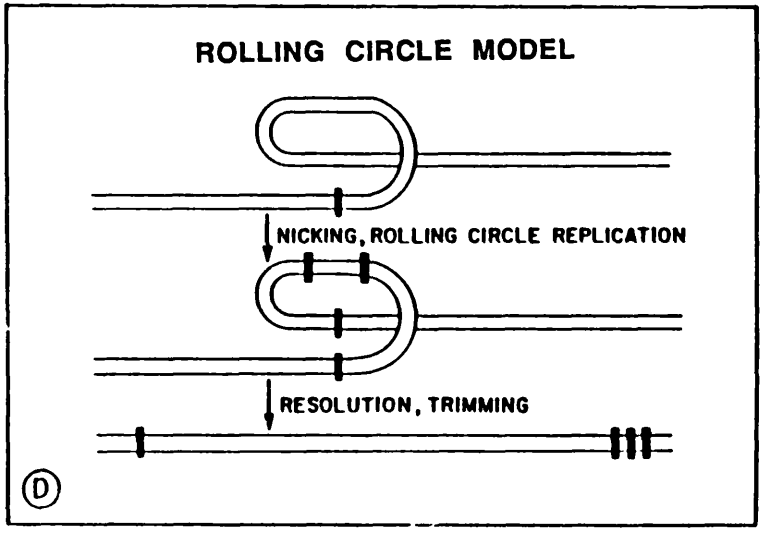
D) *Rolling Circle Replication*: The direction of one replication fork initiated at a bidirectional replication origin is reversed by recombination across an inverted repeat. The two forks continue to replicate DNA but in the same direction, producing a concatamer of the amplified unit.

Schematics pilfered from Trask and Hamlin (1989).


SISTER CHROMATID EXCHANGE
GENE
CENTROMERE
A


DISPROPORTIONATE REPLICATION
INTEGRATION TRIMMING?
AMPLIFICATION
??
EPISOME
INSERTION IN LOCO
B


DELETION-EPISOME MODEL
DELETION
AMPLIFICATION
EPISOME
??
INTEGRATION TRIMMING?
C


ROLLING CIRCLE MODEL
NICKING, ROLLING CIRCLE REPLICATION
RESOLUTION, TRIMMING
D

and a number of recombination events were proposed to lead to loss of non-selected DNA and preferential amplification of DHFR. This work was elaborated upon by Hoy *et al.* (1987) who also used aphidicolin or UV irradiation to block DNA synthesis and found that a subset of cells that were in S-phase at the time of the block contained over-replicated DNA. However, Painter and co-workers (1987) suggested that the results were artifactual and developed an alternative protocol in which inhibition of DNA synthesis did not lead to over-replication of parts of the genome. Although the question is still unresolved as far as cultured cells are concerned, this mechanism has been shown to be valid for amplification of the chorion genes of Drosophila and also for amplification of integrated SV40 and polyoma sequences. The model addresses the following features of amplified DNA: it can produce the large number of extra gene copies in the times observed; it can produce the large amounts of DNA thought to be amplified in the initial event; it can give rise to both linear and circular forms of amplified DNA through different recombination events; and finally, it can also explain the heterogeneity of co-amplified sequences between different isolates of resistant cells and also between individual cells of these isolates.

1.6.4

The Sister Chromatid Exchange Model

This model is primarily based on amplification of the CAD gene in Syrian hamster cells. Work by Giulotto *et al.* (1986) and Saito *et al.* (1989) showed that large amounts of DNA (up to 10, 000 kb) were co-amplified with the CAD gene, and that novel joints detected in single-step mutants were not amplified to the same degree as the selected gene in subsequent selection steps. The rolling circle and episome deletion models (described below) predict that the novel joints formed when inverted duplications are created should be amplified to the same degree as the selected gene. Unequal sister chromatid exchange would involve the transfer of DNA from one sister chromatid to the other without a reciprocal exchange in the opposite direction as shown in the right hand route of figure 1.4a. Such an exchange could

involve the large amounts of DNA described by Giulotto *et al.* (1986), Trask and Hamlin (1989), Smith *et al.* (1990) and Windle *et al.* (1991), but in the first step can only lead to one extra copy of the gene selected for. Further unequal exchanges in different cell divisions could generate additional copies of the selected gene, or a different mechanism could create multiple copies from an early exchange product. The outcome of such exchanges could be: (a) a normal homologue; and (b) an elongated homologue containing both the amplified genes and also the endogenous gene at its original chromosomal position. The first step of amplification would include both the unequal exchange and also a mechanism responsible for the rapid accumulation of extra copies of the selected gene.

1.6.5

The Deletion-Episome Model

This model postulates that a circular DNA molecule containing an origin of replication is formed by a deletion event. The deletion could result from illegitimate recombination within a replicon (Carroll *et al.*, 1988), perhaps containing an inverted repeat (Passananti *et al.*, 1987; Ruiz and Wahl, 1988), or from an onionskin structure (Schimke *et al.*, 1986). The circular DNA molecule (episome) would replicate semiconservatively but its acentromeric nature would allow accumulation through unequal segregation at mitosis. Recombination between episomes could increase their size to that of DMs and the relationship between episomes, DMs and ECRs has been described above. What this model cannot easily resolve is the retention of an unaltered gene and the existence of an ECR on the same arm as the original gene whilst the unaltered chromosome homologue is still present (Trask and Hamlin,1989; Smith *et al.*, 1989). Production of the episome by deletion would restrict homologous sequences to the other homologue. Reintegration by homologous recombination could occur but an unaltered homologue would not be present

1.6.6

The Double Rolling Circle Model

This model was devised to explain the replication of the yeast 2μ plasmid (Futcher, 1986). This plasmid is indigenous in several Saccharomyces cerevisiae strains and confers no known phenotype on the host cell. A plasmid-encoded gene mediates site-specific recombination across two inverted repeats which are asymmetrically located with respect to the single origin of replication. If one of the replication forks has passed the recombination site when recombination occurs then the fork is inverted and is now travelling in the same direction as the other fork. The two forks follow each other indefinitely to produce a concatamer of the plasmid. Passananti *et al.* (1987) adapted the model for DNA amplification in mammalian cells following the discovery of an amplified inverted repeat in a polyoma-transformed rat cell line. Inverted repeats can be formed from a DNA duplication and are therefore important candidates for the initial structure formed in an amplification event. In the model a circular DNA molecule which contains the gene to be amplified within an inverted duplication is excised from a chromosome, perhaps by illegitimate recombination between newly synthesized DNA strands. Replication is initiated within the circle by a replication origin and, after one round of replication, homologous recombination occurs between the newly replicated strand and the unreplicated strand. This causes one of the replication forks to invert so that both forks are moving in the same direction. Further replication produces multiple copies of the inverted repeat derived from a single initiation of replication in a single S-phase. The amplified array could either remain extrachromosomal (as a DM) or could reintegrate into the genome.

1.6.7

The Spiral Model

Hyrien *et al.* (1988) developed this model to describe the formation of an inverted duplication and its subsequent role in amplification in coformycin-resistant Chinese hamster cells. A hairpin loop can form on one strand of DNA in palindromic-

rich areas if the two strands are mismatched through partial homologies. This can lead to the formation of an inverted duplication if the replication machinery, when confronted with a hairpin loop, switches template strands and replicates the other DNA strand. If the hairpin loop is excised, the resulting structure has the appearance of a replication fork and can prime replication of the already duplicated DNA. If this occurs at both replication forks of the same replicon, two replication forks will follow one another and generate multiple copies of the inverted duplication. If the replication complex switches strands and replicates the second template from a slightly different location to the strand it left, then the duplicated elements of the inverted duplication will be separated by an asymmetric unduplicated DNA sequence. This is a common feature of inverted duplications (Ford and Fried, 1986; Looney and Hamlin, 1987; Nalbantoglu and Meuth, 1986; Saito and Stark, 1986).

1.6.8

Overview

Other models such as the uptake and utilisation of DNA fragments from dead cells and reverse transcription of RNA have been cited in the past as mechanisms of DNA amplification, but those described above are probably the most likely models for the different systems they are based on. In recent years new techniques have been developed and with time, patience and a certain amount of luck DNA can be analysed almost as soon as it is amplified (K. A. Smith and G. R. Stark, personal communication). In BHK cells resistant to low levels of PALA it is likely that homologous but unequal sister chromatid exchange is involved in the amplification of CAD genes (Smith *et al.*, 1990). Initially a large part of the chromosome arm carrying the CAD gene is transferred from one sister chromatid to the other by non-homologous unequal sister chromatid exchange. After division both daughter cells would contain a normal homologue, but one would contain an elongated chromosome carrying the transferred DNA and the other would have a chromosome carrying a deletion. The copy number is rapidly increased by further unequal but homologous exchanges between duplicated chromosomal regions. A similar pattern

of events also appears to be true for DHFR amplification in CHO cells (Trask and Hamlin, 1989). However, in other systems the same mechanisms may not be involved as a result of other factors relating to cell type, selection regime and experimental design. Some models are based on experimental systems in which the amplification of artificially introduced genes is studied. Results from these systems are difficult to interpret in the context of endogenous gene amplification as many other variables such as integration effects and *cis* effects of vector sequences are introduced with the gene.

1.7

Simian Virus 40 (SV40)

1.7.1

Introduction

SV40 is an icosahedral, double stranded DNA papovavirus with a circular genome of 5243 bp (Tooze, 1973; Fiers *et al.*, 1978). Other papovaviruses include polyoma virus and the papilloma viruses. The SV40 genome encodes two early genes, large T antigen (T-ag) and small t-antigen, and three late genes, viral proteins VP1, VP2 and VP3 (Fiers *et al.*, 1978) defined relative to viral DNA synthesis. Early genes are transcribed counter-clockwise and the late genes clockwise. Infection of permissive cells results in a lytic life cycle whereas infection of non-permissive cells can lead to cellular transformation (Martin, 1981).

1.7.2

SV40 Host Range and Life Cycle

Infection of permissive cells results in production of virus from a lytic cycle. The virus is transported to the nucleus, unmasked and early expression of T-ag initiates the viral productive cycle (Brady and Salzman, 1986). Infection induces production of host replicative enzymes and after the onset of DNA replication the late genes are transcribed. Progeny DNA is packaged into virions after which cell lysis and

death occurs (Kornberg, 1980). Infection of non-permissive cells follows the same course up to the expression of the early genes but replication, late gene transcription and capsid protein production are blocked. Infected cells are stimulated to replicate their chromosomes and enter mitosis, after which most become quiescent. A fraction of the infected cells become transformed and continue to divide; these cells were found to contain several chromosomally integrated copies of the SV40 genome or fragments of it (Kornberg, 1980).

Permissive cells usually used are African Green monkey kidney-derived cell lines (eg. CV1), and non-permissive cell lines include mouse and hamster cell lines. Replication and transcription of SV40 are dependent on T-ag and cellular co-factors that interact with T-ag. The specificity of these interactions probably define the basis of cellular permissiveness.

1.7.3

SV40 Transcription

The early region of SV40 is transcribed soon after infection and uncoating and a single mRNA species is produced (Crawford *et al.*, 1978). This is differentially spliced into transcripts for the small t and large T antigens (reviewed in Tooze, 1981; Salzman, Natarajan and Seltzer, 1986). Production of T-ag from the SV40 early promoter is autoregulated by T-ag itself, probably by binding to several defined sites close to the promoter. T-ag also initiates viral replication and the start of late gene expression is dependent on DNA replication and the activity of T-ag acting through a 72 bp repeat sequence upstream of the late promoter. Viral transcripts are poly-adenylated prior to splicing through signals located 3' to transcripts in all reading frames.

1.7.4

SV40 Replication I - Initiation

SV40 replication is mediated by T-ag through the viral replication origin. The functional origin of SV40 replication for the purposes of this thesis is 340 bp long;

from nucleotide 5171 through 0 to nucleotide 270 of the SV40 sequence registered on the EMBL and NIH DNA sequence databases, where the last G residue of the unique Bgl I recognition site is nucleotide 1. Bases 5171 and 270 correspond to the 5' bases of a Pvu II and a Hind III restriction site, respectively.

T-ag is the sole viral protein required for initiation of SV40 replication; all other factors are provided by the host cell (see reviews by Campbell, 1986; Stillman, 1989; Challberg and Kelly, 1989). A complex containing T-ag and cellular co-factors binds to the replication origin and initiates bidirectional unwinding of the viral origin sequences (reviewed by Borowiec *et al.*, 1990). There are two strong binding sites for T-ag in the origin in addition to several weaker ones. A 64 bp region from nucleotides 5210 through 0 to 31 represents the core-origin region and contains one of the strong binding sites. The presence of the other binding sites alters the binding activity of T-ag at this site and the additional presence of binding sites for the transcription factor Sp-1 greatly enhance the efficiency of DNA replication. Each T-ag binding site contains several G-rich consensus sequences which differ in number and spacing between the sites, and these act as T-ag recognition sites. Other elements critical for replication are a 10 bp region that partially overlaps an inverted repeat and a 17 bp AT-tract; both undergo structural changes in the initial stages of replication. ATP increases the binding of T-ag to the origin and up to twelve T-ag monomers have been detected clustered around the region in a two lobed structure covering all faces of the DNA helix. The ATP-dependent T-ag complex denatures the origin region and continues to unwind the DNA bidirectionally in a 3' to 5' direction through a DNA helicase activity. This is linked to the presence of a single stranded DNA binding protein (SSB) which presumably keeps the single strands separate. Replication factor A (RF-A) has been shown to contain a substitutible SSB activity but its elongation activity cannot be substituted for by any other SSB in an SV40 *in vitro* replication assay (Kenny *et al.*, 1989). Phosphorylation of T-ag at critical residues is also required for the unwinding process (reviewed by Prives, 1990) and short RNA primers synthesized in the direction of early mRNA production by a primase activity initiate replication (Hay and DePamphilis, 1982).

1.7.5

SV40 Replication II - Elongation

A multi-enzyme complex of T-ag, DNA polymerases and other accessory proteins (including RF-A , RF-C and proliferating cell nuclear antigen [PCNA]) are responsible for replication of nascent DNA strands produced at the origin (reviewed by Challberg and Kelly, 1989). There is evidence for DNA polymerases α and δ being present and being involved with leading and lagging strand synthesis, respectively (Tsurimoto and Stillman, 1991). T-ag forms a specific complex with DNA polymerase α and a DNA polymerase α-primase activity interacts with the unwound DNA complex which includes T-ag and RF-A. PCNA is closely associated with DNA polymerase δ and its presence is required for efficient SV40 replication *in vitro*, suggesting that it is involved with DNA polymerase δ in DNA chain elongation. RF-C is a primer recognition protein and has stimulatory effects on the two DNA polymerases which are increased in the additional presence of RF-A and PCNA (Tsurimoto and Stillman, 1989).

1.7.6

SV40 Replication III - Termination

Termination of viral replication occurs 180° from the replication origin and the gap is sealed by gap-filling cellular proteins (DePamphilis and Bradley, 1986). Separation of the newly synthesized daughter strands is mediated by topoisomerase II which may also be responsible for relieving superhelical tension in the SV40 DNA prior to unwinding at the origin. The latter function could also be performed by topoisomerase I (Challberg and Kelly, 1989)

1.7.7

Transformation by SV40

Infection of non-permissive cells causes transformation of a small sub-set of cells. Translation of the early genes produces the small t and large T antigens of 20 kDa and 94 kDa, respectively. Most cells undergo abortive transformation where the

early genes and some host genes are transiently expressed leading to a transient induction of cellular DNA synthesis and cell division (Martin, 1981). A small proportion of these cells become permanently transformed and contain integrated SV40 sequences. Characteristics of fully SV40 transformed cells include continuous growth as an immortal population, ability to form clones when seeded at low cell density, low serum requirements, ability to overgrow a continuous cell layer to form foci, and anchorage-independent growth (Monier, 1986). Transformed cells continuously express the early SV40 viral antigens, with T-ag being mainly located in the nucleus and small t antigen in both the nucleus and the cytoplasm. A small population of T-ag resides on the cell surface and could influence the behaviour of cellular membrane proteins. Small t appears to provide an accessory function in cellular transformation and is not thought to be essential in the maintenance of *in vitro* transformation. T-ag on the other hand, is the only important transforming protein of SV40 and its multifunctional nature has made it difficult to determine how it affects the cell.

1.8

Large T Antigen

1.8.1

Introduction

T-ag is a multifunctional phosphoprotein with an apparent molecular weight of 94 kDa (reviewed by Rigby and Lane, 1983; Livingston and Bradley, 1987; Butel, 1986 and Stillman, 1989). It is produced from the single early transcript of SV40 by differential splicing and is the only viral protein required for SV40 replication. Functions mediated by T-ag include specific binding to the SV40 origin of replication, initiation of viral DNA replication, auto-regulation of SV40 early gene transcription, induction of SV40 late gene expression, initiation of cellular DNA synthesis and replication, and initiation and maintenance of the transformed phenotype in non-permissive cells. Several forms of T-ag have been described which are differentiated

by apparent differences in molecular mass, and perhaps the different functions of T-ag can be attributed to these different forms.

1.8.2

Structure and Function of T-ag

T-ag is 708 amino acids long and is described as oncogenic because it can induce tumours in new born hamsters. Ninety five to 98% of newly synthesized T-ag is transported to the nucleus and about 2% is destined for the plasma membrane. T-ag does not contain any signal peptide and the method of transport to the membrane is unknown. It is an integral membrane protein and when associated with an MHC class I product, it can act as a tumour-specific transplantation antigen. As described above, T-ag in the nucleus acts as an initiator of both viral and host DNA synthesis and also mediates late viral gene expression.

The different functions of T-ag have been mapped to different regions of the amino acid sequence. Site-specific origin binding is mediated by the region containing amino acids 132-246 (from the N- terminus) and the helicase activity has been mapped to the C-terminal half of the protein (Stillman, 1989). The N-terminal half has limited ability to initiate and maintain transformation (Rigby and Lane, 1983) and amino acids 418-528 contain the ATP binding site (Prives, 1990). The nuclear localization domain is located at amino acids 126-132, and residues 673-708 are involved with the host range of the virus (Prives, 1990).

The functions of T-ag are regulated by phosphorylation of two clusters of phosphorylation sites (reviewed by Prives, 1990). One cluster is close to the N-terminal and the other close to the C-terminal. Mutations at amino acid positions 120 (serine), 123 (serine) or 124 (threonine) which render these sites unphosphorylated completely abolish viral DNA replication. Removal of phosphate groups from serines and not from threonines by calf intestinal phosphatase increases the activity of T-ag in initiating viral replication. The effect was more marked in the early stages of replication, implying that dephosphorylation stimulated T-ag initiation activities at the SV40 replication origin. Phosphatase treatment did not affect the ATPase activity of

T-ag, suggesting that the dephosphorylation alters a domain involved in replication and not the overall protein structure. Mutation of ser-679 to an alanine residue appears to enhance SV40 replication *in vivo*. Phosphorylation of the threonine residue at position 124 *in vitro* by cdc2 kinase leads to more efficient binding of T-ag to the major binding site of the SV40 core origin. Phosphorylation at this single threonine residue of E. coli-synthesized T-ag, in the absence of serine phosphorylations, provides a T-ag which is more efficient at initiating DNA replication than T-ag produced in HeLa cells (McVey *et al.*, 1989). The model proposed from these data suggests that phosphorylation of the serine residues suppresses initiation of replication whilst phosphorylation of the threonine residue stimulates the process.

1.9

The Fate of Chromosomally Integrated SV40 Sequences

1.9.1

Introduction

In non-permissive cells that have been stably transformed with SV40, the viral sequences are integrated within chromosomal sequences (Sambrook *et al.*, 1968). If such cells are fused with permissive cells, viral sequences are excised using the integrated viral copy as a replication template. Excision appears to occur in three phases: first an over-replication of the viral DNA and flanking host DNA is initiated by T-ag and the replication origin (the 'onionskin'); secondly, recombination resolves the structure produced; and thirdly, excision products may continue to replicate as extrachromosomal elements.

1.9.2

Excision of SV40 Sequences

In transformed cells the chromosomal site of SV40 integration appears random, as does the break point(s) in SV40 which produces linear integrated viral DNA (Botchan, Topp and Sambrook, 1979). The only SV40 DNA sequence required

for the excision of extrachromosomal circular DNA is the SV40 origin of replication (Conrad, Liu and Botchan, 1982). A model was proposed in which integrated viral sequences, in the presence of T-ag, replicated in situ and that excision was a direct consequence of this replication (the 'onionskin' model - see above; Botchan, Topp and Sambrook, 1979; Bullock and Botchan, 1982). Multiple rounds of replication are initiated at the SV40 origin by T-ag and the resulting structure of replication forks following one another along the same DNA sequence can be resolved in several ways. Gaps produced by replication combined with a confined space could increase the probability of homologous excisional recombination. Cellular DNA repair mechanisms could also play a role in the removal of the 'onionskin' structure which it may see as an aberration of normal chromosomal replication.

Excision products can be grouped into three size-related categories: (1) precise and constant in size, (2) random in size, and (3) a mixture of both. Cell lines that produce precise excision products usually contain tandem duplications of DNA including the replication origin, and also produce the largest number of excised molecules (Conrad, Liu and Botchan, 1982). Cell lines containing a single integration without any duplications produce a heterogeneous population of excised DNA (Botchan *et al.*, 1980). Bullock and Botchan (1982) observed that for each product in a heterogeneous population of excised molecules, small amounts of homology pre-existed at the excision sites. A possible source of the tandemly duplicated arrays found in transformed cell lines is replication prior to the integration event or intra-chromosomal resolution of early replication *in situ*. Dora, Schwartz and Knippers (1989) observed that host sequences flanking the integrated viral sequences were invariably excluded from the excision products, arguing that homologous recombination is involved in the excision event.

1.9.3

Excision of Polyoma Sequences

Another system in which integration-excision studies have been carried out is that of polyoma virus transformation. Polyoma virus (Py) is similar to SV40 in many

ways (see Tooze, 1981). It is of the same class of virus as SV40 and also expresses a T-ag (Py T-ag) upon which viral replication is dependent. Excision of Py sequences is dependent on a functional Py replication origin, Py T-ag and duplicated Py sequences (Pellegrini, Dailey and Basilico, 1984). Unscheduled replication from the Py replication origin and recombination between homologous Py sequences could give rise to circular excision products. Colantuoni *et al.* (1982) showed that excision of Py can produce an integral number of viral genomes, leaving the flanking viral-host DNA junctions intact. This implied that homologous intramolecular recombination was responsible for the excision event in this system.

1.9.4

Summary

In summary, integrated SV40 sequences can be excised from their chromosomal locations as circular products under certain conditions. These are that a functional viral origin of replication and active T-ag are present, as well as duplications in the integrated viral sequences. Homologous recombination between duplicated sequences during over-replication of the viral DNA can produce excision products. The fate of the integrated template DNA is unclear, but in some cases it is unaltered. During excision studies, amplification of viral sequences at the site of integration has also been observed and are dealt with in the next section.

1.9.5

Amplification of Integrated Viral Sequences

Rat cells can be transformed by integration of Py sequences. These sequences generally integrate in head-to-tail tandem arrays which have been shown to be unstable and can undergo amplification (Colantuoni, Dailey and Basilico, 1980). Cell lines containing partially duplicated Py sequences in a tandem array can amplify integrated DNA at a high frequency; cell lines carrying single, unduplicated integrations do not appear to amplify the viral sequences (Colantuoni *et al.*, 1982).

Integrated Py sequences have been shown to be amplified *in situ* (Pellegrini *et al.*, 1984; Pellegrini and Basilico, 1987), although the genetic arrangement of the selectable gene was artificially created in these system. This may have been important in influencing the behaviour of the viral sequences. For example, certain transfection procedures have been shown to cause tandem integrations of transfected DNA and this may provide a structure with special properties relating to rearrangements and amplification. In the case of amplification of an introduced DHFR minigene attached to Py sequences (Pellegrini and Basilico, 1987), the endogenous DHFR gene was never seen to be amplified after selection with MTX, perhaps providing indirect evidence of positional and structural effects of surrounding sequences on introduced DNA.

1.10

The Aims of Projects Described in This Thesis

1.10.1

The Model System for DNA Amplification

Amplification of the CAD gene has always been intrachromosomal and until recently, very little had been discovered about the mechanism responsible. The main hurdle in uncovering the mechanism of amplification in the CAD and in other amplification systems is one of looking early enough at the amplified structure, before secondary rearrangements have occurred. If a population of cells is selected with a drug for cells that have amplified a particular gene, resistant cells are only visible when sensitive cells have detached from the plate and resistant cells have formed colonies. A colony needs to be expanded in order to obtain enough cells for molecular analyses. During this expansion many secondary rearrangements, amplifications and deletions may have occurred, and the structure of the amplified array is not a reflection of the initial mechanism that produced it.

The aim of the main project of this thesis was to construct a cell line in which a single copy of the SV40 origin of replication was integrated into DNA sequences

flanking the endogenous CAD gene of a monkey cell line. The cell line chosen would constitutively produce T-ag but the protein would only be active with respect to replication at a permissive temperature. Thus, the replicative activity of the SV40 replication origin could be modulated by temperature. In the presence of active T-ag the origin would initiate DNA synthesis and amplify the surrounding DNA through over-replication. The advantage of such a model would be that a particular locus (CAD) could be synchronously amplified in all cells of a population. The nature of PALA would allow selection of cell lines which had successfully amplified the CAD gene and probes for the amplified area could be isolated before the amplification event. The effect of short pulses of T-ag activity could be analysed in the population as a whole and secondary amplification of amplified CAD genes could also be followed. With the development of *in situ* hybridization techniques, a time course of amplification could be observed.

In summary, the model system would allow visualisation of very early and secondary amplification events at a specific locus and it would provide a test for the disproportionate replication ('onionskin') model in this particular system of DNA amplification. The model is described in more detail in the introduction to Chapter 3.

1.10.2

Evolution of Amplified DNA

Amplified DNA is lost from cells in the absence of the selective agent particular to the amplified gene and DNA co-amplified with such a gene is lost when the cell line is passaged in selective conditions over an extended period of time. Also, cells that have been selected in a single step can increase the copy number of already amplified genes by large amounts in second and third steps of selection. All these observations suggest that amplified DNA is unstable and has special properties. One aspect of this is the loss of amplified genes in the absence of selection. If, during passage at a constant drug concentration, amplified DNA is arranged into more stable and condensed arrays then how does this effect the loss of amplified genes when the cells are grown out of selection? The aim of the work

described in Chapter 4 was to investigate this question. Another aim was to measure the ability of various single-step resistant cell lines to produce second-step resistant colonies. This work contributed to the paper by Saito *et al.* (1989) which is included in this thesis.

1.10.3

In Situ Analysis of PALA-Resistant Cell Lines

The last section of Chapter 4, and Chapter 5 describe *in situ* experiments in PALA-resistant BHK and human cell lines. The aim of the work in BHK cells was to determine the resolution of the *in situ* hybridization technique described in the enclosed paper by Smith *et al.* (1990) using probes that had been co-amplified with CAD in previous experiments. Amplification of CAD had not been described in human cells and the aim of other work in this chapter was to investigate PALA-resistant human cell lines and determine the pattern of amplified CAD genes for comparison with the patterns seen by Smith *et al.* (1990) in PALA-resistant BHK cells. The human cell line system also has several advantages over the BHK system. For example, using probes which have been mapped to the same chromosome or chromosomal arm as CAD, it would be possible to analyse co-amplified DNA from all other parts of the chromosome. Such probes are available from various sources and gene banks and hybridised probes could be visualised using a two-colour detection protocol to distinguish the hybridisation pattern of each probe. Furthermore, many probes have been finely mapped to the chromosomal location of the gene encoding them and this makes the system highly accessible for analysis of not only CAD amplification but also for amplification of many other genes for which probes exist.

Chapter II

Materials and Methods

2.1

Nucleic Acid Preparation

2.1.1

Genomic DNA

Eukaryotic cells were trypsinised from semi-confluent monolayer cultures, washed once with PBS, the cell pellet suspended in PBS. SDS, sodium acetate and EDTA were added to 0.5%, 0.6 M and 5 mM, respectively (final volume is 5-10 ml/10^6-10^7 cells). When homogeneous, the lysate was gently mixed with an equal volume of TE-equilibrated phenol for at least 5 minutes. Chloroform (24:1 mixture of chloroform and iso-amyl alcohol), of half the volume, was added to the emulsion, followed by gentle mixing for at least five minutes. The organic and aqueous phases were separated by centrifugation for 10 minutes. The aqueous phase was recovered and extracted with an equal volume of phenol-chloroform (1:1) by gentle mixing for at least 5 minutes. The aqueous phase was separated and extracted as before, with an equal volume of chloroform. High molecular weight genomic DNA was precipitated by the addition of 2.5 volumes of ethanol at room temperature and collected by spooling onto a plastic rod or a bent and sealed Pasteur pipette. The DNA was dissolved in 2-5 ml water containing 5 µg/ml RNase and incubated at 37°C for 30 minutes. The DNA was ethanol precipitated (addition of sodium acetate to 0.3 M and 2.5 volumes of ethanol), spooled, washed with 70% ethanol and dissolved in TE to a concentration of 300-600 µg/ml.

2.1.2

Large Scale Plasmid Preparation (Maniatis *et al.*, 1982)

E. coli containing the plasmid were grown overnight at 37°C with vigorous shaking in LB medium containing an appropriate antibiotic. Cultures were harvested by centrifugation and the cells were suspended in 10 ml of an ice-cold solution of 50 mM Tris.HCl/10 mM EDTA. After 10 minutes incubation on ice, 20 ml of a freshly

made solution containing 0.2 N NaOH/0.5% SDS was added and the lysate was mixed by gentle inversion and left on ice for 10 minutes. Protein debris and denatured chromosomal DNA were precipitated with 15 ml of 3 M potassium acetate, pH 4.8 on ice for 10 minutes. The precipitate was gently broken up and cleared from the solution by centrifugation in a benchtop centrifuge. Plasmid DNA was precipitated from the supernatant by the addition of an equal volume of iso-propanol on ice. Plasmid DNA was pelleted by centrifugation and dissolved in 4 ml of TE. One gram of CsCl was added for each ml of solution and the mixture was sealed in a Beckman Quickseal tube with 200 µg/ml of ethidium bromide. After centrifugation at 400 000 g for 4 hours in a vertical rotor (VTi65 series), the band of supercoiled plasmid DNA was collected using a syringe and needle. The plasmid DNA was diluted, CsCl added as before, and spun again for 18 hours. The DNA was collected as above, diluted to 5 ml with TE and extracted at least five times with water-saturated butan-1-ol to remove the ethidium bromide. The plasmid DNA was ethanol precipitated, pelleted by centrifugation, washed with 70% ethanol and dissolved in TE at 300-600 µg/ml.

2.1.3

Large Scale Preparation of Phage Lambda (λ) (Maniatis *et al.*, 1982)

Four hundred millilitres of NZCYM medium was inoculated with 4 ml of E. coli strain LE392 plating bacteria (prepared as described in section 2.4.3) and 1-50 x 10^7 pfu (plaque forming units) of λ. The culture was shaken at 37°C until the bacteria lysed (8-16 hours). RNase A and DNase, both at 5 µg/ml, were added to the lysed cultures which were incubated at 37°C for 30 minutes. NaCl was added to 1 M and the culture cooled and allowed to stand on ice for at least 60 minutes, after which precipitated proteins were removed by centrifugation. PEG 6000 was added to the supernatant to 10% and dissolved using a magnetic stirrer. The solution was left on ice for at least 60 minutes, after which the phage particles were collected by centrifugation and suspended in 10 ml of SM. The PEG was extracted by vortexing

the suspension for 30 seconds with an equal volume of chloroform. CsCl was added at 0.75 g/ml and the suspension was spun at 285 000 g at 4°C for 24 hours in a Beckman SW40 rotor. The phage particles appeared as a blue/white band midway up the tube and were collected with a syringe and needle through the side of the tube.

2.1.4

Small Scale Plasmid Preparation (Maniatis *et al.*, 1982)

Single bacterial colonies were cultured overnight in 2.5 ml LB + antibiotic at 37°C. Approximately 1.5 ml was removed, the cells pelleted and resuspended in 100 µl of ice-cold 50 mM Tris.HCl/5 mM EDTA. Two hundred microlitres of 0.2 N NaOH/0.5% SDS was added and mixed by gentle inversion. Protein debris and chromosomal DNA were precipitated with 3 M potassium acetate (pH 4.8) and extracted with 100 µl chloroform. The aqueous phase was extracted once each with an equal volume of phenol, phenol/chloroform and chloroform. Plasmid DNA was ethanol precipitated, washed with 70% ethanol and dissolved in 30 µl of TE containing 20 µg/ml RNase.

2.1.5

Small Scale λ Preparation (Maniatis *et al.*,1982)

Single λ plaques were individually picked with a Pasteur pipette into 500 µl of SM. A drop of chloroform was added and the tube was briefly vortexed and spun. Fifty microlitres of E. coli LE392 plating bacteria were mixed with 50 µl of phage particles and the mixture was incubated overnight in 5 ml of NZCYM at 37°C with vigorous shaking. RNase and DNase (both to 5 µg/ml) were added to each lysed culture. After 30 minutes at 37°C, λ particles were precipitated with an equal volume of 2 M NaCl/20% PEG 6000 on ice for at least 60 minutes. λ particles were spun down and resuspended in 500 µl SM. SDS and EDTA were added to 0.1% and 5 mM, respectively, and the mixture incubated at 68°C for 30 minutes. After phenol

extraction, the λ DNA was precipitated with an equal volume of iso-propanol, pelleted, washed with 70 % ethanol and dissolved in 30 µl of TE.

2.1.6

Extraction of Episomal DNA (Hirt, 1967)

Cells were grown to 10^6 cells per 100 mm dish and washed extensively with PBS. Two millilitres of a solution containing 10 mM EDTA and 0.6% SDS was added directly to the plate and left at room temperature for 10 minutes. The lysate was poured into Beckman Ultraclear centrifuge tubes with 500 µl of 5 M NaCl and placed at 4°C for 48 hours. Precipitated protein was pelleted by centrifugation at 18, 000 rpm for 60 minutes at 4°C and the supernatant was phenol extracted. DNA was ethanol-precipitated and dissolved in an appropriate amount of TE.

2.2

DNA Manipulations (Maniatis *et al.*, 1982)

2.2.1

Restriction Endonuclease Digestion of DNA

DNA was generally restricted in 10 mM Tris.HCl (pH 7.5), 10 mM $MgCl_2$ and 1 mM DTT supplemented with NaCl from 0-150 mM as recommended by the supplier, or in the buffer supplied by the manufacturer. For some enzymes and for genomic DNA digestion, BSA was included at a final concentration of 100 µg/ml. Restriction enzymes were normally used at a 5- to 10-fold unit excess and incubation was at 37°C for 1-4 hours.

2.2.2

Recovery of DNA from Agarose Gels

DNA fragments were recovered from gel slices by one of two methods: (a) by using a Geneclean kit (Bio 101, Inc.) as directed by the manufacturer; or (b) by cutting

the fragment from the gel and electroeluting it within Spectropore dialysis membrane in 1x TAE running buffer. For fragments larger than 8 kb the latter procedure was used.

2.2.3

Ligation of DNA fragments

Plasmid DNA was digested with restriction enzymes and incubated for 10 minutes at 37°C with 1 unit of calf intestinal alkaline phosphatase (CIP; Boehringer-Mannheim). The DNA was agarose gel-purified and mixed with a 5-fold excess of insert DNA. Ligation with 10 units of T4 DNA ligase was performed in 50 mM NaCl, 10 mM Tris.HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT and 1 mM of freshly added ATP for 4-18 hours at 14°C. After incubation, the ligase was inactivated at 65°C for 10 minutes.

2.2.4

Oligo-Labelling (Adapted from Feinberg and Vogelstein, 1983)

Fifty to two hundred nanograms of DNA were added to 2 µl of random hexamer primers ($p(dN)_6$, Pharmacia; 90 OD units/ml) and 10 µl of 2.5x HTM buffer (0.5 M Hepes pH 6, 0.125 M Tris.HCl pH 8,12.5 mM $MgCl_2$ and 20 mM β-mercaptoethanol). The mixture was boiled for 3 minutes, cooled on ice and 166 µM each of dATP, dGTPand dTTP, 70 µCi of deoxycytosine 5'-[α-^{32}P] triphosphate (radio-labelled dCTP; Amersham) and water were added to a final volume of 24 µl. The mixture was incubated at 37°C for 1-3 hours with 4 units (1 µl) of the Klenow fragment of DNA polymerase (Amersham). The reaction was stopped by adding EDTA and SDS to 5 mM and 1%, respectively, and unincorporated nucleotides were removed through a Sephadex ® G50 (medium) spin column.

2.2.5

Nick Translation (Rigby *et al.*, 1977)

Cosmids used as probes for *in situ* hybridisation were labelled either with biotin-11-dUTP using a BRL Nick Translation kit or with biotin-14-dATP using a Bio-nick kit (Bethesda Research Laboratories) as instructed by the supplier.

2.2.6

DNA Sequencing (Sanger *et al.*, 1977)

Double stranded plasmids were sequenced using a Sequenase kit (U. S. B. Corp.) with adenosine 5'-α-[^{35}S] thiotriphosphate as instructed by the manufacturer.

2.2.7

Polymerase Chain Reaction

DNA fragments were amplified in 10 mM Tris.HCl pH 8.3, 50 mM KCl, 0-6 mM $MgCl_2$ and 0.001% gelatin, 0.5-1 µM of each primer, 200 µM of each dNTP and 2.5 units of Taq DNA polymerase (Amplitaq, Cetus Corp.) with 10 pg-1 µg of template DNA. Optimum concentrations of primers and $MgCl_2$ and annealing temperatures of the primers were determined using a positive control template. Thirty cycles of denaturation (94°C,1.5 minutes), primer annealing (37-65°C, 3 minutes) and extension (72°C, 5-10 minutes) were typical. Amplified fragments were visualised by agarose gel electrophoresis.

2.3

Gel Electrophoresis and Transfer to Solid Supports

2.3.1

Agarose Gel Electrophoresis

Agarose was dissolved in 1x TAE by boiling, cooled to 50°C before adding Ethidium Bromide to 0.5 µg/ml and cast to a depth of 5-10 mm in a gel tray. Samples

were mixed with 6x loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol and 25% type-400 Ficoll in water) and run at 2-4 V/cm in 1x TAE running buffer against 400 ng of a 1 kb ladder (BRL) or Hind III-digested λ DNA as a size reference. Photographs were taken using a U.V. light box (360 nm) and Polaroid 654 instant black and white positive film.

2.3.2

Polyacrylamide Gel Electrophoresis

The products of sequencing reactions were run on vertical, 0.5 mm thick, 8% polyacrylamide (19:1 acrylamide:bis-acrylamide) gels containing 7M urea, 0.1% ammonium persulphate, and 1x TBE polymerised with 40 µl TEMED (N, N, N',N'-tetramethyl-ethylenediamine) per 50 ml gel. Reactions were stopped using 2.5x stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol green) and denatured at 80°C for 5 minutes before loading. Gels were run in 1x TBE and pre-run for 20 minutes prior to loading.

2.3.3

Alkaline Southern Blotting (Reed and Mann,1985)

One to fifty nanogrammes of plasmid DNA or 5-20 µg of genomic DNA were restricted and fractionated on agarose gels. Size marker lanes were removed after photography and the gel was depurinated in 0.25 M HCl for 10 minutes. The gel was rinsed with water before being neutralised in 0.4 M NaOH for 5 minutes. The DNA was transferred to a pre-wet nylon membrane (Zetaprobe [Biorad]or Genescreen plus [Dupont]) by capillary action in 0.4 M NaOH for at least 4 hours. After transfer the membrane was neutralised in 2x SSC for 5 minutes and then air-dried for at least 1 hour.

2.3.4

Genomic Slot Blots (McIntyre and Stark, 1988)

Ten thousand-10^5 eukaryotic cells in 50 µl of PBS were added to 200 µl of 0.5 M NaOH and heated to 80°C for 10 minutes. One hundred microlitre aliquots were applied to duplicate slots of a slot blot manifold containing a nylon membrane and were drawn through by capillary action. Each slot was washed through with 100 µl of 0.4 M NaOH. When all the liquid had been absorbed, the blotter was dismantled and the membrane neutralised in 2x SSC for 5 minutes before being air-dried for at least an hour.

2.3.5

Nucleic Acid Hybridisation to Solid Support-Bound DNA

Pre-wet (2x SSC) nylon membranes were pre-hybridised in 50% formamide, 5x SSPE, 5x Denhardt's solution, 0.1-1% SDS and 100-200 µg/ml of sonicated, single stranded herring sperm DNA (300-500 bases long) at 42°C for at least 1 hour. One to two million dpm/ml of denatured ^{32}P-labelled DNA was added directly to the prehybridising membrane and left at 42°C for a further 12 hours. Membranes were rinsed twice in 2x SSC for 5 minutes each, and then washed in 2x SSC/1% SDS followed by 0.2x SSC/1% SDS for 30-60 minutes each at 50-68°C. Washing temperatures depended on the stringency of hybridisation required. Bound probe was visualised by autoradiography using Kodak XAR5 X-ray film and two Cronex intensifying screens (Dupont) at -80°C.

2.3.6

Densitometry

The signals of hybridisation on Southern blots and slot blots were quantified where necessary using an LKB UltraScan II densitometer. For BHK Southern blots the signals were normalised for loading differences with a repeat-free 5 kb EcoR I fragment from the cosmid c981. This cosmid contains genomic DNA present at a

single copy in the BHK genome (Saito *et al.*, 1989). Whole cell slot blots were corrected for loading differences by hybridisation to a total genomic DNA probe. In each case the slots were internally controlled with wild type cells of the relevant species.

2.4

Bacteriology (Maniatis *et al.*, 1982)

2.4.1

Competent Bacteria

A single colony of E. coli JM109 or DH5α bacteria was grown to OD_{550} = 0.3 in 7 ml of ψ broth. 5 ml was sub-cultured into 200 ml of ψ broth and grown to OD_{550} = 0.5. The bacteria were transferred to pre-chilled centrifuge tubes, left on ice for 15 minutes and pelleted. The pellet was suspended in 20 ml of TfB I (100 mM RbCl, 50 mM $MnCl_2$, 10 mM $CaCl_2$, 35 mM KOAc pH6.8 and 15% glycerol) and left on ice for 20 minutes. The cells were pelleted, resuspended in 7 ml of TfB II (10 mM Mops pH 6.8, 75 mM $CaCl_2$, 10 mM RbCl and 15% glycerol) and left on ice for 20 minutes. Two hundred microlitre aliquots were frozen in a methanol/dry ice bath and transferred to -80°C until required. Transformation efficiency was 10^6-10^7 transformants/μg of plasmid.

2.4.2

Bacterial Transformation

One to fifty nanograms of DNA were added to 200 μl of competent E. coli and left on ice for 30 minutes. Bacteria were heat shocked at 42°C for 90 seconds and returned to ice for 5 minutes. One millilitre of prewarmed LB was added and the cells incubated at 37°C for 30-60 minutes. Cells were pelleted in a microfuge, resuspended in 100 μl of LB, and 10 μl and 90 μl were spread on LB plates (LB + 1.5% Bactoagar) containing antibiotic(s). Plates were incubated overnight at 37°C.

2.4.3

Preparation of Plating Bacteria

A single colony of E. coli strain LE392 bacteria was grown overnight at 37°C in NZCYM medium supplemented with 0.2% maltose. The cells were pelleted, resuspended at half the original volume in 10 mM $MgSO_4$ and stored at 4°C for up to 7 days.

2.4.4

Bacterial Infection with λ

E. coli strain LE392 plating bacteria were incubated at room temperature for 30 minutes or at 37°C for 10 minutes with λ. Cells were then mixed with molten NZCYM/0.7% agarose (at 42°C), poured onto a prewarmed NZCYM plate and incubated at 37°C.

2.4.5

***In Vitro* Packaging of Phage Particles and Cosmids**

One to fifty nanograms of phage or cosmid DNA were added to a 1:1.5 mixture of the packaging extracts supplied as part of a Giga-Pack Gold Packaging kit (Stratagene) and left at room temperature for 90 minutes. The packaged DNA was mixed with 500 µl of SM and 1 µl, 50 µl and 450 µl aliquots were used to infect separate 500 µl aliquots of E. coli LE392 plating bacteria . For packaged phage the bacteria were mixed with 3.5 ml of NZCYM medium containing 0.7% molten agarose (at 42°C) and plated onto NZCYM plates which were then incubated at 37°C overnight for plaque formation. Bacteria infected with packaged cosmids were mixed with prewarmed NZCYM and incubated at 37°C for 60 minutes. Bacteria were then pelleted, resuspended in 100 µl of NZCYM medium and 10 µl and 90 µl were plated on to separate NZCYM plates containing the appropriate antibiotic.

2.4.6

Antibiotic Concentrations

Ampicillin (Ap; Sigma) was used at 50 µg/ml in LB or NZCYM media. Tetracycline (Tc; Sigma) was used at 12.5 µg/ml in media that did not contain $MgSO_4$, which antagonises Tc.

2.5

Tissue Culture Techniques

2.5.1

Medium

All cell lines were grown in Dulbecco's modified Eagle's medium - (DMEM; Flow Laboratories) supplemented with 10% foetal calf serum (FCS), 100 U/ml penicillin, 100 µg/ml streptomycin and 2 mM glutamine.

2.5.2

Serum

FCS was purchased from Gibco or Bockneck and was either heat-inactivated for 60 minutes at 55^{o}C (HiFCS) or dialysed (Swyryd, Seaver and Stark, 1974) in order to remove uridine (dFCS). One hundred millilitres of serum was serially dialysed for 24 hours each against 4 L of 0.15x PBSA, 0.5x PBSA and 0.15x PBSA at 4^{o}C. All serum was filtered through a 0.22 µm filter, aliquoted and frozen until required.

2.5.3

Transfection I - Electroporation

This method for the introduction of exogenous DNA into cells is based on the procedure of Chu, Hayakawa and Berg (1987). DNA enters the cell via holes in the plasma membrane induced by the discharge of an electrical current through the medium in which the cells are suspended. The method is said to be biased towards

the integration of individual copies at a single site per cell due to the non-precipitation and concatamerisation of input DNA.

Approximately 10^5-10^6 cells were suspended in 0.8 ml of ice-cold PBS and placed in a Gene Pulser (BioRad) electroporation cuvette. The cuvette was left on ice for 10 minutes, shocked and returned to ice for a further 10 minutes. Cells were then plated in medium containing 10% HiFCS for 48 hours before the appropriate drug was added.

2.5.4

Transfection II - Calcium Phosphate Precipitation

Cells were seeded at 10^6 per 100 mm dish and allowed to attach overnight. DNA was prepared in sterile TE at a concentration of 2-10 µg per 40 µl. For a single 100 mm dish, 40 µl of the DNA solution was added to 440 µl of 2x Hebs buffer (20 mM Hepes, 137 mM NaCl, 10 mM KCl and 2 mM Na_2HPO_4; pH 7.05) and air was gently bubbled through the mixture while 400 µl of CC mix (250 mM $CaCl_2$ in 0.1x TE) was added dropwise. The solution was left at room temperature for 20-40 minutes and the culture medium on the cells was replaced with 6 ml of fresh DMEM. Eight hundred and eighty microlitres of the $CaPO_4$ precipitate was added to each dish and thoroughly mixed with the medium. The cells were left undisturbed for 8 to 16 hours and then glycerol-shocked with 2.5 ml of pre-warmed DMEM containing 25% glycerol. The cells were rinsed twice with 10 ml of prewarmed DMEM and incubated for 48 hours in DMEM + serum before selection with the appropriate drug.

2.5.5

Selective Concentrations of Drugs

Stable ts-COS transfectants were selected in 400 µg/ml of G418 (neomycin) or 200 µg/ml of hygromycin B and BHK stable transfectants were selected in 500 µg/ml of G418. All selections were carried out at the restrictive temperature (where applicable) in DMEM containing 10% HiFCS.

2.5.6

Frequency and Rate of PALA Resistance

The frequency of PALA resistance for a cell line was determined by plating 1-5 x 10^5 cells on a 100 mm dish in the concentration of PALA required. At least five dishes were usually set up. The number of colonies obtained for that drug concentration reflects the steady state number of resistant mutants present in the population. This frequency is useful for determining the sensitivity of a cell line to a particular drug and is different to the rate of amplification, which is the number of amplification events per cell per division.

For the analysis of amplification events, cells were selected using a Luria-Delbrück fluctuation analysis (Luria and Delbrück,1943) so that only independent and non-pre-existing amplification events are studied. A small population of cells (100-1000; dependent on the frequency of resistance for the particular cell line) were grown in non-selective conditions until they numbered about 10^6. The cells were dispersed and placed in medium containing a concentration of PALA which was close to the minimum required to select for amplification of the CAD gene (usually about 3 times the LD_{50} of PALA for the cell line). If a large number of resistant colonies survived then a pre-existing mutant was most likely present in the starting population and was expanded during the selection. If only a few resistant colonies arose then the amplification event probably occurred within a few cell divisions before selection.

2.5.7

Isolation and Storage of Cell Lines

Isolated colonies were ring cloned when they reached approximately 300-500 cells in size and were expanded in successively larger dishes. Aliquots of about 10^5 cells were suspended in 50% HiFCS, 40% DMEM and 10% DMSO and placed in vials which were wrapped in tissue. After 24 hours at -80°C the vials were transferred to liquid nitrogen.

2.5.8

Retrieval of Cells

Frozen vials were removed from liquid nitrogen and rapidly thawed in a 37°C water bath. Cells were plated in DMEM containing 10% HiFCS without selection. After allowing time to attach, the cells were washed several times with PBSA and refed with fresh DMEM and serum. Selective medium was added (if required) after the first passage.

2.5.9

Colony Staining

Plates were washed at least once with PBSA and cells were fixed in formol saline for at least 20 minutes. After fixing, the cells were washed with distilled water and stained with 20% giemsa stain in PBSA for at least 20 minutes. Plates were rinsed in water, inverted and air-dried. If necessary, colonies were examined under a light microscope to distinguish colonies from cell clumps.

2.6

In Situ Hybridisation (Smith *et al.*, 1990)

2.6.1

Chromosome and Slide Preparation (Wahl *et al.*, 1982)

Cells were seeded in T-flasks and grown to 80% confluence overnight. Colcemid (BRL) was added to each culture (BHK cells received 150 ng/ml and human cells 50 ng/ml) and the flasks were returned to an incubator for 1.5 to 2 hours (BHK cells) or 90 minutes (human HT 1080 cells). Metaphase-arrested cells were dislodged from the flask by vigorous shaking and were removed from the medium by centrifugation at 1000 rpm for 5 minutes. Most medium was aspirated off and the cell pellet was gently resuspended in the remaining drops of medium. Five millilitres of prewarmed (37°C) 75 mM KCl was gently added to the cells and incubated at 37°C for

13 minutes. The cells were very gently pelleted at 800 rpm for 5 minutes, the KCl was removed and the cells were fixed on ice for 5 minutes in pre-cooled methanol/acetic acid (3:1, v/v) fixative. The cells were very gently pelleted for 5 minutes and resuspended in approximately 1 ml of fixative.

Glass microscope slides were cleaned by immersion in 30% RBS (Pierce and Warriner, Ltd) for 1 hour then rinsed with distilled, deionised water and stored in alcohol. Chromosomes were dropped onto the slides from a siliconised Pasteur pipette equipped with a rubber bulb. Each slide was checked with a light microscope for areas with abundant metaphase spreads of chromosomes, which were marked with a diamond etcher. Slides were matured at room temperature for at least 24 hours before use.

2.6.2

RNase A and Proteinase K Treatment

Slides were incubated with 100 µg/ml of DNase-free RNase A (in 2x SSC) at 37°C in a moist chamber and under a 50 mm x 22 mm coverslip for 1 hour. Coverslips were removed and the slides were dehydrated by sequential 5 minute washes in 70%, 85% and 100% alcohol. Dehydrated slides were air-dried for at least 1 hour before proteinase K treatment.

Proteinase K (Boeringher) was diluted from a stock solution to 600 ng/ml in 20 mM Tris.HCl and 2 mM $CaCl_2$ (pH 7.4) and pre-warmed to 37°C in a coplin jar. RNase A-treated slides were immersed in the jar for 7.5 minutes and then sequentially rinsed at room temperature with 20 mM Tris.HCl and 2 mM $CaCl_2$ (pH 7.4) and with PBSA containing 50 mM $MgCl_2$. Slides were fixed in 4% paraformaldehyde in PBSA for 10 minutes at room temperature, rinsed with PBSA and dehydrated as before.

2.6.3

Denaturation of Chromosomes

Slides previously treated with RNase A and proteinase K were denatured in 70% deionised formamide and 2x SSC (pH 7) at 76°C for 7.5 minutes (BHK chromosomes) or at 73°C for 5 minutes (human chromosomes) in small batches to maintain accurate denaturation times and temperatures. Denatured slides were dipped through ice-cold 2x SSC and washed for 5 minutes in ice-cold 70% alcohol before being dehydrated as before, except that the slides were not allowed to dry out after the final wash.

2.6.4

Hybridisation

For a single slide,100-200 ng of nick-translated probe was added to 10-50 µg of competitor DNA and the volume was adjusted to 3 µl. Seven microlitres of a hybridisation master mix was added to give a final solution containing the probe and competitor in 50% formamide, 10% dextran-sulphate, 2x SSC and 1% Tween-20 (Surfact-Amps 20; Pierce). The probe and competitor were denatured at 70°C for 10 minutes and repeat sequences in the probe were suppressed at 37°C for 20 minutes. After the final dehydration wash following denaturation, slides were drained until the last of the alcohol had evaporated (about 1 minute) and 10 µl of hybridisation solution containing the competed probe was added to the marked area of each slide. A coverslip was placed over the area, bubbles were expelled and the edges were sealed with rubber solution. Slides were incubated in a moist chamber at 37°C for 18 hours.

2.6.5

Detection of Hybridisation

The rubber seal and coverslip were removed from each slide and the slides were washed three times in 50% deionised formamide, 2x SSC (pH 7) at 42°C for 5

minutes each, followed by three washes in 2x SSC (pH 7), also at 42°C for 5 minutes each. All further washes were at room temperature. After a 5 minute wash in 4x SSC, 0.05% Tween-20, slides were blocked with 100 µl of blocking solution (4x SSC, 0.05% Tween-20 containing 5% non-fat rehydrated milk (Carnation; Nestlè))under a coverslip for 10 minutes. One hundred microlitres of 5 µg/ml avidin-FITC in blocking solution was added under a coverslip and the slides were left for 20 minutes, after which they were washed three times for 5 minutes each in 4x SSC, 0.05% Tween-20. The slides were drained and 100 µl of 5 µg/ml biotinylated goat anti-avidin in blocking solution was added to each slide under a coverslip. The slides were given three 5 minute washes in blocking solution and avidin-FITC was applied as before. Excess avidin was washed off with blocking solution followed by two 5 minute washes with PBS. Chromosomes were counterstained for 5 minutes in 0.1 µg/ml propidium iodide in PBS followed by a rinse in PBS. Forty microlitres of Citifluor Antifade mountant was added to the marked area of each slide and a 50 mm x 22 mm coverslip was placed over the area. Air bubbles and excess mountant were expelled and the edges of the coverslip were sealed with rubber solution. Slides were stored in a desiccated box at 4°C until analysis using a MRC-600 confocal microscope.

Chapter III

A Model System for Gene Amplification in Primate Tissue Culture Cells

3.1

Overview

3.1.1

Introduction

Recurring problems in gene amplification studies have been the rarity of amplification events at a specific locus, and the time that lapses between an event and the earliest time that it can be recognised as such and studied. When these 'early' amplified structures are dissected by standard molecular techniques it is likely that rearrangements will have taken place and this confuses the mechanistic implications of the structures. A system was devised in which a single defined locus could be amplified throughout a cultured cell population at the same time in each cell. This would facilitate analysis of early events before secondary rearrangements could occur. Central to the proposed system was the 'onionskin' model (see Chapter I), in which a cellular replication origin is thought to initiate replication more than once in a single cell cycle, thereby generating either intra-chromosomal or episomal amplified arrays following recombination of newly replicated strands. The project was to test the 'onionskin' mechanism and to determine the outcome of over-replication from a single origin and its consequences on the surrounding the DNA. This chapter describes the progress in setting up a cell line with a viral replication origin targeted to the CAD locus by homologous recombination. Also described are the effects of over-replication from a viral origin integrated into a cellular genomic background.

3.1.2

The Model System

In order to over replicate a defined locus, a viral origin of replication was chosen as the sequence requirements of a mammalian replication origin are not known. The viral origin would be delivered to the locus with minimal disruption of the surrounding DNA and it was hoped that this could be achieved by homologous replacement of an existing region of DNA with a modified fragment containing the

viral replication origin. A restriction fragment from the locus, which did not contain coding sequences would be modified by insertion of the viral origin and following transfection of this modified fragment, a certain number of integrations would replace the fragment's cellular homologue.

The SV40 replication origin initiates bidirectional replication multiple times per permissive host cell cycle when SV40-encoded T-ag is present. The model required integration of the SV40 origin into sequences flanking the CAD gene in permissive cells. Previous studies (Mark Rolfe, unpublished results) using CV1 cells (African green monkey kidney cells) and a temperature-sensitive SV40 early mutant virus as a source of T-ag showed that single copy integrations of an SV40 origin-containing plasmid did not give rise to episomal DNA at the permissive temperature. To simplify the system in terms of the number of variables, ts-COS cells (Rio, Clark and Tjian, 1985) were used. These cells have a single integrated copy of the SV40 tsA 1609 gene under the transcriptional regulation of the Rous sarcoma virus long terminal repeat (RSV LTR). T-ag is constitutively produced but is inactive with respect to replication at a restrictive temperature (40^{o}C) and active at a permissive temperature (33^{o}C). Therefore, replication of the SV40 origin is temperature-dependent in these cells. The SV40 origin would be targeted to the CAD gene by homologous recombination using DNA isolated from the CAD locus. Once integrated, replication from the SV40 origin could be switched on and off by a temperature shift to and from the permissive temperature with consequent replication of the flanking DNA, including CAD.

3.2

Preliminary Experiments

3.2.1

The Monkey CAD Locus

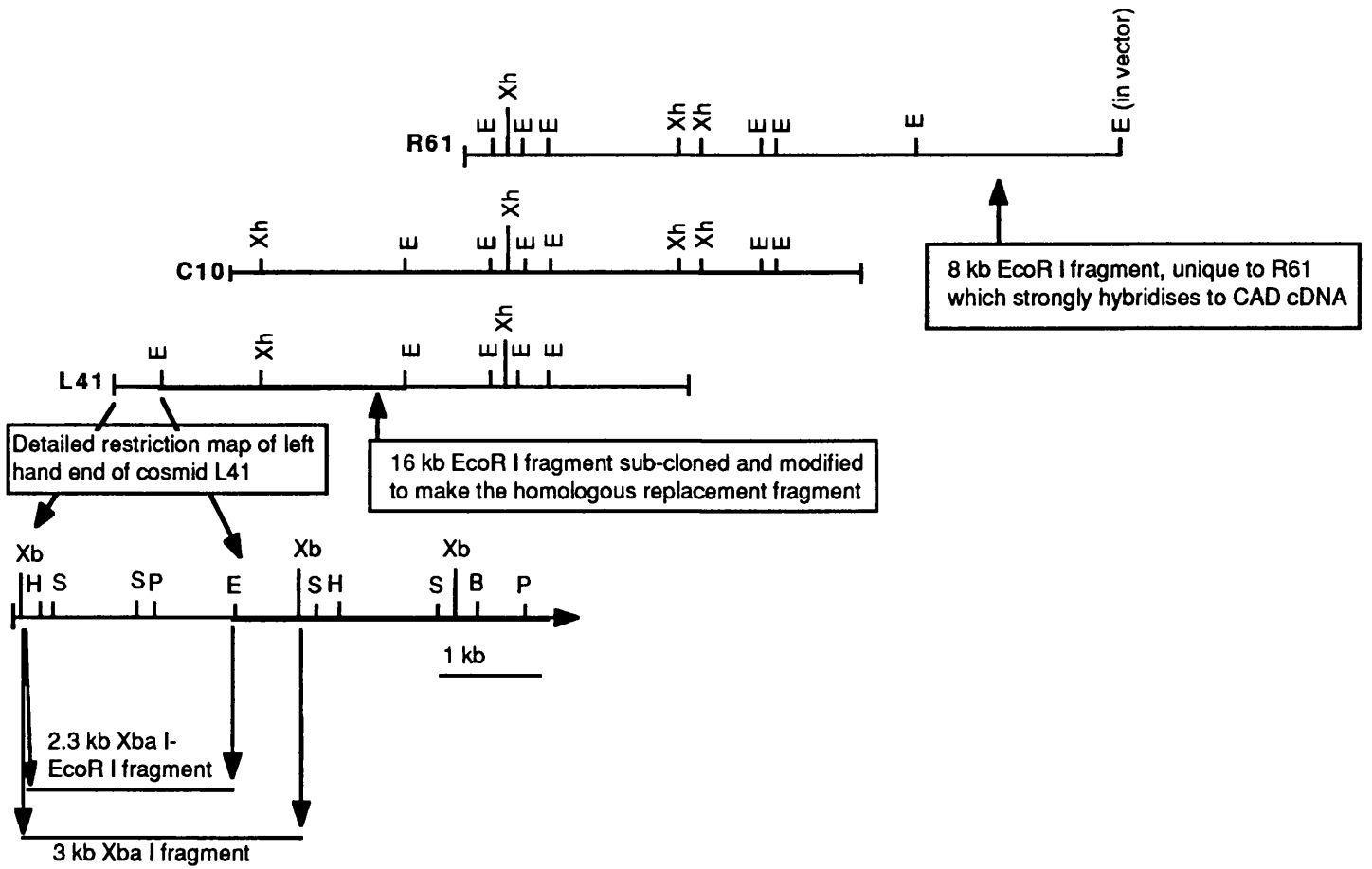
Cosmid C10 was isolated from a monkey genomic library probed with a hamster CAD cDNA at low stringency . Terminal fragments from the insert of C10

Figure 3.1

Restriction Map of the Cosmids L41, R61 and C10

The central Cosmid C10 was isolated from a monkey genomic cosmid library using a hamster cDNA probe. Terminal fragments of the C10 insert were used to isolate L41 and R61. A major coding region was located, using the CAD cDNA, on an 8 kb EcoR I fragment unique to the cosmid R61. The restriction sites shown on each cosmid were mapped by partial and mixed restriction digests (M. Rolfe, unpublished data). The left hand end of cosmid L41 was sub-cloned into a bacterial plasmid and mapped in more detail using mixed restriction digests. The heavy line between two EcoR I sites on L41 was the region selected as CAD homology in the homologous recombination fragment. The two fragments shown beneath the fine map of the left hand end of cosmid L41 were extensively used to analyse the integration site of the homologous replacement fragment in transfected cells.

E=EcoR I; H=Hind III; P=Pst I; S=Stu I; Xb=Xba I; Xh=Xho I


R61
C10
L41
E (in vector)
8 kb EcoR I fragment, unique to R61
which strongly hybridises to CAD cDNA
16 kb EcoR I fragment sub-cloned and modified
to make the homologous replacement fragment
Detailed restriction map of left
hand end of cosmid L41
Xb
H S
S P
E
Xb
S H
Xb
S
B
P
1 kb
2.3 kb Xba I-
EcoR I fragment
3 kb Xba I fragment

were used to isolate the left and right flanking cosmids L41 and R61. The three cosmids were mapped using mixed restriction digests and partial restriction digests, and the partial map that was provided by M. Rolfe is shown in figure 3.1.

To determine where the major CAD coding sequences were located in the three cosmids, each was digested with EcoR I, Southern blotted and hybridised with a CAD cDNA probe (Shigesada *et al.*, 1985). The hybridised filter was washed under high stringency conditions and the autoradiogram is shown in figure 3.2b. The probe hybridised to a 8 kb EcoR I fragment unique to cosmid R61 as shown by comparison with the ethidium bromide stained gel prior to Southern blotting. This shows that a large amount of coding sequence is located in the region of cosmid R61 that does not overlap with the central cosmid C10. Although the cosmid C10 was isolated by hybridisation to the same probe as that used above, these sequences were not seen on the Southern blot as the filter was washed under more stringent conditions than those used to screen the monkey library (M. Rolfe, personal communication). No coding sequences are seen in the L41 lane, although it may be possible to see hybridisation in blots washed at lower stringency. The coding sequences of C10 with low homology to the hamster CAD cDNA and the more homologous sequences in R61, indicate that a large coding region for CAD exists in the area of overlap between these two cosmids; this makes it less likely that coding sequences are present in cosmid L41.

The strongly hybridising 2.9 kb band in all the lanes is a vector fragment bound to labeled pUC fragments which are present in the probe preparation. The differences in intensity between the vector band and the CAD-specific band are probably due to the differences in homology between the probe sequences and the hybridising band sequences. One hundred percent homology exists between the pBR322 sequences of the cosmid and those contaminating the probe. On the other hand, the hamster CAD gene has been shown to contain a large number of small introns separated by small exons and spread over 30 kb. If this is the case for the monkey locus also, then the amount of non-homology between the cDNA

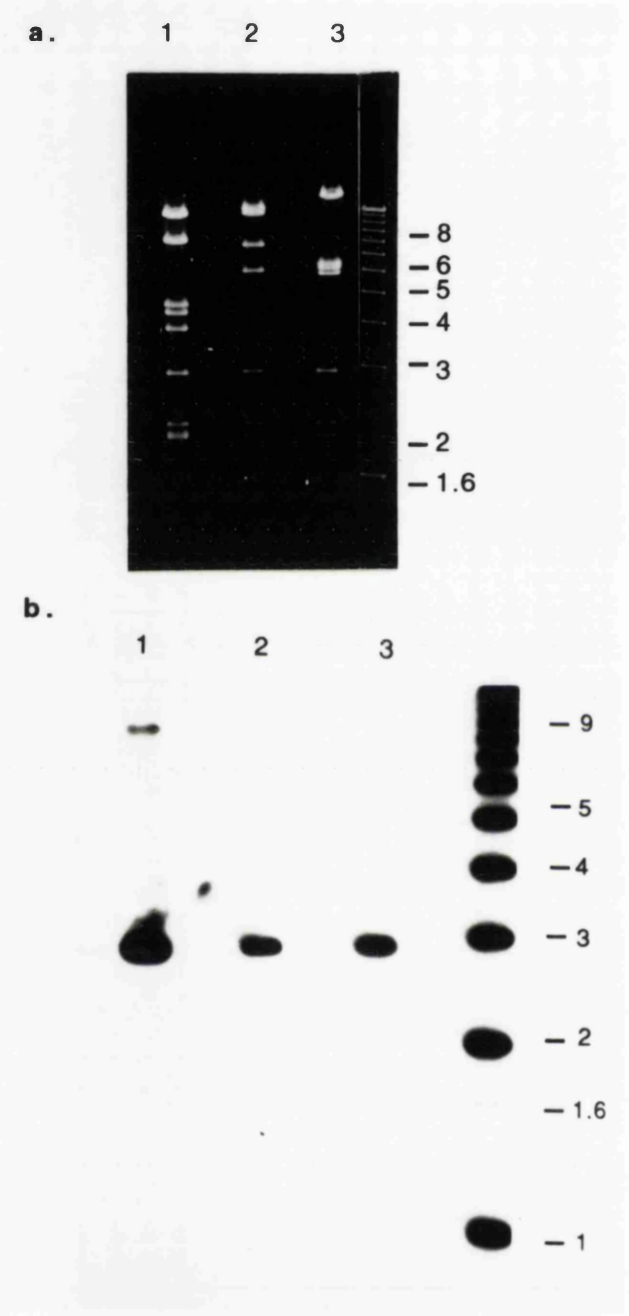
Figure 3.2

CAD Coding Sequences within Cosmids R61, C10 and L41

a) Each of the cosmids provided by M. Rolfe was digested with EcoR I and size-fractionated by agarose gel electrophoresis in the presence of ethidium bromide. The lanes are as follows: (1) cosmid R61; (2) cosmid C10; (3) cosmid L41.

(b) After depurination the DNA was transfered to a nylon membrane under alkaline conditions and probed with a CAD cDNA. The filter was stringently washed in 0.2x SSC/1% SDS at 65^{o}C for 30 minutes. Preflashed X-ray film was exposed for 15 hours at -70^{o}C between two intensifying screens.

The lane containing R61 DNA shows a fragment of 8 kb which hybridises to the CAD probe, indicating that coding sequences reside on this cosmid. Although the cosmid C10 was isolated using the same probe fragment, the washing conditions were less stringent and the coding sequences in C10 are not seen. The 8 kb band of hybridisation is unique to the cosmid R61 as the 8 kb band seen in the agarose gel is only seen in the R61 lane. Finally, no coding sequences are seen in the L41 lane although it may be possible to see hybridisation in blots washed at lower stringency. The low homology-coding sequences of C10 and the more homologous sequences in R61 indicate that a major coding region for CAD exists in the area of overlap between these two cosmids and this makes it less likely that coding sequences are present in cosmid L41.


a.
1
2
3
8
6
5
4
3
2
1.6
b.
1
2
3
9
5
4
3
2
1.6
1

probe and the genomic sequence of CAD could be substantial and this could give rise to the relatively poorly hybridising CAD band.

3.2.2

Initial Attempts at Homologous Recombination

A initial attempt to target an SV40 replication origin to the CAD locus was tried using two plasmids which were provided by M. Rolfe, each of which contained a sub-cloned fragment from the insert of either L41 or R61 as homology for targeted integration into the monkey CAD locus. The cloned fragments were terminal regions most distant from the overlap with C10, cloned into a pSV2neo-derived plasmid containing an SV40 replication origin and the dominant selectable gene for hygromycin B resistance. pCXba.L41 contained a 2.7 kb Sal I-EcoR I fragment from the extreme left hand end of L41 and was linearised within the insert by Xba I. pCHind.R61 contained a 2 kb EcoR I fragment from the extreme right hand end of R61 and was linearised within the insert by Hind III. Terminal fragments had been chosen by M. Rolfe as they were unlikely to contain CAD coding sequences which would be interrupted by homologous recombination.

Each plasmid was separately linearised with the respective restriction enzyme and transfected into ts-COS cells using the $CaPO_4$ precipitation technique. Forty eight hours after the glycerol shock, the cells were selected at the restrictive temperature with 200 µg/ml of hygromycin B. When colonies were visible, those on each plate were pooled and expanded. An aliquot of each pool was frozen and the remainder were plated at 1-5 x 10^5 cells per 100 mm dish and incubated either at the permissive or restrictive temperature with 150 µM PALA and hygromycin B. Medium was changed every 5 days for four weeks. No PALA-resistant colonies were observed at either the restrictive or permissive temperature for any of the pools of hygromycin resistant clones. The reason for this was not entirely due to the lack of homologous recombination and the experimental design is discussed in the following section.

3.2.3

Conclusion

The absence of PALA-resistant colonies in the pool of hygromycin B-resistant clones, following exposure to active T-ag, could be the result of several factors: (i) the amount of CAD homology used in the vectors may be too small to give a detectable frequency of homologous recombination; (ii) the PALA concentration used after exposure of the cells to active T-ag may have been too high; and (iii) excessive incubation of the cells at the permissive temperature, in the presence of an SV40 replication origin, could be fatal for the cells. A paper published at the time of this experiment (Thomas and Capecchi, 1987) demonstrated that, as the amount of homology increased so did the frequency of homologous recombination. I also investigated for myself the lowest concentration of PALA that would be required to remove the background of spontaneous PALA-resistant mutants, and found it to be 75 μM (the frequency of PALA resistance for ts-COS cells at this concentration is $<1 \times 10^{-6}$). With these points in mind the experimental design was improved in the following ways: (i) the amount of homology used in a homologous replacement fragment was increased to improve the chances of homologous recombination following transfection of the fragment; (ii) cells were selected in 75μM PALA; and (iii) cells were incubated at the permissive temperature for a maximum of 6 days.

3.3

Improved Design and Strategy

3.3.1

Construction of Homologous Replacement Vector λ16neo

To increase the frequency of homologous recombination, a larger amount of homology was used than before in the fragment that was to be modified by inclusion of a SV40 replication origin. To reduce the amount of non-homologous sequence in the modified fragment, all bacterial plasmid sequences were removed from the SV40 replication origin and neomycin resistance gene used. Thus a bacteriophage lambda

(λ) vector was chosen in which the fragment would be cloned and modified. Another advantage of using a λ vector was that large DNA fragments are more easily manipulated in such a system than they are in a bacterial plasmid. A 16 kb EcoR I fragment from L41 was chosen as homology with the CAD locus because: (i) it is large, increasing the frequency of homologous recombination, (ii) it does not contain coding sequences, and (iii) the fragment contains a unique BamH I site, useful for modification by insertion of an SV40 origin. This 16 kb EcoR I fragment was cloned into the λ vector EMBL3a to form λ16. The construction of the homologous replacement vector is shown in figure 3.3. For transfection, the λ arms were removed by digestion with EcoR I and the 18 kb EcoR I insert was purified by agarose gel electrophoresis.

3.3.2

Transfection of ts-COS Cells with the Homologous Replacement Fragment

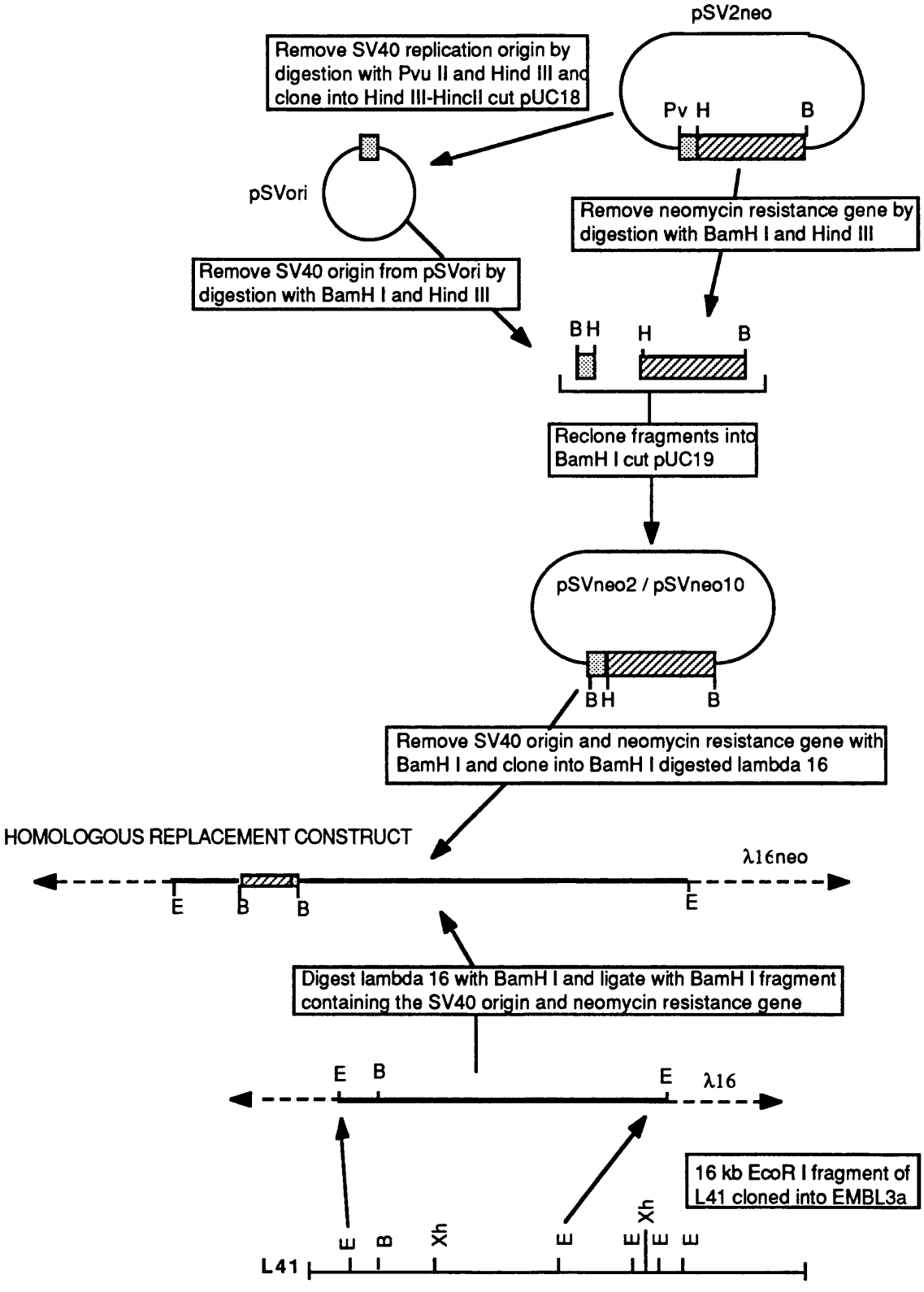
Electroporation was the transfection method of choice as it favours single copy integrations. The efficiency of stable transfection using electroporation with a BioRad GenePulser was determined using the pSVneo2 and pSVneo10 intermediate plasmids. This also served to check that the neo gene was functional in these constructs and also in ts-COS cells. Stable transfectants, following electroporation with linearised plasmids and selection with G418, were stained and counted, and the best conditions (250 μF, 450 V in PBS) were used to transfect ts-COS cells with the homologous replacement fragment.

Figure 3.3

Construction of the Homologous Replacement Fragment

A 16 kb EcoR I fragment from L41 was selected as the CAD homology for the homologous replacement fragment and was subcloned into the EMBL3a bacteriophage vector. An SV40 replication origin (stippled box) and a gene for G418 resistance (hatched box) were removed from the plasmid pSV2neo (Southern and Berg, 1982) and were cloned into the unique BamH I site of the cloned 16 kb EcoR I fragment of L41. As these sequences could not be removed directly from the plasmid, each was removed as shown in the diagram. The SV40 replication origin was removed by digestion with Hind III and Pvu II. Pvu II creates a blunt-ended DNA end which could be exploited during cloning as the vector was cut with Hinc II (as well as Hind III), which also creates a blunt DNA end. To remove the SV40 origin from the plasmid intermediate pSVori, the BamH I site in the vector polylinker was used to release Pvu II/Hinc II end of the insert; Hind III was used to release the other end. The BamH I-Hind III SV40 origin fragment was cloned, simultaneously with the Bam H I-Hind III neomycin resistance gene fragment, into the vector pUC19. The SV40 origin and neomycin resistance gene could be removed from this plasmid as a BamH I unit to facilitate modification of the subcloned 16 kb EcoR I fragment. To generate the homologous replacement fragment from the resulting construct, the λ arms were removed by EcoR I digestion and the 18 kb insert was gel-purified before transfection.

B=BamH I; E=EcoR I; H=Hind III; Pv=Pvu II; Xh=Xho I


pSV2neo
Remove SV40 replication origin by digestion with Pvu II and Hind III and clone into Hind III-HincII cut pUC18
Pv H
B
pSVori
Remove neomycin resistance gene by digestion with BamH I and Hind III
Remove SV40 origin from pSVori by digestion with BamH I and Hind III
B H
H
B
Reclone fragments into BamH I cut pUC19
pSVneo2 / pSVneo10
B H
B
Remove SV40 origin and neomycin resistance gene with BamH I and clone into BamH I digested lambda 16
HOMOLOGOUS REPLACEMENT CONSTRUCT
λ16neo
E
B
B
E
Digest lambda 16 with BamH I and ligate with BamH I fragment containing the SV40 origin and neomycin resistance gene
E
B
E
λ16
16 kb EcoR I fragment of L41 cloned into EMBL3a
E
B
Xh
E
Xh
E
E
E
L41

3.4

Analysis of G418-Resistant Pools

3.4.1

Screening by PALA Resistance

Following transfection of the homologous replacement fragment a large number of G418-resistant colonies would be obtained. It was not known what the ratio of random integrations of the transfected fragment to homologous replacement of the endogenous CAD locus sequences would be. A functional screening method was employed to determine if an increase in the frequency of PALA resistance could be observed in a pool of G418 resistant clones following deliberate exposure of the cells to active T-ag. If such an increase could be seen, then the number of founding clones in the pool would indicate an approximate frequency at which the homologous replacement fragment was integrating at the CAD locus.

Stable G418-resistant colonies from 7 independent transfections of the homologous replacement fragment were separately pooled and designated G418-resistant pools RNp1 to RNp7. To determine the frequency of PALA resistance before and after exposure to active T-ag, six 100 mm plates were seeded with aliquots of 10^5 cells each for each pool. Cells were allowed to attach for 24 hours at the restrictive temperature, after which 4 plates were moved to the permissive temperature. Two plates were maintained at the restrictive temperature for 24 hours before being selected with 75 μM PALA. After 3 days two of the plates at the permissive temperature were returned to the restrictive temperature, followed by the remaining two plates after 6 days of exposure to active T-ag. The temperature-shifted cells were selected at the same concentration of PALA 24 hours after their return to the restrictive temperature. Pools were maintained in G418 throughout. After 4 weeks the plates were fixed, stained and colonies were counted.

Results of the PALA selections are shown in table 3.1. The frequency of PALA resistance in ts-COS cells at 75 μM PALA was determined to be $<1 \times 10^{-6}$ and

Table 3.1

The Number of PALA-Resistant Colonies Obtained from G418-Resistant Pools Exposed to T-ag

One hundred thousand (10^5) cells were plated per 100 mm dish, in duplicate, and half the plates were moved to the permissive temperature after 24 hours to attach at the restrictive temperature. After a temperature shift of 3 or 6 days the cells were allowed to recover at the restrictive temperature for 24 hours before the culture medium was replaced with medium containing dFCS and 75 μM PALA. Unshifted cells were selected in the same concentration of PALA after 48 hours incubation at the restrictive temperature. The medium in each plate was changed every 5 days and the cells on each plate were fixed and stained after 4 weeks of selection. The table shows the number of PALA-resistant colonies scored per plate in each of the G418-resistant pools. The two columns beneath each length of exposure to active T-ag represent the duplicate plates, and the number of founding clones refer to the number of G418 -resistant transfectants which were pooled.

G418^r pool	Founding clones	Exposure to active T-ag 0 days		3 days		6 days	
RNp1	15	0	0	0	0	0	0
RNp2	15	0	0	0	0	0	0
RNp3	30	2	10	24	-	100	70
RNp4	50	1	1	1	6	70	74
RNp5	200+	11	0	13	8	23	10
RNp6	250+	10	0	51	39	100	79
RNp7	100+	nd		nd		nd	
control							
ts-COS	-	0	0	0	0	0	0

nd = not determined.

this frequency was not increased by the temperature shift and corresponding change in T-ag activity as the frequency of PALA resistance did not increase following exposure to active T-ag. The G418-resistant pools RNp1 and RNp2 showed no T-ag activity-dependent increase in PALA resistance, but the G418-resistant pools with more founding colonies (RNp3 to RNp6) did. This experiment shows that some clones in RNp3 to RNp6 appear to amplify the CAD gene in response to T-ag activity, and this allows survival and growth of these clones in the presence of PALA after incubation at the permissive temperature. The G418-resistant pool with the fewest founding clones that showed a T-ag activity-dependent increase in PALA-resistance frequency was RNp3 which had 30 founding clones. The G418-resistant pool RNp4 had 50 founding clones and showed a similar increase in the PALA-resistance frequency following exposure to T-ag activity, but gave fewer PALA-resistant colonies in the unshifted population. A tentative figure for homologous recombination was determined to be approximately one homologous replacement event per 50 integrations. As the G418-resistant pool RNp4 had the cleanest increase in PALA-resistance following exposure to active T-ag and also had relatively few founding clones, further analysis was restricted to this pool of transfectants.

3.4.2

CAD Copy Number Analysis of RNp4

To confirm that amplification of CAD sequences had occurred in the pools of G418-resistant cells that gave rise to PALA-resistant colonies after exposure to active T-ag, the experiment was modified using cells derived from the G418-resistant pool RNp4. After exposure to active T-ag for 0, 3 or 6 days, small aliquots of cells were removed and analysed by whole cell slot blots for the copy number of a CAD-specific sequence. Filters were hybridised with a probe specific to the CAD locus which flanks the target sequences but has no homology to the transfected fragment. The fragment used as a probe was a 2.3 kb Xba I-EcoR I fragment which was cut from the cosmid L41.This fragment is located immediately to the left of the target for

homologous recombination and is indicated on the restriction map of the CAD locus shown in figure 3.1. After exposure to X-ray film, the filters were rehybridised with a probe made by oligo-priming total genomic monkey DNA to normalise the signal between the slots. The autoradiograms are shown in figure 3.4. The relatively faint hybridising signal obtained for cells maintained at the restrictive temperature was caused by fewer cells being present on the plates used for lysates than were present on the shifted plates. The signals were quantified by densitometry and corrected for loading differences. The copy number of the 2.3 kb Xba I-EcoR I fragment increased 2-fold in the pool as a whole after 3 days exposure to active T-ag. This increased to 5-fold after 6 days exposure. However, it was not possible to determine the state of the amplified DNA because the slot blot method cannot differentiate between extra- and intra-chromosomal arrays. This experiment confirmed that concurrent with an increased frequency of PALA-resistance following exposure to active T-ag, sequences at the CAD locus were increased in copy number.

3.5

Analysis of G418-Resistant Cell Lines

3.5.1

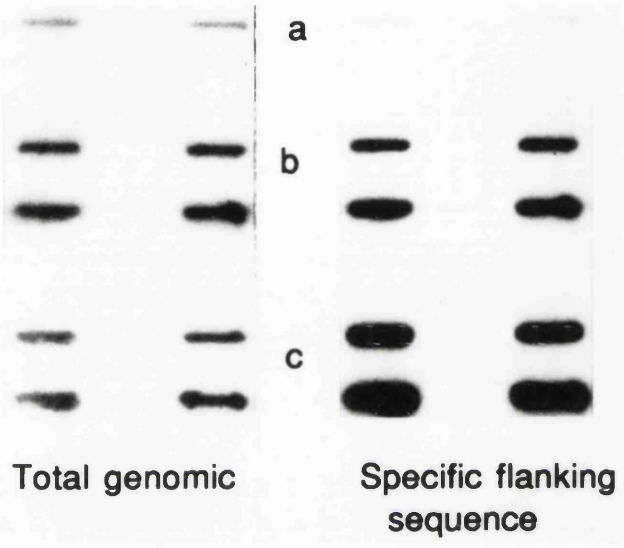
Whole Cell Slot Blot Screening of Lines Cloned from RNp4

A T-ag activity-dependent increase in PALA resistance was observed in the G418-resistant pool RNp4. This pool was founded by 50 G418-resistant transfectants and an approximate frequency of homologous recombination was thought to be 1 homologous recombinant cell line per 50 transfectants. With such an apparently high frequency of homologous recombination, it was decided that instead of sub-dividing the pool of transfectants and repeating the temperature shift and PALA selection, individual cell lines derived from the pool would be separately screened for homologous recombination of the transfected fragment. The cell lines would be screened using the whole cell slot blot technique of cells before and after

Figure 3.4

Whole Cell Slot Blots of RNp4 Cells

Aliquots of cells were lysed in 0.4 M NaOH and applied through duplicate slots on to a nylon filter. Slots were loaded as follows: (a) cells maintained at the restrictive temperature; (b) cells cultured at the permissive temperature for 3 days; and (c) cells cultured at the permissive temperature for 6 days. The filter was first hybridised with a 2.3 kb Xba I-EcoR I fragment from L41 which has no homology with the transfected fragment but is located immediately next to the target site of integration of the homologous replacement fragment. After hybridisation with the 2.3 kb Xba I-EcoR I fragment the filters were hybridised to a total genomic DNA probe to allow normalisation of the slots for loading differences.


a
b
c
Probe:
Total genomic
Specific flanking sequence

exposure to active T-ag. A potentially positive homologous recombinant cell line would show an increase in copy number of a CAD locus-specific probe.

An aliquot of cells from the G418-resistant pool of transfectants, RNp4, that had always been maintained at the restrictive temperature, was plated at low density and 96 individual colonies were ring-cloned. Fifty-six clones survived cloning and each was separately expanded. When sufficient cells had grown, an aliquot was frozen in liquid nitrogen and the remaining cells were plated into duplicate 24-well dishes. After 24 hours at the restrictive temperature, one dish was moved to the permissive temperature for six days and the other remained at the restrictive temperature until the cells reached semi-confluence. Clones at the permissive temperature were given 24 hours to recover at the restrictive temperature before harvesting. Genomic slot blots were made for each clone and the filters were probed with the 2.3 kb Xba I-EcoR I fragment that flanks the target site for the transfected homologous replacement fragment. Examples of the screening are shown in figure 3.5. Twenty-three out of fifty-six clones showed some form of increase in copy number of the probe sequence after exposure to active T-ag. This was a surprisingly high number of potentially positive homologous recombinant cell lines but over-representation of one particular clone within the G418-resistant pool could give such a result. Also, if homologous recombination occurs in more than one transfectant, the independent transfectants could not be told apart as the integration site of the transfected DNA would be the same in each clone.

3.5.2

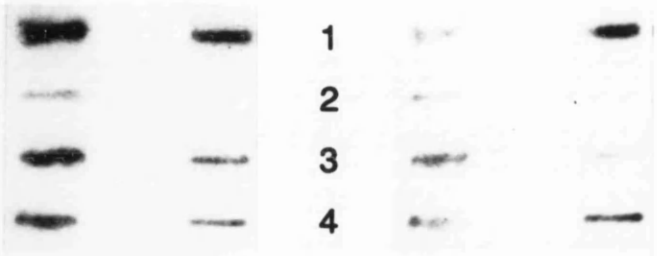
Nomenclature of Cell Lines

Cell lines bearing a RN suffix , eg. 3RN, are G418-resistant clones that were randomly picked from dishes following transfection of cells with the homologous replacement fragment and selection with G418. These cell lines were used as control cells as is was unlikely that any were homologous recombinant cell lines. Cell lines which bear a numerical prefix and suffix to a letter are clones that were deemed to be potential homologous recombinant cell lines on the basis of the screening

Figure 3.5

Representative Whole Cell Slot Blots of Cell Lines Cloned from RNp4

Aliquots of cells from each cell line isolated from the pool of G418-resistant transfectants, RNp4, were maintained at the restrictive temperature or were exposed to active T-ag for 6 days. The cells were lysed in 0.4 M NaOH and applied to a nylon membrane which was hybridised with the 2.3 kb Xba I-EcoR I fragment that contains no homology with, but immediately flanks, the target integration site of the transfected homologous replacement fragment. The filter was then hybridised with a total genomic monkey DNA probe in order to normalise the slots for loading differences. Slots were loaded as follows: rows 1-4 represent four different cell lines isolated from the pool of G418-resistant transfectants, RNp4; the cells in first column had been maintained at the restrictive temperature and those in the second column had been grown at the permissive temperature for 6 days. Clones 1 and 4 are potentially positive clones; in clone 1 the loading is approximately equal for shifted and unshifted cells, but the copy number of the CAD-specific sequence is increased in cells exposed to active T-ag; a similar situation is observed for clone 4. Clone 3 shows the same loading ratios for shifted and unshifted cells with both of the probes, indicating that the copy number of the 2.3 kb Xba I-EcoR I fragment has not increased after exposure to active T-ag in these cells.



Probe: Total genomic Specific flanking sequence

procedures performed upon them after isolation from the G418-resistant pool of transfectants, RNp4.

3.6

Further Analysis of Potential Homologous Recombinants

3.6.1

Southern Analysis

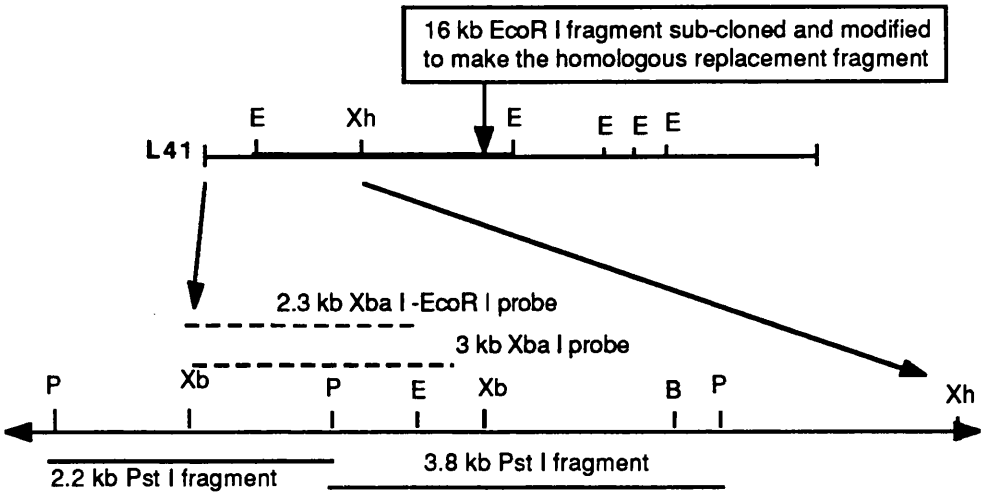
Several of the cell lines which had been isolated from the pool of G418-resistant transfectants, RNp4, were chosen for further analysis. To determine whether these cell lines were homologous recombinants, Southern blot analysis was performed on genomic DNA isolated from these cell lines which had been maintained at the restrictive temperature. The basis of the Southern blot analysis is shown in figure 3.6. The left hand end of the cosmid L41 was finely mapped by mixed restriction enzyme digests to produce the map shown in the figure. By comparison of the restriction pattern for the unaltered region with the restriction pattern predicted should homologous recombination occur, it can be seen that the pattern of Pst I sites in the region is changed by the presence of the neomycin resistance gene. If a Southern blot of Pst I-restricted genomic DNA from the potential homologous recombinant cell lines is probed with a restriction fragment that can hybridise to the Pst I fragments that are affected by the changes in the Pst I site pattern, then the nature of the integration site can be examined. For such an analysis the 3 kb Xba I fragment shown was chosen as a hybridisation probe. This fragment overlaps the EcoR I site that marks the left hand end of the sequence from which the homologous replacement fragment was derived. This 3 kb Xba I fragment therefore contains homology with the transfected fragment and can indicate, by hybridisation, all the integration sites of the transfected DNA, including homologous recombination events. As the Pst I sites of the homologous recombination target site and the flanking region have been mapped, the size of the Pst I fragment(s) to which the 3 kb Xba I fragment hybridises are indicative of the site where the homologous

Figure 3.6

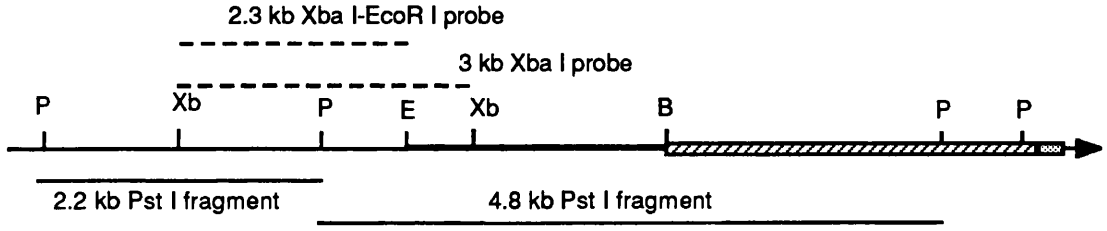
Predicted Pst I Restriction Patterns for the CAD Locus Before and After Homologous Recombination

The cosmid L41 is shown at the top of the diagram and the subcloned 16 kb EcoR I fragment, used in the construction of the homologous replacement fragment, is shown as a heavy line. Diagram (A) shows the predicted pattern of Pst I fragments for the left hand end of the endogenous, unaltered CAD region. The two probe fragments used on Southern blots are shown as broken lines above the map. Below the map, the Pst I fragments which are hybridised by these probes are shown. If either of the two probes were used on a Pst I-digested genomic Southern blot, the 2.2 kb and 3.8 kb Pst I fragments would be seen. Part (B) shows the predicted arrangement of Pst I sites following homologous integration of the homologous replacement fragment (shown as a heavy line). The integrated neomycin resistance gene sequences introduce additional Pst I sites to the region and this is reflected by the different size of one of the Pst I fragments which hybridise to the two probes shown. The size of the novel Pst I fragment is predictable from the anticipated restriction map. Part (C) shows the situation at a random integration site of the homologous replacement fragment. The 2.3 kb Xba I-EcoR I fragment has no homology with the replacement fragment and therefore random integrations cannot be observed with this probe. The 3 kb Xba I probe can hybridise to left hand end of the transfected DNA but the size of the hybridised Pst I fragment is unpredictable as it depends on the distance to the nearest Pst I fragment in the flanking DNA.

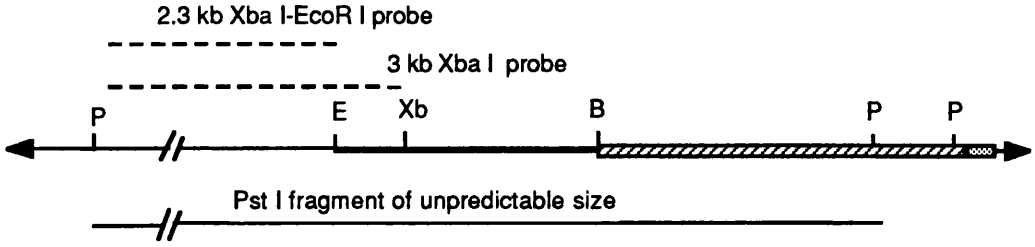
The stippled box = the SV40 replication origin; the hatched box = neomycin resistance gene; B=BamH I; E=EcoR I; P= Pst I; Xb=Xba I; Xh=Xho I



A. The predicted Pst I restriction pattern for an unaltered CAD locus



B. The predicted Pst I restriction pattern for a CAD locus in which the homologous replacement fragment has integrated by homologous recombination



C. The unpredictable Pst I restriction pattern for an random integration event of the homologous replacement fragment. No homology exists between the 2.3 kb Xba I-EcoR I fragment and the integrated DNA

replacement fragment has integrated. If the hybridisation pattern obtained with this probe comprises a novel band in addition to those generated by an unaltered target region (2.2 kb and 3.7 kb), then the size of such a novel band would be predictable (ie. 4.8 kb) if targeted integration had occurred at the other allele of the cells. The size of such a novel band of hybridisation would be unpredictable if the transfected DNA had randomly integrated into the genome.

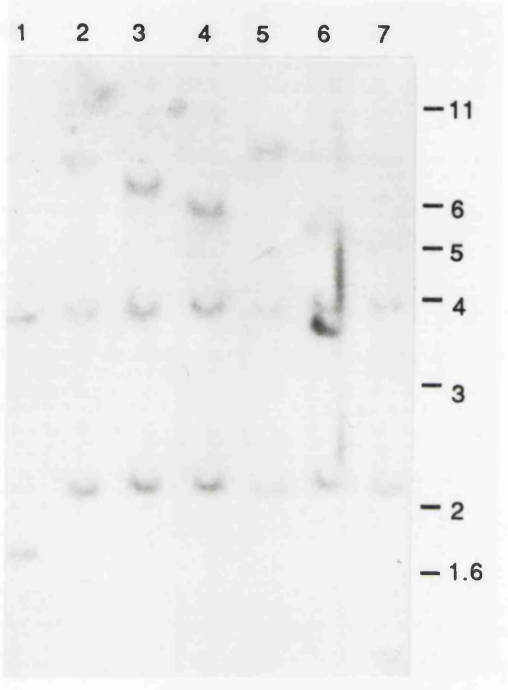
Genomic DNA was prepared from several clones that had shown T-ag-dependent amplification of the 2.3 kb Xba I-EcoR I fragment probe by slot blot analysis and also from several G418-resistant clones that had been chosen at random from another transfection experiment. The DNA was digested with Pst I, Southern blotted and hybridised with the 3 kb Xba I fragment which had been labeled with ^{32}P-dCTP. A 2.2 kb Pst I fragment was always seen. This is in agreement with the mapping data and represents the Pst I fragment to the left of the Pst I site present in probe fragment. This fragment is not affected by homologous recombination of the transfected homologous replacement fragment and will therefore be present in both homologous recombinants and random integrants. The autoradiograms shown in figures 3.7 and 3.8 also show a band of 3.7 kb - this band is the size expected from an endogenous, unmodified CAD locus as demonstrated by the pattern of hybridisation for ts-COS DNA (fig. 3.7, lane 7). It is assumed that, if homologous recombination occurred, only one target site would be modified and the wild type pattern would always be present in addition to any novel bands of hybridisation. An additional, novel band of a predictable size would only be seen following a successful targeting event. Random integration events would be characterised by a novel fragment of unpredictable size when probed with this 3 kb Xba I fragment. All the cell lines analysed in this manner showed at least one further band of hybridisation when probed, indicating that the homologous replacement fragment had integrated somewhere in the genome (figure 3.7). In several cell lines previously screened by slot blot analysis after isolation from the G418-resistant pool of transfectants, the additional band is 4.8 kb - in agreement with the predicted fragment size if homologous recombination had occurred (figure 3.8). One cell line (2B2, lane 3) has

Figure 3.7

Southern Blot of Transfectants Containing a Non-Homologous Integration of the Homologous Replacement Fragment

Genomic DNA was isolated from independent G418-resistant colonies and digested with Pst I. The fragments were fractionated by agarose gel electrophoresis and transferred to a nylon membrane. The filter was hybridised with the 3 kb Xba I fragment described in figure 3.6. Approximately 10 µg of genomic DNA was loaded per lane except for lane 1 where 50 pg of Pst I-digested L41 DNA was loaded with 10 µg of sheared herring sperm DNA. The lanes are loaded as follows: (1) cosmid L41, (2) clone 1RN, (3) clone 3RN, (4) clone 4RN, (5) clone 5RN, (6) clone 6RN, (7) ts-COS genomic DNA; sizes marked are in kilobases. The pattern of Pst I fragments which hybridise to this probe show that each cell line contains a single, random integration of the transfected homologous replacement DNA.

During cloning, the smaller Pst I fragment of L41 that hybridises to the 3 kb Xba I fragment was truncated when ligated into the cosmid polylinker and therefore runs faster than its genomic equivalent.


1
2
3
4
5
6
7
11
6
5
4
3
2
1.6

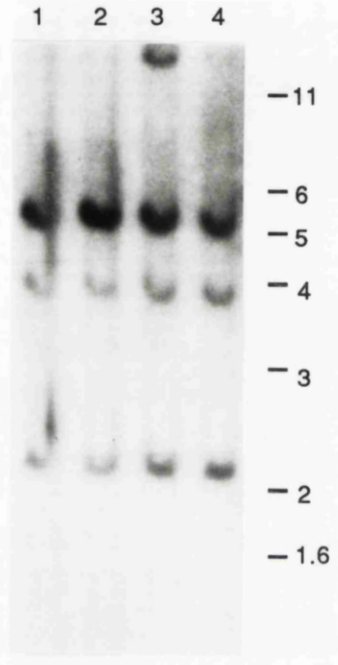
an additional band, indicating that a second integration site is present. In cell lines isolated at random (see figure 3.7), the third band is of variable size (3.5 kb to about 8 kb) between the cells in lanes 2-6, and in the cell line 6RN (lane 6) there are 2 extra bands of different sizes, suggesting that there are at least 2 integration sites of the transfected fragment in this cell line. In figure 3.8 the novel band seen in all of these cell lines was more intense than the other bands, and a possible cause of this was amplification of the sequences by the SV40 origin due to leakage of T-ag activity. This implies that during expansion of these cell lines amplification of the region has occurred. To test this possibility the earliest frozen aliquots of these clones were expanded taking great care to maintain the cells at the restrictive temperature. Genomic DNA was isolated and the Southern blots repeated. In this second attempt, the same pattern of hybridisation was seen but the bands were of similar intensity (figure 3.9a) thus showing that during expansion of these cell lines prior to the first Southern analysis, amplification of the sequences flanking the SV40 replication origin has occurred.

In conclusion, G418-resistant clones that had been chosen at random following transfection of the homologous replacement fragment, showed mostly single integrations at random locations. This was demonstrated by the variable and unpredictable size of an additional and novel fragment seen after hybridisation with the 3 kb Xba I fragment as a probe. Cell lines that had been isolated from the G418-resistant pool of transfectants all showed an additional and novel fragment which was the same size as that predicted for an accurately targeted homologous replacement fragment. During initial expansion of these cell lines the predicted fragment had been amplified but earlier freezings of the cells which had been minimally expanded did not appear to have amplified these sequences.

Figure 3.8

A Southern Blot of Potential Positive Homologous Recombinants Isolated from the Pooled G418-Resistant Transfectants in RNp4

Cell lines, isolated from the pool of G418-resistant transfectants, and which amplified the 2.3 KB Xba I-EcoR I fragment probe after exposure to active T-ag, were expanded and genomic DNA was prepared. Ten micrograms of DNA were digested with Pst I, fractionated by agarose gel electrophoresis and transferred to a nylon filter. The filter was hybridised with the 3 kb Xba I fragment probe from L41. This probe can detect all integrations, both targeted and random, of the homologous replacement fragment. The lanes are loaded as follows: (1) cell line 1D6, (2) cell line 2D1, (3) cell line 2B2 and (4) cell line 2B1; sizes marked are in kilobases. Cell line 2B2 (lane 3) appears to have more than one integrated copy of the transfected fragment as at least two bands are visible. All the other cell lines appear to have the same integration pattern which is consistent with an accurately targeted 18Rneo molecule. These cell lines could either be siblings or independent homologous recombinants.


1
2
3
4
11
6
5
4
3
2
1.6

3.6.2

Further Analysis of the Integration site of the Homologous Replacement Fragment in the 1D6 Cell Line

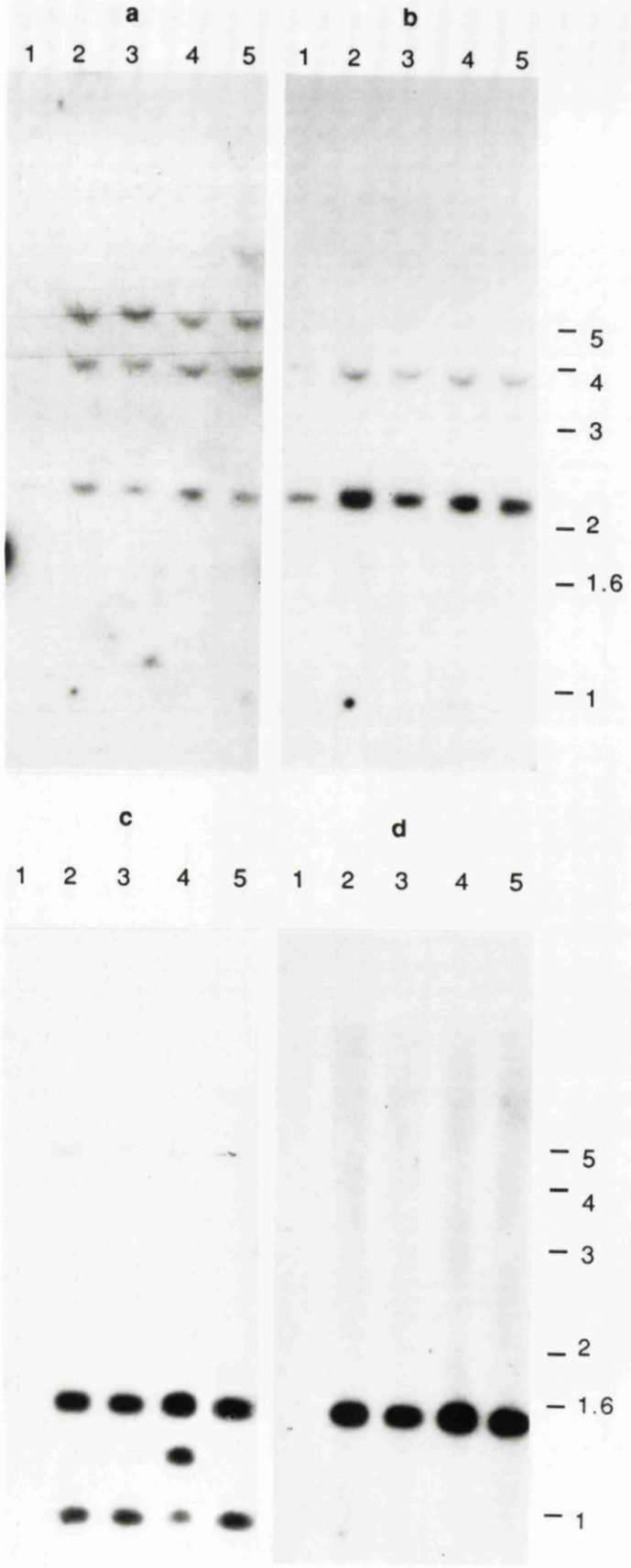
From Southern blot analysis of several G418-resistant cell lines isolated from the pool of transfectants, RNp4, it was shown that integration of the homologous replacement fragment produced a hybridisation pattern which was consistent with a modified CAD locus. The analysis also showed that the integration site in these cell lines was susceptible to amplification during growth at the restrictive temperature. A possible cause of this amplification could be leakage of T-ag activity at the restrictive temperature and its interaction with the SV40 replication origin. As the SV40 replication origin would be involved in such events, these sequences were examined in the potential homologous recombinants shown in figure 3.9a.

The Southern blot filter used in figure 3.9a was stripped of the hybridised 3 kb Xba I probe and a probe comprising the SV40 replication origin and the neomycin resistance gene was hybridised in its place. The result is shown in figure 3.9c. The autoradiogram shows three bands which hybridise to the probe. The 1.6 and 0.9 kb bands agree with restriction map data as being fragments of the sequences cloned into λ16 to make λ16neo. The 4.8 kb band previously seen with the 3 kb Xba I fragment probe is also seen on this blot, indicating that this fragment contains both neo sequences and sequences in common with the probe fragment. However, the relative intensities of the 4.8 kb band and the neomycin resistance gene indicate that the copy number of the SV40 origin and neo sequences are much higher than the genomic sequences. This is not due to a very small amount of overlap between the neo/SV40 origin probe and the novel 4.8 kb fragment, as the restriction mapping data shows a 1.5 kb overlap between the two sequences. It is likely that the neo/SV40 origin sequences have been amplified and may reside on excised DNA in the form of episomes or are integrated amplified sequences. The filter was stripped once again and was hybridised to a probe comprising the SV40 origin only. This probe hybridised to the 1.6 kb band previously seen with the SV40 origin/neomycin resistance gene probe. This information implies that the 4.8 kb Pst I fragment that

Figure 3.9

Hybridisation of Various Probes to a Genomic Southern Blot of Potential Homologous Recombinants

The potential homologous recombinants shown in figure 3.8 appear to have amplified the novel fragment hybridising to the 3 kb Xba I fragment probe. A frozen aliquot of each cell line, which had been minimally grown before freezing, was thawed and expanded with care to maintain the cells at the restrictive temperature. Genomic DNA was prepared from each cell line and approximately 10 μg was digested with Pst I and Southern blotted. The nylon filter was hybridised with the following probes: panel (a) 3 kb Xba I fragment from L41; panel (b) 2.3 kb Xba I-EcoR I fragment from L41; panel (c) mixed probe containing SV40 origin and neomycin resistance gene sequences; panel (d) SV40 replication origin. Lanes are loaded as follows: (1) ts-COS genomic DNA, (2) cell line 1D6, (3) cell line 2D1, (4) cell line 2B2, and (5) cell line 2B1.


a
b
1 2 3 4 5
1 2 3 4 5
5
4
3
2
1.6
1
c
d
1 2 3 4 5
1 2 3 4 5
5
4
3
2
1.6
1

hybridised to both the 3 kb Xba I fragment and to the SV40 origin /neomycin resistance gene probe contains only neomycin resistance gene sequences and sequences in common with the 3 kb Xba I fragment.

The blot shown in figure 3.9c also shows an extra band of hybridisation in the cell line 2B2 (lane 4). Hybridisation with a probe for the SV40 origin alone showed this band to comprise neo sequences only (as shown by their absence in lane 4 of figure 3.9d). The significance of this is unclear, but it may be linked to a second integration of all or perhaps a part of the homologous replacement fragment, that is carried by this cell line (see figure 3.8). This was not investigated further as subsequent analyses were directed at cell line 1D6. This cell line was selected as it appeared to contain a single integration of the transfected fragment but any other potentially positive homologous recombinant cell line carrying a single integration could have been used in further experiments.

3.6.3

Analysis of Hirt DNA Extracted from Cell Line 1D6

The Southern blot analysis of potential homologous recombinant cell lines described above suggested that the neomycin resistance gene and SV40 origin sequences may have been amplified through leakage of T-ag activity at the restrictive temperature. The sequences amplified did not appear to include those that hybridise to the 3 kb Xba I fragment used as a probe, although previously these sequences had been seen to be amplified. Analyses of integrated SV40 sequences including the replication origin have been carried out in other laboratories (including work by Gurney and Gurney, 1989; Conrad, Liu and Botchan, 1982; Botchan, Topp and Sambrook, 1979; Botchan *et al.*, 1980) and it has been shown that excision and amplification of SV40 origin sequences and the surrounding DNA can occur, giving rise to episomal DNA. The only viral component required for excision was found to be a SV40 replication origin and active T-ag (Conrad, Liu and Botchan, 1982). Intrachromosomal events have also been described in some of these reports.

To determine if excision and amplification of episomal DNA, mediated by active T-ag and the SV40 replication origin, were occurring in the potentially positive homologous recombinants, the low molecular weight DNA (<20 kb) extraction technique of Hirt (Hirt, 1967) was applied to these cells before and after deliberate exposure to active T-ag.

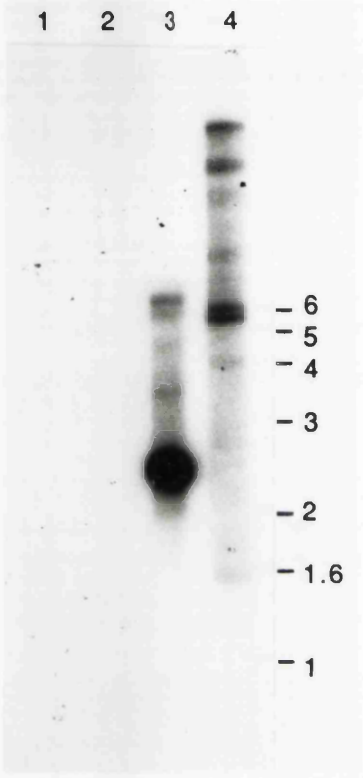
Cells from one of the potential homologous recombinant cell lines, 1D6, were plated in duplicate 100 mm dishes and incubated at the restrictive temperature for 24 hours. Hirt DNA was prepared from the plates either after the initial 24 hour incubation or after a shift to the permissive temperature for two days followed by a 24 hour recovery period at the restrictive temperature. Twenty microlitres of each Hirt DNA extract was fractionated by agarose gel electrophoresis either undigested or after digestion with the restriction enzyme Stu I, and transferred to a nylon membrane (Zetaprobe). Stu I cleaves the SV40 replication origin at a single site and no Stu I sites exist in the neomycin resistance gene. Therefore digestion with this enzyme should linearise each episomal molecule unless sites exist in flanking DNA which has been included in the episome during the excision event. The membrane was hybridised with an SV40 origin probe and washed under stringent conditions to determine if episomes were produced when the SV40 replication origin was active. An SV40 origin probe was used, as any episome capable of replicating would need these sequences to do so, and the SV40 origin is also likely to be a component of any excision event that occurs as a result of over-replication. The results are shown in figure 3.10.

The autoradiogram shows that episomes are produced when the cells are incubated at the permissive temperature. The uncut Hirt DNA runs as several bands which may represent the different forms of circular DNA, ie. supercoiled, negatively supercoiled or relaxed, or may represent several species of episome. When the Hirt DNA was digested with Stu I and then gel-fractionated the episomes were resolved into a single major band of 2.5 kb. This suggests that one species of episome prevails in these cells. The filter was stripped and rehybridised with the 2.3 kb Xba I-EcoR I fragment from L41 which has no homology with the homologous replacement

Figure 3.10

Southern blot of Hirt DNA Extracted from 1D6 Cells

Hirt DNA was extracted from 1D6 cells which had either been maintained at the restrictive temperature, or had been exposed to active T-ag for 24 hours. Approximately 20 µl (out of 100 µl) was size-fractionated by agarose gel electrophoresis, either undigested or after digestion with Stu I (this enzyme cleaves the SV40 replication origin at residue 5193 but has no sites in the neomycin resistance gene, including the SV40 splice and termination signals). The gel was Southern blotted and the nylon filter was hybridised to a probe specific for the SV40 replication origin. The lanes are loaded as follows: (1) undigested Hirt extract from cells maintained at the restrictive temperature, (2) Stu I digested Hirt extract from cells maintained at the restrictive temperature, (3) Stu I digested Hirt extract from cells shifted to the permissive temperature for 24 hours, and (4) undigested Hirt extract from cells shifted to the permissive temperature for 24 hours. ts-COS cells treated in the same manner do not produce detectable episomes at the permissive temperature (data not shown). Size markers are in kilobases.


1
2
3
4
6
5
4
3
2
1.6
1

fragment but is located immediately to the left of the transfected fragment's integration target. No hybridisation signal was observed, implying that the episome contained no CAD-related sequences. This result could explain why the SV40 origin and neomycin resistance gene sequences appear to be amplified in the Southern blots described in figures 3.9c and 3.9d. It is possible that some of the SV40 origin and neomycin resistance gene sequences could be amplified intra-chromosomally. Episomal DNA is not detectable in the cells maintained at the restrictive temperature (lanes 1 and 2, figure 3.10) and therefore the amplified neomycin resistance gene and SV40 origin sequences could be integrated into the chromosomal DNA.

3.6.4

Stability of Amplified Material in 1D6 Cells

The previous section describes the production of an episomal DNA element in the potential homologous recombinant cell line 1D6 when these cells are incubated at the permissive temperature for T-ag activity. Southern blot analysis of this cell line showed that the neomycin resistance gene, SV40 replication origin and sequences that hybridised to the 3 kb Xba I fragment could be amplified, although the cells were always maintained at the restrictive temperature. Episomal DNA could not be detected in cells when they were maintained at this temperature and therefore it was possible that the amplified DNA sequences were integrated into the genomic DNA. The persistence of episomal DNA elements at the restrictive temperature, after deliberate exposure of the cells to active T-ag was assessed.

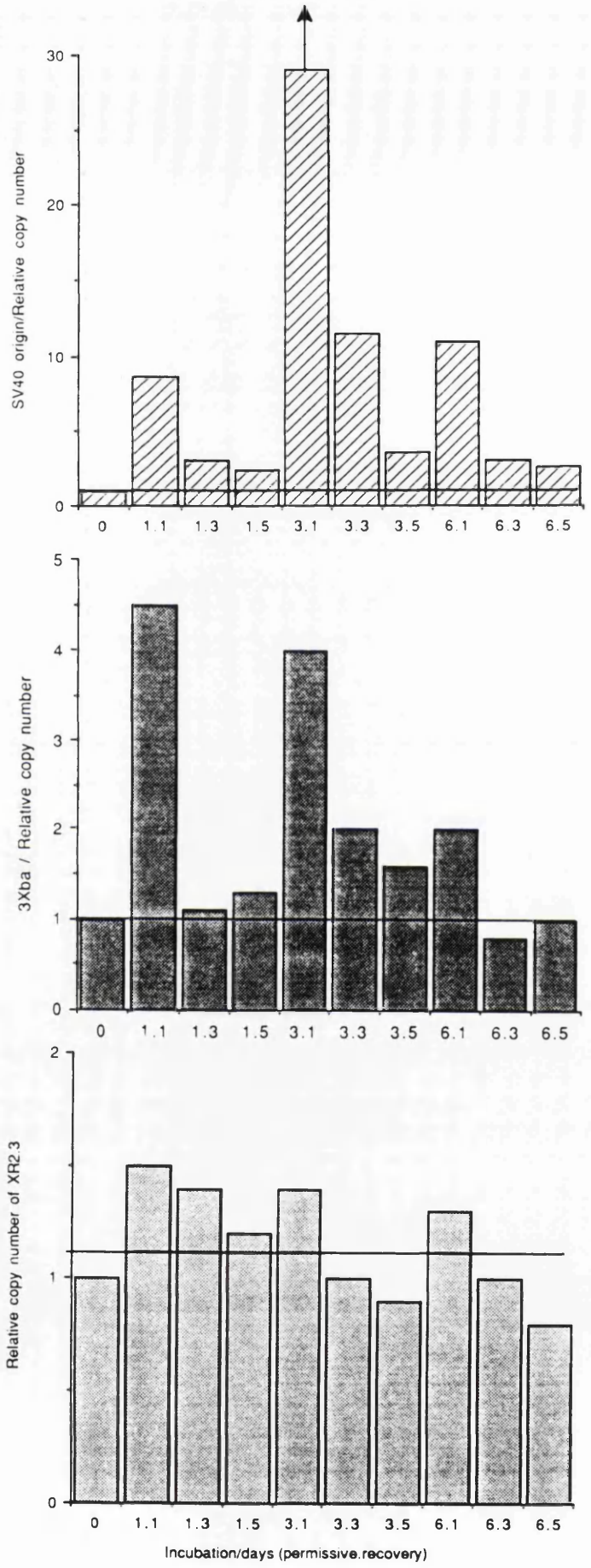
Cells from the potential homologous recombinant cell line,1D6, were seeded onto 100 mm dishes and were allowed to attach at the restrictive temperature for 24 hours. Several dishes were moved to the permissive temperature for 1, 3 or 6 days, and those remaining were maintained at the restrictive temperature for several days but were not allowed to achieve confluence. After each time point, cells were returned to the restrictive temperature and allowed to recover for 1, 3 or 5 days. After the recovery period, the cells were removed from the plate, counted and either replated at 5×10^5 cells per 100 mm dish and selected with 75 μM PALA, or washed

with PBS and frozen at -20^{o}C for whole cell slot blot analysis. When aliquots for all time points were ready, whole cell slot blots were made. The filter was hybridised with a series of probes, stripping the filter between each hybridisation. Probes used were: (i) the 3 kb Xba I fragment of L41, (ii) the 2.3 kb Xba I-EcoR I fragment of L41, and (iii) an SV40 origin probe. The 3 kb Xba I fragment would hybridise to sequences derived from the a random integration or homologous recombination event involving the transfected fragment and the Xba I-EcoR I fragment would only hybridise to sequences derived from the endogenous CAD locus. These sequences would be present in episomal DNA only if excision was occurring at the CAD locus, ie. following homologous recombination. The final hybridisation was with a total monkey genomic probe which was used as a loading control for densitometric analysis. The results of the densitometric analysis are shown in figure 3.11. It is apparent that after a single day at the permissive temperature and a single day's recovery, amplification of sequences including the SV40 replication origin has occurred. This amplification reaches a maximum at 3 days before falling to a level equivalent to 1 day's amplification after 6 days. This shows that the amplified sequences which hybridise to the 3 kb Xba I fragment of L41 and the SV40 origin probe are quickly lost from the cells following the removal of T-ag activity - the degree of amplification is halved after 3 days of recovery and is nearly normal after five days recovery at the restrictive temperature. Amplification of sequences with homology to the 3 kb Xba I fragment also occurs. Amplification is greatest - about 4.5-fold - during the first day at the permissive temperature, and returns to pre-amplified levels 6 days after exposure to active T-ag. As with the SV40 origin hybridising sequences, the sequences that hybridise to the 3 kb Xba I fragment are rapidly lost at the restrictive temperature. The 2.3 kb Xba I-EcoR I fragment, which has no homology with the transfected fragment, does not appear to hybridise to any amplified material. Slight increases in signal intensity are seen after exposure to T-ag, but this is most likely due to non-specific hybridisation between the probe and the amplified material as the signal increases and decreases in the same manner as SV40 origin-containing episomes.

Figure 3.11

Copy Number of Different Sequences In 1D6 Cells After Exposure to Active T-ag

Cells from the potential homologous recombinant cell line, 1D6, were plated at the restrictive temperature and after 24 hours several plates were moved to the permissive temperature. Plates were returned to the restrictive temperature after either 1, 3 or 6 days exposure to active T-ag and aliquots were removed for analysis 1, 3 or 5 days later. The cells from each aliquot were washed in PBS, counted and stored as cell pellets at -20^{o}C until all time points had been taken. DNA from these cells were analysed by whole cell slot blot and the nylon filters were hybridised in turn with: (i) the 3 kb Xba I fragment (middle graph), (ii) the 2.3 kb Xba I-EcoR I fragment (lower graph), and (iii) an SV40 origin probe (upper graph). The filter was stripped between each hybridisation and the signal for each duplicate slot was quantified by densitometry and plotted against the length of exposure to active T-ag and recovery time. The arrow on time point 3.1 for the SV40 origin copy number, indicates that the intensity of the signal was beyond the linearity of the film and that the real relative copy number is significantly higher than a 30 -fold increase.


SV40 origin/Relative copy number
30
20
10
0
0 1.1 1.3 1.5 3.1 3.3 3.5 6.1 6.3 6.5
3Xba / Relative copy number
5
4
3
2
1
0
0 1.1 1.3 1.5 3.1 3.3 3.5 6.1 6.3 6.5
Relative copy number of XR2.3
2
1
0
0 1.1 1.3 1.5 3.1 3.3 3.5 6.1 6.3 6.5
Incubation/days (permissive.recovery)

The aliquot of cells which were removed at each time point and cultured in PALA showed no PALA-resistant colonies after 4-5 weeks of selection at the restrictive temperature, and on staining all plates were free of attached cells. This is in agreement with the slot blot analysis of the same cells described above and implies that the CAD locus is not amplified at the permissive temperature in these cells.

3.6.5

PALA Sensitivity of 1D6 Cells

If the potential homologous recombinant cell line 1D6 did have an SV40 replication origin accurately targeted to the CAD locus, as suggested by the Southern blot analysis, then it was curious as to why incubation at the permissive temperature did not increase the frequency of PALA resistance, as would be expected. It was possible that the amplified CAD sequences could not be maintained long enough for resistant colonies to be identified. To address this possibility, the PALA sensitivity of these cells was checked before and after exposure to active T-ag by measurement of growth inhibition of the cells in the presence of PALA. This experiment would determine if any amplified CAD sequences which were produced at the permissive temperature, could render the cells less sensitive to the drug. The experiment was started before the results described in the previous section were complete, and therefore the lack of amplification of sequences specific to the CAD locus was not known.

Cells of the potential homologous recombinant cell line 1D6 and ts-COS were plated at 1.5×10^5 cells per well in 2 sets of duplicate 24-well dishes. After incubation for 24 hours at the restrictive temperature, 2 plates were moved to the permissive temperature for 48 hours followed by a 24 hour recovery period at the restrictive temperature. The other 2 plates were maintained at the restrictive temperature. The cells were then exposed to PALA concentrations of 0, 50, 100 or 200 µM for 72 hours. Cells were washed twice with PBS, lysed in 1 M NaOH and heated to 50°C for 30 minutes. One tenth dilutions were made of each lysate and the OD_{260} was measured as an indication of the number of cells that were still

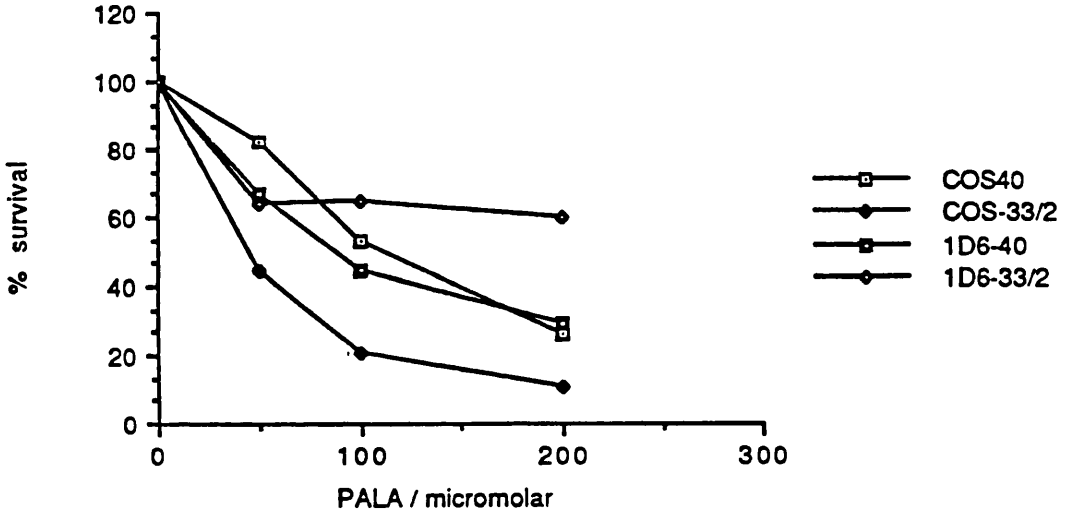
attached to the dish. Percentage survival was calculated as the ratio between the number of attached cells in the absence of drug and those in increasing concentrations of PALA (0 μM PALA = 100% survival). The plot of survival against PALA concentration is shown in figure 3.12. Cells maintained at the restrictive temperature gave better survival curves as the plating efficiency and metabolic rate of the cells are both higher than for cells incubated at the permissive temperature. It can be seen that the slope gradients for ts-COS cells from both shifted and unshifted populations and for unshifted 1D6 cells are similar. This implies that growth inhibition by PALA is similar in these cells. 1D6 cells which have been grown at the permissive temperature have a much shallower curve, suggesting that PALA does not inhibit their growth to the same degree. After an initial inhibition at the lowest PALA concentration, the cells survive at the same rate in increasing PALA concentrations. These results imply that active T-ag can decrease the inhibitory effects of PALA in the cell line 1D6 but does not have a corresponding effect on the PALA sensitivity of ts-COS cells. However, this method for assessing the the inhibitory effect of PALA is flawed in the following way. The optical density of each lysate is dependent on the amount of DNA present. In cells which do not produce episomal DNA this is dependent on the number of cells which are lysed. For the 1D6 cells, which have been shown to produce episomal DNA at the permissive temperature, the optical density of the lysate does not reflect the number of cells lysed as episomal DNA will contribute to the overall DNA content. Because of this factor, the apparent increased number of cells not inhibited by PALA is misleading as the increase in optical density reflects the higher amount of DNA in each cell and not increased cell numbers. Thus it is probable that cells from this potential homologous recombinant are equally sensitive to PALA both before and after exposure to active T-ag.

The lack of PALA resistant colonies and the lack of CAD locus-specific amplification shown in the previous section and the constant sensitivity of the cells to PALA before and after exposure to active T-ag strongly suggest that the cell line 1D6 does not contain an SV40 replication origin which has been accurately targeted to the CAD locus by homologous recombination.

Figure 3.12

Sensitivity of 1D6 Cells to PALA

Triplicate samples of cells were plated in different PALA concentrations following exposure to T-ag activity, or having been maintained at the restrictive temperature. After 72 hours the cells were lysed in NaOH and the OD_{260} was measured for each lysate. The average of the three readings was converted into survival, where the OD of cells not placed in PALA was taken as 100% survival, and plotted against the concentration of PALA used. The graph shows that the cell line 1D6 appears to be less sensitive to growth inhibition by PALA after exposure to active T-ag. However, as these cells have been shown to produce episomal DNA at the permissive temperature, this effect may be the result of increased DNA per cell rather than more cells being present on the dishes.


120
100
80
60
40
20
0
% survival
0
100
200
300
PALA / micromolar
COS40
COS-33/2
1D6-40
1D6-33/2

3.6.6

PALA Selection of 1D6 Cells

To determine in a controlled way, whether the potential homologous recombinant cell line 1D6 could give rise to PALA-resistant colonies at an increased frequency following exposure to active T-ag, the experiment described below was performed. The control cell line 3RN had previously been shown to contain a single, randomly integrated homologous replacement fragment by Southern blot analysis. This cell line was included in the experiment as it would not be expected to show an increased frequency of PALA resistance following incubation at the permissive temperature. As a positive control, an aliquot of the G418-resistant pool of transfectants RNp4 was also included as these cells had previously shown an increase in the frequency of PALA resistance following T-ag activity. This pool of transfectants was also selected as it was the source of the cell line 1D6.

Populations of 1D6, 3RN and RNp4 cells were separately plated into duplicate 100 mm dishes and incubated at the restrictive temperature for 24 hours. One set was moved to the permissive temperature for 3 or 6 days followed by a 24 hour recovery period at the restrictive temperature, and the duplicate set was maintained at the restrictive temperature as a control. Five aliquots of 5×10^5 cells each were taken from the shifted and control populations and separately plated in 75 µM PALA at the restrictive temperature. For cell lines 1D6 and 3RN, no colonies grew within the 4 weeks of selection before the plates were stained. Examination of stained plates under a microscope showed that no cells remained attached to the surface of the dishes. PALA-resistant colonies were obtained from RNp4 cells that had been shifted to the permissive temperature. PALA-resistant colonies also grew in control plates of RNp4 cells, but these were fewer than had been observed before (see table 3.1). The whole experiment was repeated with the same result. The lack of PALA resistant colonies from the shifted population of the potential homologous recombinant 1D6 confirms the results of previous experiments and again, strongly suggests that this cell line is not a homologous recombinant. Another possibility is that this cell line does contain an accurately targeted integration of the transfected

fragment, but that the arrangement of the DNA does not allow amplification of the CAD gene following exposure of the cells to active T-ag. The following section seeks to determine the stability of the integrated SV40 origin sequences in this cell line.

3.6.7

Analysis of SV40 Origin Sequences in Cell Line 1D6

The arrangement of SV40 sequences in the now dubious homologous recombinant cell line 1D6 was investigated to determine whether any amplification or rearrangement of these sequences had occurred. Data from Gurney and Gurney (1989) showed that such events can occur and that the frequency of such events was variable, possibly depending on the local environment of the integrated SV40 replication origin.

Genomic DNA isolated from 1D6 cells could be digested with a restriction enzyme that cleaves the SV40 replication origin at a single site. The pattern of restriction fragments which could hybridise to an SV40 origin probe would show the complexity, and possibly the copy number, of integrated SV40 origin sequences in this cell line.

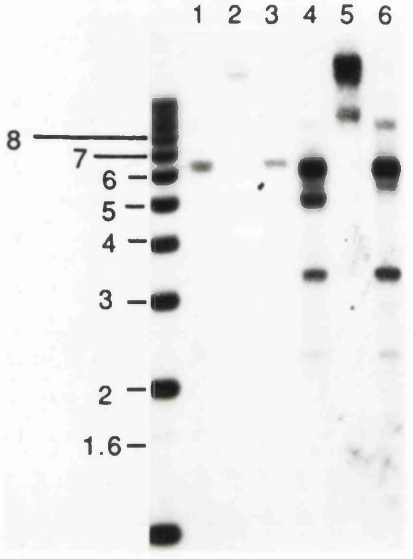
The restriction enzyme Stu I cleaves the SV40 origin at base 5190 but does not cut the flanking neo sequences. The restriction sites of this enzyme were mapped in the cosmid L41, and genomic DNA from ts-COS, 1D6 and 3RN cells was prepared and restricted with either EcoR I or Stu I, or with both of these enzymes. The DNA fragments were fractionated by agarose gel electrophoresis and transferred to a nylon membrane. The membrane was hybridised to a SV40 origin probe and washed at high stringency (figure 3.13). The autoradiogram shows that the random integrant 3RN has a simple, non-rearranged and non-amplified integration of the homologous replacement fragment. DNA from the now dubious homologous recombinant cell line 1D6 shows a more complex pattern of SV40 origin sequences and previous Southern blot analysis had shown that this cell line contained integrated material at a single site within the cell. Material at this site had been shown to be amplified during expansion of the cell line at the restrictive temperature. The simple hybridisation

Figure 3.13

The Arrangement of SV40 origin Sequences in the Cell Line 1D6

Genomic DNA was prepared from the potential homologous recombinant cell line 1D6 and from the cell line 3RN, which had previously been shown to have a single, random integration of the homologous replacement fragment.The DNA was digested with either EcoR I or Stu I or simultaneously with both enzymes.The restricted DNA was Southern blotted and hybridised to an SV40 replication origin probe. The enzyme Stu I cleaves the SV40 origin at a single site but does not cleave the neomycin resistance gene. EcoR I should cut out the transfected DNA within a genomic fragment which is larger than the transfected DNA. Stu I digestion of the transfected fragment should produce a 6.5 kb fragment which is not affected by codigestion with EcoR I. Digests with both enzymes can show whether the sequences surrounding the origin have been rearranged and/or amplified. If the Stu I fragment is affected by EcoR I digestion, then rearrangements of sequences surrounding the origin have placed an EcoR I site within the Stu I fragment, where no EcoR I site previously existed.

The lanes are loaded as follows: (1): 3RN DNA cut with EcoR I and Stu I; (2) 3RN DNA cut with EcoR I; (3) 3RN DNA cut with Stu I; (4) 1D6 DNA cut with EcoR I and Stu I; (5) 1D6 DNA cut with EcoR I; (6) 1D6 DNA cut with Stu I.


1 2 3 4 5 6
8
7
6
5
4
3
2
1.6

pattern seen for the control cell line 3RN , confirms that it has a single integrated SV40 origin and that no rearrangements have occurred during the expansion of this clone. The complex pattern of Stu I fragments and the two novel EcoR I bands, in addition to the transfected fragment, which hybridise to the SV40 origin probe in the 1D6 cell line may be the result of either secondary integrations of the transfected fragment (or fragments of it) or of intra-chromosomal rearrangements of integrated material following amplification. The 6.5 kb Stu I fragment seen in both 3RN and 1D6 DNA is the predicted size for an unrearranged homologous replacement fragment. Its size is not affected by EcoR I digestion (see the double digest lane of 3RN DNA). When 1D6 DNA was digested with both enzymes, a novel Stu I/EcoR I fragment is seen. This is consistent with the integration of an episome containing an SV40 origin or the rearrangement of an integrated homologous replacement fragment. It is not possible to determine, from this Southern blot, whether the rearranged DNA is present at the original site of integration or at a secondary integration site. The pattern observed raises the possibility that 1D6 cells may be intrinsically unstable over time with respect to integrated SV40 origin sequences when compared to other cell lines containing different integration events.

3.6.8

Sensitivity of 1D6 Cells to Leakage of T-ag Activity

Southern analysis of genomic DNA from the cell line 1D6 suggested that the SV40 sequences present at the original integration site of the transfected fragment could be amplified and rearranged during growth at the restrictive temperature for T-ag activity. It was not clear whether the events were mediated by leakage of T-ag activity or through instability of the sequences at the integration site. An experiment was designed to determine if the integrated SV40 sequences in cell line 1D6 were stable at the restrictive temperature. 1D6 cells which had previously shown no CAD sequence amplification by Southern blot analysis (figure 3.9a) were maintained at the in culture for several weeks at the restrictive temperature. At the same time cells from the cell line 4RN were grown in parallel. These cells have previously been shown to

contain a single, random integration of the transfected homologous replacement fragment and no amplification of sequences that hybridised to the 3 kb Xba I fragment of L41 was observed.

Genomic DNA was isolated from 1D6 and 4RN cells that had been in culture at the restrictive temperature for several weeks. The DNA was digested with Pst I, blotted and probed with the 3 kb Xba I fragment of L41; the result is shown in figure 3.14. An additional, novel band not seen in ts-COS DNA, is amplified in 1D6 cells but is of nearly normal intensity in DNA isolated from 4RN cells. Southern blots of Pst I-digested genomic DNA isolated from each cell line at an earlier date confirmed that, at that stage of expansion, both cell lines had a normal or low copy number of the novel fragment. An amplified novel fragment had been seen previously in 1D6 cells (figure 3.8), and a similar analysis of 1D6 cells which had been minimally expanded showed a normal or low copy number of this fragment (figure 3.9a). These results further suggest that the integrated SV40 origin in 1D6 cells is relatively unstable at the restrictive temperature when compared to the same sequences in other cell lines.

3.6.9

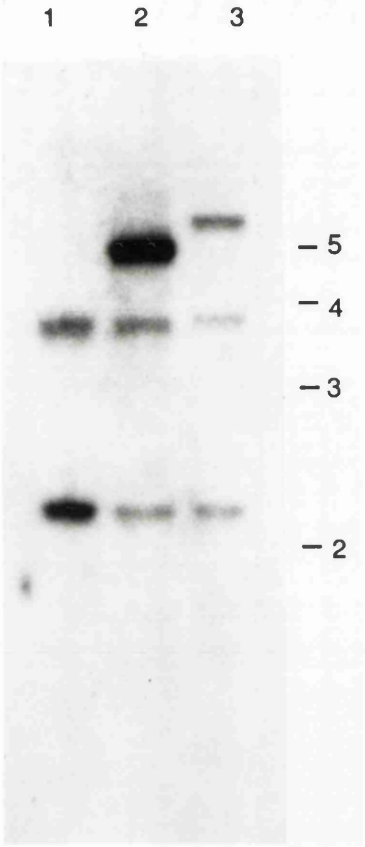
Diagnostic Southern Blots of Potential Homologous Recombinants

It had become evident from Southern blot analyses of genomic DNA isolated from the potential homologous recombinant cell line 1D6, that this cell line was intrinsically unstable with respect to the integrated SV40 replication origin sequences. Results had suggested that when these cells were maintained at the restrictive temperature the SV40 origin and sequences flanking it underwent amplification and rearrangement. Control cell lines which had been shown to contain a single random integration of the transfected homologous replacement fragment were shown to maintain these integrated sequences in a stable manner when cultured in parallel with the 1D6 cell line. Cell culture experiments showed that the cell line 1D6 could not give rise to PALA-resistant colonies, or amplify CAD-specific sequences, following deliberate exposure to active T-ag. To determine if the cell line

Figure 3.14

Southern Blot of Genomic DNA from the Cell Line 1D6 After Prolonged Culture at the Restrictive Temperature

1D6 cells and cells from a cell line which had been shown to contain a single, random integration of the homologous replacement fragment, 4RN, were expanded in parallel for several weeks at the restrictive temperature. Genomic DNA was isolated and approximately 10 μg was restricted with Pst I, Southern blotted and hybridised to the 3 kb Xba I fragment from L41. This probe had previously shown amplification of a novel Pst I fragment in the cell line 1D6. In 1D6 cells which had been minimally expanded, the novel fragment was detected at normal, unamplified levels. The lanes are loaded as follows: (1) ts-COS genomic DNA, (2) cell line 1D6, and (3) cell line 4RN. The blot shows that although the continuous growth of 4RN cells has not affected the copy number of the novel Pst I fragment in this cell line, the novel fragment detected in the potential homologous recombinant cell line 1D6 appears to have been amplified. Sizes marked are in kilobases.


1
2
3
5
4
3
2

1D6 did contain an accurately targeted homologous replacement fragment, a diagnostic experiment was designed. The approach was to probe Pst I-digested genomic DNA from this cell line with a shortened version of the 3 kb Xba I fragment from the cosmid L41. The shortened fragment has been described previously in this chapter and comprises a 2.3 kb Xba I-EcoR I fragment from L41. This fragment has no homology with the homologous replacement fragment but hybridises to a Pst I fragment from L41 which overlaps the EcoR I site which forms the boundary of homology between the transfected fragment and the CAD locus. When Pst I digested, Southern blotted DNA is hybridised to this 2.3 kb Xba I-EcoR I fragment, an additional, novel band should only be seen if homologous recombination has occurred due to the additional Pst I sites that would be introduced into the locus by integration of the neomycin resistance gene and the SV40 replication origin. The size of a novel fragment identified by this probe would be 4.8 kb for an unrearranged, but modified, locus

The Southern blot shown in figure 3.9 was stripped and rehybridised with the 2.3 kb Xba I-EcoR I fragment. The autoradiogram of the blot is shown in figure 3.9d. By the absence of any hybridisation, except that of the unmodified locus, not one of these potential homologous recombinant cell lines appear to have an SV40 replication origin, accurately targeted by homologous recombination.

3.6.10

Summary and Conclusion

Several potential homologous recombinant cell lines were identified, following isolation from a pool of G418-resistant transfectants. They were identified by their ability to amplify a CAD locus-specific fragment after deliberate exposure to active T-ag. The cell lines initially showed a Southern blot hybridisation pattern consistent with an accurately targeted integration event of the homologous replacement fragment and one of these potential homologous recombinant cell lines, 1D6, was further investigated. It was found that in this cell line, the integrated SV40 origin sequences were unstable at the restrictive temperature and underwent

amplification and rearrangement. The cell line was unable to give PALA resistant colonies following incubation at the permissive temperature for T-a activity and that during such an incubation, episomal DNA was produced by excision and amplification events. Finally, a diagnostic Southern blot of DNA from this cell line showed that an SV40 replication origin was not present at the CAD locus of this cell line and work was discontinued on these cells. However, the possibility remains that amplification and rearrangement of the SV40 sequences may have led to the loss of these sequences from the CAD locus during the course of the experiments described.

3.7

Further Transfections and Screenings

3.7.1

New Transfectants and Clones

Initial attempts to screen G418-resistant transfectants for homologous recombination of the homologous replacement fragment were confusing, ambiguous and ultimately unsuccessful (see summary above, 3.6.10). The results obtained from selecting pooled G418-resistant tranfectants in PALA, following deliberate exposure to active T-ag, gave a frequency for homologous recombination of approximately 1 in 50 integrations. It was decided that, if G418-resistant transfectants were cloned and screened as early as possible, potential homologous recombinant cell lines could be identified with minimal expansion of the clone. If the frequency of homologous recombination was close to 1 homologous integration per 50 random integrations, then the pooling of G418-resistant clones could be avoided as independent G418-resistant colonies could be analysed. This approach would keep unintentional amplification and rearrangements to a minimum as the cells would not be cultured for extended periods and also, independence of the cell lines could be maintained.

New transfections of the homologous replacement fragment were performed and independent, unexpanded G418-resistant colonies were obtained. These cell lines were screened by Southern blot, PALA selection and slot blot with minimum expansion. It was also hoped that PCR could also be used to screen such clones at an early stage. The advantages of such an approach would be that relatively small numbers of cells could be used for a template, therefore allowing minimal expansion of the G418-resistant cell lines. A second advantage of using a PCR screening technique would be the ability to process large numbers of G418-resistant cell lines at one time. If a CAD-specific primer with no homology to the transfected homologous replacement fragment was used in conjunction with a neomycin resistance gene-specific primer, then a PCR product of a predicted size could only be obtained if the two primers were correctly orientated at the correct distance apart. This arrangement would only occur if the homologous replacement fragment had integrated at the CAD locus. The possibility of using PCR was explored as described below.

3.7.2

Design of CAD-Specific and SV40/neo-Specific Primers

To screen newly selected G418-resistant transfectants by PCR, a CAD-specific primer was required that would anneal to CAD sequences not present in the transfected DNA but would be sufficiently close to the SV40/neomycin resistance gene sequences introduced into the CAD locus by homologous recombination. The smallest distance for such a sequence from the SV40 origin/neomycin-resistance gene sequences is 2.7 kb, ie. the distance from these sequences to the end of the homologous sequence in the homologous replacement fragment. The sequence which immediately flanks the homologous recombination target site at the CAD locus comprises part of the 2.3 kb Xba I-EcoR I fragment used as a probe in the Southern blot analysis of the potential homologous recombinant cell line 1D6. This fragment was cloned into the Promega transcription vector pGEM4 and sequenced from the SP6 transcription site using a SP6 site-specific sequencing primer and a Sequenase

Figure 3.15

Nucleotide Sequences of PCR Primers

The nucleotide sequences of the PCR primers were derived as follows: sequences of primers specific for the CAD locus and having no homology with the transfected homologous replacement fragment were obtained from sequencing part of the 2.3 kb Xba I-EcoR I fragment which had been previously used for Southern blot hybridisations. The DNA close to the EcoR I end was sequenced and PCR primers were made corresponding to this region; all other sequences were obtained from the EMBL or NIH nucleotide sequence databases.

CAD locus-specific PCR primers

a) 20mer;

5'- CATTCTACAATTAATATACA -3'

b) 40mer;

5'- TAATTGTAGAATGCCTCTGACTAAATTTGGACTCGTTGAG -3'

SV40 termination sequence-specific PCR primers

a) 20mer; bases 2562-2542 of SV40

5'- TCATCAATGTAACTTATCATG -3'

b) 40mer; bases 2581-2542 of SV40

5'-AGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATG -3'

Internal neomycin resistance gene PCR primers

a) 20mer; bases 135-154 of the neo gene

5'- ATGAGGATCGTTTCGCATGA-3'

b) 20mer; bases 1200-1181 of the neo gene

5'- CGGATTTGCACTGCCGGTAG-3'

SV40 replication origin-specific PCR primers

a) 40mer; bases 226-265 of SV40

5'- GGGGAGCCTGGGGACTTTCCACACCCTAACTGACACACAT-3'

kit. Primers were designed from the sequence obtained and the primer sequences are shown in figure 3.15. The Southern blot analysis shown in figure 3.9 shows that the sequences of the transfected fragment that hybridise to the 3 kb Xba I probe of L41 are located on the same Pst I fragment as the neomycin resistance gene sequences and therefore these sequences are closest to host DNA flanking the integration site. The SV40 transcription termination sequences present in the neomycin resistance gene were obtained from the EMBL database and the primers shown in figure 3.15 were designed. All primers were made using an oligonucleotide synthesiser by I. Goldsmith at ICRF, Clare Hall Laboratories.

3.7.3

Construction of a PCR-Positive Control

In order to establish the PCR conditions required for amplification of the sequences between the primers, and also as a positive control, a construct was designed in which the sequences of the CAD and SV40 primers were arranged as they would be predicted following homologous recombination of the homologous replacement fragment.

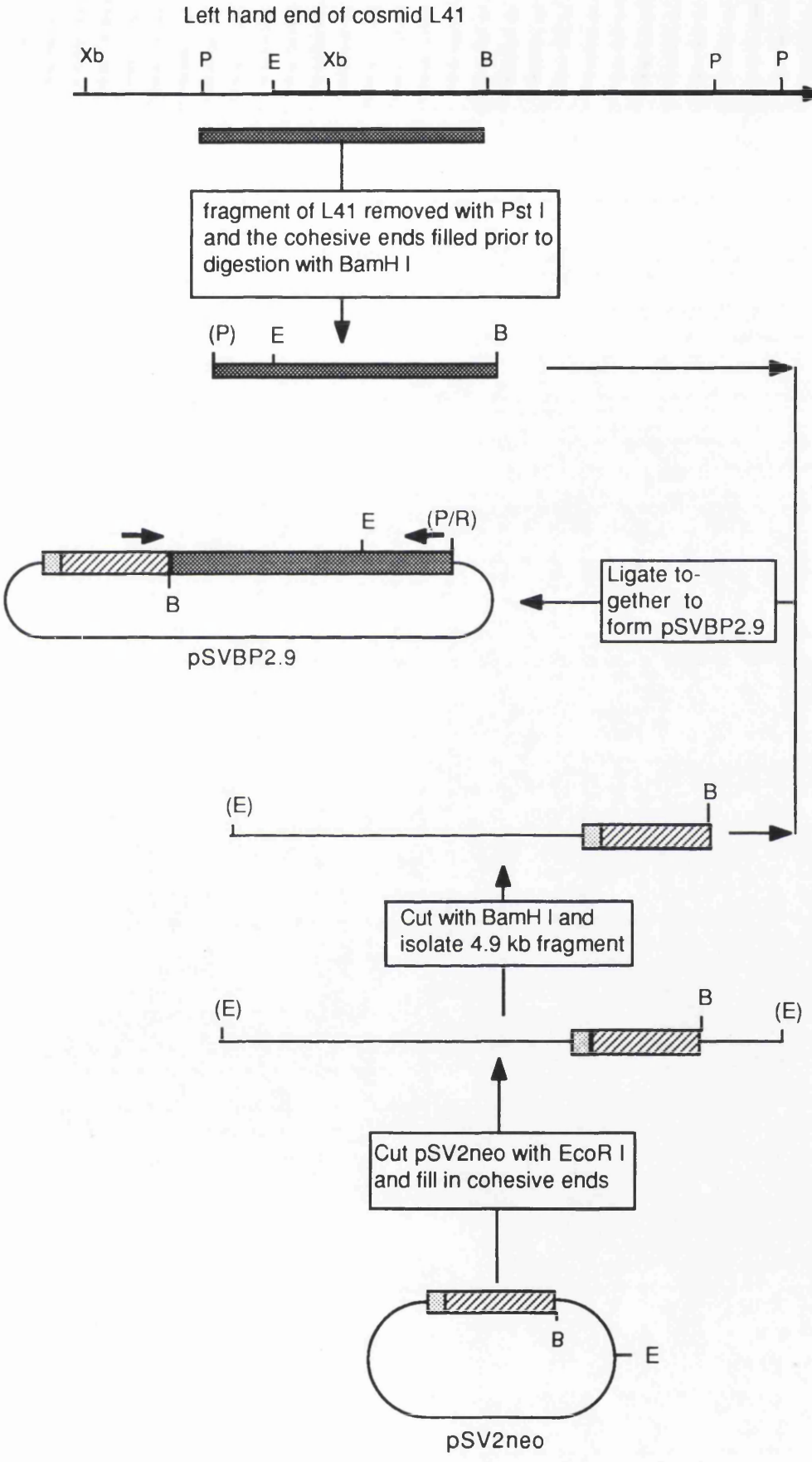
The plasmid pSVBP2.9 was constructed as shown in figure 3.16. The plasmid pSV2neo was the source of the neomycin resistance gene and SV40 replication origin which were cloned into the homologous replacement fragment. This plasmid was modified to include the DNA from the CAD locus which would be predicted to separate the PCR primers following homologous recombination of the transfected DNA . Therefore the plasmid construct pSVBP2.9 contains the genomic targets for the CAD and SV40 termination signal sequence PCR primers described above, separated by the predicted CAD sequences. The SV40 termination signal primer was tested against a primer specific to a pBR322 sequence about 700 nucleotides distant, using pSV2neo as a template. After 30 rounds of amplification a single product of the correct size was observed. The CAD-specific primer was checked against the SV40 primer using the PCR positive control construct, pSVBP2.9, as a template, and a single fragment of the correct size was produced

Figure 3.16

Construction of the PCR Positive Control Template Plasmid, pSVBP2.9

A positive control template plasmid was designed and constructed as shown. A Pst I fragment was removed the cosmid from L41. This fragment contains the sequences that would separate the neomycin gene from the flanking DNA of the target region, and also the monkey sequences which are contiguous with those in the homologous replacement fragment. This Pst I fragment was cloned into a pSV2neo-derived fragment to form pSVBP2.9. This plasmid therefore contains the neomycin resistance gene and CAD-locus sequences in an arrangement which is analogous to those that would be present at the CAD target site following homologous integration of the transfected DNA. The position of PCR priming sites in pSVBP2.9 are shown (heavy arrows) and therefore this plasmid could be used as a positive control template for PCR between the CAD-specific sequences and the neomycin resistance gene. If a similar predicted arrangement of sequences occurred following homologous integration of the transfected fragment, then PCR could be used to screen for these integration events.

The SV40 replication origin is lightly stippled, the neomycin resistance gene is hatched, the fragment from cosmid L41 is densely stippled and the pBR322 sequences of pSV2neo are shown as solid lines. The positions of PCR primers are shown with heavy arrows in the plasmid pSVBP2.9. B=BamH I; E=EcoR I; P=Pst I; Xb=Xba I; bracketed restriction sites are those which have been filled in at their cohesive termini and have been destroyed.


Left hand end of cosmid L41
Xb
P
E
Xb
B
P
P
fragment of L41 removed with Pst I and the cohesive ends filled prior to digestion with BamH I
(P)
E
B
E
(P/R)
B
pSVBP2.9
Ligate to-gether to form pSVBP2.9
(E)
B
Cut with BamH I and isolate 4.9 kb fragment
(E)
B
(E)
Cut pSV2neo with EcoR I and fill in cohesive ends
B
E
pSV2neo

after amplification. Conditions of $MgCl_2$ concentration, primer annealing temperature and primer concentration of the 20-mer (and later 40-mer) primers for the CAD locus and the SV40 termination sequence primers were determined using pSVBP2.9 as a template.

3.7.4

Limits of PCR

Conditions for the PCR with the CAD- and SV40-specific primers were determined using pSVBP2.9 as a template and at best, a product of the correct length (2.7 kb) could be detected after 25 cycles with 10 pg of template DNA (figure 3.17). The ethidium bromide stained gel shows a single product of 2.7 kb when 100 pg of the template plasmid pSVBP2.9 was used (lane 2). When larger amounts of template plasmid were used the amount of non-specific amplification increased, probably due to non-specific hybridisation of the primers to the template plasmid.

To determine the efficiency of the PCR in a cellular DNA environment the positive PCR control template plasmid pSVBP2.9 was transfected into BHK clone 13 cells. Using lysates from these transfectants or genomic DNA prepared from them, it was hoped that the PCR and template preparation conditions could be established prior to screening the homologous replacement fragment transfectants.

In preparation for transfection, the control plasmid PCR template pSVBP2.9 was linearised by digestion with the restriction enzyme Nde I and stably transfected into BHK cells by electroporation. Following G418 selection, genomic DNA was isolated from pooled G418-resistant colonies and used as a PCR template. The 20-mer CAD and SV40 signal sequence primers were initially used to attempt to amplify the integrated plasmid region. Using genomic DNA from these transfectants as a template for the 20-mer oligonucleotide primers of CAD and the SV40 termination signal, a product could not be amplified under any of a wide range of conditions. A second set of primers, 40 nucleotides long and of different sequence to the 20-mers, were designed to increase the specificity of primer/target hybridisation (see fig. 3.16) It was hoped that the extra homology and different sequence of the 40-

Figure 3.17

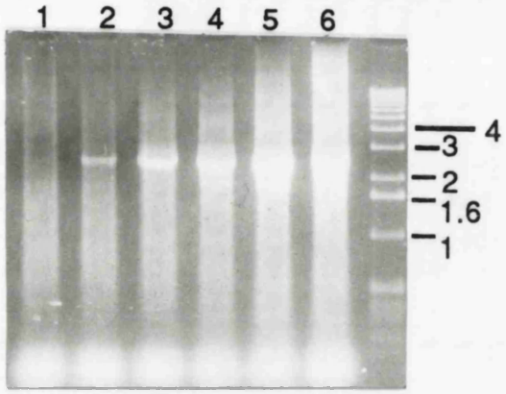
PCR with Decreasing Amounts of pSVBP2.9 Template (upper panel)

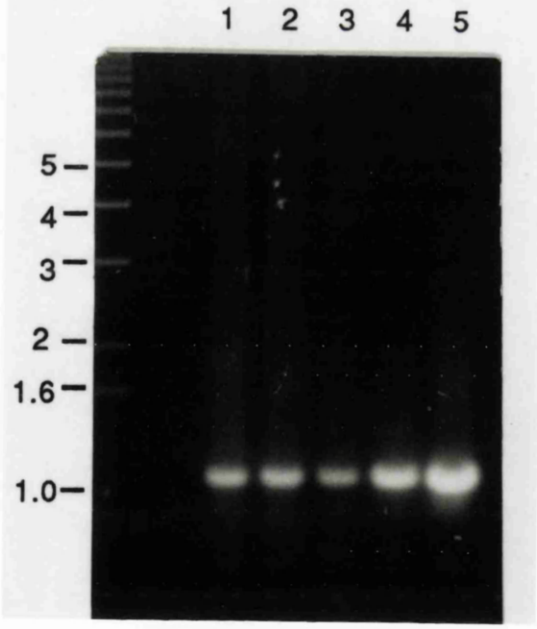
The DNA sequence between the 40-mer CAD and 40-mer SV40 termination signal primers was amplified using different amounts of pSVBP2.9 as template. Primers were used at 1 μM each in 2 mM $MgCl_2$ and amplification was from 25 cycles of 3 minutes at 94°C (denaturation), 5 minutes at 60°C (primer annealing) and 9 minutes at 72°C (primer extension). The amounts of template per lane are as follows: (1) 10 pg, (2) 100 pg, (3) 1 ng, (4) 10 ng, (5) 100 ng, and (6) 1 μg. 10 μl (out of 50 μl) were run per lane; size markers are in kilobases.

Figure 3.18

PCR of an Internal Neomycin Resistance Gene Sequence from a Genomic DNA Template (lower panel)

Primers for the neomycin resistance gene sequence (see figure 3.14) were used at 1 μM in 1.5 mM $MgCl_2$. Amplification was from 30 cycles of 94°C for 1.5 minutes, 60°C for 3 minutes and 72°C for 5 minutes. The templates used for each reaction were different;
lane 1: 1 μg of clone 3RN genomic DNA with 4.5 μg of ts-COS genomic DNA as carrier; lane 2: 500 ng of 3RN genomic DNA and 2.25 μg of ts-COS genomic DNA; lane 3: 100 ng of 3RN genomic DNA and 450 ng of ts-COS genomic DNA; lane 4: 1 μg of 3RN genomic DNA; and lane 5: 13 ng of pSV2neo (see Southern and Berg, 1982). Size markers are in kilobases.


1
2
3
4
5
6
4
3
2
1.6
1


1
2
3
4
5
5
4
3
2
1.6
1.0

mer primers would make the annealed primer/target sequences more stable and increase the efficiency of the PCR. Using the control plasmid pSVBP2.9 alone as a template, these primers were more efficient than the 20-mer primers at amplifying the target sequence, but using genomic DNA from the transfected BHK cells as template the reaction could still not be made to work. Lysates from ts-COS cells were 'spiked' with 100 pg to 1 µg of the positive control template plasmid and the PCR did not amplify the target sequence with either set of primers. However, primers made against internal neo sequences approximately 1 kb apart amplified a fragment of the correct size using the genomic DNA as a template. The amplification of this fragment is shown in figure 3.18. The amplified fragment is the predicted size and non-specific PCR products can be seen.

3.7.5

Use of PCR as a Screening Procedure

Although PCR would have be a powerful screening technique for identifying homologous recombinant cell lines, the target sequence for the PCR primers could not be amplified in a cellular environment. A n internal neomycin resistance gene target could be amplified in the presence of isolated genomic DNA suggesting that an inhibitor in the DNA preparation was not responsible for the failure of the PCR using the CAD and SV40 termination sequence primers. A control cell line containing the integrated PCR positive control plasmid showed that the target sequence could not be amplified when in a genomic DNA background. For this reason the PCR screening strategy was discontinued and the newly transfected G418-resistant cell lines were screened using Southern and slot blots of the DNA and a functional PALA selection of the cells.

3.7.6

Isolation of New Homologous Replacement Fragment Transfectants

Previous work with potential homologous recombinant cell lines had shown that the integrated DNA can be amplified and rearranged during expansion of the cell

lines at the restrictive temperature. Therefore an effort was made to expand G418-resistant clones for the shortest amount of time possible, before attempting to identify potential homologous recombinants. New, independent homologous replacement fragment transfectants were produced by electroporation of ts-COS cells which were plated at low density to give 10 to 20 G418-resistant colonies per plate after G418 selection. When large enough to clone, about five, well separated colonies from each plate were ring-cloned. When each clone had reached approximately 1 x 10^5 cells, duplicate aliquots were frozen and a third aliquot was expanded to 10^6 cells. At this point, clones were pooled and genomic DNA from each pool was digested with Pst I, Southern blotted and probed with the 2.3 kb Xba I-EcoR I diagnostic probe fragment from the cosmid L41. This probe would only hybridise to an additional, novel band if the homologous replacement fragment had integrated at the CAD locus.

Table 3.2 shows 25 pools of independent G418-resistant transfectants, named p1 to p25. When each clone reached approximately 10^6 cells, 2 aliquots of 2 x 10^5 cells each were frozen and the remaining cells were pooled with other clones at a similar stage of expansion. The pooled cells were expanded to a total of 1-2 x 10^7 cells and genomic DNA was isolated from the pool. If all clones grew at the same rate, each clone would have been expanded from 5 x 10^5 to 1-2 x 10^6 cells. Before the preparation of genomic DNA, some cells were removed and replated at the restrictive temperature. Manipulations of these cells are described in the following section.

Table 3.2

G418-Resistant Clones Screened for Homologous Recombination

Pools of G418-resistant transfectants p1 to p25, several cell lines isolated from the G418-resistant pool RNp4 (1D6, 2D1, 2B2, 2B1, 2C6, 2B3) and also several clones containing a single, random integration of the transfected fragment (3RN, 9RN, 4RN, c54 and c55) were all screened by PALA resistance and whole cell slot blot before and after exposure to T-ag activity and also by Southern blot. Other clones listed were analysed by Southern blot and whole cell slot blot only. The nomenclature of the clones is as follows: the initial number of each clone is the dish from which it was isolated and the following letter is the clone designation from a particular dish, eg. 8A and 8E are the first and fifth clones isolated from the same dish. In some pools the number and letter of clones are listed in reverse order as these clones are derived from a different set of plates; in this case the letter defines the dish and the number of the clone from that particular dish. Clones designated c1, c2 and so on were isolated from a yet another set of transfectants; several clones were analysed on their own without being pooled because each clone reached the required number of cells at a time when no other clones were ready.

Pool	Clones	Pool	Clones
p1	6D 5A 8E 9A 9C	p13	c15 c8 c3 c5
p2	5B 8A 7E M2 (2 clones)	p14	c9 c11 c12 c19
p3	6E 5C 5D 10E M1 (2 clones)	p15	c18 c16 c13 c17 c20
p4	6A 7C 7A 5E 9B	p16	c21 c10 c24 c29 c33
p5	10A 10B 2B 2A 3E	p17	c26 c25 c28 c27
p6	1C 2C 4E 3E 1E	p18	c30 c31 c 38 c 37
p7	G3 3F 2D	p19	c23 c14 c22
p8	D5 F5 A8 10D H6	p20	c35 c39 c32
p9	E10 D3 A5 4A	p21	c44 c34 c36 6C
p10	4B 1B 9E	p22	c40 c41 c42 c43
p11	c15 c8 c5 c3	p23	c39 c47
p12	c6 10E 1D	p24	c50 c51 c52 c48
		p25	c45 c46 c49 c53

The following clones were analysed as cell lines, ie. they were not included in a pool: 6C, 9D, 10C, 10D, c54 and c55.

56 other clones, isolated from RNp4, had been previously analysed (see 3.6.1) and a further 7 (1RN, 3RN, 4RN, 5RN,6RN, 7RN and 9RN) had also been screened.

Total = **170**.

3.7.7

PALA Selection of Pools p1 to p25

Each pool of G418 resistant transfectants was tested to determine if exposure to active T-ag caused an increase in the frequency of PALA resistance. If such an increase could be observed then it could be possible that one of the founding transfectants contains an accurately targeted homologous replacement fragment. Each cell line contributing to the pool could then be independently screened for such an event as all cell lines were individually frozen in liquid nitrogen prior to being pooled.

When each pool had been expanded to 1-2 x 10^7 cells, the plates were trypsinised and 10^5 cells/dish were seeded on to five 100 mm dishes. The remaining cells were washed with PBS and high molecular weight genomic DNA was isolated from them. The seeded cells were allowed to attach overnight at the restrictive temperature, after which 3 dishes of each pool were shifted to the permissive temperature for 3 days and then allowed to recover at the restrictive temperature for 24 hours. The other 2 dishes were maintained at the restrictive temperature throughout. All dishes were trypsinised and populations of shifted and unshifted cells were made for each pool. Triplicate aliquots of 5 x 10^4 cells were taken for whole cell slot blots and the remainder were plated at 5 x 10^5 cells/100 mm dish and selected in 75 μM PALA at the restrictive temperature. Medium was changed every 5 days and plates were stained after 4 weeks of selection.

The frequencies of PALA resistance for each pool before and after exposure to T-ag are shown in table 3.3. As can be seen, some pools had a detectable frequency of resistance at the restrictive temperature but this frequency was not altered by exposure to active T-ag. Pools which showed no PALA-resistant colonies have been given a upper limit to the frequency of resistance, based on the number of cells selected. Other concurrent selections using the same stocks of PALA, dFCS and medium provided PALA-resistant colonies at expected frequencies and rates, showing that the reagents being used were not preventing resistant cells from

Table 3.3

PALA Resistance Frequencies of Pooled G418-Resistant Clones Before and After Exposure to T-ag

Cells were seeded at 1-5 x10^5 cells per 100 mm dish and were allowed to attach at the restrictive temperature overnight. Several plates were moved to the permissive temperature for 3 days and then returned to the restrictive temperature for 24 hours before being counted, dispersed and selected in 75 μM PALA. Unshifted cells were also counted, dispersed and selected in the same manner. Medium was changed every 5 days and the plates were stained after 3 to 4 weeks. No increase in the frequency of PALA resistance was seen in any pool.

Pool	Frequency of PALA Resistance (x 10^{-6}) Unshifted	Shifted
p1	<10	1.25
p2	<1.6	<1.6
p3	<1.3	<1.3
p4	<2	<2
p5	<1.5	<1.6
p6	<1.3	<1.3
p7	<1.6	<1.8
p8	<1.4	<1.3
p9	1.2	2
p10	<1.25	<1.25
p11	<1.4	<1.4
p12	1.25	3
p13	<1.4	<1.4
p14	n/d	<1.25
p15	2	1.3
p16	<2	<1.4
p17	10	7
p18	12.8	9.6
p19	5.5	5.2
p20	1	3.2
p21	1	4
p22	4	4
p23	1	0.25
p24	10	7
p25	1	<1.25
c54	<1	<2
c55	<1	<1
ts-COS	**1**	**<1**

n/d = not determined.

< = no colonies for the number of cells selected.

surviving the selection. The lack of PALA resistance in the pools after exposure to T-ag is consistent with a homologous recombinant not being present in any of the pools.

3.7.8

Whole Cell Slot Blot Analysis of Pools p1 to p25

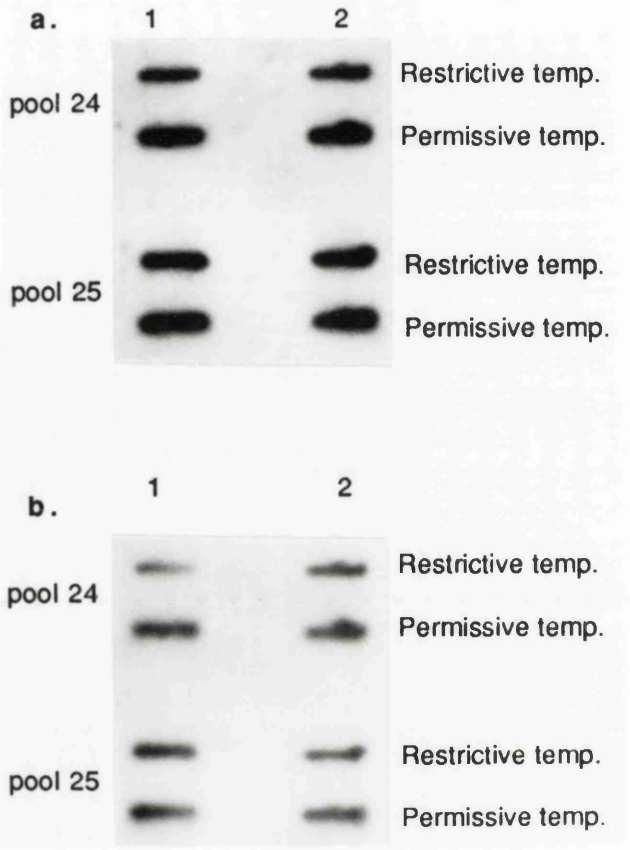
To determine if any of the pools contained amplified CAD DNA in the absence of PALA resistance the cells were analysed by whole cell slot blots as described below. It was possible that, although homologous recombination may have occurred in one of the pools of transfectants, amplification of the DNA surrounding the SV40 replication origin did not include a functional CAD gene. Alternatively, a functional CAD gene may be amplified following exposure to active T-ag but instability of the amplified sequences may prevent stable PALA resistance in the cells.

Whole cell slot blots were made from cells of each pool that had been maintained at the restrictive temperature or that had been shifted to the permissive temperature for 3 days (see see previous section). The filters were first hybridised with the 2.3 kb Xba I-EcoR I fragment of cosmid L41. This fragment is specific for the CAD locus and has no homology with the transfected homologous replacement fragment. Amplification of this sequence could only occur if the transfected fragment had integrated at the CAD locus. After satisfactory exposures were obtained for this probe, the slots were normalised for loading differences by hybridisation with a total monkey genomic DNA probe. The signal obtained with the 2.3 Xba I-EcoR I fragment was corrected for loading differences and none of the pools showed elevated levels of the CAD-specific sequence after exposure to active T-ag activity. A representative slot blot result showing no increase in the copy number of this fragment is shown in figure 3.19. There is no obvious difference in signal intensity between the two probes following incubation at the temperature permissive for active T-ag. This is consistent with there being no homologous recombinant cell lines present and is in

Figure 3.19

Representative Whole Cell Slot Blots of Pools of G418-Resistant Transfectants Exposed to Active T-ag

Whole cell slot blots were made from cells of each pool of G418-resistant clones that had been maintained at the restrictive temperature or that had been shifted to the permissive temperature for 3 days. The filters were first hybridised with the 2.3 kb Xba I fragment from L41 which has no homology with the transfected DNA (panel a) and later with a total monkey genomic DNA probe (panel b) to normalise the slots for loading differences. As can be seen, no amplification of the L41-derived probe has occurred following exposure to T-ag.


a.
1
2
pool 24
Restrictive temp.
Permissive temp.
pool 25
Restrictive temp.
Permissive temp.
b.
1
2
pool 24
Restrictive temp.
Permissive temp.
pool 25
Restrictive temp.
Permissive temp.

agreement with the constant frequency of PALA resistance seen in the pools before and after exposure to active T-ag.

3.7.9

Southern Blot Analysis of Pools p1 to p25

As final confirmation of there being no homologous recombinant cell lines in the pools of transfectants, a diagnostic Southern blot analysis was performed on each pool. Each pool comprised cells which had always been strictly maintained at the restrictive temperature. Genomic DNA was isolated from each pool, digested with Pst I, size fractionated by agarose gel electrophoresis and Southern blotted. Each filter was hybridised with the the diagnostic 2.3 kb Xba I-EcoR I fragment of L41. An additional novel fragment would only be seen using this probe if homologous recombination had occurred. A typical Southern blot is shown in figure 3.20 which shows the two expected bands consistent with an unmodified CAD locus; an additional band of 4.8 kb is not present. All pools showed the wild type pattern of two bands and a novel band of the predicted size was not seen in any pool. This confirms the previous conclusion that a homologous recombinant cell line is not present in any of the pools analysed and at this point it was decided that no future screenings would be done.

3.7.10

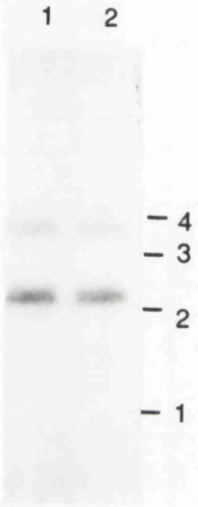
Conclusion

Following transfection of the homologous replacement fragment, a total of 170 cell lines were screened for a targeted integration event and no positive cell lines were identified. A pool of transfectants had shown a T-ag activity-dependent increase in the frequency of PALA resistance but no cloned cell line could demonstrate the same effect. CAD locus-specific sequences were shown to be amplified in the pool of transfectants and also in some cell lines isolated from the pool but following Southern blot analysis, these cell lines were shown not to contain an accurately targeted SV40 replication origin. One cell line studied in detail, showed

Figure 3.20

Representative Southern Blot of Pooled G418-Resistant Transfectants

After an aliquot of each cell line had been frozen, cell lines were pooled and expanded. Genomic DNA was isolated, digested with Pst I and Southern blotted. Filters were hybridised to the 2.3 kb Xba I-EcoR I fragment of L41 which would only hybridise to a novel, additional fragment if the homologous replacement fragment had integrated at the CAD locus. An additional, novel band was not seen for any cell line.


1
2
4
3
2
1

that the integrated SV40 sequences could be unstable, with amplification, rearrangement and excision events contributing to this instability. Bearing in mind these factors and also emerging data that strongly suggested that over-replication was not likely to be responsible for mammalian gene amplification (Trask and Hamlin, 1989; Smith *et al.*, 1990), the project was closed.

3.8

Discussion

3.8.1

Overview

This chapter has dealt with the attempt to set up a cell line in which gene amplification events at a predetermined locus could be studied simultaneously in all cells of a population. The model was also to be an experimental test for the disproportionate replication model of gene amplification. Overall, the project was not successful and possible improvements that could be made to increase the chance of success are discussed below.

3.8.2

Design of the Homologous Replacement Fragment

The homologous replacement fragment was designed with the following features in mind. The monkey sequence (insert of λ16) which was used as homology for homologous recombination was selected to be as large as possible. Thomas and Capecchi (1987) describe the features of homologous replacement vectors which increase the efficiency of homologous recombination. Generally, the greater the homology between target and input DNA, the higher the homologous integration frequency; repeated elements were not thought to effect the efficiency of the process. If the free ends of the linearised input DNA were normally non-contiguous in the genome sequence, homologous replacement would occur; if they were normally contiguous then homologous sequence integration, forming a duplication of the target sequence, would take place. The 16 kb EcoR I fragment from the cosmid L41 was the largest convenient fragment which could be used for sequence replacement, and it possessed a useful site into which the SV40 replication origin and neomycin resistance gene could be cloned. Southern blot hybridisation of the cosmids with a CAD cDNA probe demonstrated that the fragment did not contain any CAD coding sequences, and so modification of the fragment or

its homologous integration should not interfere with expression of the CAD gene product.

The SV40 replication origin was chosen as it is the best characterised origin known to function in eukaryotic cells, and the specific nucleotide sequences that constitute a mammalian replication origin have not yet been determined. The only viral protein required for SV40 replication is T-ag and this may be used to control the activity of the SV40 replication origin. In the presence of active T-ag the origin continually initiates bidirectional replication into the flanking DNA sequences, and in doing so may mimic the activity of an aberrant mammalian replication origin. Several attempts were made to control the activity of the SV40 origin before my arrival in the laboratory. Preliminary work (Mark Rolfe, unpublished results) on a CV1 cell line which contained a single, randomly integrated SV40 replication origin showed that infection of these cells with wild type SV40 virus (as a source of T-ag) caused cell death as part of the viral lytic cycle. However, an encouraging observation (M. Rolfe, unpublished data) was that replication at the integrated viral replication origin did not produce circular excision products during replication, presumably due to the absence of duplicated DNA which can act as a substrate for homologous recombination as described in Chapter I. In contrast, a CV1 cell line containing several copies of the SV40 replication origin and other vector sequences integrated in tandem produced large amounts of Hirt-sized DNA excision products after infection by SV40 (M. Rolfe, unpublished data). The replacement fragment was designed to exploit this effect as it contained a single copy of the replication origin and neomycin resistance gene. The integration was biased towards single, unduplicated events by using electroporation to transfect the cells. Excess foreign sequences were removed from the homologous replacement fragment in order to reduce the modification of the CAD locus to a minimum and also to reduce the amount of non-homology in the fragment. In particular, unnecessary bacterial sequences (eg. promoters etc.) were eliminated from the SV40 origin/neomycin resistance gene insert of the transfected fragment to avoid the possible formation of sequence duplications.

3.8.3

Temperature-Sensitive COS Cells

In order to control the activity of T-ag, ts-COS cells (Rio, Clark and Tjian, 1985) were used. These cells contain a single integration of an SV40 A gene (ts A 1609) which encodes a temperature-sensitive T-ag. The gene is constitutively expressed from the promoter of the Rous sarcoma virus LTR which is insensitive to T-ag repression. At the restrictive temperature (40^{o}C) the T-ag produced is inactive with respect to initiation of SV40 replication, and at the permissive temperature (33^{o}C) replication is initiated in the usual way. It is unknown whether the cellular effects of T-ag are also temperature sensitive.

Experiments to investigate the frequency of PALA resistance in ts-COS cells that do not contain an SV40 replication origin have shown that exposure to active T-ag does not significantly raise the frequency of PALA resistance in these cells (see table 3.1). The frequency of resistance to 75 μM PALA at the restrictive temperature is $<1 \times 10^{-6}$ and this figure does not increase after exposure of the cells to active T-ag. Upon infection of permissive cells, SV40 stimulates cellular DNA synthesis and cell division. The repertoire of cellular genes induced by T-ag are involved in DNA replication, recombination and possibly mitosis, presumably for the virus' own reproductive requirements. The developers of ts-COS cells suggest that the transcriptional activities are also temperature sensitive in these cells, and the experimental evidence described above implies that the frequency of PALA resistance is not increased by the expression of these cellular genes.

3.8.4

PALA Selection of G418-Resistant Pools

Early pools of G418-resistant colonies transfected with the homologous replacement fragment were relatively small, containing up to 15 colonies (RNp1 and 2). When aliquots of these pools were incubated at the permissive temperature for limited periods of time, no resistant colonies survived the following PALA selection.

Larger pools of transfectants (RNp3 to 6) showed a T-ag activity-dependent increase in the frequency of PALA resistance following incubation at the permissive temperature. The pool of transfectants RNp3, was founded by about 30 independent G418-resistant colonies, and PALA resistance appeared to be dependent on the length of exposure to active T-ag (see table 3.1). However, the control cells of this pool which were maintained at the restrictive temperature also gave rise to PALA-resistant cells. There are at least two reasonable explanations for this. Firstly, a pre-existing PALA-resistant mutant could be present in this particular pool of transfectants in which case the frequency of PALA resistance should be independent of exposure to active T-ag. Secondly, leakage of T-ag activity from unintentional exposure to active T-ag during cell manipulations could have caused over-replication and amplification of CAD sequences flanking an accurately targeted SV40 replication origin in one of the founding transfectants. The number of PALA-resistant colonies in the pool RNp3 is dependent on the length of exposure to active T-ag, supporting the latter explanation. Cell lines were not cloned from RNp3 for further analysis because it was not known when and for how long T-ag had been active in these cells. It was possible that frozen stocks of the pool could already contain cells that were resistant to PALA through leakage of T-ag activity. It was also known from earlier work that prolonged culture of ts-COS cells at the permissive temperature was fatal if the cells were carrying a functional SV40 replication origin. The plating efficiency and metabolic rate of the cells are also reduced at 33^{o}C. Cells incubated at the permissive and restrictive temperatures were seeded at the same density, but because of the above mentioned differences the proportion of cells surviving before selection with PALA is higher in the unshifted populations. The PALA resistance frequency following incubation at the permissive temperature is therefore probably higher than the figures suggest; the lower plating efficiency and doubling time of cells at the permissive temperature mean that less cells are selected in shifted cell populations. To minimise these effects, all cells were plated at the restrictive temperature and cells which were to be exposed to active T-ag were given 24 hours to attach at the restrictive temperature before being moved to the

permissive temperature. After the temperature shift, cells were given a 24 hour recovery period at the restrictive temperature before being detached from the plate, counted and dispersed. In this way the plating conditions and densities prior to PALA selection were the same for both shifted and unshifted populations of cells.

3.8.5

Isolation of G418-Resistant Cell Lines

The pool of transfectants RNp4 showed an increase in PALA resistance which was dependent on exposure to active T-ag. A CAD-specific sequence not present in in the transfected DNA was amplified 5-fold after 6 days at the permissive temperature. Assuming that all founding clones were equally represented in the pool, the copy number increase of the sequence could be between 250-fold in one clone or 5-fold in all clones. If one clone had grown more than any other and was also a homologous recombinant, then the fold increase per cell of that clone would be lower than 250.

The RNp4 pool of transfectants containing the homologous replacement fragment was chosen as a source of cell lines which could be screened for homologous recombination because the number of founding clones was low (approximately 50), which meant that a sib-selection screening method could be avoided. The results of screening each isolated cell line by whole cell slot blots before and after exposure to T-ag were surprising. Approximately 41% of the clones showed an increase in copy number of a CAD-specific sequence. This could be due to over-representation of one clone in the population or it could be related to the probe. If a repetitive element was present in the probe sequence, it could give misleading results when amplified in another part of the genome. The three CAD cosmids were tested for repetitive elements and none were found in the probe region (M. Rolfe, unpublished data). The increase in specific signal may also be the result of non-specific hybridisation background, arising from high amounts of nucleic acids in the cell samples (especially in the form of episomes resulting from excision events at the permissive temperature). Certain rearrangements of the SV40

sequences (for example small sequence duplications) may promote excision events, and these rearrangements could occur in some of the clones. Another explanation is the over-representation of one particular clone which has amplified a sequence which cross-hybridises with the probe, or alternatively has amplified the CAD locus. Several of the clones were expanded, and Southern blot hybridisation showed a similar restriction fragment pattern (figures 3.8 and 3.9a) which at first suggested that homologous recombination had occurred in these lines. The consistency of the slot blot and Southern data implied that these lines were homologous recombinants. It was not be possible to tell if the lines were siblings and further analysis was performed on one of the potential homologous recombinant cell lines, 1D6.

3.8.6

Analysis of the Cell Line 1D6

The potential homologous cell line 1D6 showed promising results when initially screened for homologous recombination of the transfected fragment. However, Southern blot hybridisation with a shortened probe which could only give an additional, novel band of hybridisation if homologous recombination had occurred, demonstrated that this cell line did not have an accurately targeted CAD locus. Whilst this analysis was being performed, the cell line was characterised by other means. Hirt DNA extractions from shifted and unshifted cells showed that one main species of episomal DNA was produced on exposure to active T-ag and that this DNA did not hybridise to a CAD-specific probe. Such events are in agreement with published data described in the introduction (Botchan *et al.*, 1980; Bullock and Botchan, 1982), where excision events involving SV40 sequences have been observed in transformed cells. The published data suggest that a homogeneous excision product is consistent with a duplication of the integrated SV40 sequences. During replication of the integrated SV40 sequences homologous recombination between duplicated sequences causes excision of episomal DNA. In the potential homologous cell line 1D6, a homogeneous excision product is produced which implies that the excision is mediated via a duplication and not through some other

mechanism. The nature of the excision product also suggests that some amplification (at least a duplication) has occurred intra chromosomally at the integration site of the transfected DNA. If the excision event occurs at a low frequency then a small number of episome types will be produced, and these could be amplified to high copy number by extra-chromosomal replication. Alternatively the excision event could occur at a high frequency with moderate extra-chromosomal replication. In the latter hypothesis, the excision event would have to be accurate and repeatable to produce homogeneous episomes and this implies that specific recombination sites are used. In the former case, the event need only occur once per cell to produce the original episome, but for each cell in a population to produce an identical episome, a specific recombination site is required as before. With the data described here it is not possible to say which hypothesis is the most credible. A detailed analysis of the episome population and also of individual molecules cloned in bacteria could shed some light on the origin of these excision products. Finally, because the excision product is the same in most cells in the population, then the rearrangement and duplication which mediates the excision event must have occurred soon after the integration event as most descendants of the original transfectant must have a similar rearrangement in order to produce a homogeneous episome size.

A comparison between the potential homologous recombinant cell line 1D6 and cell lines containing a single random integration of the homologous replacement fragment, suggested that the SV40 sequences in 1D6 cells were intrinsically unstable. After passaging this cell line for several weeks, the integration pattern of the SV40 sequences became more complex, but in control cell lines remained simple. Gurney and Gurney (1989) reported that SV40 sequences in transformed rodent cell lines were unstable to varying degrees. The SV40 sequences of one particular cell line were highly prone to rearrangements and it was suggested that DNA sequences within or near the integration site were responsible for this activity. Although duplications of the integrated DNA were present in this cell line, other cell lines with a similar or higher degree of duplication were relatively stable. These data

suggest that the immediate environment of an integrated sequence may be a contributory factor towards its stability at that site. It is possible that, for 1D6 cells, small leakages of T-ag activity during cell manipulations and the local genetic environment had a great effect on the stability of the integrated DNA, and with time these factors destabilised the sequences around the integration site. It was apparent that the copy number of the novel fragment in 1D6 cells became amplified during passage at the restrictive temperature, and as described earlier, the arrangement of the integrated DNA also evolved into a more complex state.

Neither of the above aspects of 1D6 cells were pursued as the primary objective remained the isolation of a homologous recombinant cell line containing an SV40 replication origin which had been accurately targeted to the CAD locus.

The behaviour of 1D6 cells in the presence of PALA was apparently contradictory. All selections for PALA-resistant cells following exposure to active T-ag were negative. In contrast to this lack of resistant colonies, 1D6 cells appeared less sensitive to PALA after exposure to active T-ag. This was probably an artifact. Following exposure to PALA, the number of attached cells per dish was quantified by measurement of the cellular DNA content by optical density; the higher the optical density, the more cells present. Episomal DNA would only be present in cells containing an SV40 replication origin and active T-ag, and this additional DNA could increase the optical density measurement and give the impression that these particular cells had grown more than control cells in the presence of PALA .

In summary, the potential homologous recombinant cell line 1D6 looked promising. The restriction pattern of the CAD locus was encouraging but eventually the cell line was shown not to have an accurately targeted SV40 replication origin. At the permissive temperature episomal DNA was produced but it did not contain CAD-specific sequences and the chromosomally located SV40 sequences in this cell line were shown to be unstable during culture at the restrictive temperature. Finally, 1D6 cells were unable to give PALA-resistant colonies after exposure to active T-ag. Because the cell lines had been isolated from a pool of transfectants which had been expanded for some time before the leakage of T-ag was fully appreciated, it was

decided that new, independent transfectants would be screened for homologous recombinant cell lines containing an accurately targeted SV40 replication origin.

3.8.7

Analyses of New G418-Resistant Clones

G418-resistant cell lines were isolated from new transfections of the homologous replacement fragment and were analysed by Southern blot hybridisation, whole cell slot blots and by PALA resistance. It was hoped that PCR could be used to screen independent clones, but this methodology was not effective. Sequence information for the region flanking the homologous recombination target sequence was obtained by sequencing part of the 2.3 kb Xba I-EcoR I fragment from L41; this fragment has no homology with the homologous replacement fragment but immediately flanks the targeted integration site. PCR primers were designed from this sequence first as 20-mers and later as 40-mers. Using a positive control plasmid template that contains the SV40 origin/neomycin resistance gene and CAD sequences in the orientation that would be predicted following homologous recombination at the CAD locus, a product of the correct size could be amplified with both sets of primers. However, neither set of primers could amplify the target sequence after it had been transfected into a BHK cellular background. The reason for this is not clear. Using a set of primers specific for the neomycin resistance gene, which is part of the positive control template, a PCR product of the correct size could be amplified in all situations. The distance between the primers was about 1 kb, approximately 1.7 kb shorter than the template between the CAD and SV40 sequences in the control plasmid, and for these primers the PCR was less efficient on a chromosomal template than a plasmid template.

The predicted distance between the CAD- and SV40-specific primers at an accurately targeted CAD locus was about 2.7 kb, towards the upper limit of efficient PCR. It was tested whether the reaction might be less efficient in a cellular background, especially when crude cell lysates are used as templates. Highly purified genomic DNA from BHK cells, transfected with the PCR control plasmid, was

used as a template in the reaction, but the same non-result was obtained over a wide range of conditions. Lysates were also prepared from ts-COS cells and these preparations were 'spiked' with the PCR control plasmid. In this system also, PCR could not amplify the target fragment. As the CAD-specific and SV40 transcription termination signal primers can work on the control plasmid template but not on templates derived from a chromosomal environment, the results suggest that the state of the DNA is different, perhaps as a result of the purification techniques used, and this difference somehow effects the efficiency of the PCR reaction. Certain templates cannot be amplified by PCR because of the nature of the templates and/or primers. Some sets of primers have been known to anneal to one another or themselves and certain templates can assume secondary structures after denaturation, eg. hairpin loops, which may impede the progress of the Taq polymerase (Frohman and Martin, 1990). The reason why PCR could not be made to work in this system may be a combination of the length of target template and the cellular environment from which it is derived. As the PCR could not amplify a chromosomally located target template, the technique was discontinued as a screening procedure for identifying homologous recombinant cell lines.

The analysis of new homologous replacement fragment transfectants did not identify a homologous recombinant cell line in the 100 or so clones examined. All were negative with respect to hybridisation patterns on Southern blots, by quantitative whole cell slot blots and by PALA resistance after exposure to active T-ag. In total, 170 clones had been screened by a combination of methods but no homologous recombinant cell lines had been detected. The frequency of homologous recombination was thought to be about 1 homologous recombination per 50 integrations, as determined from pooled G418-resistant transfectants. This figure was in rough agreement with Thomas and Capecchi (1987), albeit in a different system. These workers found that if the amount of homology was approximately doubled (in their case from 4 kb to 9.5 kb), the frequency of homologous recombination rose from 1 in 40, 000 transfectants to 1 in 950. However, Doetschman *et al.* (1989) managed to correct a deleted HPRT gene in mouse

embryonic stem cells by homologous recombination using 2.5-5 kb of normal HPRT DNA at a frequency of 1.4 corrections per 10 transfected cells. These experiments show by way of example that the amount of homology used is not the only factor which can influence the frequency of homologous recombination. For the pooled homologous replacement fragment transfectants, PALA-resistant colonies were consistently obtained after exposure of the cells to active T-ag . All pools showed PALA-resistant colonies, the number of which was dependent on the length of exposure to active T-ag. A tentative figure of 1 homologous recombinant in 50 transfectants seemed reasonable, and individual clones were screened for an accurately targeted SV40 replication origin. Most research involving homologous recombination of transfected DNA in eukaryotic cells has been performed in micro-injected or electroporated mouse embryo-derived stem cells. These cells may differ from other cells in their ability to recombine DNA, especially homologously, perhaps because of their pleuripotential nature. Thus the frequency of homologous recombination quoted above may have little meaning in terminally differentiated cell lines such as ts-COS. There is no simple explanation for the discrepancy between the apparent frequency of homologous recombination in G418-resistant pools and the failure to find a homologous recombinant, except that perhaps the cellular activity of T-ag may be responsible. If these factors have an effect in the pool then why is the same effect not seen in at least one of the isolated clones?

3.8.8

Further Experiments

The objective was to obtain a cell line in which the CAD locus had been modified by the integration of an SV40 replication origin. Although such a cell line has not been isolated, the system may still have potential for other avenues of research. It was observed that episomes produced in the once-potential homologous cell line 1D6 are of a consistent size range and this precise excision event may depend on sequence duplications surrounding or including the SV40 replication origin. The different cell lines could also be used to determine whether

episomes are produced by a high frequency consistent event with little, if any, extra-chromosomal replication, or whether episomes are produced rarely, with post-excisional replication accounting for their high copy number. Another avenue of exploration may be the integration sites of the transfected fragment in the different cell lines. Although each clone has a different integration site for the homologous replacement fragment, there may be chromosomal regions into which the fragments integrate more frequently. The sequence of DNA flanking the integration site could be determined from rescued episomes after T-ag-induced excision.

3.8.9

Epilogue

If another attempt were made to construct the model system, the following changes should be considered. The length of homology should be kept the same but certain modifications to the replacement fragment could be made. A small oligonucleotide could be cloned into the homologous replacement fragment, about 500-1000 bp from the either end. Another oligonucleotide, complementary to the cloned oligonucleotide, could be used as a PCR primer so that homologous recombinants could be screened using PCR to amplify the DNA between the cloned oligonucleotide and DNA flanking the targeted integration site. The reaction should be more efficient than before as the distance from the oligonucleotide primer to the nearest CAD-specific sequence could be reduced to 500-1000 bp. Small pools of G418-resistant transfectants could be screened for homologous recombinants at an early stage using PCR, allowing minimal expansion of the transfectants. A different strategy would be to adopt the approach of Thomas *et al.* (1989), where the replacement fragment has a thymidine kinase (TK) gene ligated to one or both ends. If integration occurs at the homologous site the TK sequences are usually lost during the cross-overs that lead to integration, and selection against TK^+ cells with gancyclovir would allow an enrichment of homologous recombinants (TK^-) which could then be screened by conventional methods. If integration occurs at a non-homologous site, the TK sequences are likely to be integrated with the rest of the

transfected DNA, and these cells could be selected against as they are TK^+. Improvements might be made to the control of T-ag activity. The A gene encoding T-ag could be more tightly controlled using a highly regulatable promoter with no basal expression, or perhaps by introducing the protein rather than the gene into the cell by microinjection, cell fusion, lipofection or spheroblast fusion. A tight control of T-ag production is probably the cleanest way of regulating the T-ag activity as protein delivery systems, with the possible exception of microinjection, would involve other biological materials which could indirectly affect the response of the cell to T-ag and/or PALA.

Chapter IV

Evolution of Amplified DNA

4.1

Introduction

4.1.1

Background

To form a clearer impression of events that occur at different stages of DNA amplification, it is desirable to study the amplification of a single gene in a single cell line. Zieg *et al.* (1983) and Giulotto, Saito and Stark (1986) extensively analysed amplified DNA in BHK cells selected in a single, low concentration of PALA (single-step PALA-resistant mutants) and observed that the amount of DNA amplified in the first step of selection was very large - up to 10,000 kb. This estimate was based upon the very low number of novel joints detected in a large number of independent first-step mutants and upon earlier work which showed amplified CAD genes to be present on large marker chromosomes found in PALA-resistant cells. In contrast, highly resistant mutants selected in multiple steps of increasing PALA concentrations showed a much smaller amount of co-amplified DNA than that found in single-step mutants. These data suggest that different mechanisms may be responsible for the primary and secondary amplification events. To determine whether different mechanisms are responsible for different stages of DNA amplification, the evolution of amplified DNA was studied over several steps of selection.

This chapter describes experiments which determine the degree and stability of amplified DNA in resistant cells selected in a single, low concentration of PALA and also in derivatives of these cells, which had been selected in higher doses of the drug. An attempt to determine the resolution of an *in situ* hybridisation technique using DNA sequences which are co-amplified with CAD in single-step PALA-resistant BHK cells is also described. The results of these experiments contributed to the papers by Saito *et al.* (1989) and Smith *et al.* (1990).

4.2

Degree of Amplification in a Second Step of Selection

4.2.1

Background

The additional copies of CAD which can be obtained by a BHK cell in the initial amplification event is usually small, with up to 6 or so copies being generated (Zieg *et al.*, 1983). Using a higher dose of PALA in the initial selection step does not produce mutants containing many more copies of the CAD gene, but rather no resistant mutants survive. If a step-wise selection method is employed using gradually increasing amounts of PALA, a large number of extra CAD gene copies per cell can be generated giving mutants with high resistance to PALA and huge CAD gene copy numbers (Ardeshir *et al.*, 1983). These two different selection protocols give rise to mutant cell lines whose amplified DNA is very different. It is therefore possible that different mechanisms are involved in secondary steps of amplification than those involved in the initial amplification event. To determine if this was true, mutants which had been selected in a single, low dose of PALA were challenged with a second, much higher dose. The number of resistant colonies obtained was measured for each mutant tested.

4.2.2

Second-Step Selections at High PALA Concentrations

Three single-step mutants, originally selected in 15 μM PALA, were plated at 10^5 cells per 100 mm dish in PALA concentrations ranging from 100 to 1000 μM. Medium was changed every five days for four weeks, after which PALA-resistant colonies were stained and counted. The results are shown in table 4.1. As the number of colonies obtained for each cell line shows, at the highest concentration of PALA (1000 μM) used, mutants 15.6 and 15.2 gave resistant colonies at a frequency of 2-3 x 10^{-5} and 15.3 cells gave no resistant colonies above 200 μM.

Table 4.1

Second-Step Selections at High PALA Concentrations

To determine the frequency at which first, single-step PALA-resistant mutants could generate second-step mutants at higher PALA concentration, cells from several single-step mutants were seeded at 10^5 cells per 100 mm dish in medium containing 10% dFCS and PALA. the medium was changed every five days and colonies were stained and counted after four weeks of selection.

The table shows that the single-step mutants 15.2 and 15.6 can give rise to PALA-resistant colonies at 1000 μM PALA. The single step mutant 15.3 cannot do so, giving resistant colonies at PALA concentrations up to 200 μM.

PALA Concn (μM)	Colonies /10^5 cells of each mutant		
	15.6	15.2	15.3
100	nd	115	4
200	34	23	5
350	15	nd	0
500	9	nd	0
1000	3	2	0

nd: not determined.

Second-step mutants 15.6-500.1 and 15.6-500.2, which were selected from 15.6 cells in a fluctuation analysis experiment, and a previously isolated second-step mutant of 15.3, 15.3-120D (Saito *et al.*, 1989), were analysed for CAD gene copy number using the whole cell slot blot method. The increase in copy number was found to be 37 extra copies in 15.6-500.1, 62 extra copies in 15.6-500.2 and 68 extra copies in 15.3-120D. Therefore, a 20- to 30-fold increase in CAD copy number was achieved in a single, second step of selection at high PALA concentrations. This shows that DNA can be amplified to a greater degree in a second step of selection than it can be in the first step. This strongly implies that there are separate mechanisms involved at these different steps of selection. It is also possible that the state of the amplified DNA, after a single step of selection, is in a more accessible state for further amplification and the initial amplification mechanism is still responsible for secondary amplification events.

4.3

Stability of Amplified DNA

4.3.1

Background

In BHK cells amplified CAD genes have only been observed on intra-chromosomal arrays (Wahl *et al.*, 1982; Zieg *et al.*, 1983) and not on extra-chromosomal structures such as DMs, although in mouse cells amplified CAD genes have been found on such elements (Meinkoth *et al.*, 1987). Maintenance of these two amplified DNA states in the absence of selection is very different; DMs are rapidly lost by random segregation at cell division, but the fate of intrachromosomally amplified DNA is not clear. Zieg *et al.* (1983) studied the stability of amplified CAD genes in multi-step PALA-resistant mutants and found them to be very stable with no reduction in copy number after continuous culture in the absence of PALA for several months. The stability of amplified CAD genes in single-step mutants resistant to low

concentrations of PALA, and second-step mutants derived from these cells in higher PALA concentrations, had not been tested. To determine the stability of amplified DNA containing the CAD gene, single-step mutants and second-step mutants isolated from them (Zieg *et al.*, 1983) were assayed for loss of amplified CAD genes and PALA resistance during growth in the absence of selection as described below.

4.3.2

Loss of Amplified CAD Genes

Five first-step and three previously isolated second-step mutants were cultured in normal serum in the absence of PALA over a 90 day period. Samples of cells were removed at intervals and analysed by genomic slot blot for loss of CAD sequences. The filters were hybridised with a BHK CAD cDNA probe (Shigesada *et al.*, 1985) and quantified by densitometry. Differences in signal intensity between cell samples caused by unequal loading were corrected for by hybridisation of a BHK total genomic probe. The results are shown in table 4.2. As the table shows, first step mutants 15.3, 15.6, 30.7, and 20.2 all reached a normal, wild-type copy number of CAD genes after about 30 days of selection. The mutant 30.7 was shown to lose its amplified DNA within 103 days and although aliquots of cells were only taken at the beginning of this period and again at the end it is likely that the additional copies had been lost early on in the experiment. This would be consistent with the three other cells lines described above and also with other single step mutants tested, but not described here. The single step mutant 10.1 had lost only half its extra copies of CAD by 36 days and this may reflect an extraordinarily stable arrangement of amplified DNA in this cell line.

Table 4.2

Stability of Amplified CAD Genes in Various PALA-Resistant Mutants Selected in One or Two Steps of Selection

The stability of amplified CAD genes was determined in several single-step or two-step PALA-resistant mutant cell lines. Cells were grown in the absence of PALA and in HiFCS. At intervals, small aliquots of cells were harvested and analysed for CAD gene copy number by whole cell slot blots. Cell lines 15.3, 15.6, 10.1, 30.7 and 20.2 are independent and were isolated in a single step of selection (Zieg *et al.*, 1983); the first number refers to the concentration of PALA used and the second is the clone number. Cell lines 30.7-70B and 30.7-70.2I are independent mutants isolated in a second step of selection (at 70 μM) from 30.7 (Zieg *et al.*, 1983). 15.3-120D was previously selected in a second step of selection (120 μM) from 15.3. All cell lines had been previously isolated and been stored in liquid nitrogen until this time. The copy number analysis shows that, in general, single-step mutants lose their extra copies of CAD within the space of 30 days. Second-step mutants, in contrast, generally take about three times as long to lose additional copies of CAD.

Mutant	Days In Culture	Extra Copies of CAD/cell
15.3	0	2
	39	0
15.6	0	3
	39	0.5
10.1	7	9
	36	4.5
30.7	0	5
	103	0
20.2	7	13.5
	21	5.5
	36	0
30.7-70B	18	31
	26	20
	38	9
	82	0.5
30.7-70.2I	18	27
	46	16
	76	13
	91	3
15.3-120D	0	68
	19	17
	32	5.5
	64	0

4.4

Resolution of *In Situ* Hybridisation

4.4.1

Introduction

Amplified DNA containing the CAD gene has been shown to be unstable in the early stages of amplification in BHK cells, and when single-step mutants are selected in higher concentrations of PALA, the copy number of the CAD gene increases as the amount of co-amplified DNA decreases (Saito *et al.*, 1989). Giulotto *et al.* (1986) isolated 6 cosmids totaling 290 kb that contained sequences co-amplified with the CAD gene in a single-step BHK PALA-resistant mutant, and these were characterised and used by Saito and colleagues to study the frequency at which these sequences were amplified in further rounds of PALA selection. The cosmids have not yet been linked to each other or with CAD but none overlap with the 90 kb of contiguous DNA that has been isolated from the CAD locus. These cosmids were used as probes to determine if the fluorescent *in situ* hybridisation technique, which had been developed to observe CAD gene amplification in BHK cells (Smith *et al.*, 1990), could resolve these cosmids when they were hybridised to BHK wild type metaphase chromosomes.

4.4.2

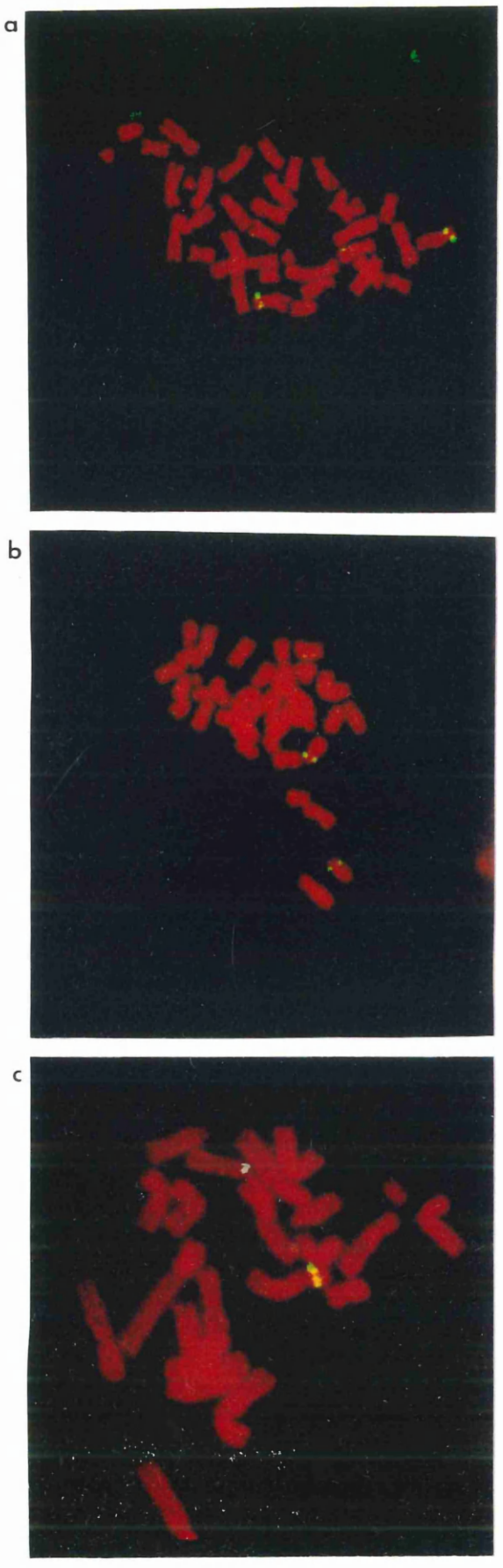
In Situ Hybridisation of Co-Amplified DNA

Cosmids representing DNA regions which had been classified by Saito *et al.* as 'near' or 'far' to the CAD gene, based upon the frequency of their co-amplification with the CAD gene in a single step PALA-resistant mutant, were hybridised to metaphase chromosomes isolated from wild type BHK cells. The 'far' cosmids are 872, 171, 671 and 971, and the 'near' cosmids are 276 and 773. Probes derived from these cosmids were used either individually or mixed with cosmid c64 which is specific to the CAD locus. Metaphase chromosomes were prepared from wild-type

Figure 4.1

The Resolution of *In Situ* Hybridisation to BHK Chromosomes

To determine if sequences coamplified with the CAD gene could be seen as separate hybridisation signals on metaphase chromosomes, a CAD-specific cosmid, cCAD64, was used simultaneously with a 'near' or 'far' cosmid, on chromosomes. These 'near' and 'far' cosmids were characterised as such based upon the frequency at which they were coamplified with CAD in a single-step PALA-resistant mutant. Metaphase chromosomes were prepared from BHK clone 9.4 cultures treated with 150 ng/ml of colcemid at 37°C for 1.5-2 hours. Metaphase chromosomes were dropped onto slides and were hybridised with either cosmid cCAD64 alone (panel a); with each 'near' or 'far' cosmid alone (panel b shows hybridisation with cosmid 171); or with a 1:1 mixture of cCAD64 and a 'near' or 'far' cosmid (panel c shows a hybridisation of mixture of cCAD64 and 171). Photographs are not shown for chromosomes probed with the other five 'near' or 'far' cosmids, either alone or with cCAD64, as all hybridisations were essentially identical to panels (a) and (b). The hybridisation pattern in panel c shows four hybridisation signals, two on each chromatid. This corresponds to one separate spot of hybridisation for each cosmid on each of the sister chromatids.


a
b
c

BHK cells (clone 9.4) and the nick-translated, biotin-labeled probes were hybridised to the chromosomes as described in section 2.6.

None of the 'far' and 'near' cosmid-derived probes hybridised to a region along the length of the chromosome distinguishable from that hybridised to by the c64-derived probe. Fluorescence microscopy of metaphase spreads hybridised to both CAD and 'near' or 'far' probes showed the same pattern of hybridisation as that seen for hybridisation of the CAD cosmid alone. This pattern is shown in figure 4.1, (panel a). However, some of the metaphase spreads that were probed with both the CAD cosmid cCAD64 and the 'far' cosmid 171, showed two spots of hybridisation on each sister chromatid s shown in panel c of figure 4.1. This was seen in approximately 20-30 % of the spreads analysed. Such a pattern of hybridisation could be consistent with the target sequences for each cosmid being sufficiently separated on the chromosomes to be seen as separated spots of hybridisation when both cosmids are simultaneously used as probes. It could also be consistent with the spatial arrangement of DNA sequences in the superhelical chromatin structure of chromosomes.

4.5

Discussion

4.5.1

Amplification in Second-Step Mutants

There is a large contrast in the degree of amplification which can be achieved in a single step of selection between first- and second-step mutants. This and work from other investigators suggests that different mechanisms are responsible for amplification at primary and secondary events. Giulotto *et al.* (1986) and recently Smith *et al.* (1990), have demonstrated that large amounts of DNA are amplified in the first step of selection. The amount of co-amplified DNA decreases in secondary steps as the number of additional CAD genes increases (Saito *et al.*, 1989). This, and the

fact that there is a large difference between the number of extra copies generated at each stage, suggests that a large amount of DNA is amplified to a small degree in the first step of selection and that secondary steps of selection involve the amplification of relatively small amounts of DNA but to a high copy number. This difference could be explained by different mechanisms being active in each situation.

On exposure to high levels of PALA, single-step mutants 15.6 and 15.2 were able to produce colonies resistant to a maximum of 1000 μM PALA. Colonies produced from another single-step mutant (15.3) were only resistant up to 200 μM, with no cells surviving at higher PALA concentrations. However, the number of additional CAD genes in second-step PALA-resistant clones previously isolated from 15.6 and 15.3, showed considerable increases in copy number which did not correlate with the PALA concentration in which the single-step mutants were reselected. For example, 15.2-120D was selected in 120μM PALA and was found to contain 68 extra copies of CAD; 15.6-500.1 and 15.6-500.2 were both selected in 500 μM PALA and have 37 and 62 extra copies, respectively.

The lack of a clear relationship between the number of extra copies of CAD and the concentration of PALA used in the second step of selection could perhaps be explained as follows: DNA containing the few extra copies of the CAD gene which are the result of the first step of selection is amplified by an unknown mechanism which is not influenced by PALA . The amplification is therefore mediated by some other regime and is terminated at a point not dependent on the PALA concentration. Only cells containing sufficient CAD genes and CAD gene product to overcome the PALA concentration will survive. It would follow that the number of additional CAD genes in different cell lines resistant to PALA would vary above a certain threshold which would be set by the concentration of the drug.

The arrangement of the amplified DNA in single-step PALA-resistant mutants may affect the frequency and extent of further steps of amplification. For example, the potential properties of inverted duplications or inverted repeats may influence the path along which the DNA is further amplified. This principle was used by Hyrien *et al.*

(1988) to develop a model for AMP deaminase amplification. Different mechanisms may amplify DNA at different rates with different efficiencies and each mechanism could be terminated at different times depending upon how that particular mechanism is regulated. These factors could all affect the degree of amplification in a second step of selection, either co-operatively or individually, leading to heterogeneity between cell lines with respect to amplified gene copy number and the concentration of inhibitor.

It is interesting that the second-step mutant 15.2-120D was found to contain 68 extra copies of the CAD gene but was selected to be resistant to 120 μM PALA. The second-step mutant 15.6-500.2 has 62 additional copies of the CAD gene but is resistant to 500 μM PALA. If 15.2-120D has a similar number of additional genes to 15.6.500.2, then why are resistant colonies not observed at concentrations above 200 μM when second-step mutants are selected from 15.2? It is possible that the arrangement of amplified DNA in 15.2 is such that a second step of amplification may occur at a frequency which is not detectable with the experimental design used, ie. at a frequency below 5×10^{-6}. If the selection was performed as a rate experiment, then a more accurate figure for the occurrence of second-step resistance at the concentrations of PALA used could be determined. Alternatively, if more cells were used in the selection, a detectable frequency of second-step resistance could be detected. The cell line 15.2-120D was isolated from a fluctuation analysis experiment, but it is not known at what frequency second-step mutants of 15.2 were observed at different PALA concentrations. If 120 μM was the highest PALA concentration in which 15.2-120D could grow it would imply that not all the additional copies of CAD were functional or, if they were, then additional factors were responsible for the high level of resistance of the 15.6 second-step mutants.

In conclusion, 15.6 and 15.3 cells were able to give rise to second-step PALA-resistant mutants at PALA concentrations up to 1000 μM. Single-step mutant 15.2 could only give resistant colonies up to 200 μM. This difference could be a reflection of the structure of the amplified DNA after the first step of selection in each

cell line. These differences could affect the frequency and degree of amplification in the second-step mutants by influencing the mechanism responsible for further amplification events.

4.5.2

Loss of Amplified DNA

The loss of amplified DNA containing additional CAD genes in the absence of PALA falls into three categories. The results shown in table 4.2 show two of these classes. PALA-resistant mutants which were selected in a single step lose their additional CAD genes within 30 days of growth in the absence of PALA; mutants derived from these cells in a second step of selection take about 90 days to return to a normal wild-type copy number of CAD. Previous experiments, described by Kempe *et al.* (1989), demonstrated that PALA resistance in serially-selected cells was very stable with no loss during one year of growth in the absence of PALA. These results are consistent with a model in which the DNA containing amplified CAD genes is continually being rearranged with the loss of sequences for which there is no selection. If PALA is removed from the growth medium then there is no selection for the maintenance of the CAD genes and these are also lost. Therefore, in single-step mutants the DNA has been rearranged to the least degree and could be the most prone to further rearrangements, ie. single-step mutants contain the least stable amplified DNA compared to second- and multiple-step mutants. Second-step mutants have undergone further amplification events and this is reflected in the reduction of co-amplified DNA and the increase in CAD copy number. The rearrangements which the amplified DNA has undergone have produced a more stable array, and this is reflected in the length of time the cell line requires to dispose of the extra copies of CAD. Loss of extra CAD genes and PALA resistance could be seen as the loss of cells which contain such DNA from a population. Cells without amplified DNA may have a growth advantage over those that contain amplified arrays, especially in the case of single-step mutants which have been seen to carry amplified

regions up to tens of megabases in size (described by Trask and Hamlin, 1989 and Smith *et al.*, 1990). Several single-step PALA-resistant mutants, including cell lines 15.3, 15.6 and 20.2, were analysed by fluorescent *in situ* hybridisation with a CAD probe, and extra copies of the CAD gene were located on extended chromosomal regions (Smith *et al.*, 1990) suggesting that the rapid loss of amplified sequences in these cell lines was not due to the loss of extra-chromosomal elements through random segregation at mitosis. Although these cell lines are mature with respect to their expansion since isolation (having reached about 10^{15} cells at the point of analysis), clones examined by Smith *et al.* at the 10^5 cell stage also showed extended chromosomal regions albeit more extended than the more mature cell lines. This suggests that from a very early stage DNA containing the CAD gene is amplified intra-chromosomally and it is from these chromosomal regions that amplified DNA is lost or further amplified.

To summarise, amplified DNA containing additional CAD genes is lost from single-step mutants within 30 days of growth in the absence of PALA, and from second-step mutants within 90 days. Previous work had shown that multiple-step mutants have extremely stable amplified arrays of the CAD gene with no loss after 1 year of growth in the absence of PALA. These data suggest that amplified DNA is being continually rearranged into more stable structures with each step of amplification. This process may be related to the loss of co-amplified DNA and the preferential amplification of CAD genes in secondary amplification events.

4.5.3

Resolution of *In Situ* Hybridisation

None of the co-amplified DNA sequences represented by the six 'near' and 'far' cosmids gave additional spots of hybridisation when used to probe the CAD locus of wild-type BHK cells. On some occasions cosmid 171 hybridised to a region distinct from the CAD signal seen as two additional spots of hybridisation across the chromosome in the same region as CAD. This phenomenon has been observed

before (Lawrence *et al.*,1988) and is thought to be caused by the superhelical arrangement of DNA in chromatin. The lack of other separate hybridisation spots with these cosmids puts the 'near' and 'far' cosmids within an estimated few megabases of the CAD gene. This estimate suggests that, as the 'near and far' sequences are fairly close to the CAD gene and are not amplified to the same degree as CAD in subsequent amplification events, the mechanisms responsible for secondary amplification events do not act on the complete structure formed in the first step of amplification, but on a sub-region close to the amplified CAD sequences.

Chapter V

DNA Amplification in Human Tissue Culture Cells

5.1

Introduction

5.1.1

Background

Several systems, including Syrian hamster cells, rat cells and Chinese hamster ovary cells, have been used to investigate the nature and the mechanism(s) responsible for amplification of the CAD gene. Most data to date has been obtained using molecular biological techniques and an important limitation of this approach is the number of divisions a cell containing a new amplification must progress through until sufficient cells exist for DNA isolation and analysis. For example, a cell may have to divide 20-30 times before sufficient DNA is present for a comprehensive Southern blot analysis and during this time the initial amplified structure can be obscured by rearrangements of the DNA. With the advent of fluorescent *in situ* hybridisation of metaphase chromosomes it is possible to visualise new amplification events at a very early stage and it is desirable to study these events in a human system. The human karyotype is well known and there are catalogues of cloned and mapped human DNA sequences. Probes that have been mapped relative to the CAD gene can be used to determine which sequences are included at different times of the amplification process. The human CAD gene has been mapped to the short arm of chromosome 2 (figure 5.1; Chen *et al.*, 1989) and clones for this region were obtained from Dr J. A. Davidson, University of Kentucky, USA.

5.2

PALA-Resistant Cultured Human Cells

5.2.1

Selection of Cell Lines

It was not known which human cell lines would be amenable to selection in PALA or at what PALA concentrations resistant mutants could be selected. Three cell lines were tested for resistance to PALA: fibrosarcoma-derived HT 1080 cells,

Figure 5.1a

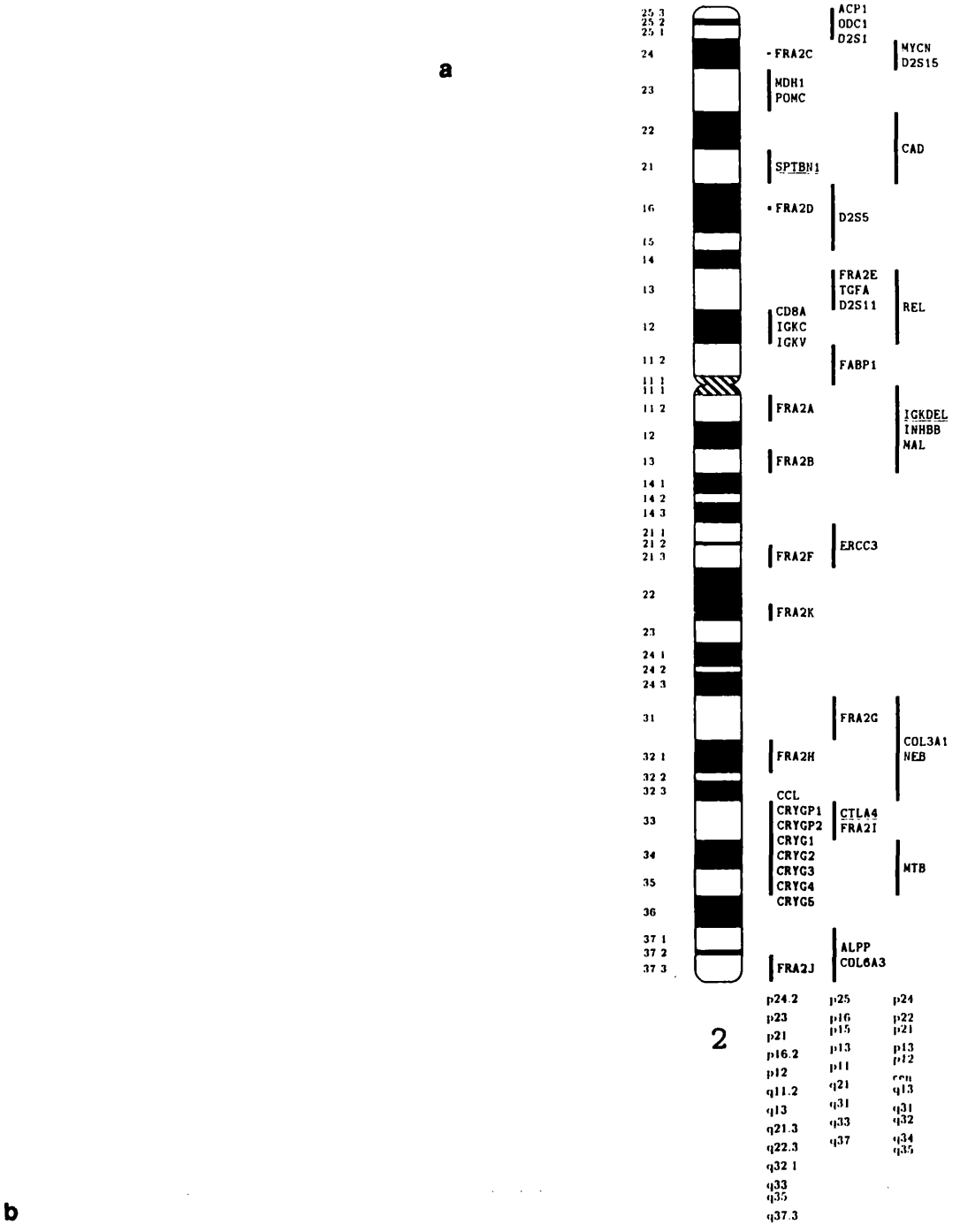
Human Chromosome 2

The human CAD gene is located at position 2p21-23 on the short arm of this chromosome. The diagram was obtained from the Yale-HHMI Gene Mapping Library from material published as part of the Human Gene Mapping Workshop 10 (Povey and Falk, 1989).

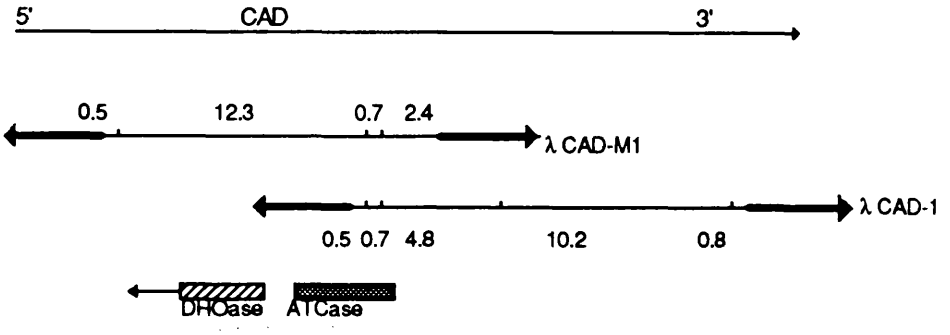
Figure 5.1b

Bacteriophage λ Constructs Containing Human CAD Sequences

Two clones containing genomic CAD sequences were provided by J. Davidson of the University of Kentucky. The EcoR I restriction map is shown and includes the coding regions for aspartate transcarbamoylase (ATCase) and dihydroorotase (DHOase).


a
25 3
25 2
25 1
24
23
22
21
16
15
14
13
12
11 2
11 1
11 1
11 2
12
13
14 1
14 2
14 3
21 1
21 2
21 3
22
23
24 1
24 2
24 3
31
32 1
32 2
32 3
33
34
35
36
37 1
37 2
37 3
FRA2C
MDH1
POMC
SPTBN1
FRA2D
CD8A
IGKC
IGKV
FRA2A
FRA2B
FRA2F
FRA2K
FRA2G
FRA2H
CCL
CRYGP1
CRYGP2
CRYG1
CRYG2
CRYG3
CRYG4
CRYG5
FRA2J
ACP1
ODC1
D2S1
D2S5
FRA2E
TGFA
D2S11
FABP1
ERCC3
CTLA4
FRA2I
ALPP
COL6A3
MYCN
D2S15
CAD
REL
IGKDEL
INHBB
MAL
COL3A1
NEB
MTB
2
p24.2
p23
p21
p16.2
p12
q11.2
q13
q21.3
q22.3
q32 1
q33
q35
q37.3
p25
p16
p15
p13
p11
q21
q31
q33
q37
p24
p22
p21
p13
p12
q13
q31
q32
q34
q35

b


5'
CAD
3'
0.5
12.3
0.7
2.4
λ CAD-M1
λ CAD-1
0.5 0.7 4.8
10.2
0.8
DHOase
ATCase

colon carcinoma-derived HT 29 cells and HeLa S3 monolayer cells. All three lines were individually plated at either 10^5 cells or 10^6 cells per 100 mm dish at 20, 50 or 80 μM PALA. Medium was changed every five days for three weeks after which the plates were fixed, stained with giemsa and examined for colonies. All the cell lines gave PALA-resistant colonies at the lower concentrations of PALA, ie. 20 and 50 μM, but no cells survived at 80 μM. HT 1080 cells were used for further selections because they appeared to be the most stable of the three cell lines and had the most 'normal' karyotype. HeLa cells have an uncharacteristic heteroploid karyotype and HT 29 cells are known to differentiate when they grow beyond the logarithmic phase of growth.

5.2.2

The Rate of PALA Resistance in HT 1080 Cells

The rate of PALA resistance was determined for HT 1080 cells to determine how often new amplification events involving the CAD gene occur in this system. This quantity is more meaningful than the frequency of PALA resistance as it excludes pre-existing amplifications, and for any colonies which are cloned from a rate experiment, it can be determined if the amplification was pre-existing or whether it has arisen within the last few cell divisions before the selection was applied. This is done by counting the number of PALA-resistant colonies present on each plate of selection. If an amplification pre-existed in the population, then during growth of a small number of cells prior to selection, a large number of resistant offspring would arise. Therefore, following selection with PALA, a large number of resistant colonies will be seen. If, for example, one PALA-resistant colony is seen at the end of a rate experiment, then the amplification event must have taken place during the cell generation immediately prior to selection as there are no descendants of the founding cell to form other PALA -resistant colonies.

Small populations of 1×10^3 HT 1080 cells each were expanded to 2.5×10^5 cells, dispersed in 100 mm dishes and selected in 30, 40 or 50 μM PALA. Medium was changed every five days for four weeks after which the dishes were examined by light microscopy for colonies. One to three colonies were ring-cloned from dishes

that appeared to contain less than five colonies. After ring-cloning, cells remaining on the dishes were fixed, stained with giemsa and the dishes were examined for colonies. Clones isolated from dishes which were found to have contained more than five colonies were discarded. The number of colonies from all the dishes were recorded and used to calculate the rates of resistance to PALA for these cells, using the tables of Capezzi and Jameson (1973). The rate of PALA resistance in HT 1080 cells was found to be 1-3 x 10^{-5} events per cell per generation. This number is in line with the rates of gene amplification observed in other cell systems where amplification has been shown mediate drug resistance. The fluctuation analysis experiment described above was repeated and instead of staining the colonies observed, several PALA-resistant colonies were cloned. These cell lines were analysed by *in situ* hybridisation and by Southern blotting.

5.2.3

Southern Blot Analysis of PALA-Resistant HT 1080 Mutants

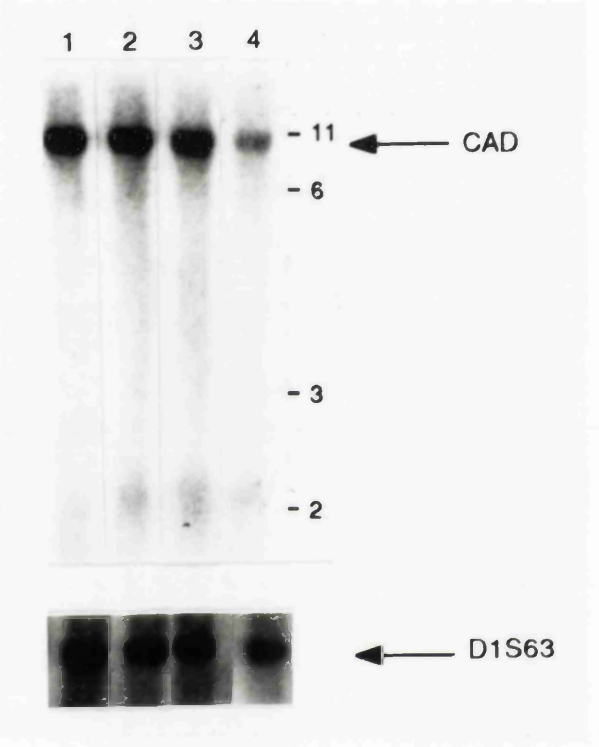
To confirm that gene amplification of the CAD gene was responsible for the PALA resistance seen in these colonies, several newly arisen PALA-resistant mutants were analysed by Southern blot. If CAD amplification had occurred in these cells then an increase in CAD copy number should be evident.

Several HT 1080 PALA-resistant colonies were cloned following selection in a fluctuation analysis experiment. When each clone reached approximately 4 x 10^5 cells, two aliquots of 1 x 10^5 cells each were frozen in liquid nitrogen and the remaining cells were expanded and used as a source of genomic DNA for Southern blot analysis. Genomic DNA was isolated, restricted with EcoR I and Southern blotted after electrophoresis in an agarose gel. The nylon filter was hybridised with a hamster CAD cDNA probe and later with a single copy reference probe. The DNA fragment D1S63, provided by Dr. N. Spurr (ICRF, Clare Hall) and mapped to human chromosome 1, was used as a single copy control probe as it was an unlikely sequence to be coamplified with CAD DNA. The copy number of CAD in several of the PALA-resistant clones was determined by densitometry of the signals obtained

Figure 5.2

Southern blot of PALA-Resistant HT 1080 Cell lines

Genomic DNA was isolated from each cell line and restricted with EcoR I. Digested DNA was size fractionated and transferred to a nylon membrane. The membrane was hybridised to a CAD probe (pCAD142) and later to a single copy probe, D1S63. The single copy probe had been mapped to chromosome 1 and is a single copy genomic fragment. The signals for each probe were analysed by densitometry to determine the degree of CAD amplification. Lanes are loaded with DNA isolated from the following PALA-resistant HT 1080 cell lines : (1) 40.11b; (2) 40.6a; (3) 40.5a; and (4) HT 1080 wild-type cells. Metaphase chromosomes from these cell lines were also analysed by fluorescent *in situ* hybridisation. The cell line 40.5a is shown in figure 5.3, panel d. Figure 5.4, panel a, shows the cell line 40.6a, and panel c of the same figure shows the cell line 40.11b. The number of fluorescent spots correlates with the CAD copy number derived from the Southern blot.


1
2
3
4
- 11
- 6
- 3
- 2
CAD
D1S63

on the Southern blots shown in figure 5.2. The cell lines 40.5a, 40.6a and 40.11b were found to have 1.5-, 1.1- and 1.6-fold amplifications of the CAD gene, respectively. This is in agreement with the number of fluorescent spots seen by *in situ* hybridisation, using a CAD probe, on these cell lines (see below).

5.2.4

In Situ Analysis of PALA-Resistant HT 1080 Clones

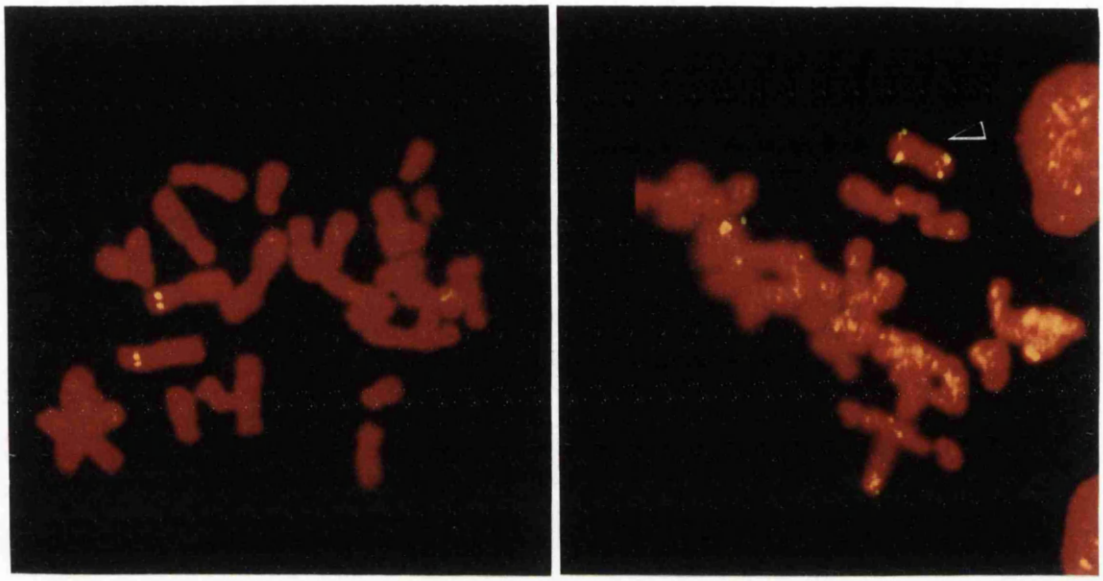
To determine the arrangement amplified CAD genes in the PALA-resistant HT 1080 mutants, several cell lines were analysed by fluorescent *in situ* hybridisation. Each PALA-resistant clone was recovered from liquid nitrogen storage and expanded to approximately 5 x 10^6 cells. Colcemid was added to 50 ng/ml and the cultures were maintained at 37°C for 45 minutes. Cells arrested in metaphase were dislodged from the flasks by shaking and metaphase chromosomes were dropped on to glass microscope slides. Slides from each cell line were hybridised with a human CAD probe derived from the bacteriophage clones λCAD-1 and λCAD-M1 shown in figure 5.1b. The results of the *in situ* analyses are shown in figures 5.3 and 5.4.

The hybridised spreads from 9 independent cells lines showed an increase in the copy number of CAD hybridising sequences. Eight of these cell lines showed a similar arrangement of amplified CAD genes. The extra copies of CAD reside near the telomeres of each arm of a novel chromosome. This chromosome was seen in over 90% of all hybridised spreads of each line and when absent, only a single wild-type chromosome was observed. Cell line 50.4a does not have this arrangement. Instead almost every hybridised spread of this cell line had the additional CAD genes located at regular intervals on what could be an extended p arm of chromosome 2.

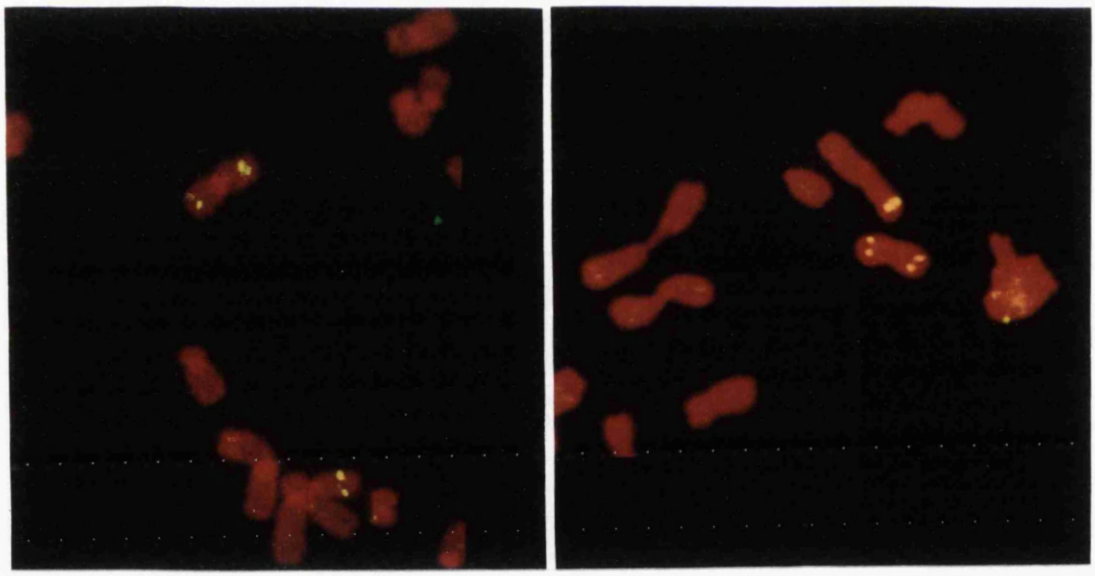
Figure 5.3

***In Situ* Analysis of PALA-Resistant HT 1080 Clones (I)**

To determine the structure of the amplified DNA in PALA-resistant HT 1080 clones, metaphase chromosomes from each cell line were prepared and probed with a mixture of biotinylated λCAD1 and λCAD-M1. The results show that eight out of nine independent cell lines, the same pattern is seen. In the exceptional cell line (mutant 50.4a), a ladder-like structure is observed. This mutant was also the only resistant cell line to show more than one additional spot of hybridisation. The eight similar mutants all carry what looks to be a peri-centromeric inversion of the short (p) arm of human chromosome 2. Each panel shows a representitive metaphase spread from one of the following mutant cell lines (except panel a): panel (a) wild-type HT 1080 cells; panel (b) 20.2a (chromosome carrying the amplified CAD genes is arrowed); panel (c) 30.6a; panel (d) 40.5a. Nomenclature is as follows: the initial number indicates the concentration of PALA in which the clone was selected, the number after the decimal point indicates the dish from which the colony was cloned and the ending letter designates the sequence in which the colony was cloned. For example, 50.5b (figure 5.4) was selected in 50 μM PALA and was the second colony to be cloned from plate number 5.



a b

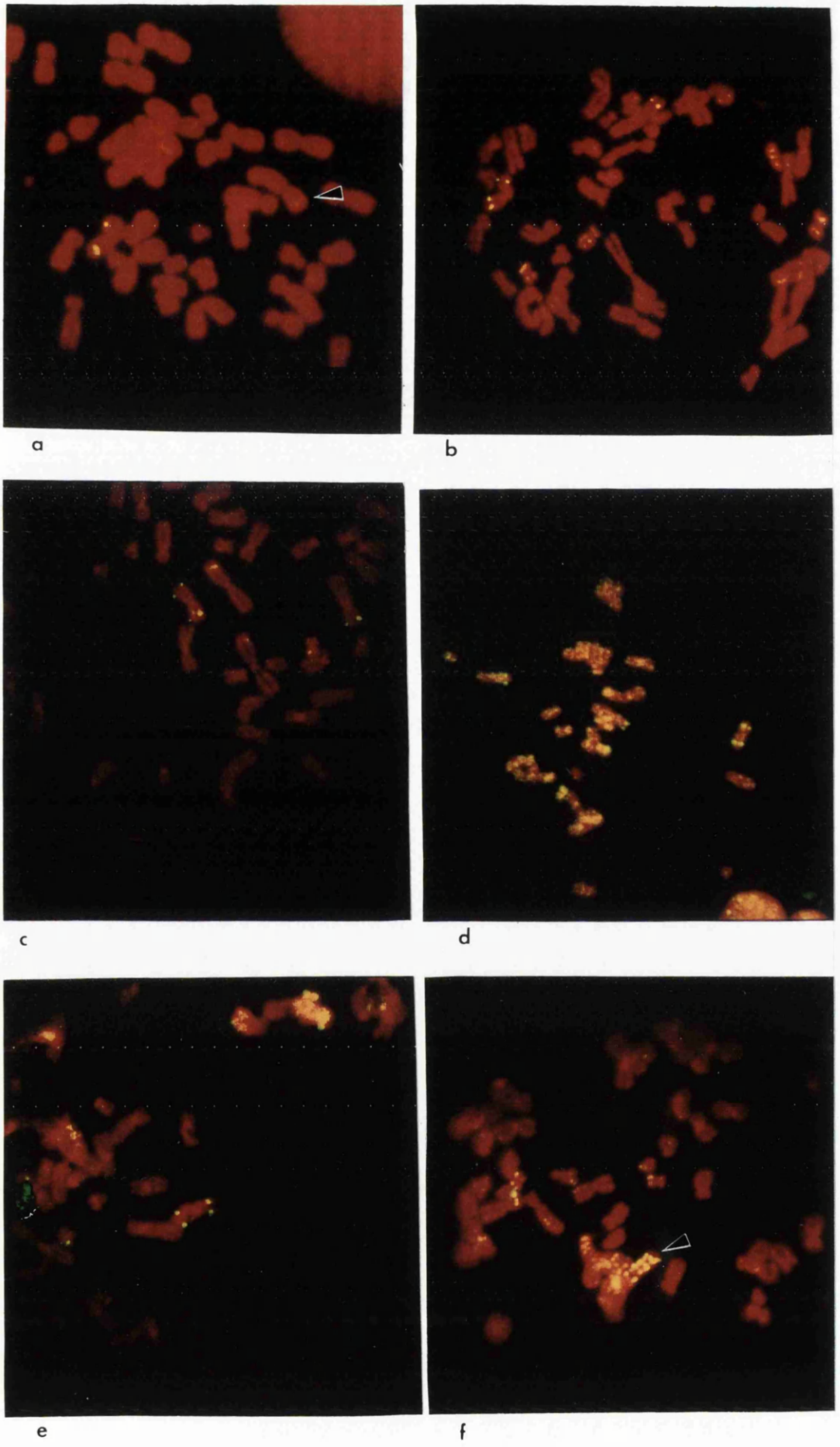


c d

Figure 5.4

***In Situ* Analysis of PALA-Resistant HT 1080 Clones (II)**

M taphase chromosomes from each cell line were prepared and probed with a mixture of biotinylated λCAD1 and λCAD-M1as described in the legend of figure 5.3. Chromosome spreads are from the following PALA-resistant clones: panel (a) 40.6a (unaltered chromosome 2 is arrowed); panel (b) 40.10a; panel (c) 40.11b; panel (d) 50.5b; panel (e) 50.3a; panel (f) 50.4a ('ladder' of amplified CAD genes is arrowed).


a
b
c
d
e
f

Discussion

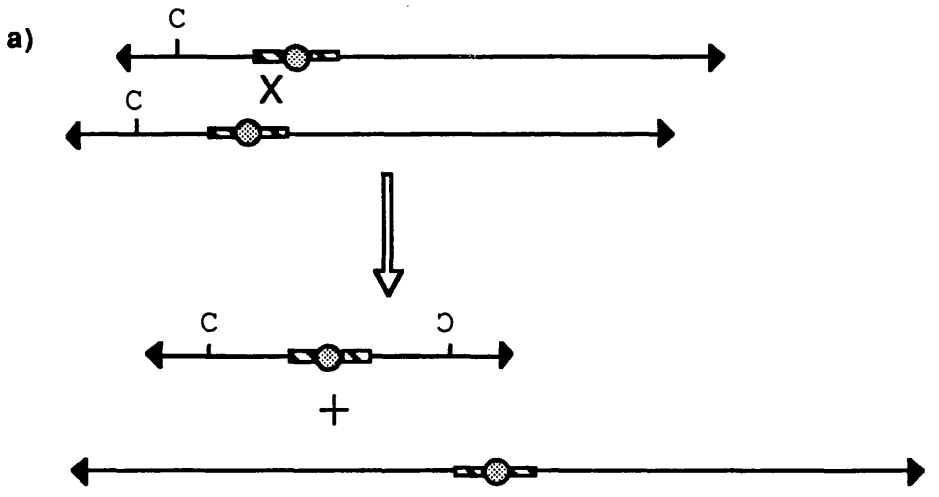
The *in situ* data and Southern blot analysis both confirm that PALA resistance in HT 1080 cells is mediated by gene amplification. The level of CAD amplification is low at 1 extra copy per cell as estimated by Southern blot and by the number of spots of hybridisation on metaphase chromosomes hybridised to a CAD probe.

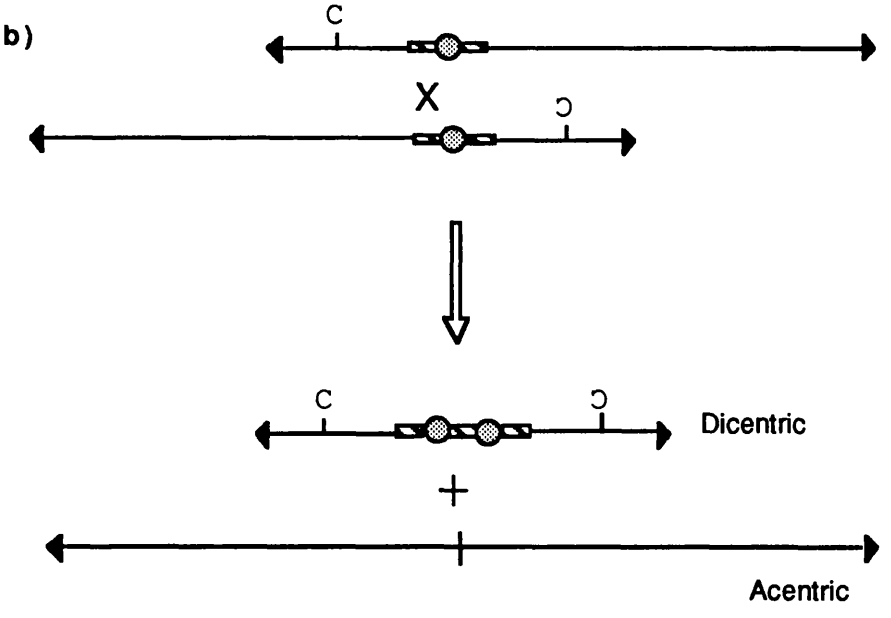
As panel (a) of figure 5.3 shows, HT 1080 cells are diploid for CAD and the genes are located in a position consistent with the mapping data (see figure 5.1). Resistant HT 1080 clones isolated in PALA concentrations of 20, 30, 40 and 50 μM all showed the same simple pattern of amplified CAD genes with the exception of one cell line which is discussed below. Amplification appears to involve the fusion of two p arms of chromosome 2, with one centromere separating the two arms. The length of each arm of this novel chromosome is similar to that of the p arm of chromosome 2 and in all spreads this structure was never seen in the presence of both chromosome 2 alleles. Therefore the novel chromosome would appear to be derived from human chromosome 2. Panel (c) of figure 5.4 shows two wild-type chromosomes and the novel chromosome, but one chromosome 2 has been contributed by a neighbouring spread. A model for the fusion of two chromosome 2 p arms can be developed based on the model of Smith *et al.* (1990) who describe the amplification of CAD in BHK cells at an early stage after a single step of selection. The extra copies of CAD are regularly spaced along an extended arm of the BHK chromosome B9, the usual location of the unamplified CAD gene. The model proposed by Smith *et al.* (1990) involves the transfer of an entire chromosome arm from one B9 sister chromatid to the other via recombination between telomeric and peri-centromeric sequences on the same arm of two adjacent sister chromatids. An adaptation of this model can give the novel chromosome seen in PALA-resistant HT 1080 cells. If recombination were to occur between centromeric sequences on the long arm of one sister chromatid and on the short arm of the other sister chromatid (ie. on either side of the centromere), the outcome would be two novel chromosomes. One would comprise two long arms and the other two short arms, and both would retain a centromere. A schematic diagram is shown in figure 5.5. If recombination were to occur between the centromeric

Figure 5.5

A Model for Gene Amplification in Human HT 1080 Cells

This schematic shows the outcome of recombination between peri-centromeric sequences of each sister chromatid. Panel (a) shows recombination between these sequences on opposite sides of the centromere, resulting in the formation of two novel chromosomes. Each new chromosome contains a duplication and inversion of either the long or short chromosomal arm and at cell division each daughter cell will receive only one of these chromosomes. The daughter cell containing the duplicated short arm, therefore receives one extra copy of the CAD gene. Panel (b) shows the outcome of recombination between these sequences on the same arm of each sister chromatid. A dicentric and an acentric chromosome are produced with each chromosome containing an inverted duplication of either the long or the short chromosomal arm. Recombination between sequences on the other side of the centromere would result in an acentric chromosome containing a duplicated short arm and a dicentric chromosome containing a duplicated long arm.


a)
C
X
C
C
C
+


b)
C
X
C
C
C
Dicentric
+
Acentric

sequences of the long arm or short arm of each sister chromatid (ie. on the same side of the centromere) then the outcome would be a dicentric chromosome and deletion of the chromosome arms distal to the recombination region. A dicentric chromosome can also be formed if a break occurs in one chromosome arm, causing loss of the telomere and resulting in the fusion of the free chromosome ends (Kaufman, Sharp and Latt, 1983; Cowell and Miller, 1983). In the cell division following the formation of a dicentric chromosome a break can occur at some point between the two centromeres, leading to the transfer of chromosomal material from the arm of one chromosome to the arm of another. This can form heterogeneous amplification products, including inverted duplications, and chromosomal deletions. An acentric chromosome can be formed by deletion of a chromosomal region or by non-homologous sister chromatid exchange mediated by recombination between centromeric and telomeric sequences on the same chromatid arm. The acentric chromosome may segregate unequally at mitosis, increasing the copy number of the resident sequences, and integrate into the genome (Windle *et al.*, 1991).

The novel chromosome, possibly containing two inverted chromosome 2 p arms, has been observed in 8/9 independent cell lines selected at different PALA concentrations. This implies that the same mechanism is responsible for amplification in these eight lines and also that the structure formed in the first amplification step is stable, at least up to the 5×10^6 cell stage.

Had more time been available the proposed model could have been tested in several ways. The nature of the novel chromosome could be confirmed as being from chromosome 2 by at least two experiments. First, it is possible to band chromosomes after *in situ* hybridisation with giemsa or other stains and this methodology could be employed to determine the banding pattern of the novel chromosome and compare it to that of chromosome 2p. Second, additional probes could be selected from the p arm of chromosome 2 at varying distances from CAD and could be used in conjunction with a CAD probe in *in situ* hybridisation experiments on PALA-resistant cells. The probes could be used in a two-colour fluorescence detection system so that each could be separately identified, thus the relationship between CAD and these 'near' or 'far' sequences in the novel chromosome could be determined.

Cell line 50.4a was an exception in which the inverted novel chromosome was not seen. Instead, the amplified CAD genes in this cell line were regularly spaced along an extended chromosome arm with the unamplified gene located at its normal position with respect to the centromere. This pattern of amplified genes is reminiscent of that observed by Smith *et al.* (1990), who interpreted their observations in terms of unequal non-homologous sister chromatid exchanges. Possibly the same mechanism is responsible for the structure observed in 50.4a cells.

In conclusion, it would appear that amplification of the CAD gene in human HT 1080 cells is mediated by one particular mechanism in the majority of cases observed to date. A model for this event involves recombination between peri-centromeric sequences on opposite sides of the centromeres of each sister chromatid. Such an event would generate one chromatid comprising the two short arms and a second chromatid comprising the two long arms. Both would contain a centromere and after mitosis each daughter cell would receive one or the other chromosome - only the daughter cell containing the inverted short arm chromosome would survive selection with PALA. If recombination were to occur between the centromeric sequences of one sister chromatid and the telomeric sequences of the same arm of the other sister chromatid then the regularly spaced array seen in 50.4a cells could be achieved according to the model of Smith *et al.* (1990). It is likely that in the HT 1080 system DNA amplification is mediated by non-homologous unequal sister chromatid exchange between the centromeric regions of sister chromatids. Depending on the relative location of each of the participating regions, the outcome of the exchanges can differ. However, in HT 1080 cells recombination between the short arm centromeric sequences of one sister chromatid and the long arm centromeric sequences of the other sister chromatid appears to be favoured, giving rise to a metacentric chromosome comprising the duplication and inversion of chromosome arm 2p.

References

Ardeshir, F., Giulotto, E., Brison, O., Liao, W. S. L., and G. R. Stark (1983). Structure of amplified DNA in different Syrian hamster cell lines resistant to N-(phosphonacetyl)-L-aspartate. Mol. Cell. Biol. *3*, 2076-2088.

Biedler, J. L., Chang, T-d., Scotto, K. W., Malera, P. W., and Spengler, B. A. (1988). Chromosomal organization of amplified genes in multidrug resistant Chinese hamster cells. Cancer Res. *48*, 3179-3187.

Beitel, L. K., McArthur, J. G., and Stanners, C. P. (1991). Nucleotide sequence requirements for the stimulation of gene amplification by a mammalian genomic element. Gene (in press).

Botchan, M., Topp, W., and Sambrook, J. (1979). Studies on Simian Virus 40 excision from cellular chromosomes. Cold Spring Harbor Symposium on Qualitative Biology *43*, 709-719.

Botchan, M., Stringer, J., Mitchison, T., and Sambrook, J. (1980). Integration and excision of SV40 DNA from the chromosome of a transformed cell. Cell *20*, 143-152.

Borowiec, J. A., Dean, F. B., Bullock, P. A., and Hurwitz, J. (1990). Binding and unwinding - how T antigen engages the SV40 origin of replication. Cell *60*, 181-184.

Brady, J. N., and Salzman, N. P. (1986). In The Papovaviridae, N.P. Salzman, ed. (New York: Plenum Press.), pp 1-25.

Bullock, P., and Botchan, M. (1982). In Gene Amplification, R. T. Schimke, ed. (New York: Cold Spring Harbour Laboratory), pp 215-224.

Butel, J. S. (1986). SV40 large T antigen: dual oncogene. Cancer Surveys *5*, 343-365.

Campbell, J. L. (1986). Eukaryotic DNA replication. Ann. Rev. Biochem. *55*, 733-771.

Capezzi, R. L., and Jameson, J. W. (1973). A table for the estimation of the spontaneous mutation rate of cells in culture. Mutat. Res. *17*, 147-148.

Carroll, S. M., DeRose, M. L., Gaydray, P., Moore, C. M., Needham-Vandevanter, D. R., Van Hoff, D. D., and Wahl, G. M. (1988). Double minute chromosomes can be produced from precursors derived from a chromosomal deletion. Mol. Cell. Biol. *8*, 1525-1533.

Carroll, S. M., Gaudray, P., DeRose, M.L., Emery, J. F., Melnkoth, J. L., Nakkim, E., Subler, M., Van Hoff, D. D., and Wahl, G. M. (1987). Characterization of an episome produced in hamster cells that amplify a transfected CAD gene at high frequemcy: functional evidence for a mammalian replication origin. Mol. Cell. Biol. *7*, 1740-1750.

Challberg, M. D., and Kelly, T. J. (1989). Animal virus replication. Ann. Rev. Biochem. *58*, 671-717.

Chen, K-C, Vannais, D. B., Jones, C., Patterson, D., and Davidson, J. N. (1989). Mapping of the gene encoding the multifunctional protein carrying out the first three steps of pyrimidine biosynthesis to human chromosome 2. Hum. Genet. *82*, 40-44.

Chu, G., Hayakawa, H., and Berg, P. (1987). Electroporation for the efficient transfection of mammalian cells with DNA. Nuc. Acid. Res. *15*, 1311-1326.

Colantuoni, V., Dalley, L., and Basilico, C. (1980). Amplification of integrated viral DNA sequences in Polyoma virus transformed cells. Proc. Natl. Acad. Sci. USA. *77*, 3850-3854.

Colantuoni, V., Dalley, L., Della Valle, G., and Basilico, C. (1982). Requirements for excision and amplification of integrated viral DNA molecules in Polyoma virus transformed cells. J. Virol. *43*, 617-628.

Collins, K. D., and Stark, G. R. (1971). Aspartate transcarbamylase. Interaction with the transition state analogue *N*-(phosphonacetyl)-L-aspartate. J. Biol. Chem. *246*, 6599-6605.

Conrad, S. E., Liu, C.-P., and Botchan, M. R. (1982). Fragment spanning the SV40 replication origin is the only DNA sequence required in *cis* for viral excision. Science *218*, 1223-1225.

Cowell, J. K. (1982). Double minutes and homogeneously staining regions: gene amplification in mammalian cells. Ann. Rev. Genet. *16*, 21-59.

Cowell, J. K., and Miller, O. J. (1983). Occurance and evolution of homogeneously staining regions may be due to breakage-fusion-bridge cycles following telomere loss. Chromosoma *88*, 216-221.

Crawford, L. V., Cole, C. N., Smith, A. E., Paucha, E., Tegtmeyer, P., Rundell, K., and Berg, P. (1978). Organization and expression of the early genes of Simian Virus 40. Proc. Natl. Acad. Sci. USA *75*, 117-121.

Deldakis, C., and Kafatos, F. C. (1989). Amplification enhancers and replication origins in the autosomal chorion gene cluster of Drosophila. EMBO. J. *8*, 891-901.

Delidakis, C., Swimmer, C., and Kafatos, F. C. (1989). Gene amplification: an example of genome rearrangement. Current Opinion in Cell Biology *1*, 488-496.

DePamphilis, M. L., and Bradley, M. K. (1986). In The Papovaviridae, N. P. Salzman, ed. (New York, Plenum Press) pp. 99-246.

Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S., and Smithies, O. (1989). Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature *330*, 576-578.

Dora, S., Schwarz, C., and Knippers, R. (1989). Excision of integrated Simian Virus 40 DNA involving homologous recombination between viral DNA sequences. J. Mol. Biol. *206*, 81-90.

Feinberg and Vogelstein (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. *132*, 6-13.

Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., van de Voorde, A., van Heuverswyn, H., van Herreweghe, J., Volckaert, G., and Ysebert, M. (1978). Complete nucleotide sequence of SV40 DNA. Nature *273*, 113-120.

Ford, M., Davies, B., Griffiths, M., Wilson, J., and 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., and Fried, M. (1986). Large inverted duplications are associated with gene amplification. Cell *45*, 425-430.

Frohman, M. A., and Martin, G. R. (1990). In: PCR Protocols, M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White, eds. (San Diego. Acedemic Press). pp 228-238.

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

Giulotto, E., Saito, I., and Stark, G. R. (1986). Structure of DNA found in the first step of CAD gene amplification. EMBO J. *5*, 2115-2121.

Gurney jr, T., and Gurney, E. G. (1989). Spontaneous rearrangement of integrated Simian Virus 40 DNA in nine transformed rodent cell lines. J. Virol. *63*, 165-174.

Hamlin, J. L., Milbrandt, J. D., Heintz, N. H., and Azizkhan, J. C. (1984). DNA sequence amplification in mammalian cells. Annu. Rev. Cyto. *90*, 31-82.

Hay, R. T., and DePamphilis, M. L. (1982). Initiation of SV40 DNA replication *in vivo*; location and structure of 5' ends of DNA synthesized in the *ori* region. Cell *28*, 767-779.

Hirt, B. (1967). Selective extraction of Polyoma DNA from infected mouse cell cultures. J. Mol. Biol. *26*, 365-369.

Hoy, C. A., Rice, G. C., Kovacs, M., and Schimke, R. T. (1987). Over-replication of DNA in S-phase Chinese Hamster Ovary cells after DNA synthesis inhibition. J. Biol. Chem. *262*, 11927-11934.

Huberman, J. A. (1987). Eukaryotic DNA replication: a complex picture partially clarified. Cell *48*, 7-8.

Hyrien, O., Debatisse, M., Buttin, G., and Robert de Saint Vincent, B. (1987). A hotspot for novel amplification joints in a mosaic of Alu-like repeats and palindromic A+T rich DNA. EMBO J. *6*, 2401-2408.

Hyrien, O., Debatisse, M., Buttin, G., and Robert de Saint Vincent, B. (1988). The multicopy appearance of a large inverted duplication and the sequence at the inversion joint suggest a new model for gene amplification. EMBO J. *7*, 407-417.

Kaufman, R. J., Sharp, P. A., and Latt, S. A. (1983). Evolution of chromosomal regions comtaining transfected and amplified dihydrofolate reductase sequences. Mol. Cell. Biol. *3*, 699-711.

Kempe, T. D., Swyryd, K. A., Bruist, M., and Stark, G. R. (1976). Stable mutants of mammalian cells that overproduce the first three enzymes of pyrimidine nucleotide biosynthesis. Cell *9*, 541-550.

Livingston, D. M., and Bradley, M. K. (1987). The Simian Virus 40 large T antigen. A lot packed into a little. Mol. Biol. Med. *4*, 63-80.

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

Luria, S. E., and Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. Genetics *28*, 491-511.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular cloning:a laboratory manual. (New York: Cold Spring Harbor Laboratory).

Mansour, S. L., Thomas, K. R., and Capecchi, M. R. Disruption of the proto-oncogene *int-2* in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature *336*, 348-352.

Mariani, B. D., and Schimke, R. T. (1984). Gene amplification in a single cell cycle in Chinese Hamster Ovary cells. J. Biol. Chem. *259*, 1901-1910.

Martin, R. G. (1981). The transformation and transmogrification of DNA synthesis by Simian Virus 40. Adv. Cancer Res. *34*, 1-68.

Matera, A. G., Weiner, A. M., and Schmid, C. W. (1990). Structure and evolution of the U2 small nuclear RNA multigene family in primates: gene amplification under natural selection. Mol. Cell. Biol. *10*, 5876-5882.

Maurer, B. J., Lai, E., Hamkalo, B. A., Hood, L., and Attardi, G. (1987). Novel sub-microscopic extrachromosomal elements containing amplified genes in human cells. Nature *327*, 434-437.

Meinkoth, J. A., Killary, A. M., Fournier, R. E. K., and Wahl, G. M. (1987). Unstable and stable CAD gene amplification: the importance of flanking sequences and nuclear environment in gene amplification. Mol. Cell. Biol. *7*, 1415-1424.

Monier, R. (1986). Transformation by SV40 and Polyoma virus. In The Papovaviridae, N. Salzman, ed (New York: Plenum Press).

McArthur, J. G., and Stanners C. P. (1991). A genetic element which increases the frequency of gene amplification. J. Biol. Chem. *266*, 6000-6005.

McIntyre, P., and Stark, G. R. (1988). A quantitative method for analyzing specific DNA sequences directly from whole cells. Anal. Biochem. *174*, 209-214.

Ohta, T. (1988). Time for acquiring a new gene by duplication. Proc. Natl. Acad. Sci. USA *85*, 3509-3512.

Oshelm, J. N., Miller, O. L., and Beyer, A. L. (1988). Visualization of Drosophila melanogaster chorion genes undergoing amplification. Mol. Cell. Biol. *8*, 2811-2821.

Padgett, R. A., Wahl, G. M., Coleman, P. F., and Stark, G. R. (1979). N-(phosphonacetyl)-L-aspartate resistant hamster cells overaccumulate a single mRNA coding for the multifunctional protein that catalyses the first step of UMP synthesis. J. Biol. Chem. *254*, 974-980.

Padgett, R. A., Wahl, G. M., and Stark, G. R. (1982) Structure of the gene for CAD, the multifunctional protein that initiates UMP synthesis in Syrian hamster cells. Mol. Cell. Biol. *2*, 293-301.

Painter, R. B., Young, B. R., and Kapp, L. N. (1987). Absence of DNA over-replication in Chinese Hamster cells incubated with inhibitors of DNA synthesis. Cancer Res. *47*, 5595-5599.

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

Pellegrini, S., Dalley, L., and Basilico, C. (1984). Amplification and excision of integrated Polyoma DNA sequences require a functional origin of replication. Cell *36*, 943-949.

Pellegrini, S., and Basilico, C. (1987). Amplification and expression of foreign genes in cells producing Polyoma Virus Large T antigen. Oncogene Res. *1*, 23-41.

Prives, C. (1990). The replication functions of SV40 T-antigen are regulated by phosphorylation. Cell *61*, 735-738.

Prody, C. A., Dreyfus, P., Zamir, R., Zakut, H., and Soreq, H. (1989). *De novo* amplification within a 'silent' human cholinesterase gene in a family subjected to prolonged exposure to organophosphorus insecticides. Proc. Natl. Acad. Sci. USA *86*, 690-694.

Reed, K. C., and Mann, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. *13*, 7207-7221.

Rigby, P. W. J., Dieckman, M., Rhodes, C., and Berg, P. (1977). Labeling deoxyribonucleic acid to high specificity *in vitro* by nick-translation with DNA polymerase I. J. Mol. Biol. *113*, 237-251.

Rigby, P. W. J., and Lane, D. P. (1983). Structure and function of Simian Virus 40 large T-antigen. In: Advances in Virology, G. Klein, ed (New York. Raven Press), pp 31-57.

Rio, D. C., Clark, S. G., and Tjian, R. (1985). A mammalian host vector system that regulates expression and amplification of transfected genes by temperature induction. Science *227*, 23-28.

Roberts, J. M., Buck, L. B., and Axel, R. (1983). A structure for amplified DNA. Cell *33*, 53-63.

Roninson, I. B., Abelson, H. T., Housman, D. E., Howell, N., and Varshavsky, A. (1984). Amplification of specific DNA sequences corelates with multidrug resistance in Chinese hamster cells. Nature *309*, 626-628.

Ruiz, J. C., and Wahl, G. M. (1988). Formation of an inverted duplication can be an initial step in gene amplification. Mol. Cell. Biol. *8*, 4302-4313.

Ruiz, J. C., Choi, K., Von Hoff, D. D., Roninson, I. B., and Wahl, G. M. (1989). Autonomously replicating episomes contain mdr1 genes in a multidrug resistant human cell line. Mol. Cell. Biol. *9*, 109-115.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Science *239*, 487-491.

Saito, I., Groves, R., Giulotto, E., Rolfe, M., and Stark, G. R. (1989). Evolution and stability of chromosomal DNA co-amplified with the CAD gene. Mol. Cell. Biol. *9*, 2445-2452.

Salzman, N. P., Natarajan, V., and Selzer, G. B. (1986). Transcription of SV40 and Polyoma virus and its regulation. In: The Papovaviridae, N. P. Salzman, ed (New York. Plenum Press) pp 27-98.

Sambrook, J., Westphal, H., Srinivasan, P. R., and Dulbecco, R. (1968). The integrated state of viral DNA in SV40 transformed cells. Proc. Natl. Acad. Sci. USA. *60*, 1288-1295.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. *74*, 5463-5467.

Schimke, R. T., Sherwood, S. W., Hill, A. B., and Johnston, R. N. (1986). Over replication and recombination of DNA in higher eukaryotes: potential consequences and biological inplications. Proc. Natl. Acad. Sci. USA. *83*, 2157-2161.

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

Schimke, R. T. (1985). Gene amplification in cultured cells. J. Biol. Chem. *263*, 5989-5992.

Sharma. R. C., and Schimke, R. T. (1989). Enhancement of the frequency of methotrexate resistance by γ-irradiation in Chinese hamster ovary and mouse 3T3 cells. Cancer. Res. *49*, 3861-3866.

Sherwood, S. W., Schumacher, R. I., and Schimke, R. T. (1988). Effect of cyclohexamide on development of methotrexate resistance of Chinese hamster ovary cells treated with inhibitors of DNA synthesis. Mol. Cell. Biol. *8*, 2822-2827.

Shigesada, K., Stark, G., Maley, J. A., Niswander, L. A., and Davidson, J. N. (1985). Construction of a cDNA to the hamster CAD gene and its application toward defining the domain for aspartate transcarbamylase. Mol. Cell. Biol. *5*, 1735-1742.

Smith, K. A., Gorman, P. A., Stark, M. B., Groves, R. P., and Stark, G. R. (1990). Distinctive chromosomal structures are formed very early in the amplification of CAD genes in Syrian hamster cells. Cell *63*, 1219-1227.

Spradling, A. C., de Cicclo, D. V., Wakimoto, B. T., Levine, J. F., Kalfayan, L. J., and 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.

Stark, G. R., and Wahl, G. M. (1984). Gene amplification. Ann. Rev. Biochem. *53*, 447-491.

Stark, G. R. (1986). DNA amplification in drug resistant cells and tumours. Cancer Surveys *5*, 1-23.

Stark, G. R., Debatisse, M., Giulotto, E., and Wahl, G. M. (1989). Recent progress in understanding mechanisms of mammalian gene amplification. Cell *57*, 901-908.

Stillman, B. (1989). Eukaryotic DNA replication. Ann. Rev. Cell Biol. *5*, 197-245.

Stryer, L. (1981). Biochemisrty, Second edition. (San Francisco. Freeman Press) pp 519-521.

Southern, P. J., and Berg, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under the control of the SV40 early promoter. J. Mol. Biol. *1*, 327-341.

Swyryd, E. A., Weaver, S. S., and Stark, G. R. (1974). N-(phosphonacetyl)-L-aspartate, a potent transition state analogue inhibitor of aspartate transcarbamylase, blocks proliferation of mammalian cells in culture. J. Biol. Chem. *249*, 6945-6950.

Thomas, K. R., and Capecchi, M. R. (1987). Site directed mutagenesis by gene targeting in mouse embryo derived stem cells. Cell *51*, 503-512.

Tlsty, T. D., (1990). Normal diploid human and rodent cells lack a detectable frequency of gene amplification. Proc. Natl. Acad. Sci. USA. *87*, 3132-3136.

Tooze, J. (1981). In: DNA Tumour Viruses, Second edition (revised), J. Tooze, ed (New York, Cold Spring Harbour).

Tooze, J. (1973). In: Molecular Biology of Tumour Viruses, J. Tooze, ed (New York. Cold Spring Harbour). pp 269-299.

Trask, B. J., and Hamlin, J. L. (1989). Early dihydrofolate reductase gene amplification events in CHO cells usually occur on the same chromosome arm as the original locus. Genes. Dev. *3*, 1913-1925.

Tsurimoto, T., and Stillman, B. (1991). Replication factors required for SV40 replication *in vivo*: II. Switching of DNA polymerase α and γ during initiation of leading and lagging strand synthesis. J. Biol. Chem. *266*,1961-1968.

Von Hoff, D. D., Needham-Vandevanter, D. R., Yucel, J., Windle, B. E., and Wahl, G. M. (1988). Amplified human MYC oncogenes localized to a replicating sub-microscopic circular DNA molecules. Proc. Natl. Acad. Sci. USA. *85*, 4804-4808.

Wahl, G. M., Padgett, R. A., and Stark, G. R. (1979). Gene amplification causes over-production of the first three enzymes of UMP synthesis in *N*-(phosphonacetyl)-L-aspartate resistant hamster cells. J. Biol. Chem. *254*, 8679-8689.

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

Wahl, G. M., Robert de Saint Vincent, B., and Derose, M. L. (1984). Effect of chromosomal position on amplification of transfected genes in animal cells. Nature *307*, 516-520.

Wahl, G. M. (1989). The importance of circular DNA in mammalian DNA amplification. Cancer Res. *49*, 1333-1340.

Windle, B., Draper, B. W., Yin, Y., O'Gorman, S., and Wahl, G. M. (1991). A central role for chromosome breakage in gene amplification, deletion formation and amplicon integration. Genes Dev. *5*, 160-174.

Wright, J. A., Smith, H. S., Watt, F. M., Hancock, M. C., Hudson, D. L., and Stark, G. R. (1990). DNA amplification is rare in normal human cells. Proc. Natl. Acad. Sci. USA *87*, 1791-1795.

Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L., and Trieber, G. (1970). Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale plasmid purification Virol. *40*, 734-744

Zehnbauer, B. A., Small, D., Brodeur, G. M., Seeger, R., and Vogelstein, B. (1988). Characterization of N-myc amplification units in human neuroblastoma cells. Mol. Cell. Biol. *8*, 522-530.

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

APPENDIX

Distinctive Chromosomal Structures Are Formed Very Early in the Amplification of CAD Genes in Syrian Hamster Cells

Kathleen A. Smith, Patricia A. Gorman, Mary B. Stark, Richard P. Groves, and George R. Stark
Imperial Cancer Research Fund
Box 123
Lincoln's Inn Fields
London, WC2A 3PX
England

Summary

As visualized by in situ hybridization with fluorescence detection, newly amplified CAD genes in 10^5 cell colonies are contained in multiple copies of very large regions of DNA, each tens of megabases long. The extra DNA is usually linked to the short arm of chromosome B9, which retains CAD at its normal site. The widely spaced genes are often interspersed with new G-negative regions. Individual cells within a clone have highly variable numbers of CAD genes (range 2–15). When resistant clones are examined later, at the 10^{15} cell stage, the amplified genes are usually found in much more condensed structures. We propose that, in the initial event of CAD gene amplification, much of the short arm is transferred from one B9 chromosome to another. In subsequent cell cycles this initial duplication expands rapidly through unequal but homologous sister chromatid exchanges. Relatively rare secondary events lead to more condensed structures.

Introduction

Knowledge of the basic mechanisms of DNA amplification is important for a full understanding of both tumorigenesis and drug resistance (reviewed by Stark et al., 1989, 1990; Wahl, 1989; Schimke, 1988). Although amplified oncogenes are a common feature of many tumors, most of our current ideas concerning mechanisms come from studies carried out with cells in tissue culture, which often achieve resistance to cytotoxic drugs through amplification of specific genes. About 20 such examples are known (Stark and Wahl, 1984; Stark, 1986), and several have been examined in considerable detail. We have studied Syrian hamster BHK cells, which achieve resistance to the aspartate transcarbamylase inhibitor N-phosphonacetyl-L-aspartate (PALA) through amplification of the CAD gene (Wahl et al., 1979). CAD is a trifunctional protein carrying carbamyl-P synthetase, aspartate transcarbamylase, and dihydro-orotase, the first three enzymes of UMP biosynthesis (Padgett et al., 1979). In contrast to other systems, amplification of the CAD gene is the only mechanism of achieving resistance to PALA that has been detected. Furthermore, amplified CAD DNA has always been found within chromosomes; in contrast, double-minute chromosomes and other extrachromosomal forms are a common feature in other situations (Wahl, 1989).

Some of our recent observations have indicated that amplified CAD DNA is unstable in BHK cells. For example, Saito et al. (1989) observed that first-step mutants, selected in a single low concentration of drug, lost their amplified DNA rapidly when grown without selection, despite the fact that the amplified DNA was chromosomal. Furthermore, second- and third-step mutants, selected sequentially in higher concentrations of PALA, lost much of the DNA that had been coamplified with CAD in the first step, leading to the conclusion that only a part of the initial array was coamplified later. A common implicit assumption in studying the nature of amplified DNA has been that the structures seen reflect initial or at least early events. However, if the structures formed very early are very unstable, they may well have been missed in most previous studies. In such a situation, it is crucial to examine new and independent colonies, selected in a low concentration of drug, as soon as possible after they can be identified.

The initial event is rare (probability $\sim 10^{-5}$ per cell per generation) and is recognized only retrospectively when a drug-resistant colony can be seen. Most molecular techniques useful for studying amplified DNA (Southern blotting, for example) require DNA from 10^6–10^7 cells, but in situ hybridization allows analysis of individual resistant cells. By using biotin-labeled probes and fluorescent techniques for detection, it has now been possible to see single-copy genes rapidly and reliably in most metaphase spreads, and individual genes on sister chromatids are resolved well (Pinkel et al., 1986; Lawrence et al., 1988). Using this technique, we have examined the structures present very soon after the first step of CAD gene amplification (about 10^5 cells per colony) and those present later, at about 10^{15} cells per colony. The results lead to a model with three major phases. In the initial step, a very large region, perhaps an entire chromosome arm, is transferred from one B9 chromosome to the other. Rapid expansion of this duplicated region gives very long arrays containing many copies of the initial structure. Finally, relatively infrequent secondary events generate more highly condensed structures containing the amplified CAD genes.

Trask and Hamlin (1989) studied Chinese hamster cells selected in steps of gradually increasing concentrations of methotrexate by using in situ hybridization with fluorescence detection. It is striking that some of the clones they observed have very large regions of amplified DNA with widely spaced dihydrofolate reductase genes, attached to the end of the same chromosome arm that carries the single-copy gene in unselected cells, with the gene still present in its normal position in the marker chromosomes. Other clones show more condensed structures containing amplified dihydrofolate reductase genes that have been translated to other chromosomes. The variety of structures observed by these workers at a later stage in the amplification process may well be derived from long arrays of amplified genes similar to those we observed in all very early events.

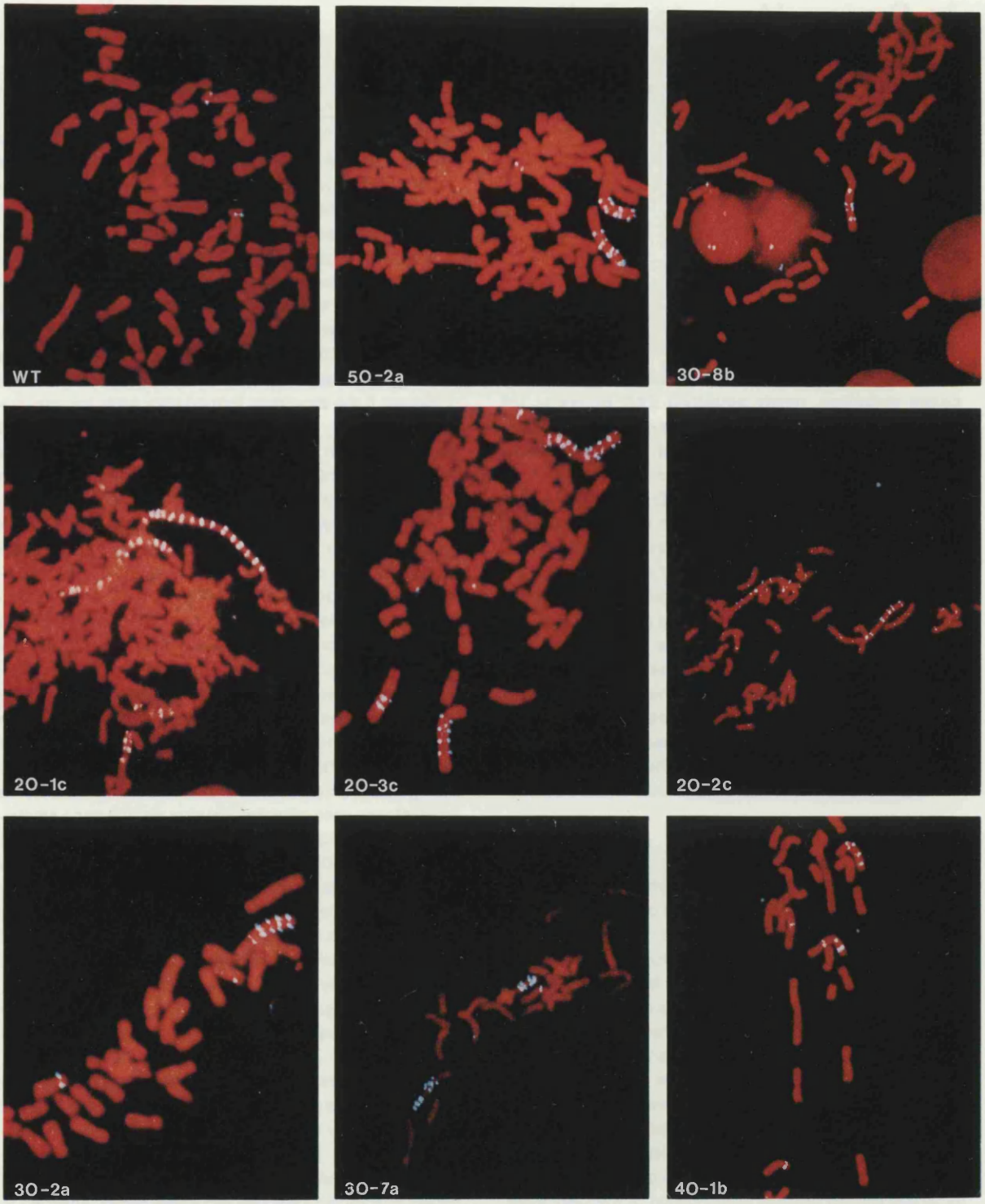


Figure 1. Examples of Chromosomes Carrying Amplified CAD Genes in Eight Independent PALA-Resistant Clones at the 10^5 Cell Stage

An example of an unselected cell (WT) is also shown. The CAD genes are revealed by the binding of fluorescein-labeled avidin to biotin-labeled DNA probes after in situ hybridization. The chromosomes are stained with propidium iodide.

Results

In Situ Analyses at the 10^5 Cell Stage

In the selection scheme used (see Experimental Procedures), the history of each clone is known, with about 15 cell divisions separating the initial event from analysis. About 90% of the metaphase spreads gave a signal with the CAD probe and were scored. Representative examples are shown in Figure 1. Analyses carried out on 195 spreads from eight independent clones are summarized

Table 1. Summary of Analyses of Clones Studied at the 10^5 Cell Stage

Clone	Number of Cells							Bands of Hybridization per Cell[b]		
	Analyzed	With B9[a]	With B9p$^+$	With a Non-B9p$^+$ Marker	With an Unscorable Marker	With a Dicentric B9p$^+$	With More Than One Marker	Range	Mean	Standard Deviation
20-1bcde	44	25	35	5	16	3	7	2–15	5.5	3.4
20-2ab	24	14	16	2	13	2	7	2–13	5.6	2.3
20-2c	5	3	1	5	0	0	3	3–12	6.0	3.9
20-3bc	17	6	6	6	8	1	7	2–10	5.2	1.7
30-2ab	29	17	11	11	9	0	3	3–15	5.3	2.5
30-7a	13	9	1	7	3	0	0	3–7	4.6	1.4
30-8ab	27	11	18	4	7	0	3	3–10	5.8	2.3
40-1ab	20	10	8	8	8	0	10	3–10	5.2	2.4
50-2a	16	13	11	2	2	1	0	2–8	4.8	1.4
Total	195	108	107	50	66	7	40	–	–	–

[a] There was almost always one copy of B9 per cell. Many cells scored as lacking B9 also have an obviously incomplete set of chromosomes.
[b] Number of discrete regions of hybridization along the chromosome, usually seen as a separate spots on each sister chromatid. Such a pair of spots counts as one band. The band on the normal B9, if present, is included.

in Table 1. Clones are named as follows: the first number is the concentration of PALA (μM) used for selection, the second is that of the plate, and the letter identifies a specific clone from that plate. Clones isolated from separate plates must be independent, but clones from the same plate are likely to be siblings. Except for 20-2c, the properties of clones from the same plate were similar and were considered together, as indicated in Table 1.

The major findings are as follows: First, in every cell examined the CAD genes were chromosomal, and hybridization to extrachromosomal DNA was not observed. Second, in each independent clone, marker chromosomes carried several distinct regions of CAD hybridization, separated from each other by tens of megabases, almost always seen as two spots of hybridization, one on each sister chromatid. Third, in unselected cells, both copies of the CAD gene were on the short arms of the two B9 chromosomes. The newly amplified regions often appeared at or near the end of the short arm of one copy of chromosome B9, to generate a B9p$^+$ marker chromosome. Clones 20-2c and 30-7a had a majority of marker chromosomes that did not appear to be derived from B9, and such markers were also present in all the other clones, on average about half as often as B9p$^+$. Only a few dicentric marker chromosomes were seen, in several different clones. About 20% of all cells examined contained more than one marker chromosome, but the percentage was especially high for clones 20-2c, 20-3b, and c, and 40-1a and b. Fourth, the number of distinct locations to which the CAD probe hybridized in each cell, representing the minimum number of copies of CAD, varied widely within each clone (range 2–15). Six examples from clone 20-1c are shown in Figure 2. Fifth, the length of chromatin carrying the amplified CAD genes was seldom less than 1 or more than 2.5 times the length of the entire B9 chromosome, with an average of about 1.5 (data not shown). Sixth, one copy of CAD was in its normal position within the former short arm of the B9p$^+$ chromosomes in 104 of the 110 cases examined (95%). Seventh, most parental cells had two copies of a normal B9 chromosome (see Figure 1); 55% of the mutant cells examined had a normal B9 chromosome, usually only one copy. This is probably a low estimate for most clones because, as a technical problem, many spreads had fewer than the normal number of chromosomes per cell and because clones 20-2c and 30-7a, included in the average, rarely had a normal B9 chromosome.

The average copy number of CAD in five mutant clones (20-1c, 20-2b, 30-2a, 30-8a, 50-2a) was determined by analyzing Southern transfers of genomic DNA cut with EcoRI, with normalization to a single-copy EcoRI fragment, as described by Giulotto et al. (1986). The average number of copies of CAD per cell was 6.3 (data not shown), in reasonably good agreement with 5.4, the average estimate made by counting bands of hybridization per cell for the same five clones (Table 1). The correlation indicates that each spot of hybridization usually represents an individual CAD gene.

In situ hybridizations for the CAD gene were carried out with two cosmid probes representing about 80 kb of contiguous DNA. Usually, this pair of probes gave two spots across the width of the pair of B9 sister chromatids, but occasionally three or four spots were seen (see Figure 2c for an example). Hybridizations together with CAD in normal BHK cells were also carried out with cosmids 276 and 773 ("near") and 171, 671, 872, and 971 ("far"), characterized by Saito et al. (1989) as corresponding to six separate regions of DNA often coamplified with CAD in first-step mutants. In no case was the site of hybridization distinguishable from that of CAD along the length of the chromosome (data not shown), demonstrating that these six regions all lie within a few megabases of CAD. When cosmid 171 was used with CAD, four spots of hybridization were often seen across the width of the pair of B9 sister chromatids. Resolution of two relatively near sequences across the width of a chromosome probably reflects the

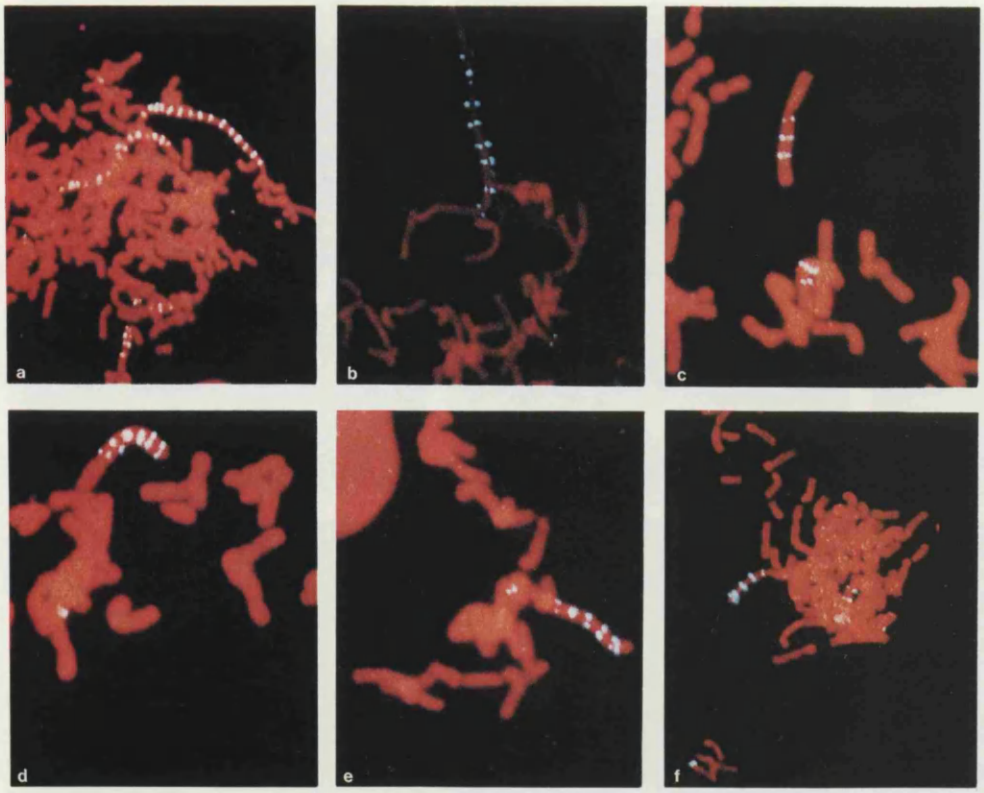


Figure 2. Examples of Six Different Spreads from Clone 20-1c, Examined at the 10^5 Cell Stage
Note the highly variable number of CAD genes on the marker chromosomes of different cells.

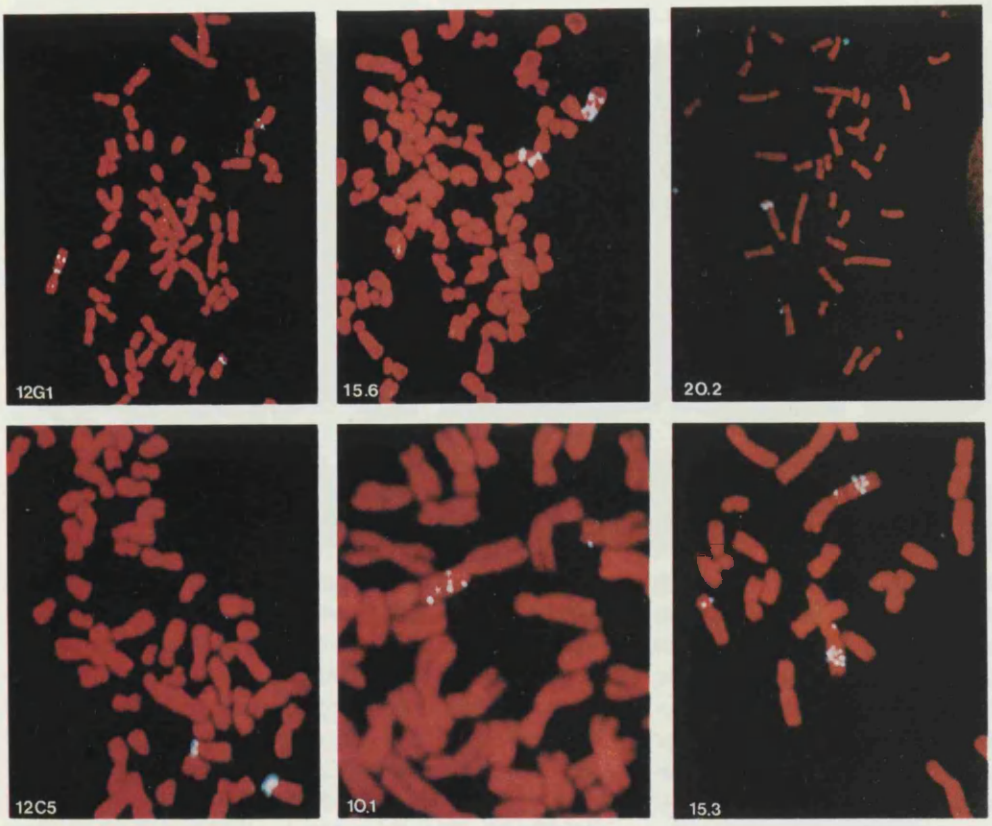


Figure 3. Examples of Chromosomes Carrying Amplified CAD Genes in Six Independent PALA-Resistant Clones at the 10^{15} Cell Stage

Table 2. Summary of Analyses of Clones Studied at the 10^{15} Cell Stage

Clone	Number of Cells					Comments
	Analyzed	With B9	With B9p+	With a Non-B9 Marker	With More Than One Marker	
10.1	15	11	15	1	2	Two to three hybridization bands per marker on a region 1× B9 length.
12C5	17	0	0	17	10	Characteristic truncated B9p, with intense hybridization at the very end.
12G1	20	16	18	0	0	Characteristic B9p+ with one band of hybridization on a region 0.75× B9 length. All have obvious G-negative gap at the position of the normal B9 end.
15.3	18	12	15	0	3	Gap visible on almost all B9p+ at normal position of B9 telomere. No hybridization in gap. Three to four hybridization bands per marker on a region ~1× B9 length.
15.6	15	4	0	15	4	One to three hybridization bands per marker on a region 0.75× B9 in length.
20.2	16	1	0	16	5	Single band of hybridization on a characteristic marker chromosome with a banding pattern quite distinct from that of B9.

superhelical winding of DNA within chromatin and has been observed before (Lawrence et al., 1988).

In Situ Analyses of Single-Step Clones at the 10^{15} Cell Stage

The six independent clones analyzed were isolated previously by Zieg et al. (1983); the results are summarized in Table 2 and illustrated in Figure 3. With the exception of clone 10.1, the widely separated pattern of amplified CAD genes was not seen in these clones, which were grown in PALA for a relatively long time. The CAD genes were present in much more highly condensed arrays. In four of six clones, marker chromosomes could still be recognized as derivatives of B9 by G-banding. In clone 12C5 the p arm of B9 was truncated, with an intense band of hybridization at one end, near the normal position of CAD.

Chromosome Banding

In many cases it was possible to G-band chromosomes with Wright's stain after the in situ analysis had been completed. Thus we could compare hybridization and banding patterns on the same chromosomes, even though the patterns were often not as distinct as those observed with chromosomes that had not been subjected to the hybridization procedure. Forty B9 chromosomes, previously identified on the basis of hybridization, lengths of the p and q arms, and positions of the centromeres, were also analyzed by banding (see Figure 4 for examples). Thirty-three (83%) were confirmed to be B9, as judged by the banding pattern of the q arm (data not shown), and the other seven could not be scored. Forty B9p+ marker chromosomes were analyzed similarly. Thirty-eight (95%) were confirmed by the same criteria, often by comparison with an unchanged B9 chromosome present in the same spread (data not shown), and the other two could not be scored.

Fifty-nine B9p+ or other marker chromosomes that hybridized to the CAD probe were examined by banding. Fifty-eight showed at least one characteristic G-negative region in the area of hybridization. Examples for clones 12G1, 20-1b, 30-2b, and 40-1b are shown in Figure 4, where the number of separate regions of hybridizations is given for each chromosome. G-negative regions were not observed in unselected BHK cells or on chromosomes that did not carry amplified CAD genes in cells selected with PALA. In all but one of the clones listed in Table 1, virtually all the marker chromosomes had G-negative regions. A clear exception was clone 30-7a, which had none. There was a G-negative region at the point where the p arm would normally end in 26 of 32 (81%) of the B9p+ chromosomes analyzed (see Figure 4 for examples). The total number of G-negative regions was about half the number of regions containing CAD genes (105 of 238) in the 59 chromosomes examined. CAD genes were almost always present between the G-negative regions and not within them. In 105 examples, 75 were between and 30 were next to these regions. Two examples are shown in Figure 5. Note that chromosome 20-1b in Figure 5 is a relatively rare example of a dicentric B9p+.

Discussion

The Initial Event

Multiple initiation of replication at a single origin within a single cell cycle cannot be involved in the initial event because the rate of elongation during DNA replication, about 3 kb/min (Edenberg and Huberman, 1975), is too slow. Only about 1.5 Mb of DNA can be copied from a single origin during a single S phase 8 hr long, and to make one copy of a 30 Mb region would require 160 hr of continuous synthesis. It seems unlikely that a cell would lose replication control simultaneously for many origins within such a very large region. Even if it did, the resulting polytene structure would have to recombine specifically enough to generate the regular arrays observed, which also seems unlikely.

The initial event is proposed to be duplication of a large region by recombination between sister chromatids. Expansion of the initial duplication, discussed below, has to occur in later cell cycles. As a specific example, we consider mechanisms that involve telomere–centromere or telomere–telomere recombination. Alternatively, other specific recombinogenic sequences may be involved, and it is important to perform further experiments to define the points of recombination in different events. Experiments with biotin-labeled $(TTAGGG)_n$ probes have thus far

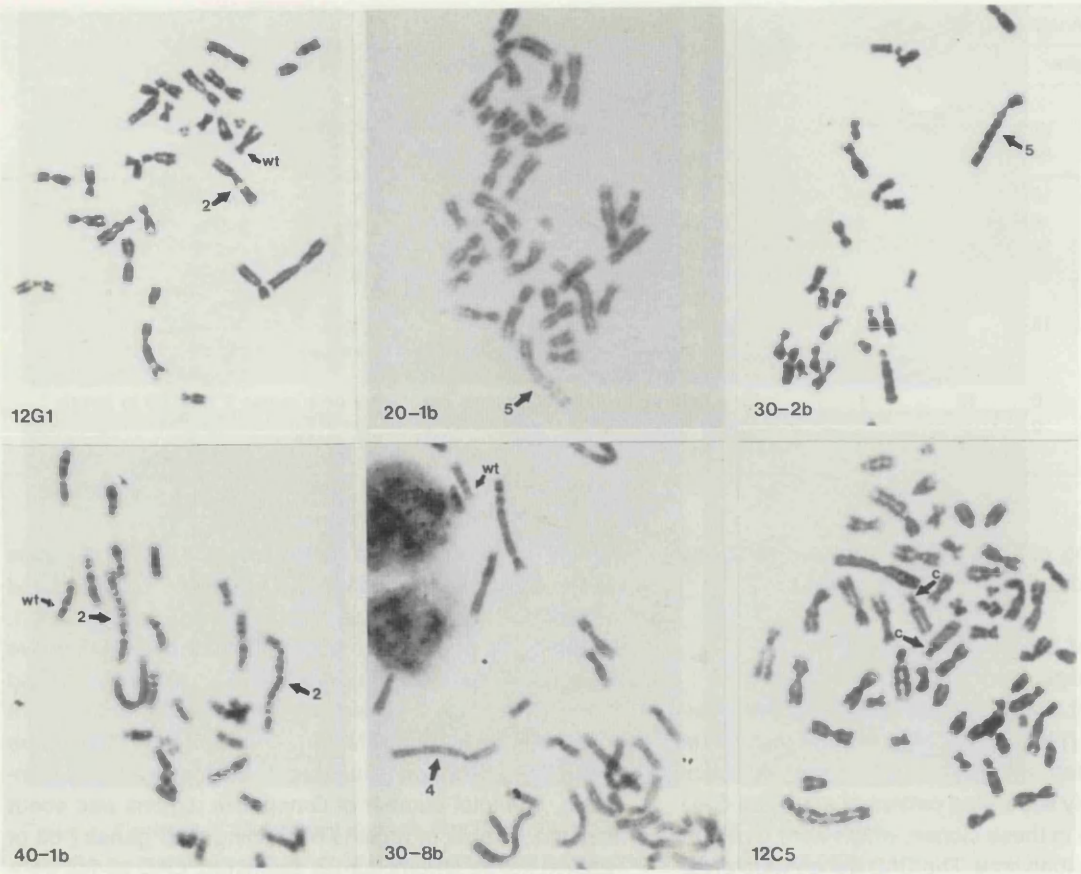


Figure 4. Examples of Chromosome Spreads Banded with Wright's Stain after Analysis by In Situ Hybridization

Normal B9 chromosomes, which show a single region of CAD hybridization, are labeled wt. Marker chromosomes are labeled to show the number of regions of CAD hybridization observed and their approximate locations. The truncated B9 chromosome carrying amplified CAD genes in clone 12C5 is labeled C.

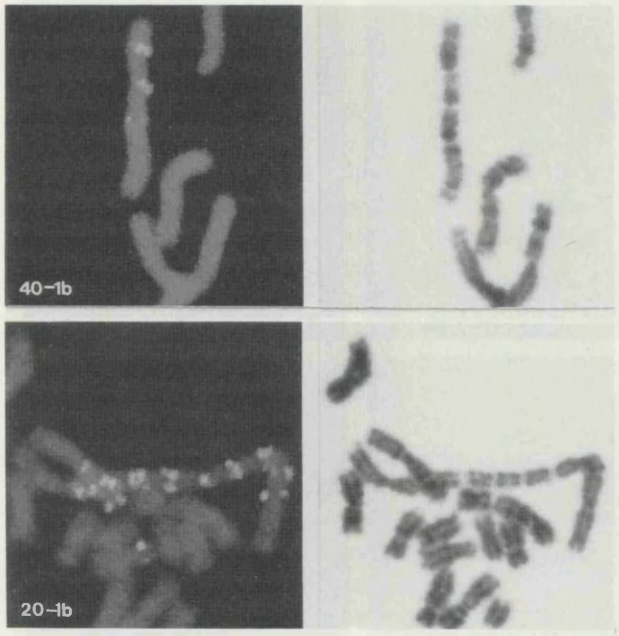


Figure 5. Correspondence of Regions of Hybridization with G-Negative Regions

Note that the areas of hybridization are well resolved from the G-negative regions in clone 40-1b and that they are near but not coincident with these regions in 20-1b. The marker chromosome in the spread of 20-1b is dicentric, with a B9q arm apparent at each end.

failed to give evidence for interstitial telomeric sequences, but many telomeres are not revealed in Syrian hamster cells by this method owing to its limited sensitivity at present. Although there is no direct evidence that telomeric sequences are involved, two observations are consistent with this possibility: First, regularly spaced arrays of interstitial G-negative regions, not seen in unamplified DNA, are a common features of newly amplified CAD DNA in BHK cells. These regions sometimes also appear as constrictions. It is known that telomeres and interstitial telomeric sequences do not stain well and are often constricted (Sutherland and Hecht, 1985; Hastie and Allshire, 1989). Second, a G-negative region is often seen in B9p$^+$ at the position of the normal B9p telomere.

Two additional observations lead us to consider that recombination between telomeric sequences and pericentromeric sequences may be important: First, there is strong homology between the -TTAGGG- repeat of mammalian telomeres and at least some pericentric sequences, as shown by their hybridization to a $(TTAGGG)_n$ probe (Moyzis et al., 1988; Meyne et al., 1989, 1990). Second, the similar spacing of the CAD genes and of the G-negative regions suggests that some common measuring device is employed. Specific recombination between telomeres and centromeres could provide such a measure. Random recombination or recombination between

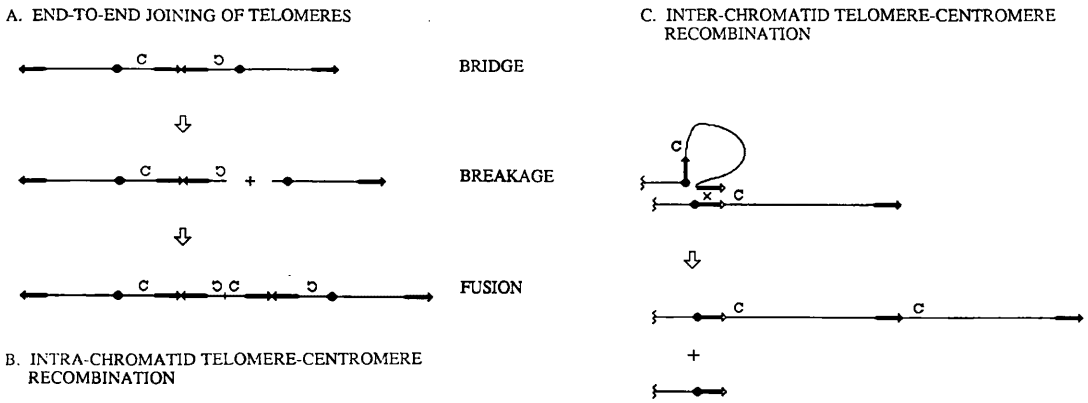


Figure 6. Models for the Initial Event of CAD Gene Amplification in BHK Cells in Which Telomeric Sequences are Involved

C, CAD gene: filled circles, centromeres; hatched arrows, telomeres or interstitial telomeric sequences.

dispersed repetitive sequences would not be expected to give such regular spacing in many independent events.

Figure 6A shows the predicted products of telomere–telomere joining to form a dicentric chromosome, which then participates in bridge–breakage–fusion cycles. A bridge might break as shown to give a fragment that has two copies of CAD but lacks a telomere. Replication and fusion of ends lacking telomeres might generate a new chromosome with four copies of CAD, as shown. Although this structure is consistent with dicentric chromosomes carrying one copy of B9q at each end, seen occasionally (Figure 5), it does not predict that the CAD genes or interstitial telomeric sequences will be spaced evenly. This model makes the testable prediction that each region containing CAD is inverted with respect to those that flank it.

Figures 6B and 6C show homologous recombinations between telomeres and pericentric DNA. We consider only events in which the orientation of the recombining sequences is the same. Intra- or interchromosomal nonhomologous recombination between sequences inverted with respect to one another (not shown) are not productive since they do not place two copies of CAD on the same piece of DNA. If a homologous event is intrachromosomal (Figure 6B), the initial structure containing CAD could be either a circular episome lacking a centromere (shown) or a circular chromosome with a centromere (not shown, formed by recombination on the other side of the centromere). Finally, to generate the structures observed, homologous reintegration into a normal B9 chromosome must be frequent and rapid, and nonhomologous integration into other sites must be very rare. These constraints make the model shown in Figure 6B unlikely but do not rule it out.

The most likely possibility is an interchromosomal event (Figure 6C), which would lead to transfer of an entire arm to the recipient sister chromatid if pericentric sequences proximal to the telomere are involved. If pericentric sequences distal to the telomere are involved (not shown), a dicentric chromosome is generated that would then be likely to participate in bridge–breakage–fusion cycles (Figure 6A).

Propagation of an Initial Duplication

Two observations give strong support to the hypothesis that the initial duplication is propagated rapidly by means of homologous, unequal sister chromatid exchanges.

First, individual cells from single small colonies show wide variation in CAD gene copy number; and second, there is no selection for copy numbers as high as 15 per cell, yet these are commonly observed. The initial events proposed give three copies of CAD in one daughter cell and one copy in the other after the first cell division, and additional copies must be generated in subsequent cell cycles. By using time-lapse cinematography, we have observed that many cells in a BHK population exposed to 20 μM PALA do go through one or two divisions before they die (Marie-France Poupon, Christopher Gilbert, and G. R. S., unpublished data). Homologous sister chromatid exchange is known to be frequent and thus is an attractive possibility (Gorden et al., 1990). Unequal exchange between duplicated chromosome arms (Figure 6D) may well be able to give as many as 15 copies by the 10^5 cell stage, since homologous recombination at any point along tens of megabases of DNA would result in net transfer of DNA from one sister chromatid to the other. In the first such event, either one or two copies can be transferred, to give a daughter cell with up to five copies of CAD (including the one on the normal copy of B9). In each subsequent step, the number of copies transferred can double. The number of possibilities quickly becomes very large, consistent with our observation that the copy number of CAD is highly variable among individual cells within a single clone of 10^5 cells (see Figure 2 and Table 1).

Cells with only a small increase in the copy number of CAD are resistant to relatively high concentrations of PALA (Kempe et al., 1976; Zieg et al., 1983). Since a large increase in resistance results from a small increase in CAD, cells don't need as many as 15 copies of CAD to resist the low concentrations of PALA employed. This important observation shows that the presence of many copies of CAD within a single chromosome (see Figures 1 and 2) is an inescapable consequence of the amplification mechanism once it has been set in motion, and can be explained by events such as those depicted in Figure 6D.

Later Events

When first-step clones are maintained in the original concentration of PALA until the 10^{15} cell stage, the amplified CAD genes become much more highly condensed (see Figure 3), with the region surrounding each CAD gene a few megabases long rather than tens of megabases as originally seen. Such structures must arise through relatively rare secondary amplifications events. Several mechnisms that might account for such events have been discussed in a recent review (Stark et al., 1989). It is interesting that the amplified CAD genes often still reside on chromosome B9, sometimes even in mutants selected stepwise to resist very high concentrations of PALA (Wahl et al., 1982). Formation of more highly condensed structures may be facilitated by an unusual feature of the extended array. Miele et al. (1989) conclude that the presence of amplified CAD genes on Chinese hamster chromosomes renders them unstable; they tend to be excluded from cells and are more prone to rearrangements. Cells with condensed arrays would have a selective advantage if cells with very long arrays were to grow more slowly.

Generality of the Features Observed in the Amplification of CAD Genes in BHK Cells

The previous data of Trask and Hamlin (1989) and preliminary results with CAD in Chinese hamster cells (K. A. S., M. B. S., G. R. S., and Elena Giulotto, unpublished data) and with AMP deaminase in Chinese hamster cells (Michelle Debatisse and K. A. S., unpublished data) reveal highly extended chromosomal structures with widely spaced amplified genes. These observations indicate that events similar to the ones described in detail here do occur in other situations. Trask and Hamlin (1989) conclude, and we agree, that it is very unlikely that such structures are generated by any mechanism involving deletion of a relatively small episome, followed by reintegration of multiple copies of the episome. If the rapid evolution of amplified CAD DNA in BHK cells is also true for other genes in other species, many conclusions concerning so-called early events will have to be evaluated again. It will be especially interesting to study cases such as amplification of dihydrofolate reductase genes in mouse or human cells at a very early time, to see whether the extrachromosomal DNA so often observed in such cells examined after an unknown number of generations (for a recent example, see Pauletti et al., 1990) is a product of the initial event or of secondary amplification.

Experimental Procedures

Selection with PALA

Dialyzed fetal calf serum (10%) was used during selection and propagation of Syrian hamster BHK cells in PALA, essentially as described by Kempe et al. (1976). To isolate new and independent resistant clones, individual populations of 10^3 cells were grown to approximately 10^5 cells in separate wells without selection. Each group of cells dispersed, seeded independently onto 9 cm dishes, and treated immediately with 15, 20, 30, 40, or 50 μM PALA. Small resistant colonies (50–200 cells), picked only from plates with fewer than eight colonies, were grown in PALA to approximately 10^5 cells before analysis by in situ hybridization. PALA-resistant clones isolated previously by Zieg et al. (1983) and grown to the 10^{15} cell stage were recovered from frozen stocks and grown for a minimal additional period in PALA at the concentration used for selection.

In Situ Hybridization with Biotinylated Cosmid Probes

Cultures of about 2×10^4 cells were incubated with 150 ng/ml of colcemid (Gibco) for 1.75 hr, and mitotic cells were shaken off, collected, swollen for 15 min in 75 mM KCl at 37°C, fixed in fresh ice-cold methanol–acetic acid (3:1), and dropped onto microscope slides. An equal mixture of two cosmids (C64 and C81; Giulotto et al., 1986) was used routinely as the probe. The DNA was labeled with biotin-11-dUTP (Enzo Diagnostics) by nick translation (Rigby et al., 1977). A trace of [^{3}H]dATP (Amersham) was added to allow analysis of incorporation. Labeled probe, approximately 50–500 bp long, was isolated by using a spin column containing Sephadex G-50 (Pharmacia).

The fluorescence in situ technique of Pinkel et al. (1986, 1988) was used, with minor modifications to optimize conditions for analysis of Syrian hamster cells. Briefly, mitotic metaphase spreads were treated with 100 μg/ml of RNAse A in 0.3 M sodium chloride, 30 mM sodium citrate (pH 7) (2× SSC) at 37°C for 1 hr, dehydrated through an alcohol series, and dried. The slides were treated with proteinase K (Boehringer), 600 ng/ml, in 20 mM Tris–HCl, 2 mM $CaCl_2$ (pH 7.5) at 37°C for 7.5 min, postfixed in 4% paraformaldehyde in phosphate-buffered saline, 50 mM $MgCl_2$ (pH 7) for 10 min at room temperature, dehydrated as described above, and dried. The chromosomes were denatured in 70% (vol/vol) formamide, 2× SSC (pH 7) at 76°C for 7.5 min, placed in ice-cold 70% ethanol, and dehydrated. Hybridization with 5 ng/μl of probe was carried out in 50% formamide, 1% Tween-20 (Pierce), 10% dextran sulfate (Pharmacia) in 2× SSC (pH 7). Sonicated Syrian hamster genomic DNA (200–600 bp) was added to the

mixture to a final concentration of 1 μg/μl to suppress hybridization of repetitive genomic sequences. The hybridization mixture was heated to 70°C for 10 min, cooled, and placed at 37°C for 20 min before adding it to the chromosomes, to allow the repetitive sequences to preanneal. After incubation with probe overnight at 37°C, the slides were washed at 42°C in three changes of 50% formamide, 2× SSC (pH 7) for 5 min each and then in three changes of 2× SSC for 5 min each.

Fluorescein-labeled avidin DCS (Vector Labs) was used at 5 μg/ml in 4× SSC, 0.05% Tween-20, 5% nonfat dried milk (Carnation) (pH 7) at room temperature for 20 min. After washing in 4× SSC, 0.05% Tween-20 (pH 7), the fluorescent signal was amplified by treating the slides with biotin-conjugated goat anti-avidin (5 μg/ml, Vector Labs.) followed by fluorescein-labeled avidin, as above. After washing with 4× SSC, 0.05% Tween -20 (pH 7) and then phosphate-buffered saline (pH 7), the chromosomes were counterstained with 0.1 μg/ml of propidium iodide (Sigma) in phosphate-buffered saline for 5 min and mounted in Citifluor AF1, an antifade preparation (Citifluor, Ltd.).

Preparations were viewed with a Nikon Microphot FX microscope equipped with optics for propidium iodide and fluorescein epifluorescence as well as for conventional bright-field microscopy. The microscope was used with an MRC-500 laser-scanning confocal imaging system (Bio-Rad Microsciences) with a filter system optimized to obtain separate images for each fluorochrome. All photographs were taken directly from the video screen using Kodak Ektar 125 color-print film.

G-Banding

After the images of individual metaphase spreads had been recorded, the slides were destained in methanol, incubated in 1× SSC at 60°C for 5–15 min, and banded in Wright's stain (B.D.H.) diluted 1:3 in 50% Sorenson's buffer (equal volumes of 30 mM KH_2PO_4 and 30 mM Na_2HPO_4).

Acknowledgments

We are especially grateful to Dr. Joyce Hamlin for stimulating discussions. Charles O'Neill was most helpful in instructing us on the use of the confocal microscope. Denise Sheer and Sarah Williams provided expert advice on cytogenetics. We thank Clare Middlemiss for preparing the manuscript.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received July 19, 1990; revised September 12, 1990.

References

Edenberg, H. J., and Huberman, J. A. (1975). Eukaryotic chromosome replication. Annu. Rev. Genet. *9*, 245–284.

Giulotto, E., Saito, I., and Stark, G. R. (1986). Structure of DNA formed in the first step of CAD gene amplification. EMBO J. *5*, 2115–2121.

Groden, J., Nakamura, Y., and German, J. (1990). Molecular evidence that homologous recombination occurs in proliferating human somatic cells. Proc. Natl. Acad. Sci. USA *87*, 4315–4319.

Hastie, N. D., and Allshire, R. C. (1989). Human telomeres: fusion and interstitial sites. Trends Genet. *5*, 326–331.

Kempe, T. D., Swyryd, E. A., Bruist, M., and Stark, G. R. (1976). Stable mutants of mammalian cells that overproduce the first three enzymes of pyrimidine nucleotide biosynthesis. Cell *9*, 541–550.

Lawrence, J. B., Villnave, C. A., and Singer, R. H. (1988). Sensitive, high-resolution chromatin and chromosome mapping in situ: presence and orientation of two closely integrated copies of EBV in a lymphoma line. Cell *52*, 51–61.

Meyne, J., Ratliff, R. L., and Moyzis, R. K. (1989). Conservation of the human telomere sequence $(TTAGGG)_n$ among vertebrates. Proc. Natl. Acad. Sci. USA *86*, 7049–7053.

Meyne, J., Baker, R. J., Hobart, H. H., Hsu, T. C., Ryder, O. A., Ward, O. G., Wiley, J. E., Wurster-Hill, D. R., Yates, T. K., and Moyzis, R. K. (1990). Distribution of non-telomeric sites of the $(TTAGGG)_n$ telomeric sequence in vertebrate chromosomes. Chromosome *99*, 3–10.

Miele, M., Bonatti, S., Menichini, P., Ottaggio, L., and Abbondandolo, A. (1989). The presence of amplified regions affects the stability of chromosomes in drug-resistant Chinese hamster cells. Mutat. Res. *219*, 171–178.

Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., and Wu, J.-R. (1988). A highly conserved repetitive DNA sequence, $(TTAGGG)_n$, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA *85*, 6622–6626.

Padgett, R. A., Wahl, G. M., Coleman, P. F., and Stark, G. R. (1979). N-(phosphonacetyl)-L-aspartate-resistant hamster cells overaccumulate a single mRNA coding for the multifunctional protein that catalyzes the first steps of UMP synthesis. J. Biol. Chem. *254*, 974–980.

Pauletti, G., Lai, E., and Attardi, G. (1990). Early appearance and long-term persistence of the submicroscopic extrachromosomal elements (amplisomes) containing the amplified DHFR genes in human cell lines. Proc. Natl. Acad. Sci. USA *87*, 2955–2959.

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc. Natl. Acad. Sci. USA *83*, 2934–2938.

Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. (1988). Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc. Natl. Acad. Sci. USA *85*, 9138–9142.

Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labeling deoxyribonucleic acid to high specificity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. *113*, 237–251.

Saito, I., Groves, R., Giulotto, E., Rolfe, M., and Stark, G. R. (1989). Evolution and stability of chromosomal DNA coamplified with the CAD gene. Mol. Cell. Biol. *9*, 2445–2452.

Schimke, R. T. (1988). Gene amplification in cultured cells. J. Biol. Chem. *263*, 5989–5992.

Shigesada, K., Stark, G. R., Maley, J. A., Niswander, L. A., and Davidson, J. N. (1985). Construction of a cDNA to the hamster CAD gene and its application towards defining the domain for aspartate transcarbamylase. Mol. Cell. Biol. *5*, 1735–1742.

Stark, G. R. (1986). DNA amplification in drug resistant cells and in tumours. Cancer Surv. *5*, 1–23.

Stark, G. R., and Wahl, G. M. (1984). Gene amplification. Annu. Rev. Biochem. *53*, 447–491.

Stark, G. R., Debatisse, M., Giulotto, E., and Wahl, G. M. (1989). Recent progress in understanding mechanisms of mammalian DNA amplification. Cell *57*, 901–908.

Stark, G. R., Debatisse, M., Wahl, G. M., and Glover, D. M. (1990). DNA amplification in eukaryotes. In Gene Rearrangement: Frontiers in Molecular Biology, B. D. Hames and D. M. Glover, eds. (Oxford: IRL Press), pp. 99–149.

Sutherland, G. R., and Hecht, F. (1985). Fragile Sites on Human Chromosomes (Oxford: Oxford University Press).

Trask, B. J., and Hamlin, J. L. (1989). Early dihydrofolate reductase gene amplification events on CHO cells usually occur on the same chromosome arm as the original locus. Genes. Dev. *3*, 1913–1925.

Wahl, G. M. (1989). The importance of circular DNA in mammalian gene amplification. Cancer Res. *49*, 1333–1340.

Wahl, G. M., Padgett, R. A., and Stark, G. R. (1979). Gene amplification causes overproduction of the first three enzymes of UMP synthesis in N-(phosphonacetyl)-L-aspartate-resistant hamster cells. J. Biol. Chem. *254*, 8679–8689.

Wahl, G. M., Vitto, L., Padgett, R. A., and Stark, G. R. (1982). Single-copy and amplified CAD genes in Syrian hamster chromosomes localized by a highly sensitive method for *in situ* hybridization. Mol. Cell. Biol. *2*, 308–319.

Zieg, J., Clayton, C. E., Ardeshir, F., Giulotto, E., Swyryd, E. A., and Stark, G. R. (1983). Properties of single-step mutants of Syrian hamster cell lines resistant to N-(phosphonacetyl)-L-aspartate. Mol. Cell. Biol. *3*, 2089–2098.

MOLECULAR AND CELLULAR BIOLOGY, June 1989, p. 2445–2452
0270-7306/89/062445-08$02.00/0

Vol. 9, No. 6

Evolution and Stability of Chromosomal DNA Coamplified with the CAD Gene

IZUMU SAITO,† RICHARD GROVES, ELENA GIULOTTO,‡ MARK ROLFE, AND GEORGE R. STARK*

Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2A 3PX, England

Received 29 September 1988/Accepted 9 March 1989

We have compared clones of Syrian hamster cells selected for the first amplification of the CAD gene with clones selected for further amplification. The large domain amplified initially was not reamplified as an intact unit. Instead, subregions were reamplified preferentially, and parts of the initial array were often lost. These events reduced the average amount of coamplified DNA accompanying each copy of the selected gene. The degree of amplification was small in the first step (about three extra copies of CAD per cell), but second-step amplifications to a high copy number (up to 60 extra copies per cell) occurred frequently. After several separate steps of amplification, highly condensed arrays that brought many CAD genes close together were formed. In striking contrast to the stability of these highly amplified arrays, the low-copy chromosomal arrays formed early were quite unstable and were often lost completely within 1 or 2 months of growth without selection. The results suggest that different mechanisms may be involved in the first step of amplification and in the later evolution of an already amplified array.

It is striking that cellular oncogenes are found to be amplified in many different human cancers, especially in highly malignant tumors, which strongly suggests a causative role for the overexpressed c-Onc proteins in tumor progression and possibly also in tumorigenesis (for reviews, see references 22, 23, 26, and 27).

Gene amplification is also a major cause of resistance to cytotoxic drugs for cells in culture. Such amplifications have been selected in distinct, operationally defined steps of increasing drug concentration. A primary event takes place in a single cell (and therefore has not been amenable to direct study so far), generating extra copies of the target gene or perhaps a form that will segregate unequally in subsequent cell divisions. Secondary events, not necessarily selected for by the drug, may then follow during growth of the initial cell to a clonal population large enough to be analyzed. If the amplified DNA is intrachromosomal, a second and distinct event, again taking place in a single cell, is required to generate the additional copies of the target gene needed for survival at a higher drug concentration. Although such a second event could involve any copy of the target gene, it seems to involve preferentially only some of the copies present in the amplified array formed after the first step (see below).

Syrian hamster (BHK) cells become resistant to *N*-(phosphonacetyl)-L-aspartate (PALA), a specific inhibitor of aspartate transcarbamylase, through amplification of the gene for the trifunctional protein CAD, which catalyzes the first three reactions of UMP biosynthesis. We have found this system to be particularly useful for carrying out a detailed analysis. Amplification is the only mechanism yet observed for resistance to PALA, and amplified CAD DNA is always found within chromosomes in BHK cells (30, 31) but not in mouse cells (17). Differential screening of bacteriophage libraries made from the DNA of the multistep mutant B5-4, which carries many copies of the CAD gene, led to isolation of a set of clones corresponding to large amounts of coamplified DNA (1, 3). Giulotto et al. (10) then screened a cosmid library from the first-step mutant 10.1 with several of the phage clones to isolate a set of larger clones corresponding to a 90-kilobase (kb) region, which includes the entire CAD gene, and also to six additional coamplified regions, totaling 290 kb, which have not yet been linked to each other or to CAD. The cosmids were used to study how often each of the six outlying regions was coamplified with CAD in 33 independent first-step PALA-resistant mutants and also to search for novel joints in the 380 kb of DNA that had been cloned. On the basis of (i) the finding that only three single-copy joints were detected in the 33 first-step mutants and (ii) the earlier observation that all of the first-step mutants tested contained large marker chromosomes carrying amplified CAD DNA (31), Giulotto et al. (10) suggested that the average amount of DNA coamplified with each copy of CAD could be as large as 10,000 kb (0.3% of a haploid genome). This surprisingly large estimate is consistent with estimates made in some other systems, which have been as large as 3,000 kb per copy of the selected gene (2, 12, 25). However, the amount of coamplified DNA is much smaller (a few hundred kilobases) when the copy number of the selected gene is as large as 1,000 per cell (6, 15).

Quite different pictures of the size and organization of amplified DNA have been arrived at by comparing situations in which different genes have been amplified to different degrees in cells of different species. To obtain a consistent picture in a single system, we have extended the analysis of first-step CAD amplifications in BHK cells begun by Zieg et al. (31) and continued by Giulotto et al. (10). The selection scheme used to isolate cells with new and independent first-step amplifications has now been applied to second and third steps of CAD gene amplification. Comparison of later steps with first steps has revealed many differences, leading us to imagine that different mechanisms may predominate in the first step, when normal DNA becomes amplified, and in later steps, when an already amplified array evolves further. Amplified DNA contains structures not present at the normal locus, such as novel joints and inverted repeats. Fea-

* Corresponding author.
† Present address: National Institute of Health, 2-10-35 Kamiosaki, Shinagawa-ku, Tokyo 141, Japan.
‡ Present address: Dipartimento di Genetica e Microbiologia, 27100 Pavia, Italy.

TABLE 1. Cosmids and probes from seven different regions coamplified in first-step PALA-resistant mutants[a]

Amplified region	Total length of amplified region cloned (kb)	Representative cosmid	Length of repeat-free probe (kb)
N1	60	276	3.0
N2	80	773	2.4
CAD	90	c64	2.1
F1	35	872	1.5
F2	40	171	1.4
F3	40	671	1.2
F4	35	971	1.1
SC		981	5.0

[a] Names of the original phage clones and the cosmids derived from them are given in Table 2 of Giulotto et al. (10). The probes are *Eco*RI fragments, free of repetitive DNA, derived from each cosmid as shown. Cosmid 981 contains a control single-copy (SC) sequence that is not amplified in PALA-resistant cells.

tures of these new structures may allow the primary event of a second step to proceed by a mechanism not available for the primary event of a first step.

MATERIALS AND METHODS

Cell culture and selection with PALA. BHK cells were grown as described by Swyryd et al. (28). Dialyzed fetal calf serum was used during drug selections. The nomenclature used for mutants is given in the legend to Table 2. All second- and third-step mutants were selected from first-step mutants isolated previously (10, 31) by using a procedure which ensured that each mutant was independent and that each initial amplification event occurred only a few generations before selection with PALA was imposed. Briefly, about 1,000 cells per dish were grown to about 5×10^5 cells per dish. The cells on each dish were dispersed and then selected with PALA. Independent clones (never more than one per dish) were picked only from those dishes with four or fewer colonies. The cells were then grown for a minimal time in the presence of PALA to small populations suitable for freezing and for preparing DNA (usually about 20 to 25 generations). When frozen stocks of cells were used for preparing more DNA, for in situ hybridizations, or for analyses of stability, the experiment was initiated within 1 week after thawing one of the original vials, during which time PALA was present at the concentration originally used for selection.

Analysis of copy number. Southern transfers of DNA were prepared and analyzed by the method of Giulotto et al. (10). Probes free of repetitive DNA were obtained from some of the cosmids described by these authors (Table 1). After autoradiography, the amount of DNA in each band was quantified by densitometry. Preflashed film was used in a linear range, and several different exposures were often used. The densitometric signals were normalized to a single-copy DNA standard. This ratio, the copy number per haploid genome (c), was used to calculate extra copies per cell ($2c - 2$) on the basis of the approximation that the cells are exactly diploid. Some experiments (see Table 5) were analyzed by the slot blot method of McIntyre and Stark (16), using the cDNA probe pCAD142 (24), with standardization to the total amount of DNA in each slot.

RESULTS

Fate of DNA amplified in the first step. We have monitored amplified DNA in second- and third-step mutants by determining copy numbers for a 90-kb region surrounding the CAD gene, for six coamplified regions that have not been linked to each other or to CAD, and for a control region. The analyses have been done simultaneously, by probing South-

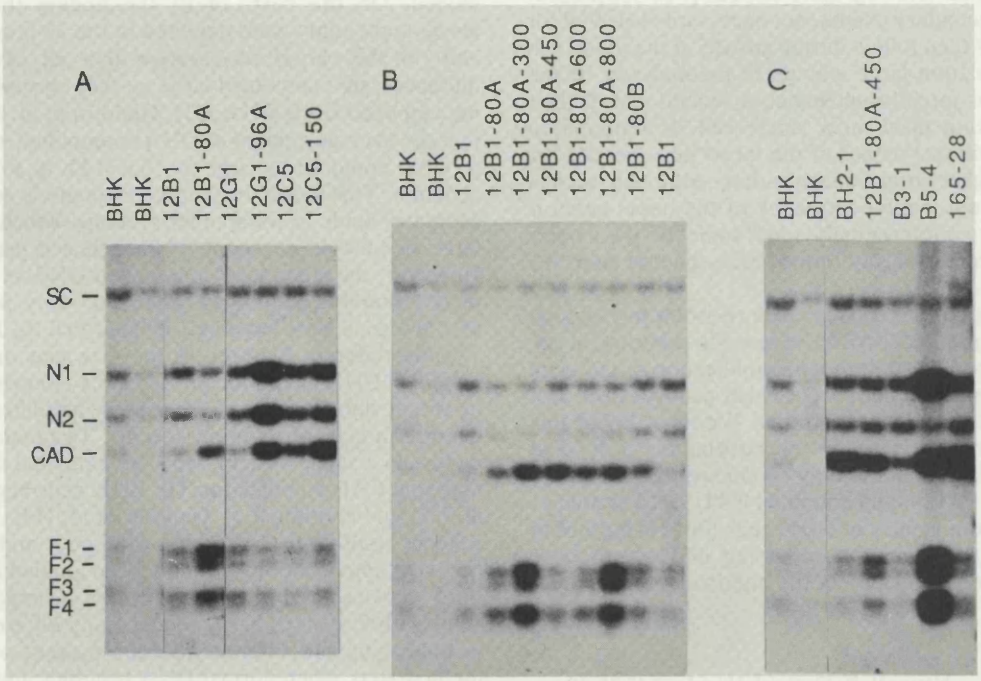


FIG. 1. Copy number analysis of mutants. *Eco*RI-digested genomic DNAs were probed with a mixture of fragments (Table 1). Half as much wild-type (BHK) DNA was loaded in the second lane of each set. (A) Second-step mutants; (B) third-step mutants; (C) multiple-step mutants.

TABLE 2. Fate of coamplified DNA in later steps of amplification[a]

Cell line	Extra copies of DNA/cell						
	CAD	N1	N2	F1	F2	F3	F4
12C5	7	2	2	0	0	0	0
12C5-150	26	5	6	1	1	1	1
12G1	1	1	1	1	1	1	1
12G1-96A	16	8	10	0	0	0	0
12G1-96B	4	4	2	1	1	1	1
12B1	3	2	2	2	2	2	2
12B1-80B	8	1	1	2	2	2	2
12B1-160	16	6	7	7	2	2	2
12B1-80A	11	1	0	9	9	9	2
12B1-80A-300	28	1	0	20	20	20	6
12B1-80A-450	22	1	0	3	3	3	1
12B1-80A-600	18	1	0	4	4	4	1
12B1-80A-800	22	1	0	20	20	20	6
B5-4 (multiple)	90	40	3[b]	40	40	40	40
BH2-1 (multiple)	30	0	0	0	0	0	0
165-28 (multiple)	400	4	5	3	3	3	3

[a] Mutants are named according to the PALA concentration (micromolar) used in each step of selection. As examples, 12B1 and 12C5 are independent first-step clones, each selected at 12 μM PALA, 12B1-80A-300 is a third-step mutant selected sequentially in 12, 80, and 300 μM PALA, and 12G1-96A and 12G1-96B are independent second-step mutants selected from 12G1 in 96 μM PALA in different plates. The names of multiple-step mutants, isolated previously and not selected in a clonal sequence, do not follow this convention. The probes used are shown in Table 1. The extra copy number per cell for each coamplified region has been normalized to values for unselected BHK cells, which are assumed to carry two copies of CAD per cell (no extra copies).

[b] Although the 2.4-kb N2 probe from cosmid 773 is amplified to only three extra copies per cell, most of the remaining DNA represented in this cosmid is amplified to 40 extra copies. An amplified novel joint, previously reported in phage 11-1700 (1), is present in B5-4 at the junction with the more highly amplified part of cosmid 773.

ern transfers of *Eco*RI-cut genomic DNAs from the PALA-resistant cell lines with a mixture of probes (Fig. 1), using repeat-free *Eco*RI fragments of different sizes derived from each region (Table 1). As shown by Giulotto et al. (10) and confirmed here, two of the coamplified regions, N1 and N2, were usually (30 of 33 mutants) amplified to the same extent as CAD in the first step, whereas the other four coamplified regions, F1 to F4, were amplified in only about half (17 of 33) of the first-step mutants. In only one case was the copy number of any coamplified region higher than that of CAD: in mutant 5.3, there was only one extra copy of CAD per cell, whereas there were four extra copies of N1 and two extra copies of N2.

Six independent second-step mutants were selected from the first-step mutants 12G1, 12C5, and 12B1. Although the copy number for CAD was always increased, the average amount of coamplified DNA was often decreased in the second step (Fig. 1 and Table 2). In the three second-step mutants derived from 12C5 and 12G1, there was no reamplification of regions F1 to F4 and, although regions N1 and N2 were reamplified in all three cases, the extent was usually less than that of CAD. In the three second-step mutants derived from 12B1, three different patterns of reamplification were seen. Only CAD was reamplified in 12B1-80B, only CAD and F1 to F3 were reamplified in 12B1-80A, and only CAD, N1, N2, and F1 were reamplified in 12B1-160. It is interesting to note that none of the six second-step mutants examined showed reamplification of CAD, N1, N2, and F1

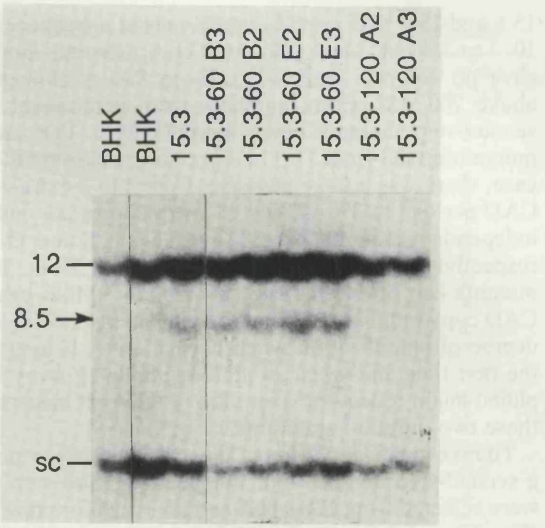


FIG. 2. Fate of novel joints formed in the first-step mutant 15.3. *Eco*RI-digested genomic DNAs from second-step mutants were probed with the 4.3-kb *Eco*RI-*Xho*I fragment purified from phage clone 11-200 and with the 5-kb single-copy fragment (SC) shown in Table 1. Half as much wild-type (BHK) DNA was loaded in the first lane. The *Eco*RI-*Xho*I fragment is fully homologous with the 8.5-kb novel *Eco*RI fragment (mapping data not shown).

to F4 together, in contrast with the first-step mutants, in which CAD and all of the more remote regions were often amplified together.

Single-copy novel joints mark specific sites of recombination in first-step mutants such as 15.3 and 30.7. The fates of two such joints were monitored in nine second-step mutants derived from 15.3, and some of the results are shown in Fig. 2. Most mutants maintained the novel joint at a single copy per cell (Fig. 2, lanes 4 to 7), but two clearly had lost the joint (lanes 8 and 9), although their CAD genes were amplified further. In some mutants, the novel joint was amplified slightly: compare the intensities of the 8.5-kb novel bands with those of the single-copy bands in Fig. 2, lanes 3 and 4. Four second-step mutants derived from 30.7 were also analyzed, and loss of the novel joint was observed in one (data not shown).

The degree of amplification that can be achieved in the first step is relatively small (10). On average, the 33 first-step mutants had 3.1 extra copies of CAD per cell, 2.6 extra copies of regions N1 and N2, and 1.6 extra copies of regions F1 to F4. In contrast, the degree of amplification achieved in the second step can be quite large. Three first-step mutants, originally selected in 15 μM PALA, were exposed to concentrations of PALA as high as 1,000 μM (Table 3). Mutants

TABLE 3. Second-step selections at high concentrations of PALA

PALA concn (μM)	Colonies/10^5 cells for given first-step mutant		
	15.6	15.2	15.3
100	ND[a]	115	4
200	34	23	5
350	15	ND	0
500	9	ND	0
1,000	3	2	0

[a] ND, Not determined.

15.6 and 15.2 both gave resistant cells at a frequency of 2×10^{-5} to 3×10^{-5} in 1,000 μM PALA, whereas mutant 15.3 gave no colonies from 10^5 cells in PALA concentrations above 200 μM. Copy numbers were determined for two second-step mutants newly derived from 15.6 and for a mutant derived from 15.3 in a previous experiment. In each case, there was a large increase, from 2 to 3 extra copies of CAD per cell to 37, 62, and 68 extra copies per cell for the independent mutants 15.6-500.1, 15.6-500.2, and 15.3-120D, respectively (slot blot data not shown). Thus, first-step mutants can often give rise to a 20- to 30-fold increase in CAD copy number in a single second step. The difference in degree of amplification when normal DNA is amplified for the first time and when an already amplified array is reamplified might reflect the operation of different mechanisms in these two different situations.

To investigate evolution of the amplified region present in a second-step mutant, four independent third-step mutants were selected from 12B1-80A and analyzed for copy number (Fig. 1B and Table 2). In 12B1-80A-300 and 12B1-80A-800, both the CAD gene and regions F1 to F3 were amplified approximately equally, just as in the step from 12B1 to 12B1-80A, which suggested that a unit similar to the one amplified in the second step was amplified again in the third step. Interestingly, in 12B1-80A-450 and 12B1-80A-600, the other two third-step mutants examined, the copy number of F1 to F3 decreased despite a further increase in the copy number of CAD. Analysis of the multistep mutants BH2-1 and 165-28, selected previously without cloning after each step (1), gave further examples in which coamplification of regions N and F with CAD was not maintained in later steps. In mutant BH2-1, regions N1, N2, and F1 to F4 were not coamplified with CAD at all, and there was only a little coamplification of these six regions in 165-28, a highly resistant mutant that has 400 extra copies of CAD (1). The multistep mutant B5-4 was the original source of phage clones for the six N and F coamplified regions, and therefore it is expected that these regions are highly amplified in this mutant.

Formation of new novel joints in later steps of amplification. The same Southern transfers of *Eco*RI-cut DNAs used for the analyses of copy number (Table 2) and two additional transfers of the same DNAs, but digested with *Kpn*I or *Apa*I, were probed with each of 10 different cosmids corresponding to 380 kb of cloned coamplified DNA (10). Only three single-copy joints were detected in 33 first-step mutants (Table 4). Low-copy joints were detected in 5 of the 15 stepwise and multistep mutants newly tested in this work (Table 4). In these cases, we did not perform the rigorous analysis (10) necessary to distinguish joints present in only one copy per cell from those present in a few copies per cell. In the second-step mutant 12G1-96B, the two low-copy joints detected were not present in the first-step parent 12G1 and therefore must have arisen during the second step of amplification. Similarly, the low-copy joint in the third-step mutant 12B1-80A-800 was not present in 12B1-80A. Thus, the types of recombination events that generated the single-copy joints detected in the first steps of amplification can also occur in later steps.

A very interesting observation was made with 12G1-96A. Two different amplified novel joints were found in this second-step mutant, and neither was present in the parental single-step mutant 12G1. One joint (Fig. 3) was 6 kb 3′ of the CAD gene, and the other (data not shown) was in group N2. There were 16 extra copies of CAD and 10 extra copies of the N2 fragment in 12G1-96A (Table 2), and the sequences beyond the joints were not amplified at all (data not shown). Since 12G1-96A was selected by using a scheme which ensured that the amplification event took place only a few cell divisions before selection with PALA was begun (Materials and Methods), and since two different amplified joints

TABLE 4. Distribution of novel joints in first-step and multiple-step mutants[a]

Mutants in which joints were detected	Amplified region			
	CAD 5′ (40)[b]	CAD 3′ (25)	Group N (140)	Group F (150)
First-step[c] (33 tested)				
15.6	S			
15.3			S (N1)	
30.7				S (F1)
Stepwise (12 tested)				
12G1-96A		H	H (N2)	
12G1-96B	L			L (F2)
12B1-80A-800		L		
Multistep (3 tested)				
BH2-1	L	L		
B5-4	H[d]	H	H (N2),[e] L (N2)	H (F2)[e]
165-28	H1,[e] H2	H[d]	L (N1)	

[a] Southern transfers of DNAs digested with *Eco*RI, *Apa*I, or *Kpn*I were probed with 10 cosmids representing 380 kb of DNA from seven different amplified regions (Table 1; 10). For representative examples of the data, see Fig. 3. S, Single copy of the novel joint per cell; L, low-copy joint, not distinguished from single copy; H, high-copy amplified joint.

[b] Numbers in parentheses are total lengths (in kilobases) of probes used.

[c] Previously reported by Giulotto et al. (10).

[d] Joint is at the center of an inverted duplication of CAD (21).

[e] Previously reported by Ardeshir et al. (1).

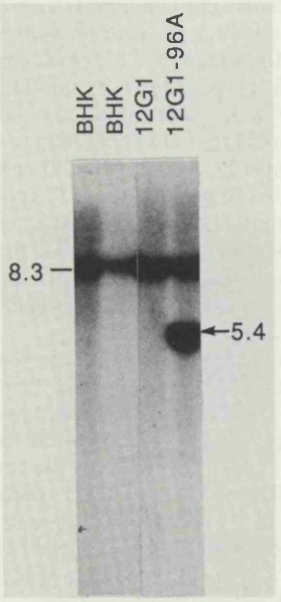


FIG. 3. Amplified novel fragments in 12G1-96A. Genomic DNAs were digested with *Eco*RI, and Southern transfers were probed with an 8.3-kb *Eco*RI fragment located next to the 3′ end of the CAD gene. The same probe also detects an amplified novel fragment in 165-28 (1). The position of the novel 5.4-kb band is indicated. Half as much wild-type (BHK) DNA was loaded in the second lane.

were present in the same mutant clone, it is very likely that both joints were formed and amplified during the same event. The de novo genesis of amplified joints has important mechanistic implications (see below).

Most of the novel joints detected in 12 stepwise and 3 multistep mutants were near the two ends of the CAD gene rather than in the six coamplified regions (Table 4). This concentration of joints near CAD is not random, since the regions flanking CAD represent only 65 kb of coamplified DNA, compared with 290 kb for the other regions. The predominance of joints near the ends of CAD and the observation that many of these joints are amplified has implications which are considered later.

Stability of the amplified DNA. Although it is clear that extrachromosomal amplified DNA (located, for example, in double minute chromosomes) is unstable in the absence of selection, much less work has been done to study the stability of chromosomally amplified DNA. Since PALA-resistant BHK cells have always been observed to carry amplified CAD DNA only on chromosomes and since double minutes have not been observed in these cells, they provide a good system in which to carry out such an analysis. In studying four multistep mutants, Wahl et al. (30) found that the amplified CAD genes of all were chromosomal and that in three of mutants the genes could be localized to the short arm of chromosome B9, the normal position of CAD. In studying nine first-step mutants, Zieg et al. (31) again found that the amplified CAD genes were chromosomal in every case and that they were usually localized to chromosome arms much longer than the short arm of B9, but their positions were not defined further. We had previously analyzed the stability of chromosomally amplified CAD genes only in multiple-step mutants and had found them to be remarkably stable. The same level of resistance was maintained for as long as 1 year in culture without PALA selection (14). Confirming these results, the copy number of CAD in mutant 165-28, resistant to 25 mM PALA, was unchanged after culture for 103 days (Table 5), as was the plating efficiency in high concentrations of PALA (data not shown). Similar results were obtained for the third-step mutant 12B1-80A-450. However, in striking contrast, the copy number of CAD was unstable during growth in nonselective medium for all seven first-step mutants and three second-step mutants tested (Table 5). As expected, loss of copy number was paralleled by loss of resistance to PALA. For first-step mutants 15.3 and 15.6, the plating efficiencies in 15 μM PALA were reduced by 90 to 95% after 60 days of growth in nonselective medium, whereas mutant 30.7 lost plating efficiency in 30 μM PALA more slowly, with a reduction of 70% after 60 to 90 days in nonselective medium. Two of the first-step mutants analyzed in Table 5, 10.1 and 20.2, had been shown previously by in situ hybridization (31) to have chromosomally amplified CAD genes, and 15.6 has recently been shown to carry all the amplified CAD genes in a single chromosome (data not shown). The remaining seven first- and second-step mutants tested have not yet been analyzed by in situ techniques, but it is important to note that we have never observed a PALA-resistant mutant of Syrian hamster cells without chromosomally amplified CAD.

TABLE 5. Stability of CAD gene amplification in BHK cells[a]

Mutant	Days in culture	Extra copies of CAD/cell
15.3	0	2
	39	0
15.6	0	3
	39	0.5
30.7	0	5
	103	0
10.1	7	9
	36	4.5
	63	1
20.2	7	13.5
	21	5.5
	36	0
15P1	5	6
	20	4
	41	0.5
15P2	5	4
	20	0.5
30.7-70B	18	31
	26	20
	38	9
	82	0.5
30.7-70.21	18	27
	46	16
	76	13
	91	3
15.3-120D	0	68
	19	17
	32	5.5
	64	0
12B1-80A-450	0	22
	103	22
165-28	0	400
	103	400

[a] Cell pellets, sampled from cultures grown without PALA for the times indicated, were dissolved in 0.4 M NaOH and analyzed by the slot blot method of McIntyre and Stark (16). Extra copies of CAD per cell are relative to numbers in unselected BHK cells, which are assumed to have two copies per cell (no extra copies). Not all time points analyzed are shown. Localization of amplified DNA in chromosomes of three of the cell lines was done by Zieg et al. (31) for 10.1 and 20.2 and recently for 15.6 (data not shown). In all cases, the cells for in situ analysis were grown from frozen cell stocks in the presence of PALA for about 1 week before each experiment, comparable to the times used for the samples analyzed for stability. Thus, cells grown for very similar lengths of time after selection are being compared. Also, the majority of 10.1 and 20.2 spreads had chromosomal amplifications. A comparable analysis has not yet been done for 15.6.

DISCUSSION

Loss of coamplified sequences and preferential reamplification. Region N was almost always coamplified with CAD (region C) in the first step, and region F was coamplified about half of the time. In contrast, regions C, N, and F were never observed to be reamplified together in subsequent steps. Regions N1 and N2 were lost in the amplification step between mutants 12B1 and 12B1-80A, and many copies of the amplified F1-F3 regions were lost in the third steps between mutants 12B1-80A and 12B1-80A-450 or 12B1-80A-600 (Table 2). The changes observed can be summarized in a model (Fig. 4) which proposes that each loss is accompanied by preferential reamplification of a subregion of the amplified array. There is no simple way that all of the N regions can be removed from a -CNFCNFCNFCNF- array as in 12B1 to give an $(\text{-FC-})_9$ array as in 12B1-80A. It seems much more likely that loss of all or most of the -CNFCNFCNFCNF- array occurred either at the same time as preferential reamplification of an -FC- subregion (required to give enough copies of C for selection) or shortly thereafter, during growth of the initial clone but before analysis of the DNA. A similar argument can be made for loss of part of an $(\text{-FC-})_9$ array in 12B1-80A together with preferential amplification of a -C- subregion in two independent third-step mutants.

Cell Line	Hypothetical structures of amplified DNA
wt-BHK	--CNF--
12B1	--CNFCNFCNFCNF--
12B1-80A	--CNF-- + $(--FC--)_9$
12B1-80A-450	--CNF-- + $(--C--)_{16}$ + $(--FC--)_4$

FIG. 4. Model for loss of amplified sequences and preferential reamplification. Shown are hypothetical structures that can account for the number of copies of C, N, and F found in the 12B1 cell lineage.

Regions N and F, which represent only a small fraction of the approximately 10,000 kb of DNA amplified initially, are usually lost together with most of the coamplified DNA in generating highly amplified arrays. Although the earlier steps have not been analyzed, events such as those depicted in Fig. 4 are likely to have taken place in the first-step ancestors of the multistep mutants BH2-1 and 165-28. In BH2-1, the sequences coamplified initially were lost completely, whereas in 165-28 a few copies of an array including C, N, and F, which must have been formed in the first step, still remained. However, most of the CAD genes are now found in a unit which includes neither N nor F. In addition to loss of amplified arrays, it was possible to observe in independent events loss or reamplification of specific sequences, marked by single-copy novel joints, in some second-step mutants derived from the first-step mutants 15.3 (Fig. 3) and 30.7 (data not shown). Gudkov et al. (11) also observed loss of initially coamplified DNA in later steps of amplification in their study of multidrug-resistant Djungarian hamster cells, as did Hyrien et al. (13) in their study of AMP deaminase amplification in Chinese hamster cells.

It is very likely that later amplification events involve the already amplified array rather than a region, for example, on a homologous chromosome, which has not yet been amplified. Thus, whenever region N or region F was not amplified in the first step, it was also not amplified later. Similarly, when the amplified copies of N were lost in mutant 12B1-80A, region N was not reamplified in four independent third-step mutants derived from 12B1-80A. In two of the four, the coamplified region containing CAD and F1 to F3 was amplified further as such. These phenomena can be understood if there is strong selection against further expansion of the already large amount of amplified chromosomal DNA present in first-step mutants (10) and if it is more likely that new amplification events occur within an already amplified array than in a region that has never been amplified.

If two copies of the CAD gene become juxtaposed during the first event, reamplification of this region should be favored, since PALA resistance can then increase twice as rapidly as it would through reamplification of a region containing only one CAD gene. The region reamplified in B5-4 fits such a model well. As shown previously (21), most copies of CAD in B5-4 lie within an inverted duplication, with their 5′ ends joined. There is also a novel joint near the 3′ end of the CAD gene in B5-4 (data not shown). The most highly amplified region of this mutant, which lies between these two novel joints, is only about 60 kb long. Similarly, amplified novel joints were detected on both sides of CAD in the highly resistant mutant 165-28, one near the 3′ end of the gene (1) and the other about 40 kb 5′ of the gene (data not shown). Therefore, the most highly amplified region in 165-28 is also quite short, about 65 kb in length. Similar small sizes have been reported for the most highly amplified regions flanking the dihydrofolate reductase gene in two cases in which the overall degree of amplification is very high (6, 15).

Novel joints were found preferentially near the ends of CAD in late stages of amplification but not in earlier stages (Table 4). There was only one single-copy joint near CAD in 33 first-step mutants, but there were six highly amplified joints near CAD in 15 later mutants. Whenever the copy number of CAD had increased more than that of N or F in a second or third amplification step (for examples, see 12G1-96A, 12B1-160, 12B1-80A-450, and 12B1-80A-600 in Table 2), it is possible that reamplification of a region in which two copies of CAD lie relatively close was selected preferentially. Knowledge that the most highly amplified region in highly resistant cells is likely to lie near the selected gene can be exploited to clone the gene, by walking from coamplified sequences that have been isolated previously. In the recent work of Debatisse et al. (5), the AMP deaminase gene was cloned in this way. It is also interesting that two striking examples of preferential reamplification were observed by Debatisse et al. (5). Two novel joints, one on either side of the AMP deaminase gene, generated independently in different first-step mutants, were amplified in every second- and third-step mutant studied by these authors.

Preferential reamplification and loss of previously amplified regions can be imagined to occur in more than one way. As suggested by Carroll et al. (4), a chromosomally amplified array might give rise to a relatively small extrachromosomal element that can be amplified further and then reinserted either into the original site or into a new site. In at least some cell lines, amplification is usually associated with chromosomal translocations and other rearrangements. (For a recent well-studied case and references to additional examples, see Biedler et al. [2].) Chromosome loss could also be involved. If a subregion of the original array were to be reamplified and moved to a new chromosome, loss of the chromosome carrying the original array would be possible.

Stability of amplified DNA. It was quite unexpected to find that chromosomally amplified CAD genes were readily lost from first- and second-step mutants in the absence of selection, in strong contrast to the stability of multistep mutants observed by us previously (14, 29) and confirmed here. From results of previous work (16, 22, 23, 27), instability has usually been associated with location of the amplified DNA in double minute chromosomes. To the best of our knowledge, there is only one other case in which an observation similar to ours has been published. Gudkov and Kopnin (12) found that amplified multidrug resistance genes in Djungarian hamster cells were unstable even when the majority of the amplified sequences were chromosomal rather than extrachromosomal.

Retention of a chromosomal array of amplified DNA may be selected against because a very large chromosome (10) is disadvantageous or because some coamplified sequence is harmful (for example, through overexpression of a coamplified gene). Eventual stabilization of an amplified chromosomal array could be achieved by deleting any destabilizing sequence, perhaps accompanied by further amplification of the selected gene. Such a process might occur by deletion and extrachromosomal amplification of a subregion containing the selected gene, followed by reinsertion by either homologous or nonhomologous recombination (4). We showed previously (30) that in three of four multistep mu-

tants the highly amplified CAD genes resided at or near the normal site of the unamplified CAD gene, the short arm of chromosome B9. Retention of the amplified array in the original location might represent either reamplification in situ or extrachromosomal reamplification followed by homologous insertion.

Are there different mechanisms in different steps? The work of several laboratories suggests that more than one process may be responsible for gene amplification in mammalian cells and that different mechanisms may predominate in primary and secondary events. In the first step of CAD gene amplification in Syrian hamster cells, each copy of CAD is amplified together with a very large amount of flanking DNA, estimated to be as much as 10,000 kb (10). Very few studies have been undertaken to analyze the structure of amplified DNA in first-step mutants, as it is a daunting prospect to attempt analysis of such a huge amplified domain in detail. Since the analyses we have performed so far were done after 20 to 30 generations (to accumulate enough cells), it is possible that the structures and properties observed represent secondary manifestations rather than primary events in first-step mutants. It is important to bring the analyses as close to the primary event as possible in the future.

After the initial observation of inverted and amplified novel joints by Ford et al. (7), several laboratories have shown that inverted joints are a common feature of highly amplified DNA (8, 13, 15, 18, 19, 21). In mechanisms proposed to account for these observations, an inverted duplication arises either before or during replication, and subsequent recombination leads to a circular structure in which two replication forks are moving in the same direction. Many copies of the circular element, including the inverted novel joint, can thus be formed. Such events are proposed to take place either extrachromosomally (19) or within replicating chromosomal DNA (13, 18). If the events are extrachromosomal, reintegration, as proposed by Passananti et al. (19) and Carroll et al. (4), could lead to eventual location of the amplified array either at the original site or at a different site. Hyrien et al. (13) have discovered an amplified inverted joint that arose de novo during the second step. In the work of Ma et al. (15) with Chinese hamster cells, it is clear that the two amplified inverted joints which characterize their type II units arose during the third or fourth steps. In the work reported here, we found two different amplified joints in the second-step mutant 12G1-96A that were not detected in the parental line 12G1. All of these examples are consistent with the newly proposed mechanisms, but it is also possible to explain them by overreplication and recombination if recombination occurs before most of the overreplication.

Whatever mechanism accounts for the majority of first events, different mechanisms, capable of giving a very large increase in copy number in a single step (see Table 3), may predominate later, perhaps dependent on some unusual feature of an already amplified array. In our work, the fact that the coamplification of regions C, N, and F often seen in the first step is not seen later suggests that different events are dominant in later steps; i.e., the amplification of a very large region occurs only during the first step in the great majority of Syrian hamster cell lineages resistant to PALA. However, the finding of low-copy-number novel joints that were not present in the parental cells in second- and third-step mutants (Table 4) does show that unique recombinations still occur in later steps.

Amplification can be stimulated by irradiating cells, treating them with chemicals, or depriving them of oxygen (20). Alternatively, it has been possible to isolate mutant amplificator cells which carry out amplification events much more often than do the parental cells (9). In the future, identifying and characterizing the gene products responsible for increasing rates of amplification and then comparing amplification events stimulated in different ways might give important clues about the mechanisms involved. It is important to examine the state and properties of amplified DNA in cells selected as soon after an amplification event as possible and to perform such analyses in as many situations as possible by examining different loci in cells of different species after different steps. In this way, we shall learn more about the various mechanisms proposed for amplification and about the situations in which each mechanism predominates.

ACKNOWLEDGMENTS

We are especially grateful to Catherine Knights for expert technical assistance and for her devotion to carrying out many complicated experiments. PALA was obtained from the Drug Synthesis and Chemistry Branch, Division of Cancer Treatment, National Cancer Institute.

LITERATURE CITED

1. **Ardeshir, F., E. Giulotto, J. Zieg, O. Brison, W. S. L. Liao, and G. Stark.** 1983. Structure of amplified DNA in different Syrian hamster cell lines resistant to *N*-(phosphonacetyl)-L-aspartate. Mol. Cell. Biol. **3**:2076–2088.
2. **Biedler, J. L., C. Tien-ding, K. W. Scotto, P. W. Melera, and B. A. Spengler.** 1988. Chromosomal organization of amplified genes in multidrug-resistant Chinese hamster cells. Cancer Res. **48**:3179–3187.
3. **Brison, O., F. Ardeshir, and G. R. Stark.** 1982. General method for cloning amplified DNA by differential screening with genomic probes. Mol. Cell. Biol. **2**:578–587.
4. **Carroll, S. M., M. L. DeRose, P. Gaudray, C. M. Moore, D. R. Needham-Vandevanter, D. D. Von Hoff, and G. M. Wahl.** 1988. Double minute chromosomes can be produced from precursors derived from a chromosomal deletion. Mol. Cell. Biol. **8**: 1525–1533.
5. **Debatisse, M., I. Saito, G. Buttin, and G. R. Stark.** 1988. Preferential amplification of rearranged sequences near amplified adenylate deaminase genes. Mol. Cell. Biol. **8**:17–24.
6. **Federspiel, N. A., S. M. Beverley, J. W. Schilling, and R. T. Schimke.** 1984. Novel DNA rearrangements are associated with dihydrofolate reductase gene amplification. J. Biol. Chem. **259**: 9127–9140.
7. **Ford, M., B. Davies, M. Griffiths, J. Wilson, and M. Fried.** 1985. Isolation of a gene within an amplified inverted duplication after "expression selection." Proc. Natl. Acad. Sci. USA **82**:3370–3374.
8. **Ford, M., and M. Fried.** 1986. Large inverted duplications are associated with gene amplification. Cell **45**:425–430.
9. **Giulotto, E., C. Knights, and G. R. Stark.** 1987. Hamster cells with increased rates of DNA amplification, a new phenotype. Cell **48**:837–845.
10. **Giulotto, E., I. Saito, and G. R. Stark.** 1986. Structure of DNA formed in the first step of CAD gene amplification. EMBO J. **5**:2115–2121.
11. **Gudkov, A. V., O. B. Chernova, A. R. Kazarov, and B. P. Kopnin.** 1987. Cloning and characterization of DNA sequences amplified in multidrug-resistant Djungarian hamster and mouse cells. Somatic Cell. Mol. Genet. **13**:609–619.
12. **Gudkov, A. V., and B. P. Kopnin.** 1987. Gene amplification and multidrug resistance. Sov. Sci. Rev. Sect. D Physiochem. Biol. **7**:95–136.
13. **Hyrien, O., M. Debatisse, G. Buttin, and B. Robert de Saint Vincent.** 1988. The multicopy appearance of a large inverted duplication and the sequence at the inversion joint suggest a new model for gene amplification. EMBO J. **7**:407–417.

14. **Kempe, T. D., E. A. Swyryd, M. Bruist, and G. R. Stark.** 1976. Stable mutants of mammalian cells that overproduce the first three enzymes of pyrimidine nucleotide biosynthesis. Cell **9:** 541–550.
15. **Ma, C., J. E. Looney, T.-H. Leu, and J. L. Hamlin.** 1988. Organization and genesis of dihydrofolate reductase amplicons in the genome of a methotrexate-resistant Chinese hamster ovary cell line. Mol. Cell. Biol. **8:**2316–2327.
16. **McIntyre, P., and G. Stark.** 1988. A quantitative method for analyzing specific DNA sequences directly from whole cells. Anal. Biochem. **174:**209–214.
17. **Meinkoth, J., A. M. Killary, R. E. K. Fournier, and G. M. Wahl.** 1987. Unstable and stable CAD gene amplification: importance of flanking sequences and nuclear environment in gene amplification. Mol. Cell. Biol. **7:**1415–1424.
18. **Nalbantoglu, J., and M. Meuth.** 1986. DNA amplification-deletion in a spontaneous mutation of the hamster *aprt* locus: structure and sequence of the novel joint. Nucleic Acids Res. **14:**8361–8371.
19. **Passananti, C., B. Davies, M. Ford, and M. Fried.** 1987. Structure of an inverted duplication formed as a first step in a gene amplification event: implications for a model of gene amplification. EMBO J. **6:**1697–1703.
20. **Rice, G. C., C. Hoy, and R. T. Schimke.** 1986. Transient hypoxia enhances the frequency of dihydrofolate reductase gene amplification in Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA **83:**5978–5982.
21. **Saito, I., and G. R. Stark.** 1986. Charomids: cosmid vectors for efficient cloning and mapping of large or small restriction fragments. Proc. Natl. Acad. Sci. USA **83:**8664–8668.
22. **Schimke, R. T.** 1984. Gene amplification in cultured animal cells. Cell **37:**705–713.
23. **Schimke, R. T.** 1988. Gene amplification in cultured cells. J. Biol. Chem. **263:**5989–5992.
24. **Shigesada, K., G. R. Stark, J. A. Maley, L. A. Niswander, and J. N. Davidson.** 1985. Construction of a cDNA to the hamster CAD gene and its application toward defining the domain for aspartate transcarbamylase. Mol. Cell. Biol. **5:**1735–1742.
25. **Shiloh, Y., J. Shipley, G. M. Brodeur, G. Bruns, B. Korf, T. Donlon, R. R. Schreck, R. Seeger, K. Sakai, and S. A. Latt.** 1985. Differential amplification, assembly and relocation of multiple DNA sequences in human neuroblastomas and neuroblastoma cell lines. Proc. Natl. Acad. Sci. USA **82:**3761–3765.
26. **Stark, G. R.** 1986. DNA amplification in drug resistant cells and in tumors. Cancer Surv. **5:**1–23.
27. **Stark, G. R., and G. M. Wahl.** 1984. Gene amplification. Annu. Rev. Biochem. **53:**447–491.
28. **Swyryd, E. A., S. S. Seaver, and G. R. Stark.** 1974. N-(phosphonacetyl)-L-aspartate, a potent transition state analog inhibitor of aspartate transcarbamylase, blocks proliferation of mammalian cells in culture. J. Biol. Chem. **249:**6945–6950.
29. **Wahl, G. M., R. A. Padgett, and G. R. Stark.** 1979. Gene amplification causes overproduction of the first three enzymes of UMP synthesis in N-(phosphonacetyl)-L-aspartate-resistant hamster cells. J. Biol. Chem. **254:**8679–8689.
30. **Wahl, G. M., L. Vitto, R. A. Padgett, and G. R. Stark.** 1982. Single-copy and amplified CAD genes in Syrian hamster chromosomes localized by a highly sensitive method for in situ hybridization. Mol. Cell. Biol. **2:**308–319.
31. **Zieg, J., C. E. Clayton, F. Ardeshir, E. Giulotto, E. A. Swyryd, and G. R. Stark.** 1983. Properties of single-step mutants of Syrian hamster cell lines resistant to *N*-(phosphonacetyl)-L-aspartate. Mol. Cell. Biol. **3:**2089–2098.