



RETURNING MATERIALS:

Place in book drop to
remove this checkout from
your record. FINES will
be charged if book is
returned after the date
stamped below.

--	--	--

PART I. STUDIES ON THE HAMSTER RIBONUCLEOTIDE REDUCTASE GENES

PART II. CONSTRUCTION OF MUTATIONS IN THE CHICKEN ADULT ALPHA GLOBIN
GENES

By

Paul F. Bates

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1985

ABSTRACT

I. STUDIES ON THE HAMSTER RIBONUCLEOTIDE REDUCTASE GENES

II. CONSTRUCTION OF MUTATIONS IN THE CHICKEN ADULT ALPHA GLOBIN GENES

By

Paul F. Bates

I. Studies were conducted in an attempt to isolate the hamster ribonucleotide reductase genes. A majority of the work described utilized DNA-mediated gene transfer techniques in an attempt to transfer and then specifically isolate the ribonucleotide reductase gene. The donor DNAs in the gene transfer experiments were from either hamster HUR-2 cells which contain an altered ribonucleotide reductase activity which confers a dominant hydroxyurea-resistant phenotype or hamster L3-107 cells which contain elevated ribonucleotide reductase levels and have an aphidicolin-resistant phenotype. Recipient cell lines in the gene transfer studies included chinese hamster ovary cells, hamster V79 tk⁻ cells, rat 3 tk⁻ cells, mouse L tk⁻aprt⁻ cells, and mouse LM cells. Neither the hydroxyurea-resistant nor the aphidicolin-resistant phenotype was transferred in these studies using any of the recipient cell lines or donor DNAs. Based on the number of cells transfected it was estimated that the transformation frequency of

the hydroxyurea-resistant phenotype is less than 2×10^{-9} . It was estimated from parallel studies with thymidine kinase phenotype transfer, that the frequency of hydroxyurea-resistant phenotype transfer is at least 1.5×10^3 fold lower than that for thymidine kinase phenotype transfer.

In other studies cloned DNAs containing the herpes simplex virus I and II ribonucleotide reductase genes were used as hybridization probes in Southern blotting experiments with hamster genomic DNA. Although many different stringencies were employed in these studies, specific hybridization of the herpes virus ribonucleotide reductase probes to the hamster genomic DNA was not seen. These experiments suggest that the hamster and herpes virus ribonucleotide reductase genes do not share sufficient homology at the nucleic acid level to be detected by hybridization.

II. Nucleotide sequence analysis has not allowed identification of the promoter elements of the chicken adult alpha globin genes, α^A and α^D . Therefore, several series of deletion mutants of the putative promoter regions of these two genes were constructed using plasmid clones. These deletions extend into the putative promoter regions from both the 5' and 3' directions, and were constructed in such a manner that they can be recombined for use in so-called linker-scanning mutagenesis studies.

-To my parents, for the genes and the environment that sparked my interest in science.

-To John Boezi, for his infectious enthusiasm and love for research.

-To Karen, for the love and encouragement that pushed me over the top when things were looking bleak.

ACKNOWLEDGEMENTS

Persons who deserve acknowledgement

Jerry Dodgson - intellectual and financial support, Friday afternoon
beer

Ed Frtisch - "the" cloning course, enthusiasm in lab

Hsing-Jien Kung - ideas and basketball scores

John Burczak - make life in E.L. shall we say "memorable"

Dan Robinson - ibid

Wynne Lewis - tissue culture help, discussions on life, "my dad
said..."

Karen Friderici - lots of general discussions, great food and parties

MaryBeth Raines - great volleyball and lots of fun in lab and out

Paul Boyer - late-night lab partner, constant supply of munchies

Michelle Fluck - only member of my committee to "stick it out" for 5
years

Sue Conrad - new insights on old problems, list of SF restaurants

"Sticky Ends" - allowed me to play shortstop (maybe short"drop")

"Fubars" - 2nd place isn't so bad

"the lab group" - Dave G., Dave B., Corrinne, Mark, JD, Sara, Dave B.,
Claudette, Larry, David H., Kristen, etc.

Julie Doll - Somehow deciphered my writing

Theresa Fillwock and Betty Brazier - cut through the red tape many
times

TABLE OF CONTENTS

	<u>Page</u>
DEDICATION.....	ii
ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	iv
LIST OF FIGURES.....	v
LIST OF TABLES.....	i

PART I

LITERATURE REVIEW.....	1
History of Project.....	1
Ribonucleotide Reductase.....	2
Gene Amplification.....	10
Gene Transfer.....	13
MATERIALS AND METHODS.....	19
Materials.....	19
Cell Culture.....	19
RNA Extraction.....	20
DNA Extraction.....	21
cDNA Synthesis.....	21
Enrichment of Aphidicolin-Resistant Cell Specific cDNA Sequences.....	21
Lamda Library Construction and Screening.....	22
DNA Mediated Gene Transfer.....	22
RESULTS.....	24
Introduction.....	24

	<u>Page</u>
cDNA Enrichment.....	24
Differential Screening of Southern Blots.....	27
Differential Screening of Lamda Libraries.....	27
Gene Transfer Experiments.....	30
Gene Transfer: Selection for APH Resistance.....	35
Gene Transfer: Selection for HU Resistance.....	37
Isolation of Higher Level HUR ^r -2 Mutants.....	40
Gene Transfer: Experiments with HUR ^r -2 (300) DNA.....	43
Serial Selection with G418 then HU.....	45
Herpes Simplex Virus Ribonucleotide Reductase Probes.....	46
DISCUSSION.....	50
REFERENCES.....	55
PART II	
LITERATURE REVIEW.....	59
Background.....	59
Chicken α Globin Genes.....	60
Eukaryotic Polymerase II Promoters.....	62
<u>In vitro</u> Mutagenesis.....	66
MATERIALS AND METHODS.....	70
Materials.....	70
Methods.....	70
RESULTS.....	71
α^D : Construction of 5' Deletions.....	71
Sequencing of α^D 5' Deletion Mutants.....	74
Location of Linkers in A5 and A10 Clones.....	79
α^D : Construction of 3' Deletions.....	86

	<u>Page</u>
α^D : Sequencing 3' Deletion Mutants.....	89
Construction of α^A 5' Deletions.....	92
DISCUSSION.....	100
REFERENCES.....	104
APPENDIX I.....	106
APPENDIX II.....	120

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
PART I		
1	Proposed model for the structure of the mammalian ribonucleotide reductase.....	4
2	Model for the allosteric regulation of ribonucleotide reductase.....	6
3	mRNA:cDNA hybridization scheme.....	26
4	Enriched cDNA: Southern blots.....	29
5	Cosmid rescue scheme.....	34
6	Comparison of mouse L tk ⁻ aprt ⁻ and CHO cell plating efficiencies in aphidicolin.....	39
7	Herpes simplex virus 1 and 2 ribonucleotide reductase maps.....	49
<u>PART II</u>		
1	Chromosomal arrangement of the chicken α globin genes.....	61
2	Typical eukaryotic RNA polymerase II transcription unit.....	64
3	Example of linker scanning mutagenesis.....	68
4	Scheme for construction of α^D 5' deletions.....	72
5	Strategy for sequencing α^D 5' deletion clones using a single sequencing reaction.....	76
6	Sequencing gel of α^D 5' deletion clones by sequencing reaction.....	78
7	Scheme for mapping α^D 5' deletion clones: One fragment length determination.....	81
8	Sequencing gel of α^D 5' deletion clone Xho I/Pvu II labelled fragments.....	83

<u>Figure</u>		<u>Page</u>
9	Summary of α^D 5' deletions.....	85
10	Construction of α^D 3' deletion mutants.....	87
11	Sequencing strategy for α^D 3' deletion clones.....	91
12	Summary of α^D 3' deletions.....	94
13	Construction of α^A 5' deletion mutants.....	96
14	Estimated locations of α^A 5' deletion clone endpoints in both directions.....	99

LIST OF TABLES

<u>Table</u>	<u>Page</u>
PART I	
1	Gene amplification in drug resistant cells.....11
2	Gene amplification in tumors and tumor cell lines.....14
3	Lamda library construction data.....31
4	DNA mediated transformation: APH selection results.....36
5	Hydroxyurea sensitivity of various cell lines.....41
6	DNA mediated gene transfer: Hydroxyurea selections.....42
7	DNA mediated gene transfer with HUR-2(300 DNA.....44

PART I

STUDIES ON THE HAMSTER RIBONUCLEOTIDE REDUCTASE GENES

LITERATURE REVIEW

History of Project

In prokaryotic systems mutations have been used extensively to dissect the processes involved in DNA replication. Similarly, it was believed that generation of mutations in the genes involved in DNA replication in eukaryotic cells would aid in analysis of their mode of DNA replication. To this end Dr. John Boezi's laboratory isolated a series of Chinese hamster ovary (CHO) cell mutants resistant to the drug aphidicolin (APH). Since APH had been shown to be a specific inhibitor of DNA polymerase α (Ikegami et al., 1978), it was thought that the APH resistant mutants would contain alterations in the polymerase enzyme. It was hoped these mutants would help clarify the role of polymerase α in replication and/or repair. Upon characterization, these APH resistant cells showed no change in polymerase α activity, but instead appeared to have increased levels of the enzyme ribonucleotide reductase (for a detailed review see Appendix 1). This led to the following suggestion regarding the mechanism of APH resistance in these cells: APH inhibits α -polymerase via competition with dCTP (Oguro et al., 1979); ribonucleotide reductase (RdRase) activity controls the cellular pool sizes of all the four dNTP's (Thelander and Reichard, 1979); thus by increasing the activity of RdRase in the cell and raising the pool sizes of the dNTP's (especially dCTP) the inhibition of α -polymerase by APH is relieved by dCTP competition. In agreement with this theory, increased resistance

to APH was correlated with larger dNTP pools and higher levels of RdRase activity (Sabourin et al. (1981).

Although there are several mechanisms whereby the enzymatic activity of RdRase could be increased, the manner in which these APH resistant mutants were selected suggested a specific mechanism might be responsible for the increased RdRase levels. CHO mutants resistant to high levels of APH were isolated by a series of selections in gradually increasing levels of the drug. Attempts to isolate such mutants by exposure to a single large dose of drug were unsuccessful (Sabourin et al., 1981). For numerous drugs such stepwise selection has been shown to cause gene amplification (see section on gene amplification for detailed review).

The intriguing possibility existed that these cells had amplified the RdRase genes; therefore it was decided to attempt to characterize these APH resistant mutants at the molecular level, specifically by isolation of the RdRase genes. When this work was initiated only two documented cases of gene amplification existed in mammalian cells and nothing was known of the mechanisms involved in amplification. Also very few genes coding for low abundance (sometimes called "housekeeping") messages had been cloned. The RdRase gene was interesting since it offered the possibility of research in both of these areas.

Ribonucleotide Reductase

Ribonucleotide reductase (EC1.17.4.1) is responsible for the conversion of all four ribonucleoside diphosphates to the corresponding deoxynucleotides required for DNA synthesis. The level of enzyme

activity is tightly coupled to DNA synthesis (Peterson and Moure, 1976) and has been shown to be cell cycle regulated (Kucera et al., 1983). Evidence suggests that this activity is the rate limiting step in DNA replication, and modulation of RdRase activity plays a key role in the regulation of cell division (Cory et al. 1975, Lewis et al., 1977). RdRase levels are very low in non-dividing cells or tissues and are elevated in rapidly growing tissues such as regenerating liver (Larsson, 1969) or malignant growths (Elford et al., 1970). The results of experiments with protein synthesis inhibitors suggest that the half-life of RdRase activity is only two hours, and the level of RdRase activity is at least partially controlled by turnover of the enzyme (Turner et al., 1968; Elford, 1972).

In nearly all organisms examined, the structure of the RdRase enzyme consists of two non-identical subunits with a molecular weight of between 250,000 and 280,000. Each subunit is a dimer of identical polypeptides, thus the enzyme is an $\alpha\alpha$, $\beta\beta$ complex (see Figure 1). The $\alpha\alpha$ dimer is generally referred to as the M1 or B1 subunit in the mammalian and bacterial systems respectively. Similarly the $\beta\beta$ dimer is called the M2 or B2 subunit (for a review see Thelander and Reichard, 1979).

Not only is the overall architecture of RdRase highly conserved, but the proposed subunit functions are also very similar between the bacterial and mammalian enzymes. The M1 subunit is thought to bind the allosteric effectors which regulate the activity and specificity of the enzyme. M2 contains the tyrosine free radical essential for enzyme activity (Thelander et al., 1980). It is unclear whether the catalytic site is contained on one of the subunits or whether it lies between

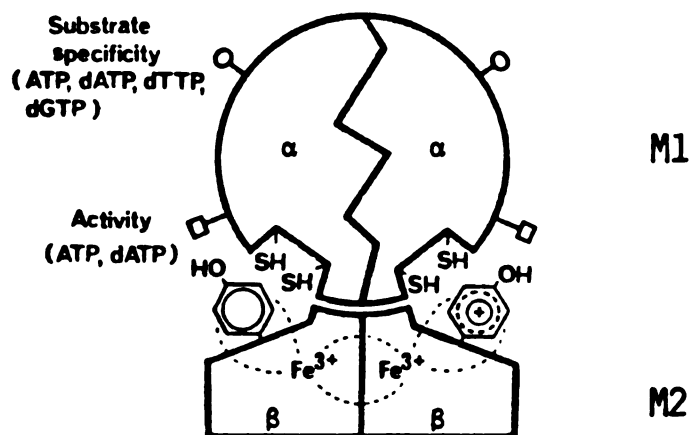


Figure 1. Proposed model for the structure of the mammalian ribonucleotide reductase. The M1 (α, α) subunit contains the two effector binding sites. The substrate specificity site controls the selection of the NDP to be reduced, while the activity site controls the overall catalytic activity of the enzyme. The M2 (β, β) subunit contains the tyrosine free radical required for enzyme activity. (From Thelander and Reichard, 1979).

them. Photoaffinity labelling of RdRase with [^{32}P]CDP shows that M1 contains a site involved in substrate binding (Caras et al., 1983). The role of M2 in the formation of the substrate binding catalytic site may be direct or indirect (for example M2 may induce a conformational change in M1). The fact that M2 contains the tyrosyl free radical required for RdRase activity argues for a more direct role of M2 in catalysis.

RdRase exhibits a complex set of strict allosteric controls based upon the levels of ribo- and deoxyribonucleotides in the cell (Figure 2). (For a review see Thelander and Reichard, 1979). Models for the allosteric effects support the theory that the RdRase contains at least two types of effector binding sites that are distinct from the catalytic site. Studies with the E. coli enzyme suggest that one of these effector sites controls overall enzyme activity while the other determines substrate specificity (Brown and Reichard, 1969). Photoaffinity labelling of the effector sites of the mammalian enzyme suggests they are located on M1 (Caras and Martin, 1982). It is noteworthy that the allosteric control mechanism of RdRase from E. coli and mammals is almost indistinguishable (Thelander and Reichard, 1979).

The E. coli ribonucleotide reductase has been purified to homogeneity (Erikson et al., 1977) and the structural genes (nrd A and nrd B) have been defined by mutational analysis (Fuchs, 1977). Recently the E. coli nrd A and nrd B genes have been cloned and their nucleotide sequence completely determined (Carlson et al., 1984). Interestingly, it has been shown that perturbations in E. coli DNA replication lead to increased levels of the polycistronic nrd mRNA

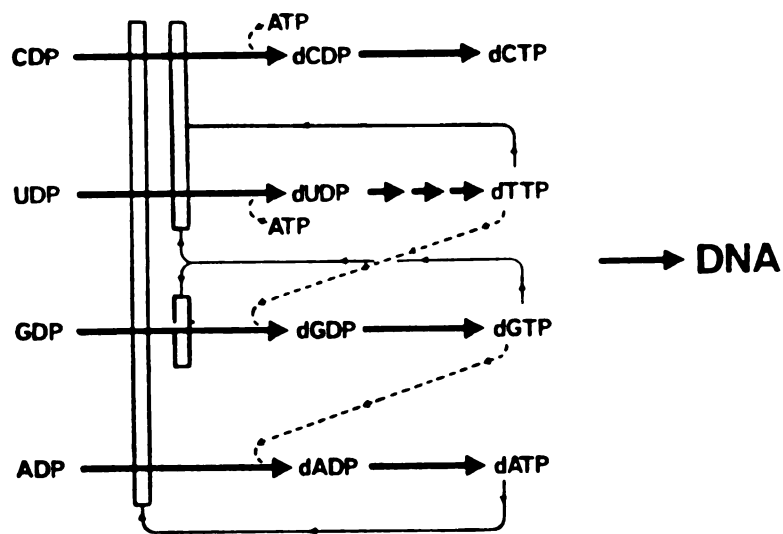


Figure 2. Model for the allosteric regulation of ribonucleotide reductase in mammalian cells. The broken lines represent positive effectors whereas the open boxes are negative effectors. For example, ATP stimulates CDP and UDP reduction while dATP has a negative effect on reduction of all four NDPs. (From Thelander and Reichard, 1979.)

suggesting that in E. coli the RdRase activity is controlled at the level of mRNA transcription (Hanke and Fuchs, 1983).

The mammalian RdRase has proven very difficult to purify. Only recently have pure preparations of the M1 subunit been obtained (Thelander et al., 1980). The M2 subunit has never been purified and appears to be very labile (Thelander et al., 1980). This instability of M2 may be due to degradation of the protein itself or destruction of the tyrosine free radical in M2 which is required for enzyme activity (Akerblom et al., 1981). Purification schemes which avoid separation of the M1 and M2 subunits always yield M2 in low, substoichiometric amounts (Thelander et al., 1980) suggesting that the M2 protein and not just the free-radical is unstable. In fact, two separate studies have suggested the M2 is the RdRase rate controlling subunit and regulates the overall levels of activity.

In one report, immunoprecipitation of M1 showed that the levels of this subunit were fairly constant over the cell cycle, while M2 activity was decreased in G1 cells and high in S phase cells (Eriksson and Martin, 1981). The other study used a mutant mouse 3T6 cell line which showed a 15-fold increase in RdRase activity. Extracts from these cells showed no increase in M1 protein when compared to wild type 3T6 cells. Only the M2 activity in the mutant cells was increased. Also the amount of M2 tyrosine free radical, as measured by electron paramagnetic resonance (EPR) spectroscopy, was greatly increased in the mutant cells (Akerblom et al., 1981).

Many viruses which require a supply of dNTP's for replication have been shown to either code for their own RdRase or induce the cellular enzyme. These viral enzymes also exhibit many of the similar

properties observed in both the bacterial and eukaryotic RdRase enzymes.

In the E. coli system, the T4 bacteriophage-encoded RdRase has been described (Berlund, 1972). Like the E. coli and the mammalian enzymes, it also has an $\alpha\alpha$, $\beta\beta$ configuration, and the subunit sizes of 160,000 ($\alpha\alpha$) and 80,000 ($\beta\beta$) are also very similar. Unlike the other enzymes, T4 RdRase responds to only positive allosteric effectors (affecting substrate specificity). It does not respond to dATP as a negative regulator (Berglund, 1972). This pattern of response can be explained if the T4-encoded RdRase lacks the activity sites on the model in Figure 1. Another similarity demonstrated by the T4 RdRase is the presence of a tyrosine free radical in the B2 subunit. Hyperfine EPR spectra show the geometry of the tyrosine free radical is very similar between T4, E. coli, and mammalian RdRase (Graslund et al., 1982).

Infection of mammalian cells with herpes simplex virus (HSV) also induces RdRase activity. This RdRase has recently been shown by genetic (Bachetti, 1983) and biochemical data to be encoded by the viral genome (Lankinen et al., 1982). Like the T4-encoded enzyme, the herpesvirus RdRase appears to lack the negative allosteric effector sites. The EPR spectra of HSV infected cells was similar to, but distinct from, that of either the T4 or mammalian enzymes (Lankinen et al., 1982) suggesting a similar environment around the tyrosine free radical. Genetic mapping studies place the HSV RdRase genes in a region of the genome which codes for two polypeptides; one of MW 140,000 and the other of MW 38,000 (Draper et al., 1982). Recently the homologous region of the Epstein-Barr virus (EBV) genome has been

322

323

324

325

326

327

328

329

330

331

332

333

334

335

sequenced. Coding regions for 35 and 93 kilodalton (Kd) proteins were found (Baer et al., 1984). These proteins were highly homologous to the 38 and 140 Kd HSV proteins (Gibson et al., 1984). Comparison of the amino acid sequences of the EBV and E. coli polypeptides has shown that the 93 Kd protein shares extensive homology with the B1 subunit (nrd A gene product). However no high degree of homology could be demonstrated between the 38Kd EBV protein and the B2 subunit (nrd B gene product) (J. Campbell, personal communication).

RdRase activity can also be induced in mammalian cells by infection of non-dividing tissue culture cells with DNA tumor viruses that stimulate DNA replication (Linberg et al., 1967) or by treatment of growth arrested cells with certain mitogens (Tyrsted, 1984). In the case of viral induction, the virus must be inducing the cellular enzyme since the viral genome contains no coding regions for its own RdRase. Induction of other enzymes involved in DNA replication, including thymidine kinase, DNA polymerase α , thymidylate synthetase, and dihydrofolate reductase, has also been seen with both viruses and mitogens (Tyrsted, 1984; Linberg et al., 1969). The mechanism by which viruses induce enzymes has not been elucidated.

Use of mutants of RdRase in mammalian cells has allowed some preliminary information on this enzyme to be deduced. In general two main classes of RdRase mutants have been described in mammalian cells. These include: a) mutants with altered allosteric control such that they are insensitive to dATP inhibition (Meuth et al., 1976; Eriksson et al., 1981), and b) mutants which are resistant to the drug hydroxyurea (HU) (Lewis and Wright, 1974; Akerblom et al., 1981). Hydroxyurea is a specific inhibitor of RdRase and is thought to act by destroying

the tyrosine free radical of the M2 subunit. Lewis and Wright have isolated a CHO cell line which by genetic and biochemical data appears to produce an altered RdRase M2 subunit which is more resistant to HU. In cell fusion studies this RdRase gene acts as a dominant marker. This cell line is called HU^r-2 (Lewis and Wright, 1978) and is resistant to low (75 µg/ml) concentrations of HU. Another type of HU^r mutant was isolated by Akerblom et al. (1981). Unlike the HU^r-2 mutants, these mutants are resistant to very high levels of HU (about 450 µg/ml). These highly resistant mutants could only be selected by continuously culturing the cells in slowly increasing HU concentrations over a period of two years. These mutants appear to overproduce a normal RdRase.

Gene amplification

In the mid 1970's mammalian cells were used increasingly for the study of gene expression. Many of these studies used resistance to specific inhibitors of enzymes essential for cell growth to select cell lines which overproduce the target enzyme. In several cases the drugs being used as selective agents were cancer chemotherapeutic drugs, and resistance to the drug was found in both cultured cells and cells originating from human neoplasias.

In 1978, Alt et al. reported that resistance of a cell line to the drug methotrexate was caused by a selective amplification of the genes encoding the target enzyme, dihydrofolate reductase (DHFR) and the resulting overproduction of DHFR. Since then numerous examples have been demonstrated of cells in which there is a substantial elevation in the levels of an enzyme in response to selective pressure (Table I). In

Table 1. Gene amplification in drug-resistant cells^a

Drug(s)	Enzyme or binding protein
Methotrexate	*DHFR (two distinct genes)
Methotrexate	*Bifunctional thymidylate synthetase dihydrofolate reductase (Leishmania)
PALA	*CAD
Cadmium	*Metallothionein I
Copper	*Copper-chelatin (yeast)
6-Azaauridine or pyrazofurin	*UMP Synthetase
Adenosine, alanosine, and thymidine ("HAT")	*Adenosine deaminase
Compactin	phosphoribosyl transferases
	*3-Hydroxy-3-methylglutaryl coenzyme A reductase
Methionine sulfoximine	*Glutamine synthetase
5-Fluorodeoxyuridine	Thymidylate synthetase
α -Methyl- or α -difluoromethyl-ornithine	Ornithine decarboxylase
Hydroxyurea or deoxynucleotides	Ribonucleotide reductase
Aphidicolin	DNA polymerase α (<u>Drosophila</u>)
Aphidicolin	Ribonucleotide reductase
Adenine and coformycin	AMP deaminase
Albizzin or β -aspartyl hydroxamate	Asparagine synthetase
Mycophenolic acid	IMP-5'-dehydrogenase
Tunicamycin	N-Acetyl glucosaminyl transferase
Borrelidin	Threonyl-tRNA synthetase
Multiple drugs	170K P-glycoprotein, 19K cytoplasmic protein
Canavanine	Argininosuccinate synthetase

*Genes have been proven to be amplified by molecular analysis.

a) Adapted from Stark and Wahl (1984).

about fifty percent of these cases gene amplification has been conclusively demonstrated as the molecular basis for the drug resistant phenotype (for a review on amplification see Stark and Wahl, 1984).

One common feature of cell lines which contain amplified genes is that all have been selected by increasing the concentration of the inhibitor in small increments or steps. In unselected populations of cells there are at any given time cells with varying levels of the target enzyme, and, it is believed, low levels of amplification of the gene coding for this protein (Zieg et al., 1983; Johnston et al., 1983). By initially using low levels of inhibitor, cells with slightly increased levels of enzyme (and presumably slightly amplified genes) can be selected. After allowing for re-amplification of these same genes, mutants with higher levels of enzyme can gradually be selected in a series of steps.

Amplification can apparently occur spontaneously at many (or all) locations in the genome; however, not all sites appear to be amplified at the same rate (Wahl et al., 1984). It has been estimated from genetic (Luria-Delbruck fluctuation analysis) and biochemical (fluorescent activated cell sorting) studies that a single DNA segment may be amplified as often as 10^{-3} to 10^{-4} times per cell per generation (Zeig et al., 1983; Johnston et al., 1983).

Another feature of gene amplification is that generally large regions surrounding the selected gene are amplified. The amount of DNA co-amplified with the selected gene can be as large as 3000 kilobase pairs (kb) and is generally at least 500 kb (Cowell, 1982). This amplified DNA may be located within chromosomes in expanded regions termed homogeneously staining regions (HSRs), or it may be on small

a
a
r
o
w
e
c
(
a
c
s
c
t
l
a
l
e
e
e
e

acentric chromosome fragments called double minutes (DMs). Generally, amplifications on HSRs are stable whereas DMs tend to segregate randomly during cell division because of a lack of a centromere and are often lost in the absence of selective pressure. It remains unclear whether DMs and HSRs are interconvertible or are distinct genetic elements (for a review see Cowell, 1982; Biedler et al., 1983).

DMs and HSRs have also been observed in numerous tumors and tumor cell lines implying that amplification may play a role in tumorigenesis (Varmus, 1984). Table 2 lists several known oncogenes which are amplified in tumors or tumor cell lines. Overproduction of normal cellular gene products is thought to be the basis for tumorigenesis in several current models (Varmus, 1984). Amplification may allow overproduction of factors which allow cells to proliferate more rapidly, escape immune surveillance, or increase motility for invasion of surrounding tissue.

In light of the randomness and size of the regions amplified, the process of "gene amplification" might more aptly be called "sequence amplification". Clearly if amplification occurs at a rate of 10^{-4} and the region amplified is 10^3 kb long, then there is a high probability of having a region amplified in a cell with a genome complexity of 3×10^6 kb. Thus it is not surprising that so many examples of amplification have been found.

Gene Transfer

A recently developed technique allows the efficient transfer of exogenous DNA into eukaryotic cells. This procedure, called DNA mediated gene transfer, relies on the ability of cells in culture to

Table 2. Gene amplification in tumors and tumor cell lines.

<u>oncogene</u>	<u>Tumor or cell line</u>	<u>Degree of Gene Amplification</u>	<u>Chromosomal Location</u>
N-myc	Kelly (N)	100-200X	HSR(13p)
"	NGP (N)	120-140X	HSR(Rp16, 12q13)
"	NLF (N)	20-25X	HSR(9q13)
"	CHP-134 (N)	20-25X	HSR(6q,7p)
"	CHP-126 (N)	100-120X	HSR(6q33) and DM
"	IMR-32 (N)	15-20X	HSR (1p34)
"	NMB (N)	100-120X	HSR(13p11) and DM
"	AR	80-100X	?
"	IMR-32	50X	HSR(1p3)
"	Primary neuroblastomas (Stage III and IV)	5-100X	DM/HSR
"	Retinoblastoma cell line, Y79, and primary tumors	10-200X	DM/HSR
"	Small cell lung carcinoma cell lines and tumor	50X	HSR
C-myc	COLO 320 (NE)	32X	DM
"	COLO 320 (NE)	16-32X	HSR
"	HL-60 (PML)	16-32X	?
"	Promyelocytic leukemia cell line, HL60 and primary tumor	20X	HSR
"	APUDoma cell line, COL0320	40X	DM/HSR
"	Small cell lung carcinoma cell lines (variants)	40X	DM/HSR
"	CSV-induced bursal lymphoma	5-10X	DM
"	Y1(ACT)	30-60X	HSR
"	Y1(ACT)	30-60X	?
"	Primary carcinomas of lung, colon, bladder and rectum	4-20X	
"	Colon carcinoma cell line, SW480		

c-Ki-ras	Adrenocortical carcinoma cell line, Y1	30-60X	DM/HSR
N-ras	Mammary carcinoma line, MCF7	5-10X	?
c-abl	Chronic myeloid leukemia cell line K562	5-10X	?
c-myb	Colon carcinoma cell lines, COL0201/205	10X	?
c-myb	Actue myeloid leukemia	5-10X	?
c-erbB	Epidermoid carcinoma cell line, A431	30X	?

Adapted from Stark and Wahl (1984) and Varmus (1984)

take up DNA which is added as a calcium phosphate-DNA precipitate. The use of this procedure to incorporate DNA into cells was first described by Graham and Van der Eb in 1973 and was used to infect cells with adenovirus DNA. More recently, total cellular DNA was used to transfer the tk^+ phenotype to tk^- mouse cells. These experiments demonstrated that single copy genes in the eukaryotic genome could be efficiently transferred with this technique (Wigler et al., 1978).

The efficiency with which single copy genes are transferred using DNA mediated gene transfer varies somewhat, but is usually 0.1 to 10 transformants per 20 μ g donor DNA per 1×10^6 recipient cells (Wigler et al., 1978; Wigler et al., 1979). In other words, with this procedure one in 10^5 to 10^7 cells become stably transformed and is isolated by biochemical selection. Since the reversion rate of many mutations is in the range of 10^{-7} - 10^{-8} , it should be feasible to use DNA mediated gene transfer for isolation of genes for which there is a biochemical selection.

One factor which has been shown to greatly influence the transformation efficiency is the cell line used as the recipient (Graf et al., 1979). Mouse L cells and mouse NIH 3T3 cells have been used extensively in transfer experiments because of their high efficiency of transfer. Another factor affecting transformation efficiency is the size of the donor DNA. Genomic DNA fragments greater than 50 kb generally yield the best results. This is apparently due to the high degree of recombination and rearrangement that occurs at the ends of the linear molecules. Perucho et al. (1980) demonstrated that upon transformation the host cell ligates the incorporated DNA into

conc

the

conc

trag

se'

poss

man

eat

mus

con

re

ei

be

re

re

'r

cc

m

a

a

e

>

t

:

concatameric structures which may be as long as 2000 kb. They coined the term "pekelasome" to describe this concatameric structure.

The fact that the transferred DNA is ligated into a single large concatamer before integration into the host genome, allows transformation of cells with unselectable genes. Inclusion of a selectable marker (often a cloned gene) with the donor DNA makes possible the transfer of genes which by themselves have no phenotypic marker (Wold et al., 1979).

The application of DNA mediated gene transfer for the isolation of eukaryotic genes relies on two criteria. First, the gene of interest must be selectable or linked to a selectable marker. Secondly, the donor DNA sequences must be discernable against the background of the recipient genome; that is, they must be "tagged" in some fashion with either a physical or biological marker. For example, the donor DNA can be ligated to or transfected with (since ligation will occur in the recipient cell) pBR322 plasmid DNA. After DNA mediated transfer, the replication origin and drug resistance genes of pBR322 may be used to "rescue" the adjacent donor DNA by their ability to replicate in E. coli (Lowy et al., 1980). Alternatively, naturally occurring repetitive sequences, such as the Alu repeats in human DNA, may be used as a hybridization probe to identify the transferred DNA (Perucho et al., 1981).

Using procedures such as those described above, numerous eukaryotic genes have been isolated with DNA mediated gene transfer. These include several human oncogenes (for a review see Bishop, 1983), the human, hamster, and chicken thymidine kinase genes (Lewis et al., 1983; Perucho et al., 1980b; S. Conrad, personal communication), the

hamster and mouse adenine phosphoribosyl transferase genes (Lowy et al., 1980; Sikela et al., 1983), human DNA repair enzyme genes (Rubin et al., 1985) and many others. Clearly this technique will continue to be useful in the isolation of eukaryotic genes.

MATERIALS AND METHODS

Materials

The aphidicolin used in these studies was supplied by the Developmental Therapeutics Program, Chemotherapy, National Cancer Institute. G418 was from Scherring Pharmaceuticals. Restriction enzymes were from New England Biolabs, Bethesda Research Labs, or Boehringer Mannheim. Mouse L tk⁻aprt⁻ cells were from Michael Wigler (Cold Spring Harbor Laboratory). V79tk⁻ cells were from William Lewis (University of Toronto). HUR-2 cells were from James Wright (University of Manitoba). Rat 3 and human 293 cells were from Sue Conrad (Michigan State University). The herpes simplex RdRase plasmid clones were from Sandy Weller (Dana Farber Cancer Institute).

Cell Culture

Growth of parental-type chinese hamster ovary cells and growth of aphidicolin resistant mutants was as described (Sabourin et al., 1981). L-tk⁻,aprt⁻ cells were cultured in Dulbecco Modified Eagle's Medium (Grand Island Biological Co., Grand Island, NY) supplemented with 10% calf serum. Maintenance of the aprt⁻ phenotype required addition of 50 µg/ml diaminopurine. Tk⁺ cells were selected in HAT medium (15 µg/ml hypoxanthine, 0.2 µg/ml aminopterin, 5 µg/ml thymidine). All other cell lines were maintained in DME with 10% calf serum except for HUR-2 and V79 tk⁻ cells which were kept in αMEM + 10% fetal calf

serum. G418 selections were done in regular media supplemented with 500 $\mu\text{g/ml}$ G418. Hydroxyurea was dissolved in sterile 10 mM Hepes pH 7.0 at a concentration of 35 mg/ml and added to the media at the concentrations indicated.

Plating efficiencies were determined by plating 2×10^2 to 5×10^5 cell per plate at various concentrations of aphidicolin. Surviving colonies were scored by staining with methylene blue. Growth in HU was determined by plating 5×10^5 cells per 100 mm plate and scoring growth on an arbitrary scale (-, +/-, +, ++).

RNA Extraction

Total cellular RNA was prepared from pelleted cells by lysis in guanidinium thiocyanate. Five volumes of 4 M guanidinium thiocyanate, 1% sodium lauryl sarcosinate, 0.15 M β -mercaptoethanol, 10 mM EDTA, 50 mM Tris, pH 7.6 were added to the cell pellet at room temperature. The cell pellet was dissolved and DNA released from the cells was sheared by homogenization with a tight fitting dounce homogenizer. One gram of CsCl was added per 2.5 ml of homogenate. The homogenate/CsCl solution was layered over 1.2 ml of 5.7 M CsCl in 0.1 M EDTA in a Beckman SW 50.1 polyallomer tube. RNA was pelleted for 12 hrs at 35,000 rpm at 20°C in a SW 50.1 rotor. The supernatant was poured off and the RNA was resuspended in 10 mM Tris pH 7.5, 5 mM EDTA, 1% SDS (TES). The cloudy mixture was extracted once with an equal volume of CHCl_3 :Butanol (4:1). The organic phase was re-extracted with an equal volume of TES. The aqueous phases were combined and RNA was ethanol precipitated. Polyadenylated RNA (Poly-A+) was prepared using

an oligo dT cellulose column. In all work, baked glassware and diethylpyrocarbonate-treated, autoclaved solutions were used.

DNA Extraction

High molecular weight DNA for use in λ library construction and DNA mediated gene transfer experiments was prepared as described by Blin and Stafford (1976). The DNA obtained was routinely at least 100 kb in length as judged by agarose gel electrophoresis in a 0.2% gel. Plasmid DNAs were prepared by the methods outlined by Maniatis et al. (1982). Fragments to be used as probes were excised with the appropriate restriction enzymes, eluted from agarose gels (Girvitz et al., 1980) and labeled by nick-translation (Maniatis et al., 1982).

cDNA Synthesis

cDNA was synthesized from poly-A⁺ RNA by standard procedures using avian myeloblastosis virus reverse transcriptase. In some synthesis reactions an RNase inhibitor RNasin (Promega Biotec, Madison, WI) or a vanadyl-ribonucleoside complex was included in the cDNA reaction mix.

Enrichment of Aphidicolin-Resistant Cell Specific cDNA Sequences

A cascade enrichment scheme similar to one described by Timberlake (1980) was used to selectively enrich for mRNA sequences found at higher copy number in aphidicolin-resistant cells than in parental-type CHO cells. Figure 3 outlines the procedure used. In short, poly-A⁺ RNA from high level aphidicolin resistant cells was used to make ³²P-labeled cDNA. This cDNA was then hybridized to a 5-fold mass excess of CHO cell mRNA in 0.42 M phosphate buffer with 0.2% SDS. At

va

va

is

tr

c

Δ

tr

va

various times corresponding to C_{0t} (nucleotide moles x seconds/liter) values of 0, 150, and 500 the non-hybridized cDNA sequences were isolated by chromatography on hydroxylapatite at 68°. These cDNAs were then used to probe either Southern blots of restriction endonuclease digestions of CHO and L3-5 genomic DNA or an L3-5 lamda library.

λ Library Construction and Screening

Genomic libraries were constructed in the λ vector Charon 28 by the method of Maniatis et al. (1982). These libraries were screened by the method of Benton and Davis (1977).

DNA Mediated Gene Transfer

The transformation protocol was as described by Wigler et al. (1979). Briefly, 24 hours prior to transformation cells (CHO or L-tk-aprt⁻) were plated at 5×10^6 cells per 100 mm dish. The medium was changed approximately 4 hours prior to transformation. Sterile, high molecular weight eukaryotic DNA was dissolved in 1 mM Tris (pH 8), 0.1 mM EDTA. The DNA solution was adjusted to 250 mM CaCl₂ with sterile 2 M CaCl₂. The DNA concentration in this DNA/CaCl₂ mix was 40 µg/ml. The DNA/CaCl₂ mix was added dropwise to an equal volume of 2X Hepes-buffered saline (280 mM NaCl, 50 mM Hepes, 1.5 mM Na₂HPO₄, pH 7.10 ± 0.05). While adding the DNA/CaCl₂, air was bubbled through the 2X HBS to gently mix the solutions. After addition of the DNA/CaCl₂ mixture the DNA-calcium phosphate precipitate was allowed to form for 40 minutes without agitation. The precipitate was then mixed by gentle pipetting and 1 ml of precipitate (20 µg DNA) was added to the 10 ml of medium on each

plate of

medium OV

later, th

kinase se

μg/ml thy

selected

resistan

medium w

in 2 to

plate of recipient cells. After 12 hours of exposure to DNA, the medium over the cells was replaced with fresh medium. Thirty six hours later, the medium was replaced with selective medium. For thymidine kinase selection, HAT (15 $\mu\text{g/ml}$ hypoxanthine, 0.2 $\mu\text{g/ml}$ aminopterin, 5 $\mu\text{g/ml}$ thymidine) medium was used. Aphidicolin resistant cells were selected with 0.4 $\mu\text{g/ml}$ aphidicolin added to the normal medium. G418 resistant cells were selected in 500 $\mu\text{g/ml}$ of that drug. The selective medium was changed every 2 to 3 days. Resistant colonies were visible in 2 to 3 weeks.

introd

S

the ha

re-as

con-

resis

gene

were

form

proce

con

ex

se

ie

a

h

s

RESULTS

Introduction

Several approaches have been utilized in our efforts to isolate the hamster RdRase genes. Initially, the putative amplification of the RdRase genes was used as a selection for the gene. More specifically, cDNA enrichment schemes and gene transfer experiments utilizing the APH resistant cell line L3-5 (and later L3-107) were used. More generalized techniques which did not rely on RdRase gene amplification were attempted subsequently. These included gene transfer of a dominant RdRase marker and screening of genomic blots with heterologous probes.

cDNA Enrichment

Schemes which utilize hybridization of cDNA and mRNA to various extents have been previously used to enrich a cDNA population for sequences expressed abundantly in one cell type and expressed to a lesser extent in a closely related one (Timberlake, 1980; Mather et al., 1981). As shown in Figure 3, the process used here involved hybridization of APH resistant cell (L3-5) ³²P-labelled cDNA to a 5-fold excess of parental type (CHO) mRNA. Hybridizations were carried out under conditions such that any sequences common to both L3-5 and CHO cells should have hybridized. Sequences unique to, or enriched in, L3-5 cells should have remained unhybridized. These unhybridized

Figure 3. mRNA:cDNA hybridization scheme. This procedure was used to selectively enrich for APH-resistant cell (L3-5) cDNA sequences. Details of the procedure are given in the text.

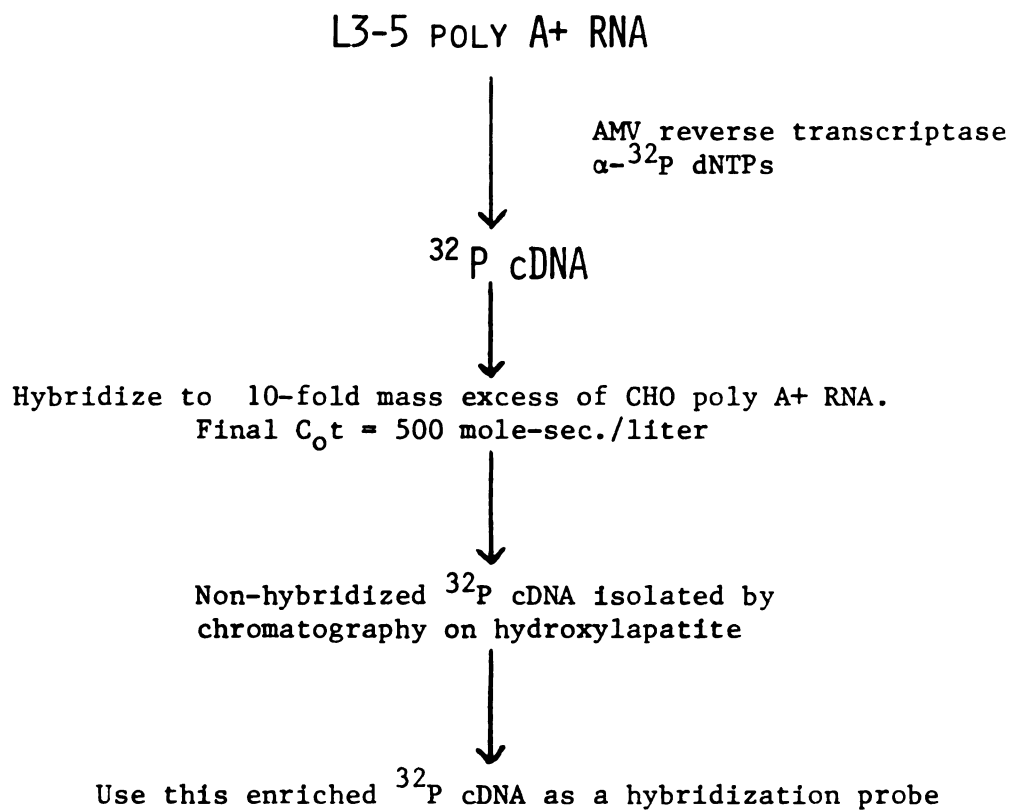


Figure 3

cDNA's were isolated by hydroxylapetite chromatography and were used in the hybridization experiments detailed below.

Differential Screening of Southern Blots

Genomic DNAs from L3-5 and CHO cells were digested with various restriction enzymes, electrophoresed on agarose gels, and blotted to nitrocellulose filters. These filters were then hybridized to the L3-5-enriched cDNA probes described in section 1a. Figure 4 shows the autoradiogram obtained using L3-5 and CHO cDNA's isolated at the indicated C_{ot} values. Any mRNA sequences which were enriched in the L3-5 cDNA should appear as bands not seen with the control CHO cDNA. Also, if gene amplification has taken place, these bands should be more intense in the L3-5 DNA lane compared to the CHO lane. As can be seen, there is no discernable difference in the banding pattern at the C_{ot} values tested. The bands which are seen in the cDNA's hybridized to moderate (175) and high (500) C_{ot} values are probably highly abundant messages which haven't been completely hybridized with the CHO driver RNA. In no case is any difference seen between the L3-5 and CHO DNA lanes.

Differential Screening of Lamda Libraries

Since the negative results obtained when screening Southern blots with the enriched cDNA probes might be due to the low sensitivity of this method, we attempted to use these cDNA probes for screening lamda libraries. Bacteriophage libraries were constructed for use in differential screening experiments with the enriched cDNA probes. It was hoped the added sensitivity of screening phage libraries would

Figure 4. Enriched cDNA:Southern Blots. Results of hybridization for enriched cDNA's to Southern blots of genomic digests. ^{32}p labeled cDNA probes (CHO or L3-5) were enriched by hybridization with CHO mRNA to a C_0t of 500 (top) or 175 (bottom) and used to probe Eco RI restriction endonuclease digests of CHO or L3-5 genomic DNA's. Each lane is labeled with the type of genomic DNA used.

29
Cot 500



Cot 175



Figure 4

allow identification of the putative amplified cDNAs in the enriched L3-5 probe. The factors contributing to added sensitivity of phage screening, compared to Southern blotting, are a) increased hybridization due to higher amounts of filter bound DNA acting as driver in the hybridization reaction, and b) more discernable isolated signals on a phage screen versus the general smear obtained when screening Southern blots with cDNA. Using a differential screening procedure similar to this, except that un-enriched probes were used, St. John and Davis were able to clone genes for messages which represented at most 0.01% of the mRNA population (1979).

Complete libraries from both CHO and L3-5 genomic DNA were constructed in the bacteriophage vector λ Charon 28. Pertinent data about these libraries is given in Table 3.

Phage from the amplified L3-5 library were plated at the density of 5×10^3 plaques per 100 mm plate. Filters replicated from these plates were hybridized with the enriched cDNA probes (Cot 500) described above. A total of 6.5×10^4 phage were screened. Many plaques which appeared to hybridize preferentially to the L3-5 cDNA probe were noted, but upon re-screening none of these were found to contain sequences enriched in L3-5 cDNA (data not shown).

Gene Transfer Experiments

Because we had been unable to obtain any positive results with the enriched cDNA probes, we decided to explore other methods for isolating the RdRase gene(s). Several genes have recently been isolated by DNA mediated transformation of eukaryotic cells. The major requirements for isolation of a gene using this procedure are a) a strong, positive

Table 3. λ Library Construction Data

Library	μ g Eukaryotic DNA Ligated	Total Number of Plaques Forming Units	Efficiency (Plaques per μ g of Eukaryotic DNA)	Mean Length of Eukaryotic Insert	Number of phage Complete Library*	Number of Phage Amplified	Total number of Amplified pfu Recovered	Fold Amplification
CHO	1.6	3.0×10^6	1.9×10^6	14.0 kb	9.9×10^5	1.36×10^6	4.2×10^{12}	3.1×10^6
L3-5	1.6	2.7×10^6	1.7×10^6	13.8 kb	1.0×10^6	1.36×10^6	5.7×10^{11}	4.2×10^5

*Complete library defined as having 99% probability of finding a particular sequence in the library. Calculated by the equation
 No. phage needed = $\frac{\ln(1-0.99)}{\ln(1-\frac{\text{insert size}}{\text{genome size}})}$

Assumes a genome size of 3×10^9 bp for CHO.

selection system for eukaryotic cells that have taken up and are expressing the gene of interest; and b) a method of distinguishing and isolating the transferred gene from the recipient genome (i.e. a "rescue" technique).

The experiments described here were conducted using two different selectable markers for the RdRase gene. In the first set of experiments, cells were selected for resistance to APH after treatment with L3-5 DNA. Since resistance to APH is correlated with increased levels of RdRase activity, this experiment relies on transfer of a sufficient number of copies of the RdRase gene to achieve APH resistance. The second set of gene transfer experiments utilized a dominant selectable marker from a HU resistant cell line obtained from James Wright (University of Manitoba). The RdRase of these cells had been shown to contain a mutated RdRase enzyme that was resistant to HU (Wright and Lewis, 1978). Presumably this mutation resides in the M2 subunit since it affects the tyrosine free radical's sensitivity to HU. Thus the HU resistant phenotype serves as a marker for transfer of the gene for the M2 subunit.

The second aspect of the research with gene transfer involved the design of a system for rescuing the transferred genes. We constructed a series of cosmid vectors, some of which contain markers which are selectable in eukaryotic cells. Our paper describing these vectors is included as Appendix 2 and will not be discussed in any further detail here. A scheme for utilizing these vectors for isolating genes after DNA mediated gene transfer is shown in Figure 5. The validity of this approach is demonstrated by the work of R. Woychik who has recently

Figure 5. Cosmid rescue scheme. This rescue scheme was developed for use with HUR^r-2 transformed eukaryotic cells, but can be adapted for use with any selectable or screenable marker. The rescue of the HUR^r gene from the recipient genome depends upon linkage of this gene with another selectable marker, namely the ampicillin resistance marker (AMP). DNA from the HUR^r transformed cells is cloned into a cosmid vector which does not contain AMP. Only those sequences that are derived from the transferred DNA and contain AMP will be present in the cosmids selected with AMP. Screening these few cosmids by a second round of DNA mediated gene transfer identifies those that contain the HUR^r gene. Also included in the example shown here is the thymidine kinase marker for selection in eukaryotic cells. Inclusion of this marker ensures that any HUR^r cells also contain the co-transformed tk, AMP DNA; i.e. this ensures that the HUR^r gene is tagged with the AMP marker.

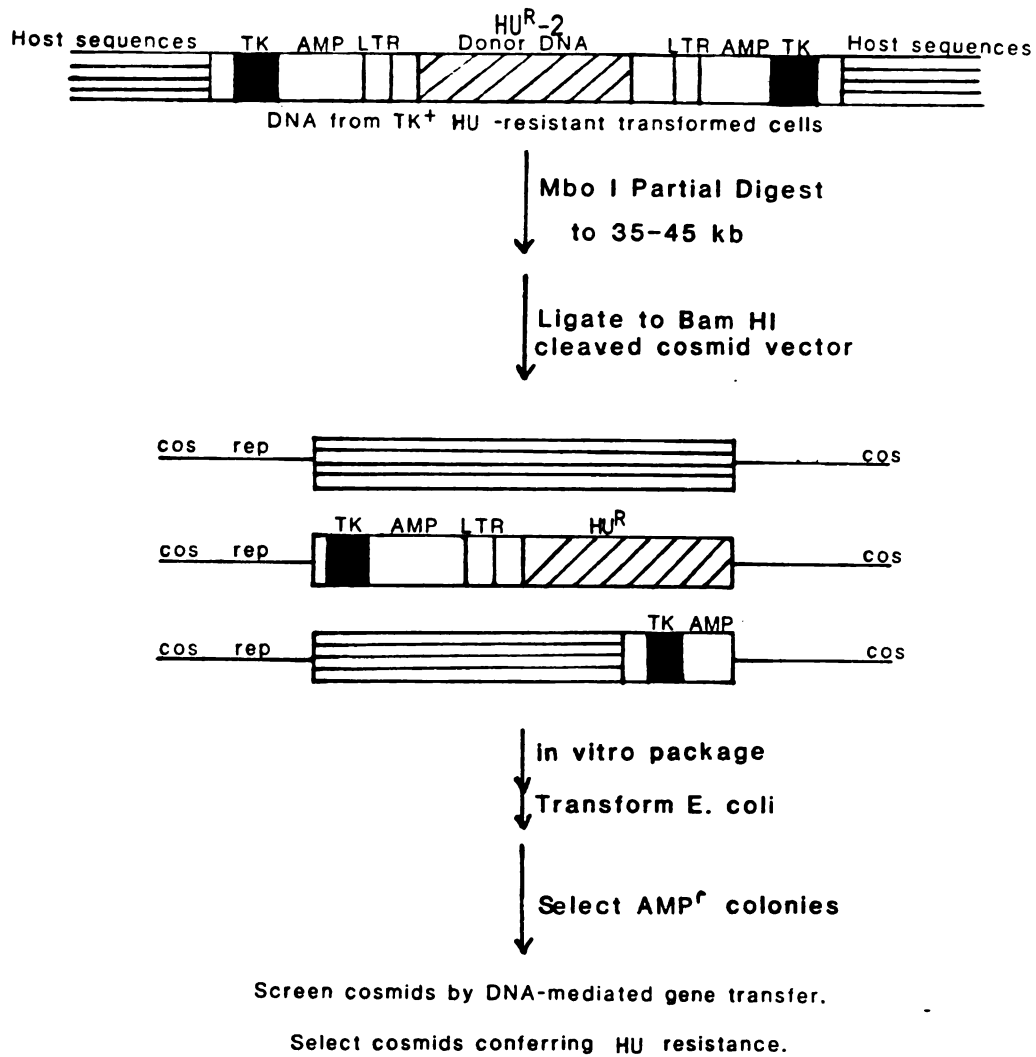


Figure 5

isolated a mutant, single-copy mouse gene affecting limb development (personal communication).

Gene Transfer: Selections for APH Resistance

We had previously shown that above a concentration of 0.3 $\mu\text{g/ml}$ APH the plating efficiency of CHO cells is less than 10^{-7} . Also our results indicated that cells with a four to eight fold increase in RdRase activity, as measured in cell extracts, were resistant to 5 $\mu\text{g/ml}$ APH (Sabourin et al., 1981). It therefore seemed feasible that transfer of a sufficient number of copies of the RdRase gene could render CHO cells resistant to a level of APH that could be used for selection. CHO cells were transformed with L3-5 DNA and cells resistant to APH were selected with the drug at 1.0 $\mu\text{g/ml}$. After 3 weeks no resistant colonies were seen on 10 plates. The same experiment was carried out a second time with identical results (Table 4).

It is known that various eukaryotic cell lines are transformed by the DNA/ CaPO_4 method with widely differing efficiencies (Graf et al., 1979). The negative results could be caused by the low transformation frequency of CHO cells. Since mouse L-cells are very efficient at incorporating exogeneous DNA whereas CHO cells are thought to be very inefficient we decided to perform the DNA transfer experiments with L-cells. In order to do this, it was necessary to characterize the growth of L cells in APH.

As shown in Figure 6, mouse L tk⁻ aprt⁻ cells are slightly more resistant to APH than the hamster CHO cell line. At 1.0 $\mu\text{g/ml}$ APH no colonies were observed when 1.0×10^7 L cells were plated. Lower

Table 4. DNA Mediated Transformation: APH Selection Results

Recipient Cell Line	Donor DNA	Total No. Colonies per No. plates transfected		Frequency	
		APH ^r	tk ⁺	APH ^r	tk ⁺
CHO	CHO	0/20	N.D.	0	ND
CHO	L3-5	0/20	N.D.	0	ND
L tk ⁻ aprt ⁻	CHO	0/20	52/5	0	2.6×10^{-6}
L tk ⁻ aprt ⁻	L3-5	0/20	48/5	0	2.4×10^{-6}
L tk ⁻ aprt ⁻	L3-10	0/20	61/5	0	3.1×10^{-6}

a) Selections were done as described in Materials and Methods with APH at 1.0 $\mu\text{g/ml}$.

b) Frequency = number of colonies/total number of cells transformed.

levels of APH (below 0.6 $\mu\text{g/ml}$) appeared to arrest or kill the L cells; however many dead cells and cell debris remained attached to the plate. Therefore 1.0 $\mu\text{g/ml}$ APH was chosen as the concentration to be used in the transformation experiments.

DNA mediated transformation experiments were carried out with L tk⁻ aprt⁻ cells as recipients using either L3-5 or L3-107 DNA. L3-107 is a cell line which was isolated from L3-5 and is resistant to 10 $\mu\text{g/ml}$ APH. No APH resistant colonies were obtained when either of these DNAs was transfected into L cells and selection was carried out with 1.0 $\mu\text{g/ml}$ APH (Table 4).

As a control, the ability to transfer the hamster thymidine kinase gene into L tk⁻ aprt⁻ cells using these same DNA's was assayed. An average of 10 tk⁺ colonies per plate was obtained in experiments run in parallel with the APH selections (Table 4), demonstrating that the transformation procedure was working.

Gene Transfer: Selection for Hydroxyurea Resistance

Since a dominant, selectable marker for the RdRase gene was available, we attempted to use this marker for gene transfer instead of the APH-resistance marker. Experiments were carried out using the hydroxyurea-resistance marker for gene transfer. High molecular weight genomic DNA (> 100 kb) was prepared from HU^r-2 cells.

Before gene transfer experiments were conducted, it was necessary to characterize the growth of the recipient cells in HU. As shown in Table 5, various cell lines tested had widely different sensitivities to HU. It was surprising that Ltk⁻aprt⁻ and their parent cell line LM had such different sensitivities to HU. In general it appeared that

Figure 6. Comparison of mouse Ltk^{-aprt-} and CHO cell plating efficiencies in aphidicolin. The cells were plated as described in the text. Relative plating efficiency was determined as described by Sabourin et al. (1981).

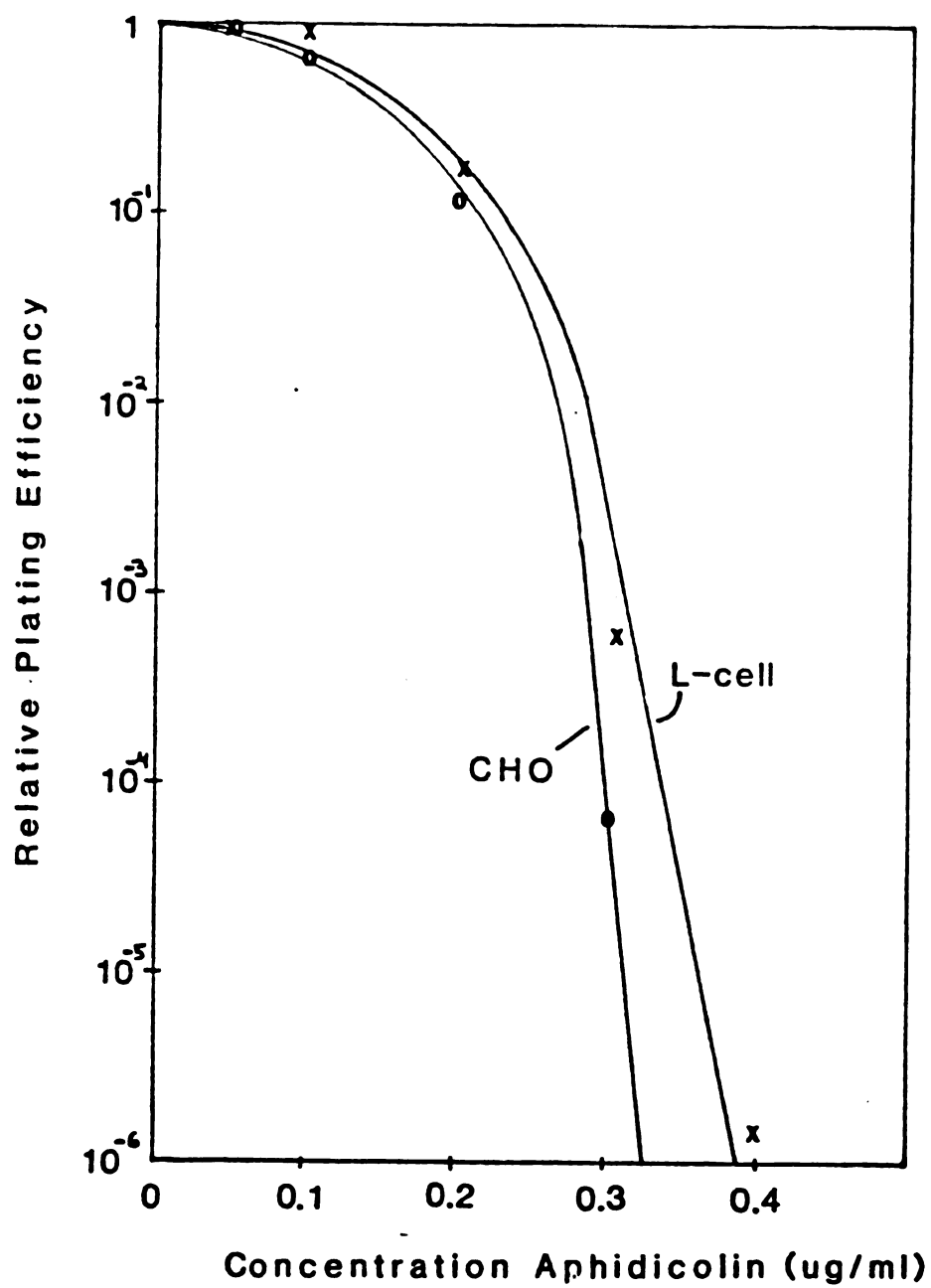


Figure 6

the two hamster lines, CHO and V79 tk⁻, were the most sensitive to HU. The HUR-2 cells appeared to grow well in HU concentrations up to 75 µg/ml.

DNA mediated transformations were carried out using both hamster and mouse cell lines as well as a rat 3 tk⁻ cell line as recipients. The donor DNA in these experiments was high molecular weight (>100kb) HUR-2 genomic DNA. All selections were done at 50 µg/ml HU, except the Ltk⁻aprt⁻ transformants which were selected with 100 µg/ml HU. The data in Table 6 shows that no HUR colonies were seen on any of the transfected plates. Each of the tk⁻ cell lines was tested for the ability to transfer the HUR-2 tk⁺ marker. As noted previously, the tk⁺ phenotype was transferred with high efficiency (Table 6). No tk⁺ colonies were seen when mouse L tk⁻ cell DNA was used as the donor DNA (data not shown).

Isolation of Higher Level HUR-2 Mutants

In order to increase the possibility of being able to transfer the mutant RdRase gene in subsequent experiments, a step-wise selection scheme was employed in an attempt to amplify this RdRase gene in the HUR-2 cells. As stated previously such step-wise selections have almost always yielded cells with amplified genes. The HUR-2 cells were selected in steps using 100, 150, 200, 250, and 300 µg/ml of HU. The cells were not cloned at each selection level, and were kept in the presence of HU continuously throughout the selection scheme. At each step, approximately 0.1-0.3% of the cells remained viable (approximate frequency of 10⁻⁴). After each selection step, cells were allowed to adapt to the higher level of drug for 2-6 weeks of logarithmic

Table

Cell 1

Cell 2

Cell 3

Cell 4

Cell 5

Cell 6

Cell 7

Cell 8

Cell 9

Cell 10

Cell 11

Cell 12

Cell 13

Cell 14

Cell 15

Cell 16

Cell 17

Cell 18

Cell 19

Cell 20

Cell 21

Cell 22

Cell 23

Cell 24

Cell 25

Cell 26

Cell 27

Table 5. Hydroxyurea Sensitivity of Various Cell Lines

<u>Cell line</u>	(HU) $\mu\text{g/ml}$						
	<u>0</u>	<u>30</u>	<u>50</u>	<u>75</u>	<u>100</u>	<u>150</u>	<u>300</u>
CHO pro	+++	-	-	-	-	-	ND
V79 tk	+++	-	-	-	-	-	ND
LM	+++	+/-	-	-	-	-	ND
Ltk aprt	+++	+++	++	+/-	-	-	ND
Rat 3 tk	+++	++	-	ND	ND	ND	ND
cos 7	+++	++	+/-	ND	ND	ND	ND
293	+++	++	+/-	ND	ND	ND	ND
Hu 2	+++	+++	+++	+++	+	-	ND
Hu 2 (300)	+++	+++	+++	+++	+++	+++	+++

Growth of cell lines was determined as described in Materials and Methods.

+++ Vigorous cell growth

++ slowed growth

+ some limited replication

+/- most cells dead, but much cell debris and enlarged cells attached to plate

- plates clean with little or no debris left

N.D. Not determined

Table 6. DNA mediated gene transfer: Hydroxyurea selections

Recipient Cells	Donor DNA	No. plates transfected	Total No. HU ^r colonies	Average No. tk ⁺ colonies per plate
CHO (hamster)	HU ^r 2	25	0	N.D.
V79 tk ⁻ (hamster)	HU ^r 2	50	0	N.D.
V79 tk ⁻	HU ^r 2	5	N.D.	3
rat 3 tk ⁻	HU ^r 2	25	0	N.D.
rat 3 tk ⁻	HU ^r 2	5	N.D.	8
LM (mouse)	HU ^r 2	50	0	N.D.
Ltk ⁻ aprt ⁻ (mouse)	HU ^r 2	50	0	N.D.
Ltk ⁻ aprt ⁻	HU ^r 2	5	N.D.	52

growth before proceeding to the next step. The highest level mutants, HUR-2 (300) appeared normal by microscopic examination and had a slightly longer doubling time than the HUR-2 cells (data not shown).

Gene Transfer Experiments with HUR-2 (300) DNA

High molecular weight DNA isolated from the HUR-2 (300) cells was used to transform the three cell lines which were most sensitive to HU. One hundred plates of LM cells, ninety-five plates of V79tk⁻ and fifty plates of CHO cells were transfected with 20 µg per plate of HUR-2 (300) DNA. Unlike the previous experiments, in this set of transformations after absorption of the DNA for 16 hours, a 30 minute shock with dimethylsulfoxide (DMSO) was used to facilitate uptake of the donor DNA. Experiments with the hamster V79tk⁻ and mouse Ltk⁻aprt⁻ cells selected for the tk⁺ phenotype showed this DMSO shock caused approximately a two fold increase in the V79tk⁻ transformation efficiency, while the efficiency in mouse cells seemed unaffected (data not shown).

As shown in Table 7 no colonies were obtained after selection with 35 µg/ml HU for 3 weeks. It should be noted that this level of HU is lower than that used in previous experiments but still yielded no HU resistant colonies. All the plates were checked for colonies first by microscopic examination then by staining with methylene blue. In all cases the HU selection appeared to be very stringent and no cell debris remained attached to the plate.

The V79tk⁻ cells were also selected for transformation to the tk⁺ phenotype. An average efficiency of 3 colonies per 20 µg DNA per

Table

Recip

V79

V79

C+O

LM

Table 7. DNA mediated transformations with HUR^r-2 (300) DNA.

Recipient Cells	Donor DNA	No. plates transfected	Total No. HUR ^r clones	Average No. tk ⁺ colonies per plate
V79 tk ⁻	HUR ^r 2 (300)	5	N.D.	3
V79 tk ⁻	HUR ^r 2 (300)	95	0	N.D.
CHO	HUR ^r 2 (300)	50	0	N.D.
LM	HUR ^r 2 (300)	100	0	N.D.

1×10^6 cells was observed. This corresponds to a frequency of approximately 3×10^{-6} .

Serial Selection with G418 then HU

Since the direct selection of the HU resistant phenotype had proven unsuccessful, it was decided to try a more indirect approach to transfer the RdRase gene. Mouse L tk⁻aprt⁻ cells were used in this procedure since our results with tk phenotype transfer indicated this cell line was 20 fold more efficient at DNA mediated gene transfer than the hamster cell lines (6×10^{-5} vs 3×10^{-6}). L cells were transfected with 20 μ g of HUR-2 (200) DNA plus 5 μ g of linearized pSV2-neo plasmid DNA. This plasmid contains the early promoter region and splicing and polyadenylation signals of SV-40 fused to the bacterial neomycin resistance gene. Incorporation of pSV2-neo into eukaryotic cells confers resistance to the drug G418 (Southern and Berg, 1982).

Transfected L cells were first selected for resistance to G418. Since the pSV2-neo DNA was added in a vast molar excess compared to the genomic DNA fragments, most cells which had taken up and expressed exogenous DNA survived the selection. Approximately 1.6×10^3 colonies were obtained per plate after G418 selection. After 3 weeks these colonies which were expressing the donor DNA were then trypsinized and replated in 80 μ g/ml HU at a density of 5×10^5 cells per 100 mm plate (since each colony had on the average about 250 cells, the trypsinized cells were split 1:2 when plating in HU). After an additional 3 weeks of selection in HU, no colonies were observed.

Herpes Simplex Virus RdRase Probes

While the gene transfer experiments were in progress, information about the possible location of the HSV-1 and HSV-2 RdRase genes became available (S. Weller, personal communication). Since the architecture of the RdRase subunits had been so well conserved throughout evolution, it was decided to attempt to use these heterologous RdRase genes as probes for the hamster gene(s).

The strategy employed in these experiments was similar to that used in the cDNA enrichment Southern blots in that DNA from CHO cells and the putative amplified cells lines (L3-5, Hu^r-2 (200)) was digested with restriction enzymes, run in parallel on agarose gels, and Southern blotted to nitrocellulose filters. RdRase gene fragments which are within the amplified region of DNA should hybridize to the HSV probe more intensely in the L3-5 and HU^r-2 (200) lanes (assuming of course that amplification has taken place). These same fragments should be present in the CHO lanes but at a decreased signal strength.

Since the degree of homology between the HSV and hamster RdRase genes was unknown, the hybridizations were carried out at several different stringencies to maximize the signal while keeping the background low. The stringencies were adjusted by varying the percentage of formamide in the hybridization solution (20 to 50 percent), temperature of hybridization (37 to 42 C°), the salt concentration of the washing solution (30 to 300 mM NaCl), and the temperature of the washes (room temperature to 50C°). Generally the filters were hybridized at various stringencies then all washed at low stringency (300 mM NaCl, room temperature). These blots were then

exposed and subsequently rewashed at higher stringency as necessary and re-exposed.

The probes used for these experiments are shown in Figure 7. Also shown is the region of HSV which was found to hybridize strongly to repetitive sequences in the hamster genome (data not shown). The Bam T probe of HSV-2 contained nearly all the sequences for the 38kd (M2) polypeptide and only the carboxyterminal end of the 140kd coding region. The Bam O probe for HSV-1 contains nearly half the 140kd coding region and all of the 40kd coding region.

Bam O and Bam T nick-translated probes were used in hybridizations with Southern blots of CHO, L3-5, and HUR-2 (200) DNAs. No bands were seen above background with either of the probes at any criteria (data not shown). Control experiments with the hamster aprt gene probe yielded the expected banding pattern with bands of approximate equal intensity in all lanes (data not shown).

Figure 7. Herpes Simplex Virus 1 and 2 ribonucleotide reductase Maps. The regions thought to contain the HSV 1 and 2 RdRase genes are shown. Both viruses code for two polypeptides in this region. The coding regions are shown as dashed boxes, and the mRNA transcripts as wavy lines. The numbers refer to the map coordinates of the respective viruses. The probes used are drawn below each map. The restriction enzyme sites are as follows: B=BamHI, Bg=BglII, P=PvuII.

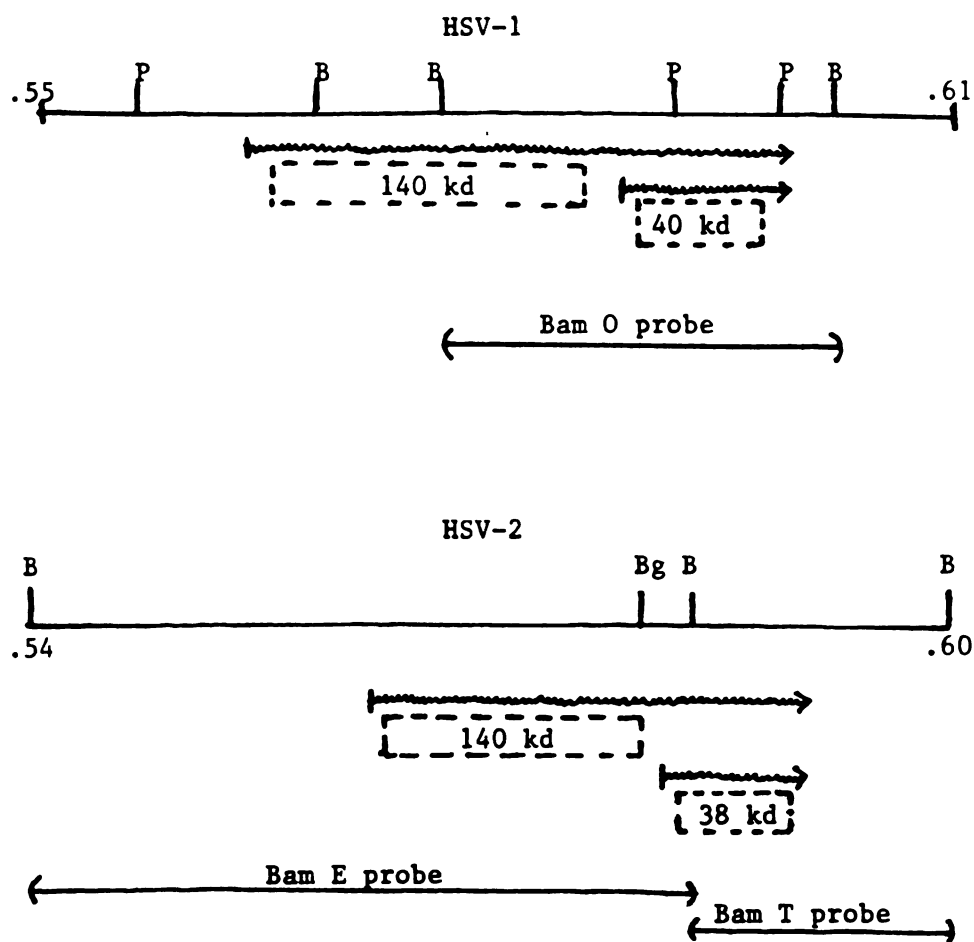


Figure 7.

DISCUSSION

The major focus of this project was the isolation of the hamster RdRase gene, or more specifically the gene for the M2 subunit of RdRase. The experiments detailed here failed to attain that goal. However several conclusions about the hamster gene can be surmized from this work.

First, the frequency of DNA mediated gene transfer with the hamster RdRase gene is very low. In our experiments nearly 500 plates of cells were transformed with DNA containing the dominant, selectable marker for the M2 subunit from HUR-2. No HU resistant colonies were found on any of these plates. Thus, the frequency of transformation with this gene must be less than 2×10^{-9} . Using the same recipient cells and donor DNA's, the tk^+ phenotype was transferred with a frequency of 5×10^{-5} or 3×10^{-6} in mouse and hamster cells respectively.

Recently Lewis and Srinivasan (1983) were able to transfer the HU resistance marker using chromosome mediated gene transfer at a frequency of 10^{-6} using the V79 tk^- cells as recipients and HUR-2 cell chromosomes. This means that the frequency with DNA mediated versus chromosome mediated transfer is at least 500 fold lower. Lewis and Srinivasan (1983) also attempted DNA mediated transfer of the HU-resistant phenotype with no success (W. Lewis, personal communication). One positive aspect of the DNA mediated transfer experiments which should be noted is the high efficiencies

which we attained in the tk transformations. Our best transformation efficiencies of 5×10^{-5} (mouse cells) and 3×10^{-6} (hamster) are 25-fold and 30-fold higher respectively than previously reported values for these cells (Abraham et al., 1982; Wigler et al., 1979).

It is intriguing to consider why we have been unable to transfer the RdRase gene using the DNA/CaPO₄ method. One possibility is that the hamster RdRase gene is too large and does not remain intact during the transfer process. Very large DNA (≥ 100 kb) was used to form the CaPO₄/DNA precipitates. If one assumes that mechanical (shearing) and biological (recombination and nuclease digestion) factors act on the DNA during the precipitation and cellular uptake, then the average size of the intact transferred sequences is probably on the order of 50kb. It is not inconceivable that the RdRase gene is larger than 50kb and cannot be efficiently transferred with the DNA/CaPO₄ method. The high efficiency seen with chromosome mediated transfer may be explained by the length of contiguous sequence transferred using chromosomes, which is generally on the order of 7000 kb (McBride and Peterson, 1980).

It is interesting to note how our perceptions of eukaryotic gene structure have changed recently. A few years ago it was generally assumed, based mainly on the globins and other abundantly expressed genes, that the length of most eukaryotic genes would be on the order of 20kb or less. Several lines of evidence have recently shown this is not the case. In somatic tissues, or before recombinational joining has taken place, the immunoglobulin genes are spread over very long distances (probably hundreds of kb, Marcu and Cooper, 1982). In Drosophila, the bithorax complex has been mapped to a region over 100

kb in length. Analysis of cDNA clones for this region suggests certain messages are derived from sequences nearly 100 kb apart (W. Bender, personal communication). Finally the human thyroglobulin gene has been cloned and was shown to extend over 200 kb. One intron in this gene is nearly 60 kb in length (G-J VanOmmen personal communication). These examples clearly demonstrate that eukaryotic genes can be much larger than originally anticipated, thus it is possible that the RdRase gene is over 50 kb in length.

A second possible reason for our inability to transfer the RdRase gene by DNA mediated transfer is that the RdRase gene remains intact during the transfer but the gene cannot be expressed. This could occur if some stable chromatin structure or transcription complex were needed to activate expression of the RdRase gene. This structure or complex would be present in the chromosome mediated transfer procedure, but would be stripped off when making DNA for CaPO_4 transfer. A similar explanation has been suggested for the unregulated expression of human α globin genes introduced into mouse erythroleukemic cells. In this case, however, it was theorized that a structure or complex was required to repress, rather than activate, α globin expression in undifferentiated MEL cells; thus the α globin genes introduced as naked DNA were always expressed (Charnay et al., 1984). In a more clear-cut example of the effect of structural features on expression, *Xenopus* 5S oocyte genes are inactive unless specifically assembled in an active, stable transcription complex. When assembled in a regular nucleosome structure in somatic cells, these genes are inaccessible to polymerase (for a review of stable transcription complexes see Brown, 1984).

Although the majority of the work described here involved attempts to transfer the RdRase gene by DNA mediated gene transfer, two other methods were used in our efforts to clone this gene. The HSV RdRase gene was unable to specifically identify any fragments in the hamster genome when used as a heterologous probe. After these experiments were completed, the nucleotide and deduced amino acid, sequences of the E. coli, HSV-2, and Epstein-Barr virus (EBV) RdRase M2 (or B2) genes were published (Gibson et al., 1984; McLauchlan and Clements, 1983; Carlson et al., 1984). Comparison of these sequences shows that although there is a high degree of conservation of the amino acid sequence among all these genes, they diverge strongly at the nucleotide level. In fact there is only a slightly discernible nucleotide homology between two of the herpes virus M2 genes (EBV versus HSV-2) (Gibson et al., 1984). In light of these results, it is not surprising that the HSV RdRase probes failed to detect the hamster gene.

A cDNA enrichment procedure was also tried as a method to isolate the RdRase gene. Using both Southern blotting and lambda library screening we were unable to identify the RdRase gene with enriched cDNA probes. There are several probable reasons for this result. First, the protocol used to synthesize the cDNA did not employ actinomycin D to inhibit hairpin loop formation in the cDNA; thus unhybridized cDNAs, which should not have bound to the hydroxylapatite, in fact may have contained small double strand loops and might have been removed in this chromatography step (M. Davis, personal communication). Second, the cells used to prepare the mRNA's were not synchronized; consequently most to them were not in S phase when harvested. Since RdRase is strongly cell cycle regulated, most of the L3-5 cells used to prepare

mRNA would probably have little or no RdRase messages. This explanation assumes that the RdRase activity reflects the amount of message in the cell. Studies with regulation of DHFR and human tk suggests that cell cycle control of these activities occurs at the transcription (or mRNA stabilization) level (Kaufman and Sharp, 1983; S. Conrad, personal communication). Finally, Srinivasan and Lewis have also been unable to clone the RdRase cDNA using enriched cDNA probes and enriched cDNA libraries from cells which overproduce RdRase (personal communication). To date no laboratory has been able to clone the eukaryotic RdRase M2 gene although several have actively pursued the project for several years.

REFERENCES

- Abraham, I., Tyagis, J.S., and Gottsman, M. (1982) Somatic Cell Genetics 8, 23-39.
- Akerblom, L., Ehrenberg, A., Graslund, A., Lankinen, H., Reichard P., and Thelander. O. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2159-2163.
- Alt, F.W., Kellems, R.E., Bertino, J.R., and Schimke, R.T. (1978) J. Biol. Chem. 253, 1357-1370.
- Baer, R., Bankier, A.T., Biggin, M.D., Deninger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Seguin, C., Tuffnell, P.S., and Barrell, B.G. (1984) Nature 310, 207-211.
- Benton, D. and Davis, R.W. (1977) Science 196, 180-182.
- Berglund, O. (1972) J. Biol. Chem. 247, 7276-81.
- Biedler, J.L., Meyers, M.B., Spengler, B.A. (1983) Adv. Cell. Neurobiol. 4, 267-307.
- Bishop, J.M. (1983) Ann. Rev. Biochem. 52, 301-354.
- Blin, N. and Stafford, D.W. (1976) Nucl. Acids Res. 3, 2303-2308.
- Brown, D.D. (1984) Cell 37, 359-365.
- Brown, N.C., Reichard, P. (1969) J. Mol. Biol. 46, 39-55.
- Caras, I.W., and Martin, D.W., Jr. (1982) J. Biol. Chem. 257, 9508-9512.
- Caras, I.W., Jones, T., Eriksson, S., Martin, D.W. Jr. (1983) J. Biol. Chem. 258, 3064-3068.
- Carlson, J., Fuchs, J.A., and Messing, J. (1984) Proc. Natl. Acad. Sci. USA 81, 4294-4297.
- Charnay, P., Triesman, R., Mellon, P., Chao, M., Axel, R. and Maniatis, T. (1984) Cell 38, 251-263.
- Cory, J.G. and Mansell M.M. (1975) Cancer Res. 35, 390-396.
- Cory, J.G., Mansell, M.M. Whitford, T.W. Jr (1976) Adv. Enzyme Regul. 14, 45-62.
- Cowell, J.K. (1982) Ann. Rev. Genet. 16, 21-59.

- Draper, K.G., Frink, R.J. and Wagner, E.K. (1982) J. Virol. 43, 1123-1128.
- Elford, H.L. (1972) Adv. Enz. Reg. 10, 19-38.
- Elford, H.L., Freese, M., Passamani, E., Morris, H.P. (1970) J. Biol. Chem. 245, 5228-33.
- Eriksson, S., Gudas, L.J., Ulman, B., Clift, S.M., Martin, D.W. Jr. (1981) J. Biol. Chem. 256, 10184-10188.
- Eriksson, S., and Martin, D.W., Jr. (1981) J. Biol. Chem. 256, 9436-9440.
- Eriksson, S., Sjoberg, B.-M., Hahne, S., Karlstrom, O. (1977) J. Biol. Chem. 252, 6132-38.
- Fuchs, J.A. (1977) J. Bacteriol. 130, 957-59.
- Gibson, T., Stockwell, P., Ginsberg, M., and Barrell B. (1984) Nucleic Acids Research 12, 5087-5099.
- Girvitz, S.C, Bacchetti, S., Rainbow, A.J., and Graham, F.L. (1980) Anal. Bioch. 106, 492-496.
- Graf, L.H., Urlaub, G., and Chasin, L.A. (1979) Somat. Cell. Genet. 5, 1031-1044.
- Graham, F.L., and van der Eb, A.J. (1973) Virology 52, 456-467.
- Graslund, A., Ehrenberg, A., and Thelander, L. (1982) J. Biol. Chem. 257, 5711-5715.
- Hanke, P.D., and Fuchs, J.A. (1983) J. Bacteriol. 154, 1140-1145.
- Hendrick, S.T., Cohen, D.I., Nielsen, E.A., and Davis, M.M. Nature (1984) 308, 149-153.
- Ikegami, S., Taguchi, T., Ohashi, M., Oguro, M. Nagano, H., and Mano, Y. (1978) Nature 275, 458-460.
- Johnston, R.N., Beverley, S.M. Schimke, R.T. (1983) Proc. Natl. Acad. Sci. USA 80, 3711-15.
- Kaufman, R.J., and Sharp, P.A. (1983) Mol. Cell. Bio. 3, 1598-1608.
- Kucera, R., Brown, C.L., Paulus, H. (1983) 117, 158-168.
- Lankinen, H., Graslund, A., and Thelander, L. (1982) J. Virol. 41, 893-900.
- Larsson, A. (1969) Eur. J. Biochem. 11, 113-21.
- Lewis, W.H., Wright, J.A. (1978) J. Cell. Physiol. 97, 73-86.

- Lewis, W.H., Kuzik, B.A., Wright, J.A. (1978a) *J. Cell. Physiol.* 94, 287-98.
- Lewis, W.H. Wright, J.A. (1974) *Biochem. Biophys. Res. Commun.* 60, 926-33.
- Lewis, W.H. and Srinivasan, P.R. (1983) *Mol. and Cell. Biol.* 3, 1053-1063.
- Lindberg, U., Nordenskjold, B.A., Reichard, P., Skoog, L. (1969) *Cancer Res.* 29, 1498-1506.
- Lowy, I., Pellicer, A., Jackson, J.F., Sim, G.K., Silverstein, S. and Axel R. (1980) *Cell* 22, 817-823.
- Maniatis, T., Fritsch, E., and Sambrook, J. in *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
- Marcu, K.. and Cooper, M.D. (1982) *Nature* 298, 327-328.
- Mather, E.L., Alt, F.W., Bothwell, A.L.M., Baltimore, D., Koshland, M.E. (1981) *Cell* 23, 369-378.
- McBride, O.W. and Peterson, J.L. (1980) *Ann. Rev. Genetics* 14, 321-346.
- McLauchlan, J. and Clements, J.B. (1983) *J. Gen. Virol.* 64, 997-1006.
- Meuth, M., Aufreiter, E., Reichard, P. (1976) *Eur. J. Biochem.* 71, 39-43.
- Oguro, M., Suzuki-Hori, C., Nagano, H., Mano, Y, Ikegami, S. (1979) *Eur. J. Biochem.* 97, 603-607.
- Perucho, M., Goldfarb, M., Shimizu, K., Lama, C., Fogh, J., Wigler, M. (1981) *Cell* 27, 467-476.
- Perucho, M., Hanahan, D., and Wigler, M. (1980) *Cell* 22, 309-317.
- Peterson, D.M., Moore, E.C. (1976) *Biochim. Biophys. Acta* 432, 80-91.
- Rubin, J.S., Prideaux, V.R., Willard, H.F., Dulhanty, A.M., Whitmore, G.F., Bernstein, A. (1985) *Mol. Cell. Bio.* 5, 398-405.
- Sabourin, C.L.K., Bates, P.F., Glatzer, L. Chang, C.-C., Trosko, J.E. Boezi, J.A. (1981) *Som. Cell Gentic.* 7, 255-68.
- Sikela, J.M., (1983) *Gene* 22, 219-228.
- Southern, P.J., and Berg, P. (1982) *J. Mol. and Applied Genetics* 1, 327-342.
- St. John, T.P. and Davis, R.W. (1979) *Cell* 16, 443-452.

- Stark, G. and Wahl, G.M. (1984) *Ann. Rev. Bioch.* 53, 447-491.
- Thelander, L., and Reichard, P. (1979) *Ann. Rev. Biochem.* 48, 133-159.
- Thelander, L., Eriksson, S. and Akerman M. (1980) *J. Biol. Chem.* 255, 7426-7432.
- Timberlake, W.E. (1980) *Developmental Biology* 78, 497-510.
- Turner, M.K., Abrams, R., and Lieberman, I. (1968) *J. Biol. Chem.* 243, 3725-3728.
- Tyrsted, G. (1984) *Mol. Cell. Biochem.* 62, 165-174.
- Varmus, H.E. (1984) *Ann. Rev. Genetics* 18, 553-612.
- Wahl, G.M. Robert de Saint Vincent, B., DeRose, M.L. (1984) *Nature* 307, 516-20.
- Wigler, M., Pellicer, A., Silverstein, S., and Axel, R. (1978) *Cell* 14, 725-731.
- Wigler, M., Sweet, R., Sim, G.K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979) *Cell* 16, 777-785.
- Wold, B., Wigler, M., Lacy, E., Maniatis, T., Silverstein, S., Axel, R. (1979) *Proc. Natl. Acad. Sci. USA* 76, 5684-5688.
- Zieg, J., Clayton, C.E. Ardeshir, F., Giulotto, E., Stark, G.R. (1983) *Mol. Cell. Biol.* 3, 2089-98.

PART II

CONSTRUCTION OF MUTATIONS IN THE CHICKEN ADULT ALPHA GLOBIN GENES

Backg

analy
to be
gene
genes
organ
and f
thall
Kazaz

As is
gene
of va
the c
are k
chron
and t
(Lars
locat
genes

LITERATURE REVIEW

Background

The recent advent of molecular cloning techniques has allowed the analysis of eukaryotic genes at the molecular level. One of the first to be isolated, and consequently one of the most thoroughly studied gene systems is that of the globins. The mammalian α and β globin genes have been extensively analyzed both in structure (gene organization, linkage, nucleotide sequence, and nuclease sensitivity) and function (evolutionary conservation of flanking sequences, thalassemias, and in vitro mutations). (For a review see Orkin and Kazazian, 1984; Efstratiadis et al., 1980).

Another well-studied system is that of the chicken globin genes. As is the case for the mammalian genes, much is known about the avian gene structure; however, not as much work has been done on the function of various elements within the chicken genes. For example, although the complete nucleotide sequences of all three chicken α globin genes are known (Dodgson and Engel, 1983; Engel et al., 1983), their chromosomal linkages have been determined (Engel and Dodgson, 1980), and their pattern of nuclease sensitivity during development is known (Larson and Weintraub, 1982), very little has been learned about the location or function of elements controlling expression of these genes.

Specific DNA sequences and/or cellular factors that control globin expression within the erythroid cell have not yet been identified; however, progress has been made at defining the elements necessary for constitutive expression (i.e. expression in heterologous cells in a nonregulated fashion) of the mammalian globin genes (reviewed in Orkin and Kazazian, 1984). Since there is little homology between the mammalian and chicken α globin promoter regions, sequence comparisons have not been useful in defining analogous sequences necessary for constitutive expression of the chicken α globin genes. One of the long term goals of this laboratory is to better understand the molecular basis for the differential expression of the chicken globin genes during erythropoiesis. As a prelude to this, we have begun studies aimed at defining the promoter elements of the chicken adult α globin genes which are necessary for constitutive expression of these genes. The work presented here describes the in vitro mutagenesis procedures which we are employing to answer this question.

Chicken α globin genes

The chicken α globin locus is shown in Figure 1. It consists of three closely linked genes which are all located within a 6.8 kb region. All three genes are transcribed in the same direction and arranged 5'- π - α^D - α^A -3' (Dodgson and Engel, 1983). The α^D and α^A genes are expressed in both primitive (embryonic) and definitive (adult) red blood cells (RBC), whereas the π gene is active only in embryonic blood cells. In adult RBCs the α^A and α^D genes are expressed at a 3:1 ratio. In other tissues these genes are inactive. Unlike some mammals which contain two adult α globin genes which are

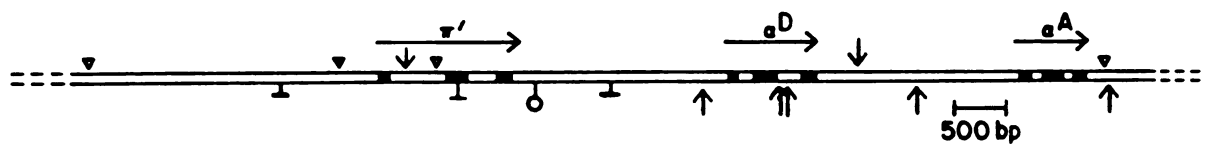


Figure 1. Chromosomal arrangement of the chicken α globin genes. The two adult (α^A , α^D) and the embryonic (α^E) genes are shown. The arrows indicate direction of transcription. Exons are shown as black boxes and introns are closed boxes. Restriction enzymes used in analysis of the locus are $\downarrow$, Eco RI; $\uparrow$, Bam HI; ∇ , Hind III; $\perp$, Taq I; $\circ$, Sac I.

almost identical in sequence, the chicken adult α globin genes are widely divergent, both in coding and flanking sequences. The α^A and α^D genes contain little homology in the 5' flanking region, except for the TATA box located 28 and 29 bp respectively upstream from the mRNA start sites.

Eukaryotic Promoters

A prototypic eukaryotic RNA polymerase II transcription unit is shown in figure 2 (for a review see Nevins, 1984). Classically, the signals which control transcription have been called promoters or promoter elements. Analysis of genes transcribed by polymerase II has revealed a complex set of elements which can interact to control the frequency and location of mRNA initiation. In general the definition of eukaryotic promoters has relied upon two methods: a) Sequences from various genes have been compared, and conserved, presumably important, sequences identified; and b) In vivo and in vitro generated mutations have been used to assay the function of these presumably important sequences.

The most clearly defined element of the polymerase II promoter is the TATA or Goldberg-Hogness box located 25-30 base pairs upstream from the transcription start site (Goldberg, 1979). This element is present in nearly all polymerase II structural genes examined. It is thought to function by identifying the location where mRNA transcription will start (also called the CAP site). Deletion of TATA results in the use of multiple start sites, but does not always decrease the rate of transcription from the region. In other words the total number of 5'

Figure 2. Typical eukaryotic RNA polymerase II transcription unit. The various elements which have been observed in the polymerase II transcription unit are indicated. In cases where the exact role of a particular element has yet to be elucidated, the element is marked with a question mark (?).

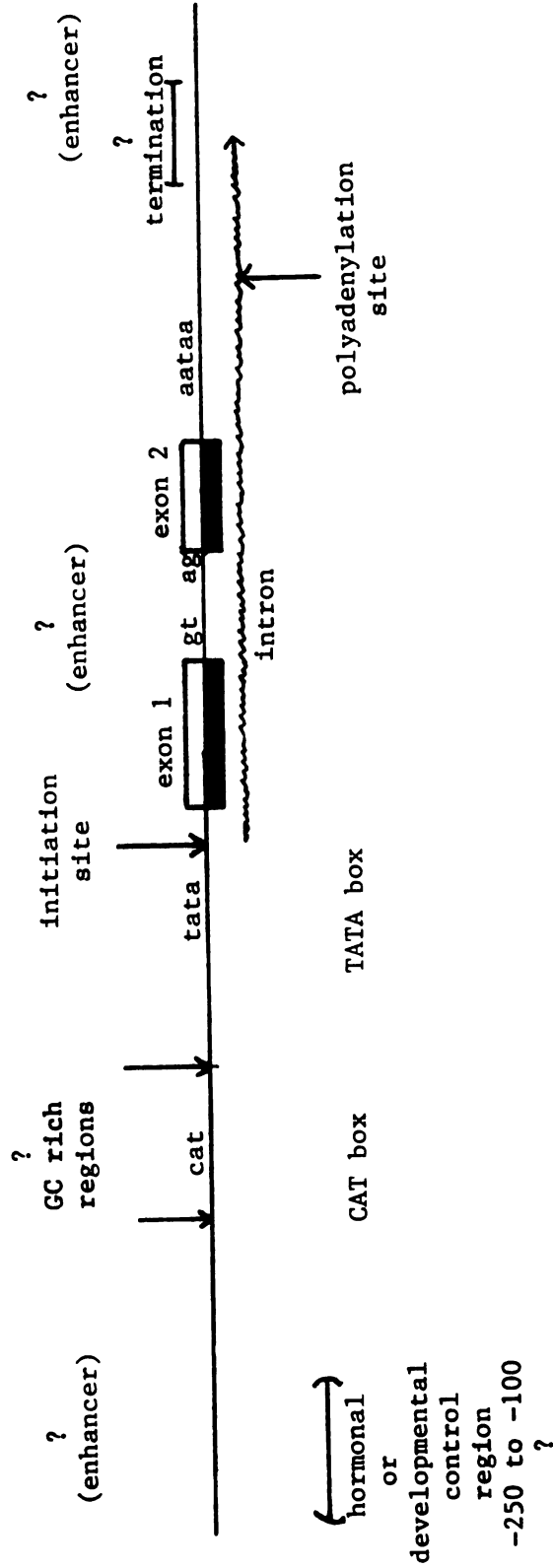


Figure 2

RNA ends may be unchanged, but the location of the ends is heterogenous in the absence of TATA (Orkin and Kazazian, 1984).

Other elements of polymerase II promoters are not so clearly defined. Sequence comparisons have identified a pentanucleotide sequence, CCAAT, located 80 to 100 base pairs upstream from the cap site of many eukaryotic structural genes, most notably all of the normal mammalian globin genes (Benoist et al., 1980; Efstratiadis et al., 1980). This so called "CAT" box is not found at a similar location in any of the avian adult α globin genes examined (Dodgson and Engel, 1983). It is also absent in several other genes and therefore its function as a promoter element is unclear.

Another distal transcription signal which has the sequence PuCPuCCC (Pu = purine) has been identified in the -80 to -100 region of the mammalian β globin gene. This element has a striking homology to transcriptional control elements defined for constitutive expression of the herpes simplex virus thymidine kinase gene and the SV 40 early region (Dierks et al., 1983). Scott and Tilghman (1983) also suggest that GC rich regions, similar to PuCPuCCC, flanking the CCAAT pentamer are important in transcriptional control. A much more detailed study of the herpes simplex thymidine kinase gene draws the same conclusion (McKnight and Kingsburg, 1982).

A final element whose importance in transcription control has recently been defined, is called an enhancer or activator. These elements seem to act in a position-independent, orientation-independent manner to increase transcription in cis from promoters located at distances up to at least 10 kb (for a review see Khoury and Gruss, 1983). Many eukaryotic viruses, including DNA viruses and

retroviruses, have been shown to contain enhancers. It has also been shown that enhancers often act in a cell-type or tissue - specific manner suggesting they may be intimately involved in differential control of gene expression. For example, the immunoglobulin heavy chain gene contains an enhancer in the major intron. Upon rearrangement of the immunoglobulin gene, this enhancer, which is active only in lymphocytes, is brought into a region near enough to the immunoglobulin promoter to activate transcription (Gillies et al., 1983).

In vitro Mutagenesis

The ability to construct and analyze specific changes in a defined DNA sequence is central to the study of transcriptional regulatory elements. This procedure of in vitro construction and subsequent analysis of specific mutations has often been called "reverse genetics" (for a review on directed mutagenesis, see Shortle et al., 1981).

In general, "reverse genetic" analysis of regulatory elements is accomplished in a two step procedure. Initially, defined sections of DNA are deleted from or inserted in the region of interest to roughly map the sections important in regulation. Secondly, specific base substitutions are used to precisely define the elements localized by the first set of mutants.

Recently McKnight and Kingsbury (1982) developed a system for studying eukaryotic promoters called "linker scanning" mutagenesis. In this system two series of deletions, terminating with a synthetic restriction endonuclease site (linker), are constructed from either end of the promoter region (see Figure 3). By using the linker for recombination of the appropriate 3' and 5' deletions, clusters of point

Figure 3. Example of linker scanning mutagenesis. The region to be mutagenized is depicted as the hatched box in A). Various deletions are constructed using restriction enzyme sites X and Y as starting points. Each deletion has a linker (boxed region) at its terminus. For example, the deletion of clone Y4 starts at Y and extends just into the target region. The X1 deletion extends from X exactly the length of the linker past the Y4 deletion endpoint. Recombination of X1 and Y4 at the linker gives the clone shown in B). Recombination of two mutants whose deletion endpoints overlap (Y2 + X2) results in an internal deletion in the target area as shown in C).

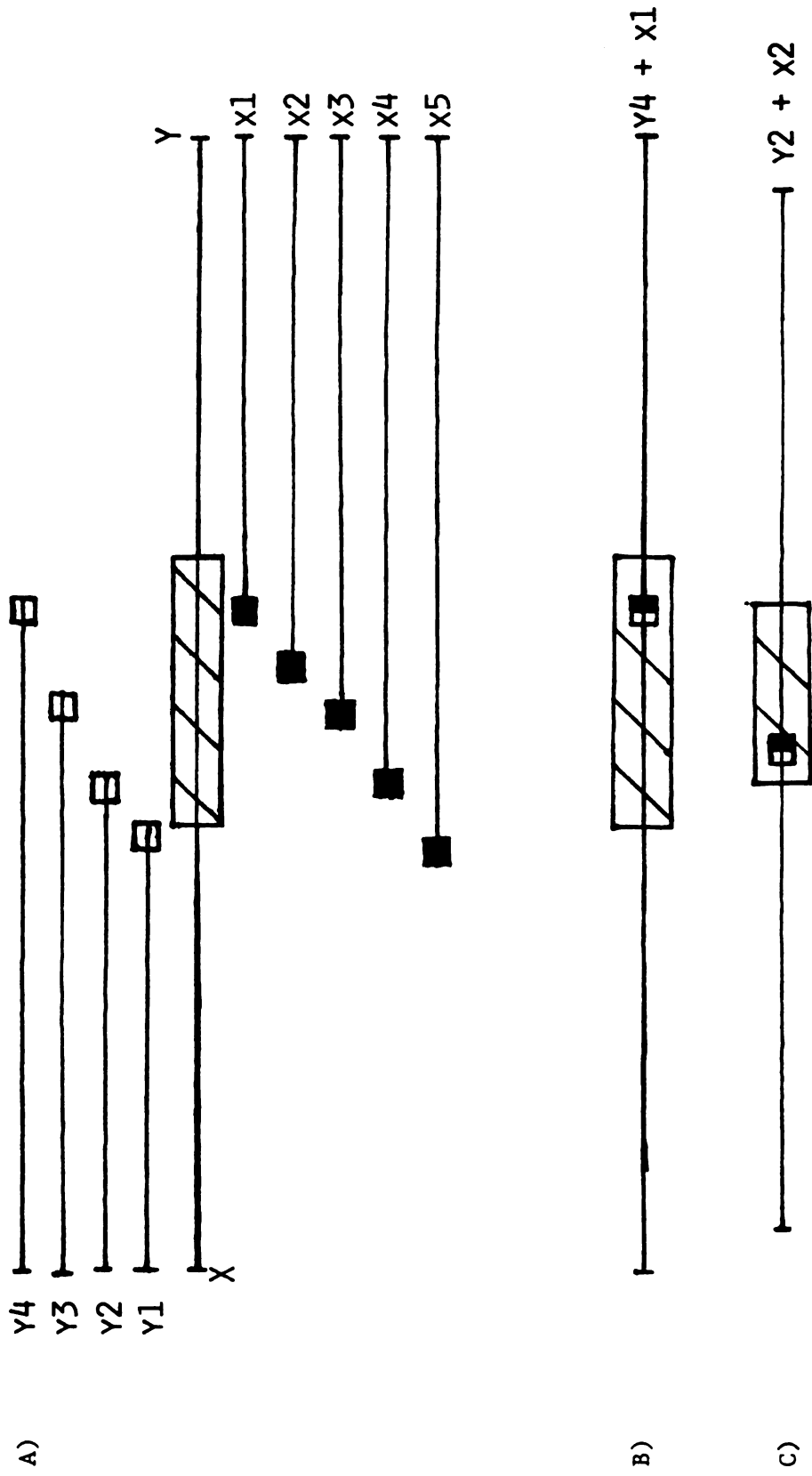


Figure 3

muta

exac

of p

end

be

lin

the

ad

de

ge

z4

c

r

r

a

a

mutations can be generated. Mutants with endpoints that are spaced exactly the length of the linker will recombine to give only a cluster of point mutations in the linker region. (Of course some of the endogenous nucleotides in the region where a linker is substituted may be identical to the corresponding bases in the linker, or the complete linker sequence may differ from the normal sequence.) The remainder of the reconstructed gene is identical to the natural gene. A major advantage of linker scanning mutagenesis is, that unlike simple deletion mutations, the orientation of sequences in the reconstructed gene, outside of the linker itself, is identical to the natural gene. If one recombines deletions with endpoints which are further apart or closer together than the length of the linker, then the resultant recombinant will yield point mutations plus deletions or insertions respectively. As with other mutagenesis procedures, initially the 3' and 5' deletions are used to define the specific area(s) of interest by assaying their promoter activity, and then the appropriate deletion mutants are recombined.

Mate

Dodg

base

pt²3i

which

two s

deriv

to 12

conta

restr

intes

Bioch

Biote

isome

from

Metho

Mania

using

modif

metho

were

modifi

treate

MATERIALS AND METHODS

Materials. Plasmids pHR α 5-4.3, and pBR α 7-1.7 have been described by Dodgson and Engel (1983). Plasmid pB α 5-0.9 is a subclone of the 900 base pair Bam HI fragment of α^D constructed by J. Dodgson. p Δ TBR α 7-1.7 Δ NI20 is a clone of the α^A gene derived from pBR α 7-1.7b in which the internal Nae I fragment has been deleted and replaced with two synthetic Xho I linkers. This α^A gene is cloned in a pBR322 derivative in which the Nae I fragments between base pairs number 401 to 1283 (relative to the Eco RI site) have been deleted and thus contains no Nae I or Nar I sites (constructed by J. Dodgson). All restriction enzymes, T4 polynucleotide kinase, T4 DNA ligase, and calf intestinal alkaline phosphatase were from the following sources: PL Biochemicals, BRL, New England Biolabs, Promega Biotec, International Biotechnologies, or Boeringer Mannheim. Nuclease Bal 31 fast and slow isomers were from International Biotechnologies, Bal 31 (mixed) was from BRL.

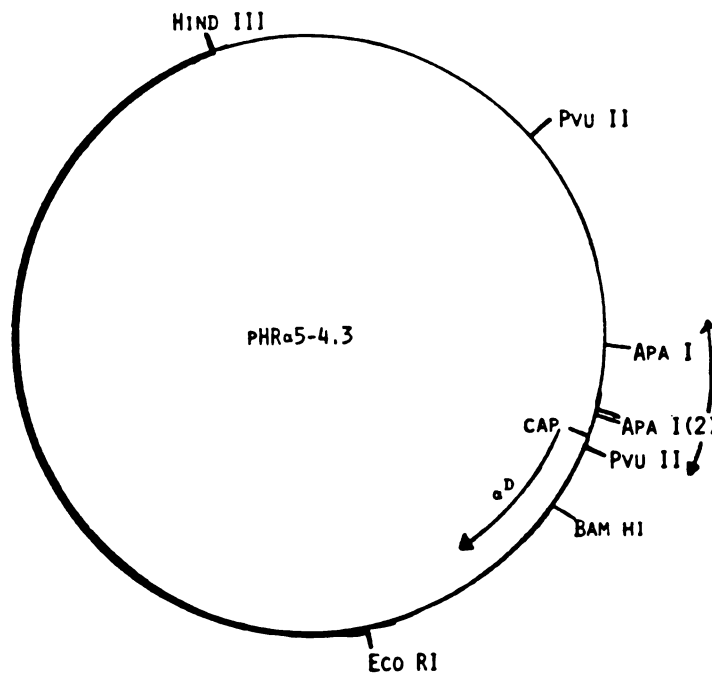
Methods. In general all cloning techniques followed those outlined by Maniatis, Fritsch and Sambrook (1982). Sequence analysis was performed using the chemical degradation method of Maxam and Gilbert (1980) as modified by Smith and Calvo (1980). Plasmid DNA minipreps used the method of Ish-Horowicz and Burke (1981). Nuclease Bal 31 digestions were generally performed according to Maniatis et al., (1982) but modified slightly as follows: Five micrograms of linearized DNA was treated with the amount of Bal 31 indicated at 30° for the length of

time indicated. Digests were stopped by chelation of calcium with EGTA, then the samples were deproteinized with phenol. The DNA was ethanol precipitated and resuspended in 10 mM Tris, 1 mM EDTA. One half of the DNA from each timepoint was digested with the appropriate restriction enzyme and run on a 5% acrylamide gel to assay the extent of nuclease digestion. The DNA of appropriate time points was made blunt-ended with DNA polymerase Klenow fragment and ligated with T4 DNA ligase to synthetic XhoI linkers. Excess linkers were removed with XhoI and the plasmids were recircularized by ligation in dilute solution. The plasmids were then transformed into E. coli HB101.

RESULTS

α^D : Construction of 5' deletions.

The strategy outlined in Figure 4 was used to generate the deletions extending into the 5' end of the putative α^D promoter region. This assumes that like other genes studied to date, the α^D promoter lies within 120 bp of the mRNA initiation site. Plasmid pHR α 5-4.3 contains 3 ApaI sites at the positions indicated. The 3' most site is 120 bp upstream (-120) of the α^D mRNA start site (the numbering system employed designates the mRNA start site as zero, sequences preceeding it are negative, and those following it are positive). Five micrograms of ApaI digested pHR α 5-4.3 was treated with a very low level of nuclease Bal 31 (0.2 unit/5 μ g) for times varying between 0.5 and 10 minutes. The extent of nuclease digestion was assayed by cleaving the DNAs with BstXI and analyzing the products on a 5% acrylamide gel. The BstXI site is located at +352 and the ApaI site is at -120; therefore the BstXI digest will generate fragments between



- 1) Linearize pHR 5-4.3 with Apa I
- 2) Digest with nuclease Bal 31 for various lengths of time
- 3) Assay extent of deletions with Bst XI digests
- 4) Klenow polymerase fill the DNA ends & ligate with synthetic Xho I linkers
- 5) Remove excess linkers & recircularize plasmids by ligation in dilute solution
- 6) Transform E. coli & screen mini DNA preps with Xho I + Bst XI digests

Figure 4. Scheme for construction of α^D 5' deletions. Plasmid pHR α 5-4.3 contains a 4.3 kb chicken DNA insert (thin line) cloned in the vector pBR322. The α^D transcription start site and direction of transcription are shown. The deletions were constructed as described above and in the text. The area deleted is indicated by the arrows outside the plasmid.

350 and 470 bp in size if the endpoints of the deletion fall in the putative promoter region.

Gel analysis showed the DNAs from the 5 and 10 minute time points contained deletions of approximately 0 to 30 and 50 to 120 base pairs respectively. Deletions of this length would be expected to lie in the -120 to -40 and -70 to CAP regions. All other time points contained no detectable deletions. One feature of the Bal 31 digestions which was noticed was the appearance of distinct bands imposed on the smear expected from the nuclease digestions. These bands suggest that there are preferred stop sites for the Bal 31 nucleolytic digestion.

After attachment of linkers and plasmid recircularization the 5 and 10 minute time point DNA samples were used to transform E. coli HB101. Approximately 500 and 62 colonies were obtained with the 5 and 10 minute DNAs respectively. Some of the 10 minute time point DNA was presumably lost during the manipulations. One hundred and eighty of the five minute time point colonies and all 62 of the ten minute digest colonies were picked and plasmid minipreps were prepared. The five and ten minute time point clones were designated A5 and A10 clones respectively. The approximate location of the linkers in these clones was determined by digestion of one fifth of each of the DNAs with Bst XI and XhoI followed by analysis on 5% polyacrylamide gels versus pHR α 5-4.3 ApaI plus BstXI digested DNA. Based on the size fragments obtained from XhoI + BstXI digests, 88 A5 and 32 A10 clones were chosen for further study. The A5 clones which were discarded all appear to have no detectable deletions, whereas the thirty A10 clones which were discarded were deleted well beyond the CAP site.

Sequencing α^D 5' deletion mutants.

The strategy used for determining the exact location of the XhoI linker in the α^D 5' deletions is shown in Figure 5. Plasmid DNA was isolated from 50 milliliter cultures of various clones using a scaled up miniprep procedure. These DNAs were then digested with Pvu II and the fragments slightly smaller than 1.2 kb were gel purified. These fragments were prepared for sequencing by treatment with calf intestinal alkaline phosphatase followed by labeling with polynucleotide kinase. The labeled DNA was digested with Sin I to cleave the 5' labelled end, and the large fragment was purified by electrophoresis on a 1.2% agarose gel before being subjected to chemical degradation sequencing using all four reactions (A>G, G, C + T, C). The sequences of four A5 clones were determined using this method (data not shown).

Determination of the linker location by performing all four sequencing reactions was not favorable for two reasons; first, the sequencing reactions themselves were laborious and also required four lanes per clone on a sequencing gel, thus decreasing the number of clones that could be sequenced at one time. Secondly, the labelled sample was divided into four reactions; thus a significant (>100 cps) number of counts were needed to obtain the sequence of each clone. Since such a large number of counts were needed, large (50 milliliter) cultures had to be prepared for each clone to be sequenced. For these reasons we attempted to simplify the sequencing procedure by only performing one reaction for each cloned DNA, namely the AG or formic acid reaction. These AG reactions were compared to a sequencing ladder of the wild type (or parental) DNA labelled at the same Pvu II site to

Figure 5. Strategy for sequencing α^D 5' deletion clones using a single sequencing reaction. Miniprep DNAs of clones containing deletions starting from the Apa I site in pHR $\alpha 5-4.3$ were subjected to the sequencing scheme outlined above and discussed fully in the text. The restriction enzyme sites are as follows: A = Apa I, P = Pvu II, S = Sin I, X = Xho I linker. The numbers in parenthesis are the locations relative to the α^D CAP site.

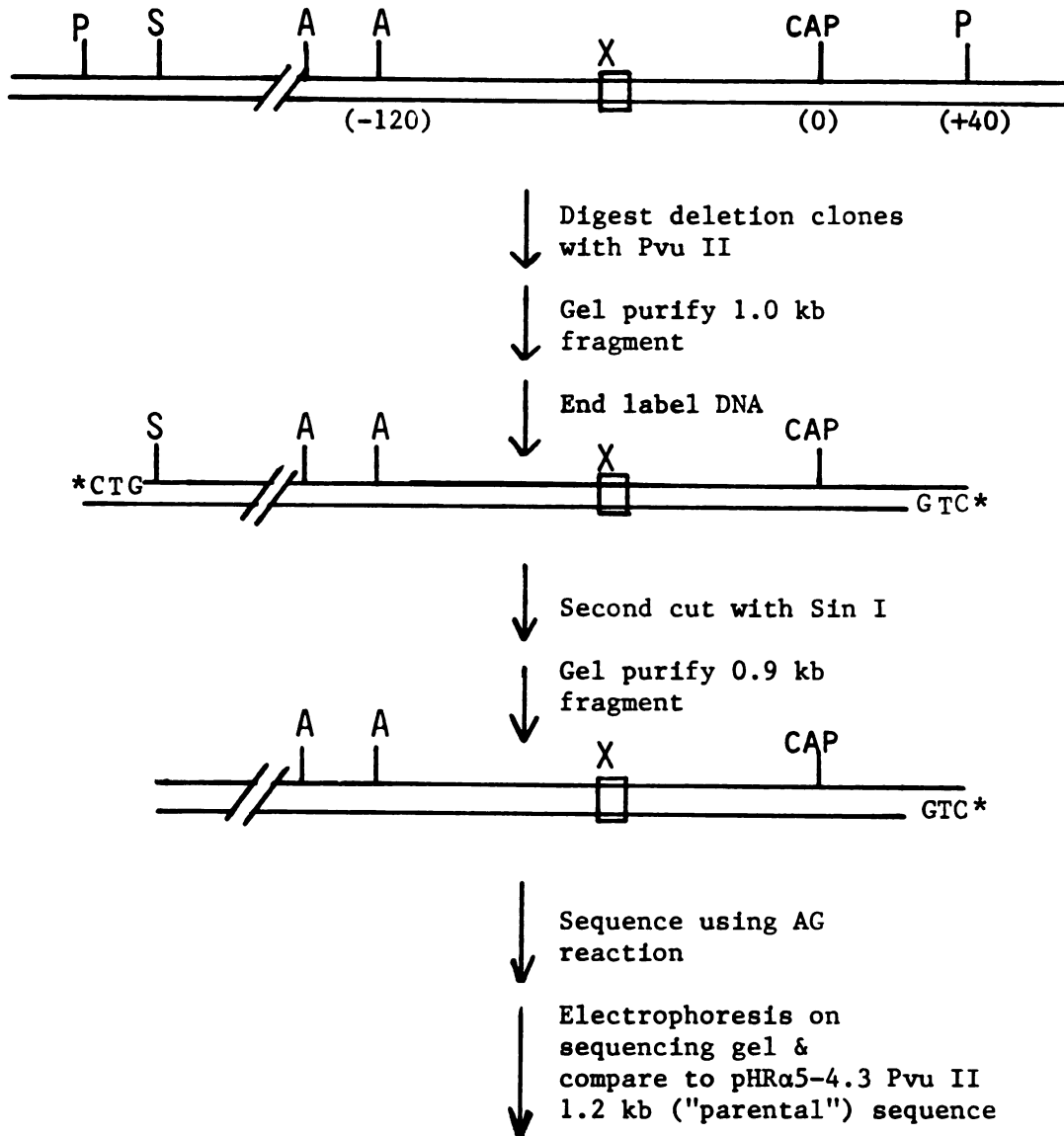


Figure 5

Figure 6. Sequencing gel of α^D 5' deletion clones: One sequencing reaction. Autoradiograph of a 6% acrylamide, 7M urea sequencing gel of 11 different α^D 5' deletion clone DNAs prepared for sequencing as shown in Figure 5. The control lanes marked G, A + G, C + T, C are the "parental" pHR α 5-4.3 Pvu II 1.2 kb DNA sequencing reactions. Deletion endpoints are determined by comparing clone and parental DNA AG reactions. The A + G and C + T reactions on the right are both reactions run with clone A5-46 DNA to confirm the validity of this method for determining deletion endpoints.

A C
+ +
G G T C

A C
+ +
G G T

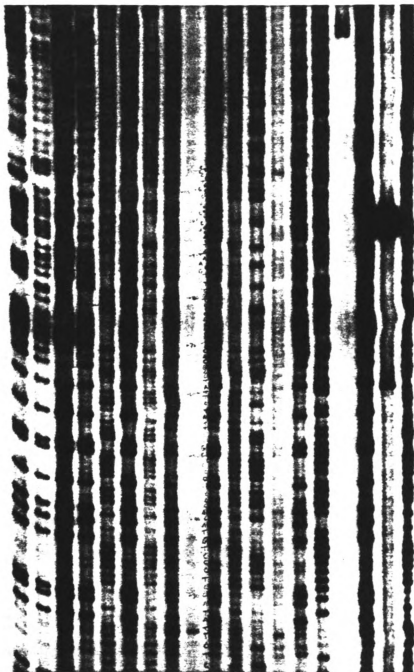


Figure 6

determine the location of the linker (Figure 6). Since the sequence of the XhoI linker was known (5'-CCTCGAGG-3'), it was relatively easy to place the linkers by scanning the sequencing autoradiograph for the 4 base gap in the AG lane left by the CCTC sequence of the linker. Sequencing reactions using this procedure could routinely be performed using the DNA from a five milliliter culture.

Although the single reaction method for placing the linkers was fairly simple, it still required a significant amount of time and labor to perform each of the AG sequencing reactions. An alternative scheme was devised to make placement of the linkers even simpler. As outlined in Figure 7 this method takes advantage of the fact that the restriction enzymes Pvu II and Xho I leave blunt and 5'-protruding ends respectively, and that after labelling at the PvuII/XhoI sites, the resulting fragments which differ in length by 4 bp can easily be resolved on a sequencing gel. As shown in Figure 8, comparison of the labelled fragment from each strand of the clone DNA with a sequencing ladder of the parental DNA labelled at the Pvu II site allows simple placement of the linkers in each clone. Subtraction of 2 nucleotides from the shorter fragment (labelled at the Pvu II end) or 6 nucleotides from the longer fragment (labelled at the Xho I end). Using this method, up to 12 deletion mutants can be analyzed on a single sequencing gel.

Location of linkers in A5 and A10 clones.

A summary of the locations of the linkers in all of the α^D 5' deletion clones sequenced to date is shown in Figure 9. The exact location of the Xho I linker has been determined for forty six of the 5' deletion mutants; thirty five of the A5 clones and eleven of the A10

Figure 7. Scheme for mapping α^D 5' deletion endpoints by fragment length determination. The position of the deletion endpoint was mapped by accurately determining the length of both strands of the fragment obtained by digesting α^D 5' deletion clone DNAs with Xho I and Pvu II. Symbols are as in Figure 5. Details are given in the text.

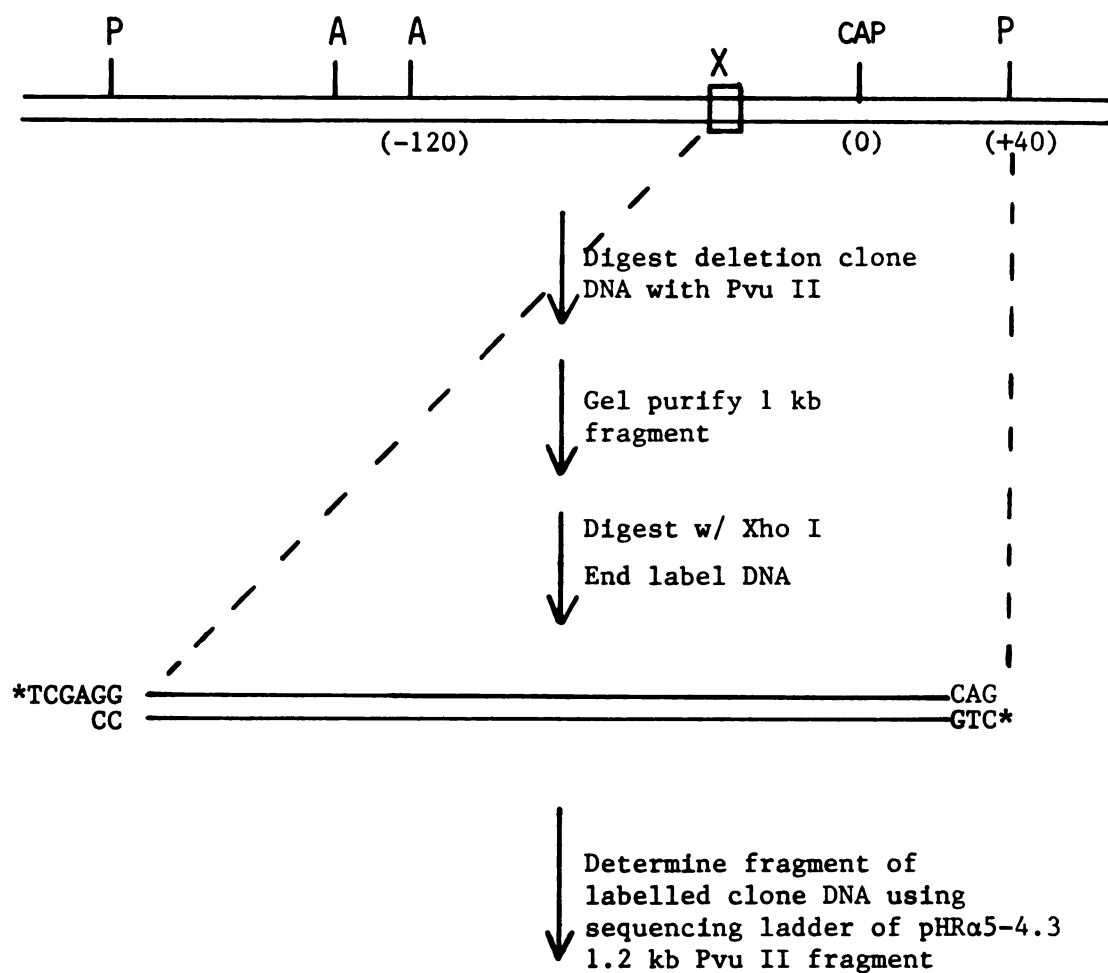


Figure 7

Figure 8. Sequencing gel of α^D 5' deletion clone DNAs Xho I/Pvu II labelled fragments. Clone DNAs prepared as in Figure 7 were analyzed on a sequencing gel versus of sequencing ladder of the "parental" pHR α 5-4.3 Pvu II 1.2 kb fragment labelled at the 3'-most (+40) Pvu II site. Additional A + G parental sequencing reactions were run to accurately size the two strands of each clone.

$$\begin{array}{c} \text{A} \\ + \\ \text{G} \end{array}$$

$$\begin{array}{cc} \text{A} & \text{C} \\ + & + \\ \text{G} & \text{T} \end{array} \quad \text{C}$$

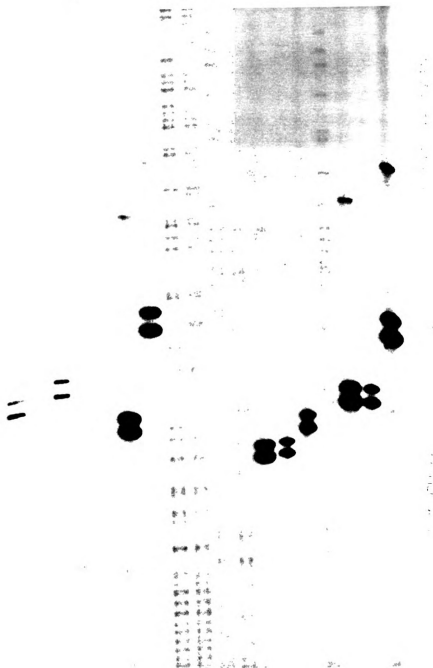
$$\begin{array}{c} \text{A} \\ + \\ \text{G} \end{array}$$


Figure 8

Figure 9. Summary of α^D 5' deletions. The locations of the linkers, and consequently the deletion endpoints for the α^D 5' deletions are summarized. In each case the sequence of the linker is shown in the position relative to the wild type sequence. Wild type sequences are designated by a line. The top line is the wild type sequence of the -120 to CAP region of α^D .

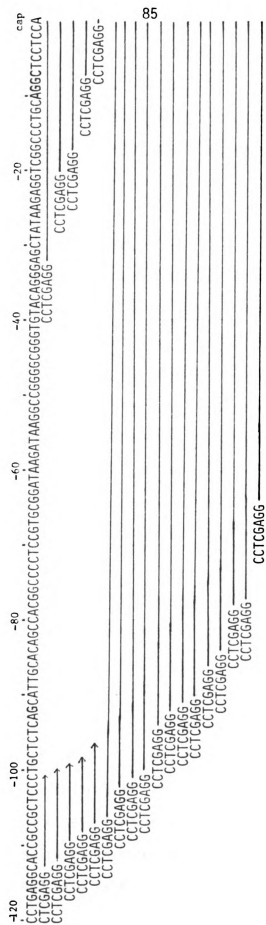


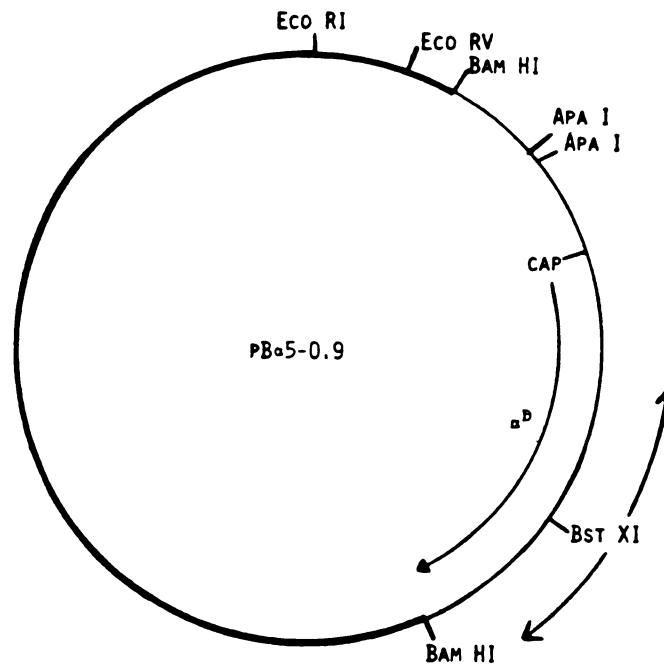
Figure 9

clones. The endpoints of these deletions fall at thirty-six different locations in the region between -120 and CAP. Several mutants have linkers located at precisely the same nucleotide. For example, there are four independently isolated mutants that have a deletion endpoint at -94, four at -77, and three at -109.

α^D : Construction of 3' deletions.

The strategy employed for construction of α^D mutants with deletions extending into the promoter region from the 3' end is shown in Figure 10. Deletions were created using plasmid pB α 5-0.9 which had been linearized with the restriction enzyme Bst XI. Bal 31 digestions were carried out using 0.5 units of the mixture of isomers of the enzyme per 5 μ g DNA for periods of time ranging from 3 to 18 minutes. Since the Bst XI site is located at position + 352, the amount of enzyme and time of digestion were optimized to remove between 350 and 470 bp of DNA, thus placing the deletion 5' endpoint in the putative promoter region.

The extent of nuclease digestion was determined by cutting the Bal 31 treated DNA with Bam HI restriction enzyme, and sizing the products on a 5% polyacrylamide gel. The 5' Bam HI site is located 254 bp upstream of the CAP site; therefore 130 to 250 base pair fragments represent deletions extending into the putative promoter region. The 12 and 15 minute Bal 31 digest yielded fragments in this size range, however in both cases the fragment size was very heterogeneous. For example the 15 minute digest fragments were spread over a range of at least 350 to 150 base pairs in size, and the 12 minute digest fragments ranged from 220 to 400 base pairs.



- 1) Linearize pBα5-0.9 with Bst XI
- 2) Digest DNA with nuclease Bal 31 for various lengths of time
- 3) Assay extent of deletion with Bam HI digests
- 4) Fill ragged DNA ends with polymerase Klenow fragment & blunt end ligate to synthetic Xho I linkers
- 5) Digest with Xho I to remove excess linkers & recircularize plasmid by dilute ligation
- 6) Transform *E. coli* & screen minipreps by Xho I + Eco RI digests

Figure 10. Construction of α^D 3' deletion mutants. The thin lines represent cloned chicken sequences. Thick lines are plasmid pBR 322 DNA. α^D transcription start site (CAP) and direction and region of transcription are indicated by arrow inside plasmid. The arrow outside the plasmid denotes the direction of deletion from the Bst XI site. The restriction sites are as indicated and cloned sequences are to scale.

XhoI linkers were added to the 12, 15, and 18 minute Bal31 digests and the recircularized plasmids were transformed into E. coli. The 12, 15 and 18 minute digests yielded 500, 1000, and 100 colonies respectively. It appears that fewer colonies were obtained with the 18 minute digest DNA because the 3' endpoints of some of the Bal 31 deletions extended into the adjacent pBR322 sequences and disrupted the origin region needed for plasmid replication.

Miniplasmid DNA preparations were made from 336 and 24 of the 15 and 18 minute colonies respectively. These 1.2 ml culture plasmid DNA preparations were very crude and contained large amounts of RNA and chromosomal DNA; therefore restriction enzyme fragments below 200 bp in size were very difficult to visualize with ethidium bromide staining of 5% acrylamide gels. Because of this, the Bam HI site at -254, which would be expected to yield fragments of 130 to 250 base pairs when used in a double digest with XhoI, could not be used to screen these deletion mutants. Instead the Eco RI site located in the pBR322 DNA 630 base pairs 5' to the CAP site was used. Fragments extending from the Eco RI site to the linkers should range from 500 to 630 base pairs in length if the linker falls in the putative promoter region. These size fragments are fairly difficult to size accurately on acrylamide gels since this is near the nonlinear portion of the gel's partitioning range. Also, the large amount of RNA in the miniprep can make the fragments run anomalously. Therefore, the linker locations obtained in these experiments could only be considered approximations.

Double digests of miniprep DNA with Xho I + Eco RI showed that all of the 18 minute colonies contained deletions well past the -120 region. These clones were not further characterized. Eighty three of

the 15 minute digest colonies were shown to have deletions which appeared to fall in or near the region between -120 and the CAP site. These clones were designated B15-1 through B15-83.

Although the EcoRI + XhoI digest had roughly positioned the linkers in the B15 clones, it was decided to more precisely map these linkers before proceeding with sequence analysis. DNA was prepared from 5 milliliter cultures (called "midipreps") of all 83 B15 clones. Care was taken while preparing these DNAs to rid the samples of chromosomal DNA and RNA and thus avoid the problems described previously. Gel analysis of Xho I + Bam HI digests of a small sample of the midiprep DNAs confirmed that 42 of the 83 B15 clones contained deletions in the -120 to CAP region. The other 41 clones contained deletions slightly outside the putative promoter region, usually within 20 base pairs of the -120 to CAP region.

α^D : Sequencing 3' deletion mutants.

The strategy outlined in Figure 11 was used to exactly position the linkers in the B15 clones. This scheme took advantage of the fact that the Bal 31 deletions had extended beyond the Bam HI site at the 3' end of the α^D gene. Therefore the B15 mutants contain only one Bam HI site at position -254. Linearized midiprep DNA (three quarters of each sample) was sequenced from this Bam HI site.

Similar to the procedure used for the A5 clones only one sequencing reaction, the AG reaction, was performed on the labelled B15 DNAs. The B15 reactions were run in parallel on a long 6% sequencing gel with a sequencing ladder of the parental Bam HI labelled pB α 5-0.9 DNA. The linker positions were located by examining the autoradiograph

Figure 11. Sequencing strategy for α^D 3' deletion clones. The steps shown above were used to map the location of the linker in the α^D 3' deletion clones (designated B15 clones). The deletions have removed the Bam HI site 3' to the Bst XI site in Figure 10 leaving only one Bam HI site in each clone. The sequenced DNAs were run on long 6% sequencing gels.

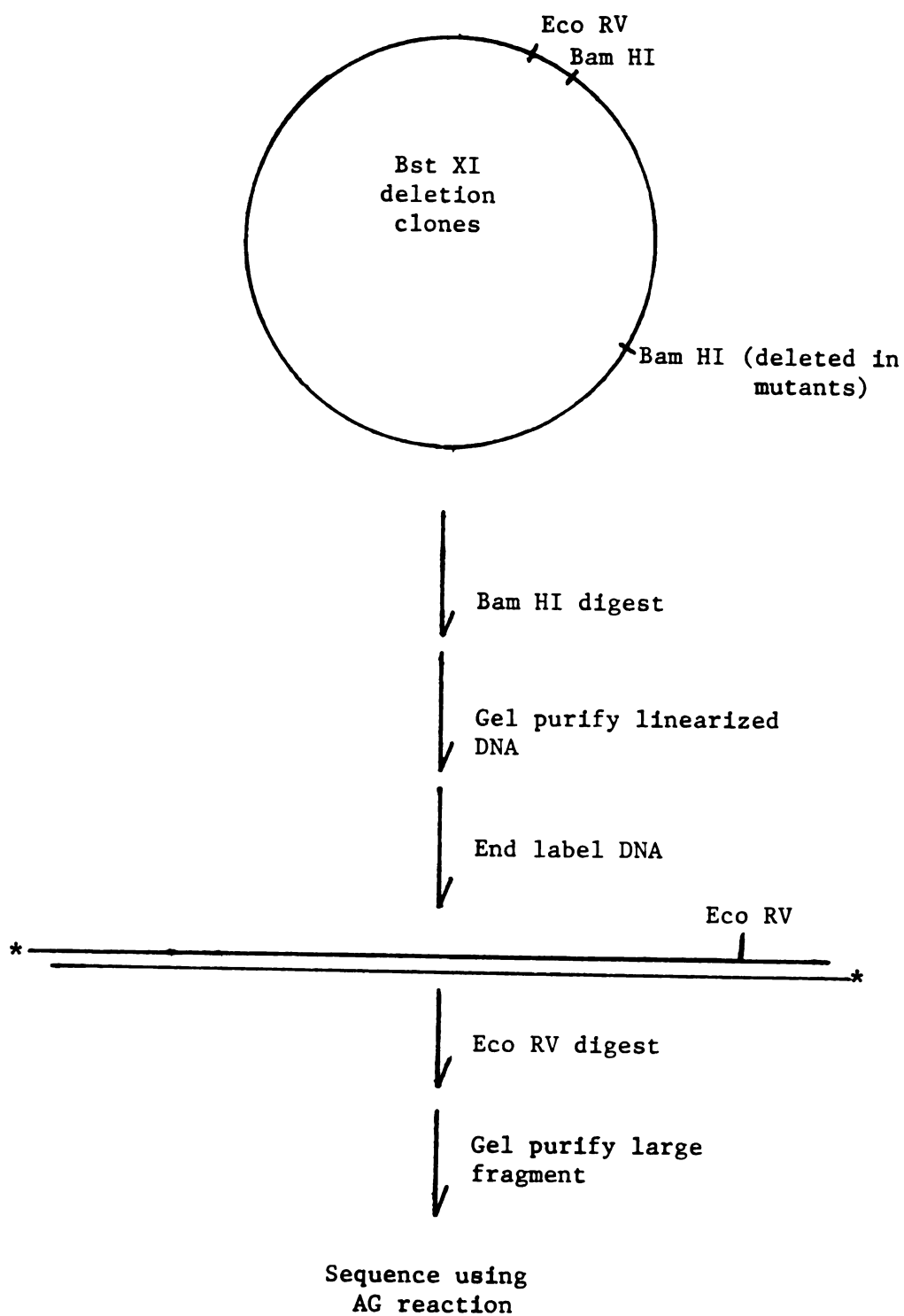


Figure 11

for the location where the mutant and parental DNAs diverged (data not shown). Using this procedure the deletion endpoints summarized in Figure 12 were mapped.

A procedure similar to the fragment sizing protocol used for the A5 clones (Figure 7) was attempted with the B15 clones. In this case the B15 clone DNAs were digested with Bam HI and Xho I and labelled at both 5' ends. The labelled DNAs were run on a sequencing gel versus the parental DNA sequencing reactions. This procedure could not be used to exactly locate the linkers in the B15 clones for two reasons: First since Bam HI and Xho I both leave a 4 bp 5' protruding end, the labelled fragments from both strands are the same length and migrate as a single, somewhat diffuse, band on the sequencing gel. Unlike the A5 clones where there are two bands separated by 4 bp, the DNA fragment lengths of the B15 clones cannot be double checked to ensure proper positioning of the linker. The second, and probably most critical, reason this method could not be used with the B15 clones was the distance between the labelled BamHI site and region where the linkers were located. The Bam HI site lies 130 to 250 base pairs upstream of the putative promoter region; thus the labelled fragments being sized are located in a region of the sequencing gel where the bands are very close together (less than one millimeter between bands). Therefore it was not possible to definitively align the labelled B15 clone fragment bands with a single band in the parental sequencing reaction lanes.

Construction of α^A 5' deletions.

Plasmid pBR α 7-1.7 DNA was used to construct deletions extending into the α^A promoter region from the 5' direction using the protocol

Figure 12. Summary of α^D 3' deletions. The location of the α^D 3' deletions are shown. Details are as in figure 9.

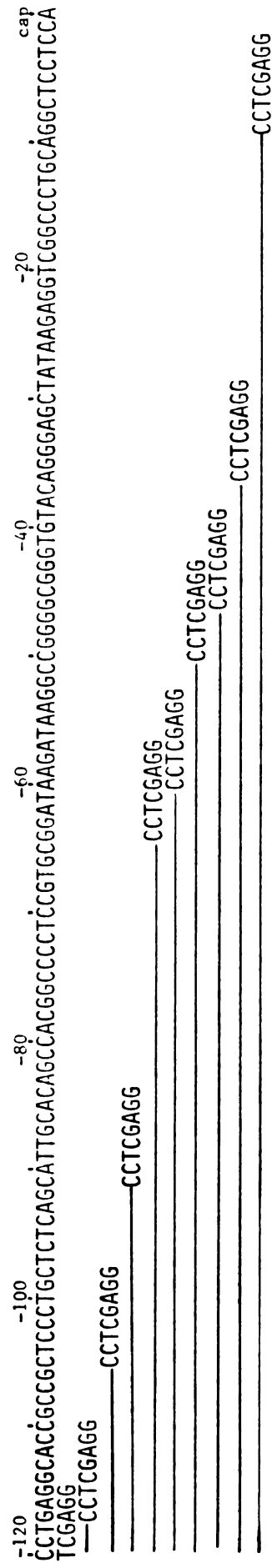
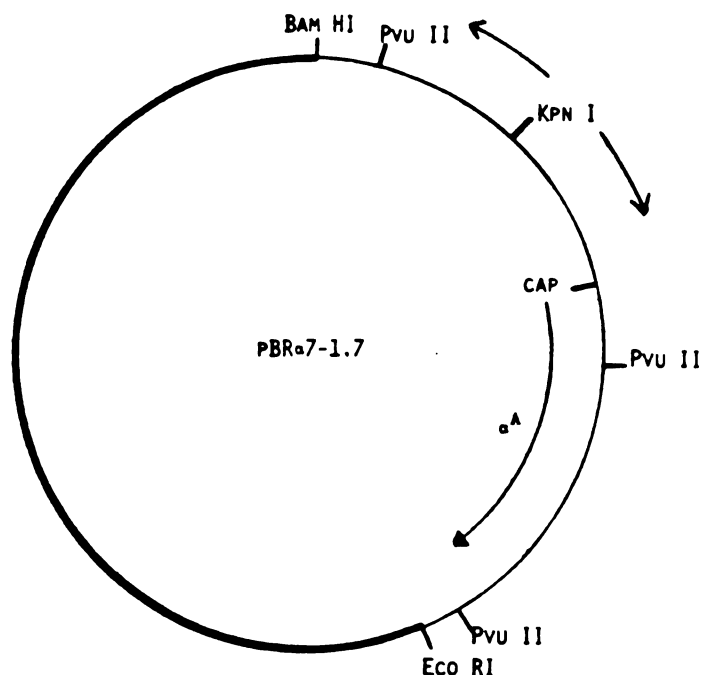


Figure 12

in Figure 13. The Kpn I site used to linearize the plasmid is positioned 340 bp upstream of the CAP site of α^A . A seven minute Bal 31 digestion (0.13 units/ μ g DNA) using the fast isomer of the enzyme was found to produce deletions of between 140 and 260 base pairs. The extent of digestion was assayed using an Eco RI digest of the nuclease treated DNA, followed by electrophoresis on a 1.2% agarose gel. The rate of digestion with the fast isomer of Bal 31 appeared to be much more synchronous than previous digests using the mixture of Bal 31 fast and slow isomers.

After addition of linkers, plasmid recircularization, and transformation, twenty four minipreps were prepared from the K7 colonies (K = Kpn site, 7 = seven minute Bal 31 digest). The linkers in these clones were approximately located by Eco RI plus Xho I digestion and analysis on 1.2% agarose gels. More accurate mapping of 12 clones, which appeared to have deletions in the appropriate region, was done using Pvu II plus Xho I. Since there are Pvu II sites both 5' and 3' to the region where the deletions were constructed (see Figure 13), this procedure mapped both the 5' and 3' endpoints of the deletions in the K7 clones. To distinguish the fragments produced from the 5' and 3' Pvu II sites, a Bam HI plus Xho I digest was also performed. This digest positioned the 5' deletion endpoints and thus by elimination allowed assignment of the Pvu II/Xho I fragment derived from the 3' deletion endpoint.

The location of the 5' and 3' deletion endpoints in the K7 and several K6 clones is shown in Figure 14. The 3' endpoints range from -290 (K6-1) to -70 (K7-5). The range of 3' deletion endpoints in the K7 clones was -262 (K7-8) to -70 (K7-5) with various deletions



- 1) Linearize pBR 7-1.7 with Kpn I
- 2) Digest with Bal 31 for various times
- 3) Assay extent of deletion with Eco RI digest run on a 5% acrylamide gel
- 4) Fill ends with DNA polymerase Klenow fragment & ligate to Xho I linkers
- 5) Remove excess linkers by Xho I digest followed by spermine precipitation. Circularize plasmids by ligation in dilute solution
- 6) Transform *E. coli* & screen minipreps of plasmid DNA with Xho I + Pvu II and Xho I + Bam HI digests.

Figure 13. Construction of α^A 5' deletion mutants. The constructions were as described in the text. pBR α 7-1.7 contains a 1.7 kb fragment containing the α^A gene in the orientation shown above (thin line) cloned in pBR 322 (thick line). The orientation and start site for α^A transcription are shown by the arrow inside the plasmid. The arrows outside the plasmid indicate the direction and regions of Bal 31 digestion.

scattered throughout this region. A surprising observation in these experiments was the strong clustering of 5' deletion endpoints around the -415 region. Of nine 5' endpoints which were accurately mapped six were located in the 26 bp region from -405 to -431.

Since none of the K7 clones had deletions extending beyond -70 and since very few of the K7 deletions even extended into the putative promoter region for α^A (i.e. -120 to CAP), we used the synthetic Xho I linker in one of these clones (K7-3) to construct deletions which extended further towards the α^A CAP site. K7-3 DNA was digested with Xho I then treated with the slow isomer of Bal 31 at a concentration of 0.1 unit Bal 31 slow per microgram DNA. This enzyme to DNA ratio is expected to remove approximately 2 bp per DNA end per minute. After treatment with DNA polymerase Klenow fragment, ligation to Xho I linkers and plasmid recircularization, plasmid DNA from twenty four colonies was prepared from ten and twenty minute nuclease digestion time points. None of these plasmids contained deletions beyond approximately -60. Other experiments using K7-3 DNA to produce deletions which extend further into the putative α^D promoter region have been conducted, but to date none of the resulting deletions have been mapped.

Figure 14. Estimated location of α^A 5' deletion endpoints in both directions. The locations of both the endpoints of the α^A deletions starting at the Kpn I site were determined by XhoI + BamHI and Xho I + Pvu II restriction enzyme digests. A) 5' (relative to Kpn I) endpoints B) 3' end points. The numbers in parenthesis are positions relative to the α^A CAP site. The position of each clone's deletion endpoint is shown with a line and the clone number (for example, clone 7-7 5' endpoint at -660 in A)).

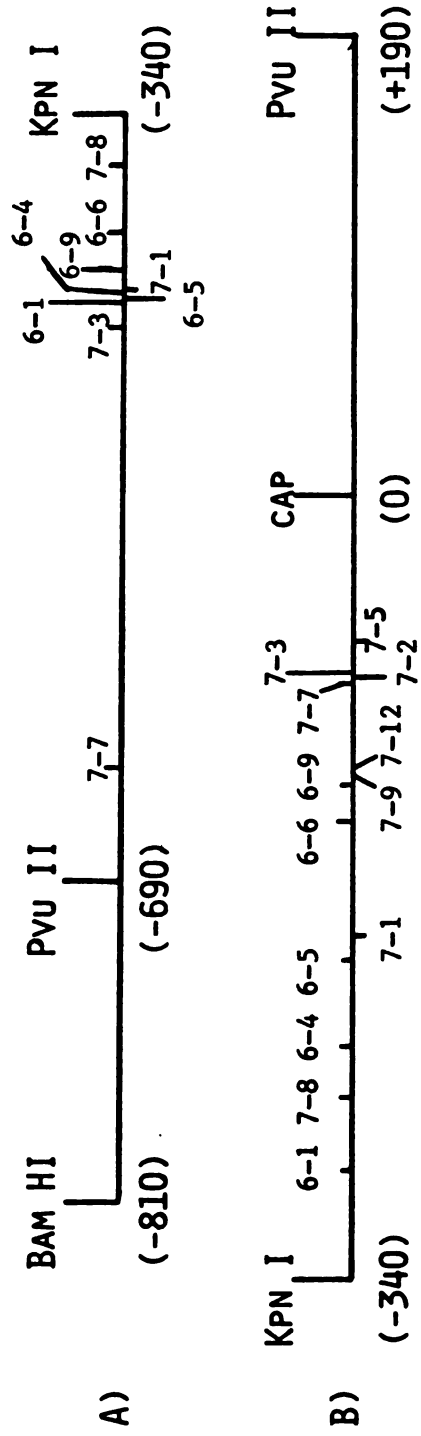


Figure 14

DISCUSSION

The experiments described here are designed to produce mutants of the putative chicken adult alpha globin promoters in vitro. The regions on which this study has concentrated lie within 120 bp upstream of the mRNA transcription start sites of the α^D and α^A genes.

The reasons for confining our work to these regions of the α^D and α^A genes stems from the detailed study of three other eukaryotic promoters, namely the rabbit β globin, herpes simplex virus thymidine kinase, and chicken ovalbumin promoters (Dierks et al., 1983; McKnight and Kingsbury, 1982; Knoll et al., 1983). In all three cases the DNA sequences necessary for initiation of transcription were defined by creation of specific mutations in vitro followed by characterization of the effects of these mutations in vivo. All three of these studies used heterologous host cells in which the genes were expressed in a nonregulated fashion, and in all three cases the DNA sequences necessary for constitutive expression were located within 110 bp of the mRNA start site. In fact, besides the TATA box at -25 to -30, the sequences which had the most dramatic effects on the level of expression were located between -60 and -110.

We have chosen to use the linker scanning method of McKnight and Kingsbury (1982) to generate the α^A and α^D promoter mutations. Initially, we have concentrated on construction of 5' deletions mutants which should allow us to roughly define the distal sequences necessary

for transcription. Other studies have found the 5' deletions are generally the most useful for the preliminary mapping of the promoter elements (for example see Dierks et al., 1983).

The results presented for the locations of the 5' and 3' deletion endpoints in the α^D promoter region in Figure 9 and 12 clearly demonstrate that we are now in a position to use these mutations to identify the promoter elements of this gene. The collection of 5' deletions nearly saturates the area between -120 and CAP with only a region between -60 and -40 lacking any mutants. We are presently constructing deletions using the slow isomer of Bal 31 and extending the deletion from the Xho I linker at -78 in one of the mutants. These new deletions should cover this -60 to -40 region. Also, twenty additional A10 clone DNAs which have deletions which map in the -60 to CAP area have been prepared and are being labeled at the Xho I and Pvu II sites for analysis on sequencing gels. It is expected that among the new deletions being constructed and these A10 clones, 5' deletions covering every base between the -120 and CAP sites of α^D will be available.

Although the collection of 3' α^D deletions presented in Figure 12 is not as extensive as the 5' deletion collection, it should be adequate for definition of the α^D promoter. In addition to the clones shown in Figure 12, twenty-four more B15 clone DNAs have been prepared for sequencing and are presently being analyzed. The linkers in all of these new clones have been mapped to the -120 to CAP region, and hopefully they will add significantly to the regions in α^D covered by the 3' deletions.

The reconstruction experiments, in which the appropriate 5' and 3' deletions will be recombined to generate an intact α^D gene with a cluster of point mutations, have not yet been initiated. Before proceeding with the reconstructions, in vivo expression experiments using various 5' deletions spaced throughout the -120 to CAP region will be conducted. These experiments should define which areas of the -120 to CAP region are most important in transcriptional control, and, thus, where the linker reconstruction experiments should be done. Studies on the expression of deletion mutant and normal control α globin genes are presently underway using both transient and long term expression systems in mammalian cells. Preliminary results with the intact α^D and α^A genes suggest that it may be necessary to introduce an enhancer element (for example, a retrovirus long terminal repeat) into the cloned DNAs to get measurable expression of the chicken α globin genes in these mammalian cell systems.

Construction and sequencing of the α^A gene deletions has not progressed as far as the α^D mutants. Several deletions extending into the putative α^A promoter region from the 5' end have been mapped. These are all located at least 70 bp upstream of CAP. One of these (K7-3) has been used to extend the deletions further into the α^A promoter region; however, these new mutants have not yet been mapped. Technical problems have precluded construction of the 3' α^A deletions, although recent experiments may have solved these problems.

Another aspect of the research described here which deserves some mention is the development of the strategies for using small culture (5 ml) for precise localization of the deletion endpoints. These methods use either a single sequencing reaction (Figure 5) or alternatively, precise fragment length determination on sequencing gels (Figure 7) to

locate the linker in the deletion mutants. A major requirement for either of these procedures is that a restriction enzyme site which can be used for labeling the DNA is located within approximately 250 bp of the deletion endpoints. This restriction site must lie on the opposite side of the region to be mutated from the restriction site used for the Bal 31 deletions. For example, in Figure 3 site X can be used to sequence (or size) the deletions which are initiated at site Y (Y1, Y2, Y3, etc.). Another requirement of the sequencing scheme, but not of the fragment length determination scheme, is that there is another restriction enzyme site available on the labelled DNA fragment which can be used for the second cut before DNA sequencing (this yields the uniquely labelled DNA end for sequencing). In fact the linker can probably be used as the second cut site. Since both fragments are utilized in the fragment length determination method. There is no need for a second restriction enzyme site.

With the deletion clones which have already been sequenced, the clones which have been constructed but not sequenced, and the methods developed for rapid localization of deletion endpoints, we may soon be ready to use these mutations to define the constitutive promoters of the chicken adult α globin genes.

REFERENCES

- Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) *Nuc. Acids. Res.* 8, 127-142.
- Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349-383.
- Dierks, P., Ooyen, A.V., Chochran, M.D., Dobkin, C., Reiser, J., Weissmann, C. (1983) *Cell* 32, 695-706.
- Dodgson, J.B. and Engel, J.D. (1983) *J. Biol. Chem.* 258, 4623-4629.
- Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeReil, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Bechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) *Cell* 21, 653-668.
- Engel, J.D. and Dodgson, J.B. (1980) *Proc. Natl. Acad. Sci. USA* 77, 2596-2600.
- Engel, J.D., Rusling, D.J., McCune, K.C., and Dodgson, J.B. (1983) *Proc. Natl. Acad. Sci. USA* 80, 1392-1396.
- Gillies, S.D., Morrison, S.L., Oi, V.T., Tonegawa, S. (1983) *Cell* 33, 717-728.
- Girvitz, S.C., Bacchetti, S., Rainblow, A.J., Graham, F.L. (1980) *Anal. Biochem.* 106, 492-496.
- Goldberg, M. (1979) Dissertation (Stanford Univ, Stanford, CA).
- Ish-Horowicz, D. and Burke (1981) *Nucl. Acids. Res.* 9, 2989-2998.
- Khoury, G. and Gruss, P. (1983) *Cell* 33, 313-314.
- Knoll, B.J., Zarucki-Schulz, T., Dean, D.C., O'Malley, B.W. (1983) *Nucl. Acids Res.* 11, 6733-6754.
- Larsen, A. and Weintraub, H. (1982) *Cell* 29, 609-622.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory, NY).
- Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
- McKnight, S.L. and Kingsbury R. (1982) *Science* 217, 316-324.
- Nevins, J.R. (1983) *Ann. Rev. Biochem.* 2, 441-466.
- Orkin, S.H. and Kazazian, H.H. (1984) *Ann. Rev. Genetics* 18, 131-171.

- Orkin, S.H. and Kazazian, H.H. (1984) *Ann. Rev. Genetics* 18, 131-171.
- Scott, R.W. and Tilghman, S.M. (1983) *Mol. Cell. Biol.* 3, 1295-1309.
- Smith, D.R. and Calvo, J.M. (1980) *Nucleic Acids Res.* 8, 2255-2274.
- Shortle D., DiMaio, D., and Nathans, D. (1981) *Ann. Rev. Genetics* 15, 265-294.

APPENDIX I

Selection of Aphidicolin-Resistant CHO Cells with Altered Levels of Ribonucleotide Reductase

Carol L. K. Sabourin,¹ Paul F. Bates,¹ Louis Glatzer,^{1,3} Chia-Cheng Chang,²
J. E. Trosko,² and John A. Boezi¹

¹Department of Biochemistry and ²Department of Pediatrics and Human Development,
Michigan State University, East Lansing, Michigan 48824

Received 10 September 1980—Final 25 November 1980

Abstract—Chinese hamster ovary cells were initially selected for resistance to aphidicolin at 0.3 µg/ml. Serial cultivation with aphidicolin at concentrations up to 5.0 µg/ml yielded a series of mutants with increasing resistance. The most resistant mutant isolated was 44 times more resistant to aphidicolin than the parental CHO. The α -polymerases, assayed in the cytoplasmic extracts of the mutants, did not increase in specific activity or differ from the parental CHO in their sensitivity to aphidicolin. When cultured in the presence of deoxythymidine, deoxyadenosine, and 1- β -D-arabinofuranosyl cytosine (araC) the mutants showed considerably more resistance to these inhibitors than did the parental CHO. The intracellular pools of all four deoxynucleoside triphosphates (dNTPs) in the mutants increased with increasing resistance to aphidicolin. The elevated dNTP pools in the mutant most resistant to aphidicolin appear to be the result of a 4- to 8-fold increase in the level of ribonucleotide reductase (2'-deoxyribonucleoside diphosphate:oxidized thioredoxin 2'-oxidoreductase, EC 1.17.4.1).

INTRODUCTION

Aphidicolin, a tetracyclic diterpenoid produced by the fungus *Cephalosporium aphidicola*, is cytotoxic to eukaryotic cells. The cytotoxicity results from a block in cell division (1, 2). Aphidicolin is a potent inhibitor of DNA synthesis, but it does not inhibit RNA or protein synthesis (1–3).

One of the target enzymes for aphidicolin appears to be DNA polymer-

³Permanent address: Department of Biology, University of Toledo, Toledo, Ohio 43606.

ase α . Purified α -polymerase, but not β -polymerase or γ -polymerase, is effectively inhibited by aphidicolin (1, 2, 4-6). With the purified α -polymerase aphidicolin functions as a competitive inhibitor of dCTP (7). Reactions catalyzed by the purified enzyme that do not use dCTP as a substrate, for example the poly(dA) · oligo(dT) directed polymerization of dTTP, are not effectively inhibited. The mode of inhibition by aphidicolin on DNA synthesis in isolated nuclei, however, appears to be somewhat different from that with the purified α -polymerase. Noncompetitive inhibition with dCTP, rather than competitive inhibition, is observed (8, 9). It is assumed that the α -polymerase is still the target enzyme but that in the nucleus the polymerase is part of a replication complex with somewhat different catalytic properties.

Mammalian cell mutants of DNA metabolism are needed for studies of the mechanism of replication, repair, and recombination. In this report aphidicolin was used to isolate resistant mutants of Chinese hamster ovary cells. Characterization of the mutants isolated indicates that the basis of resistance to aphidicolin is through oversynthesis of the dNTPs brought about by overproduction of ribonucleotide reductase.

MATERIALS AND METHODS

Reagents. Aphidicolin was supplied by the Developmental Therapeutics Program, Chemotherapy, National Cancer Institute. The aphidicolin was dissolved in dimethyl sulfoxide (DMSO) and sterilized by filtration using Millipore-type FG filters. The concentration of DMSO added to the culture was always less than 0.1%. Other reagents were from sources previously described (10) or from the usual commercial sources.

Cell Culture. The parental Chinese hamster ovary (CHO) cell line which is auxotrophic for proline was obtained from Dr. Louis Siminovitch, University of Toronto. The experiments described here were carried out with one subclone. The methods used in culturing were as previously described (11). The cells were routinely maintained in suspension culture at 37°C in MEM alpha medium, containing ribonucleosides and deoxyribonucleosides (α +) and 10% fetal calf serum (FCS).

The cells were mutagenized with ethyl methane sulfonate (EMS) at 150 μ g/ml. This concentration of EMS was sufficient to reduce the plating efficiency by about 50%. After 16 h the cells were washed twice with phosphate-buffered saline and resuspended in fresh medium. After allowing two days for recovery, the cells were plated in medium containing 0.3 μ g/ml aphidicolin. The colonies which appeared were picked with a Pasteur pipet and put into 24-well dishes.

The relative plating efficiency for Fig. 1 was determined by plating $2 \times$

10^2 – 2×10^4 cells per 60-mm dish or 2×10^4 – 5×10^5 cells per 100-mm dish at various concentrations of aphidicolin. The concentration of drug reducing the plating efficiency to 10% of the control (LD_{10}) was determined by plating 1×10^2 – 2×10^3 cells in the medium indicated in Table 1, in 35-mm dishes. Dishes were incubated at 37°C for 7–14 days and then stained with 1% methylene blue in 70% isopropanol.

The sensitivity to ultraviolet irradiation was determined by plating 2×10^2 – 8×10^3 cells per 100-mm dish. The cells were allowed 3.5 h for attachment. With the medium removed, the attached cells were exposed to ultraviolet light from a germicidal lamp which was positioned to deliver 2.0 J/m²/sec. After one week the surviving colonies were stained as previously described.

Deoxyribonucleoside Triphosphate Pool Measurements. The intracellular dNTPs were extracted from cell monolayers in the log phase of growth with 60% ice-cold methanol. The pools were measured using the defined copolymers poly d(I-C) · poly d(I-C) and poly d(A-T) · poly d(A-T) (P-L Biochemicals) and *E. coli* DNA polymerase I (Boehringer Mannheim) as previously described (12, 13). The amount of DNA per culture was measured in the methanol precipitate (14).

Assay of DNA Polymerization Reaction. The cytoplasmic extract was prepared and assayed for the α -polymerase as previously described (10). The concentration of DMSO added to each assay was 2.5%. Protein determinations were done using Coomassie brilliant blue G-250 as described by Bradford (25).

Assay of Ribonucleotide Reductase Activity. The ribonucleotide reductase assays were performed with crude extracts of parental CHO or mutant cells. These extracts were prepared by rupturing 0.035 g of cells per 1 ml of extraction buffer, 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), pH 7.2, and 2 mM dithiothreitol, with 20 strokes of a dounce homogenizer at 4°C. The homogenate was centrifuged at 40,000 rpm in a type-50 Beckman rotor for 1 h. The clear supernatant was decanted off. This fraction, called the crude extract, generally contained 10 mg protein per ml. Before assaying, the crude extract was extensively dialyzed against the extraction buffer to remove endogenous nucleotides.

Assays of the crude extract contained, in a final volume of 300 μ l, 50 mM HEPES, pH 7.2, 6 mM dithiothreitol, 6 mM ATP, 10 mM MgCl₂, 100 μ M [¹⁴C]CDP (6.7 mCi/mmol) plus the designated amount of protein. The samples were incubated for 1 h at 37°C, after which the reactions were stopped by boiling for 3 min. The reaction was determined to be linear with time. The precipitated protein was pelleted by centrifugation, and the samples were digested from the nucleotide to nucleoside forms with 25 μ g/assay *Crotalus atrox* venom phosphodiesterase (Sigma) and 25 μ g/assay

bacterial alkaline phosphatase (Sigma). The [^{14}C]deoxycytidine and [^{14}C]cytidine were assayed by the method of Steeper and Steuart (15). After digestion the samples were applied to $0.7 \times 4.5\text{-cm}$ Dowex 50-borate columns. The deoxycytidine was first eluted with 4.5 ml of H_2O and then cytidine was eluted with 5 ml of saturated $\text{K}_2\text{B}_4\text{O}_7$. The radioactivity in each peak was determined by scintillation counting in Formula-963 (New England Nuclear).

Electrophoresis. Electrophoretic analysis of extracts was carried out by the method of Laemmli (16), using an 8.75% acrylamide separating gel and 3% acrylamide stacking gel. Gels were stained with Coomassie blue. Densitometric scans were performed on a Ortec densitometer model 4310 at 540 nm, 0–0.3 OD full scale.

RESULTS

Selection of Aphidicolin-Resistant CHO Cells. In the presence of $0.3 \mu\text{g/ml}$ of aphidicolin the plating efficiency of wild-type CHO was reduced to 8×10^{-5} . Resistant colonies were selected when mutagenized cells were plated in $0.3 \mu\text{g/ml}$ aphidicolin. One of these mutants, L3, was used in a stepwise manner to isolate mutants that have increasingly greater resistance to the drug. Without additional mutagenesis L3-2, which was about 10 times more resistant to aphidicolin than the parental CHO, was isolated from L3; L3-3a, which was about 19 times more resistant than the parental CHO, was isolated from L3-2; and L3-5, which is about 44 times more resistant to the drug than the parental CHO, was isolated from L3-3a (Table 1). Figure 1 shows the effect of increasing concentrations of aphidicolin on the plating efficiency of parental CHO, L3, L3-2, L3-3a, and L3-5. Mutants with high resistance to aphidicolin ($>0.5 \mu\text{g/ml}$) could not be isolated in single-step isolations. The phenotype of the resistant clones remained stable over a period of many months when the mutants were maintained in the absence of aphidicolin.

Characterization of Aphidicolin-Resistant CHO Cells. The doubling time of L3, L3-2, L3-3a, and L3-5 increased compared to the parental CHO when grown in suspension culture in the absence of aphidicolin (Table 2). When L3-3a was grown in $2.5 \mu\text{g/ml}$ aphidicolin and L3-5 in $5.0 \mu\text{g/ml}$ aphidicolin, the concentrations at which they were selected, the doubling time increased twofold. In addition L3-3a and L3-5 in suspension culture doubled once and then stopped dividing, although the cultures retained their viability for at least two days, as determined by trypan blue staining. The doubling times in monolayer culture of L3-3a grown in $2.5 \mu\text{g/ml}$ aphidicolin and L3-5 grown in $5.0 \mu\text{g/ml}$ aphidicolin were the same as those for the suspension cultures but the monolayer cultures grew from a low density until the cultures were confluent.

Table 1. Properties of Parental CHO and Aphidicolin-Resistant Mutants^a

Cell line	Concentration of aphidicolin for selection (μg/ml)	Aphidicolin ^b		AraC ^c		Deoxythymidine ^d		Deoxyadenosine ^e	
		LD ₅₀ (μg/ml)	Relative resistance	LD ₅₀ (μg/ml)	Relative resistance	LD ₅₀ (mM)	Relative resistance	LD ₅₀ (mM)	Relative resistance
CHO		0.18	1.0	0.13	1.0	3.5	1.0	0.30	1.0
L3	0.3	0.45	2.5	0.25	1.9	10	2.8	4.0	13
L3-2	1.0	1.7	9.6	2.0	15	15	4.3	3.0	10
L3-3a	2.5	3.4	19	2.0	15	15	4.3	3.0	10
L3-5	5.0	8.0	44	2.0	15	12	3.4	4.0	13

^aLD₅₀, the concentration of drug reducing the plating efficiency to 10% of the control was determined as described in Materials and Methods. Relative resistance is expressed as the ratio of the LD₅₀ value for the resistant line to the LD₅₀ for the parental CHO.

^bα+ supplemented with 10% FCS was used.

^cα- supplemented with 10% dialyzed FCS was used.

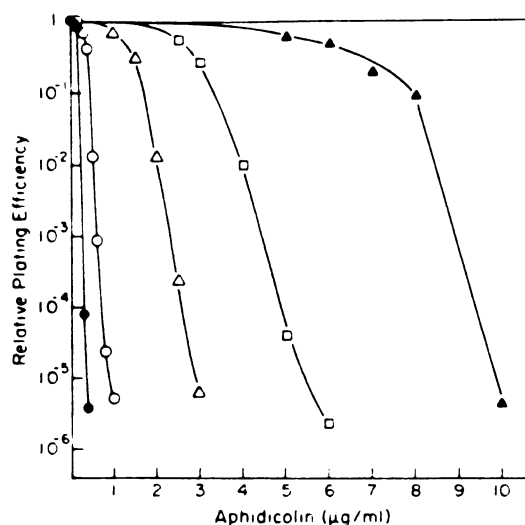


Fig. 1. Relative plating efficiency of parental CHO (●) and the aphidicolin-resistant mutants, L3 (○), L3-2 (Δ), L3-3a (□) and L3-5 (▲) with varying concentrations of aphidicolin

L3, L3-2, L3-3a, and L3-5 were examined to see if the basis of drug resistance was due to a change in the α -polymerase. Using cytoplasmic extracts of these cells, no differences in specific enzymatic activities of the α -polymerase or sensitivities of the α -polymerase to aphidicolin were observed compared to parental CHO enzyme (Table 3). Thus, the resistance to aphidicolin in the mutants is probably not due to a structurally altered α -polymerase or because of an overproduction of the α -polymerase.

The sensitivity to ultraviolet irradiation of L3 and L3-2 was examined.

Table 2. Effect of Aphidicolin on Doubling Time of Parental CHO and Aphidicolin-Resistant Mutants^a

Cell line	Doubling time (h) in aphidicolin			
	0 µg/ml	1.0 µg/ml	2.5 µg/ml	5.0 µg/ml
CHO	13	— ^b	—	—
L3	15	—	—	—
L3-2	17	20	—	—
L3-3a	27	36	62	—
L3-5	20	27	29	44

^aCells were grown in suspension culture in α + supplemented with 10% FCS. The number of cells/ml was determined using a Coulter counter model ZBI.

^b— indicates that the cells will not grow at this concentration of aphidicolin.

Table 3. Characterization of α -Polymerase Activity in Cytoplasmic Extract of Parental CHO and Aphidicolin-Resistant Mutants

Cell line	Specific enzymatic activity ^a (units/mg protein)	50% inhibition ^b (μ M)
CHO	9.4	6.5
L3	8.9	7.5
L3-2	9.3	6.0
L3-3a	16	7.8
L3-5	12	8.8

^aNanomoles of [³H]dTMP incorporated into activated calf thymus DNA/30 min/mg protein. The α -polymerase was assayed as described in Materials and Methods.

^bConcentration of aphidicolin which results in a 50% decrease in the picomoles of [³H]dTMP incorporated into activated calf thymus DNA/30 min in the absence of aphidicolin.

L3 showed no significant differences from parental CHO in the ultraviolet sensitivity in α + supplemented with 10% FCS. The results were similar for L3 and L3-2 when compared to parental CHO in MEM alpha medium lacking ribonucleosides and deoxyribonucleosides (α -), supplemented with 10% dialyzed FCS. Therefore the defect in the mutants which brings about resistance to aphidicolin is probably not related to the enzymes involved in DNA repair.

The effect of added araC on colony formation of the aphidicolin-resistant and the parental CHO was examined. The mutants showed greater resistance to araC than the parental CHO (Table 1). There are two known mechanisms by which a cell can become resistant to araC: a deficiency in the deoxycytidine kinase activity would not phosphorylate araC to its toxic form, araCTP, or an expanded pool of dCTP would confer resistance by diluting the araCTP thereby reducing the inhibition of the DNA polymerase (17).

To determine the mechanism whereby the mutants became resistant to araC, the effect of deoxythymidine and deoxyadenosine was investigated. The mutant cell lines were more resistant to both deoxynucleosides than the parental cell line (Table 1). Exogenous deoxythymidine and deoxyadenosine are converted to their respective triphosphates which in turn act allosterically to diminish the activity of CDP reductase. The supply of deoxycytidine nucleotides then becomes inadequate for DNA synthesis. The resistance to exogenous nucleosides demonstrated by the mutants could be related to modifications of the ribonucleotide reductase activity that would increase the supply of deoxycytidine nucleotides.

The above results, coupled with the knowledge that aphidicolin is a competitive inhibitor of dCTP with the purified α -polymerase (7), suggest a possible mechanism of resistance to aphidicolin involving enhanced synthesis of dCTP. The dNTP pools of L3-2, L3-3a, and L3-5 were measured and

Table 4. Deoxyribonucleoside Triphosphate Pools in Parental CHO and Aphidicolin-Resistant Mutants*

Cell line	Aphidicolin ($\mu\text{g/ml}$)	pmol/ μg DNA			
		dCTP	dTTP	dATP	dGTP
CHO		0.3	9.8	1.9	1.5
L3-2		2.1	7.5	6.6	1.2
L3-3a		16.6	9.4	26.1	4.8
L3-3a	2.5	53.2	36.1	152	15.1
L3-5		85.7	15.8	57.7	10.2
L3-5	5.0	186	69.9	163	36.8

*Cultures grown exponentially in α + supplemented with 10% FCS at the indicated concentration of aphidicolin and assayed as described in Materials and Methods. The values of each determination are the average of four separate determinations.

compared to those of the parental CHO in α + supplemented with 10% FCS. The results are given in Table 4. The pools of the four dNTPs increased with increasing resistance to aphidicolin. The dNTP pools of L3-3a and L3-5 grown in the presence of aphidicolin were larger than their pools when grown in the absence of aphidicolin. Pool expansion was also obtained for L3-3a when compared to parental CHO for cultures grown in α - supplemented with 10% dialyzed FCS.

Ribonucleotide Reductase Activity of Aphidicolin-Resistant CHO Cells. Since the increased pools of dNTPs observed in the aphidicolin-resistant mutants could result from qualitative or quantitative changes in the ribonucleotide reductase, crude extracts of L3-5 and parental CHO cells were assayed for CDP reductase activity. The L3-5 mutant showed a marked increase in CDP reduction compared to the parental CHO (Fig. 2). There was a 4- to 8-fold increase in the specific activity of the CDP reductase. Both the L3-5 and parental CHO extracts showed a nonlinear dependence of CDP reduction on protein concentration. This nonlinear protein dependence has previously been shown for the ribonucleotide reductase of mammalian systems (18). Also, there was a further increase in CDP reductase activity when the L3-5 cells were grown in the presence of aphidicolin. This increased activity parallels the increase in dNTP pool sizes when the cells were grown in the presence of aphidicolin. The increased CDP reductase activity in the crude extracts was stable when the cells were cultured for several months in the absence of aphidicolin.

Loss of allosteric regulation by dATP would result in an overproduction of all four dNTPs as is seen in the mutants. In order to determine whether the increased activity of the reductase in the mutant cells was due to loss of allosteric control, the inhibition of the CDP reductase by dATP was assayed. Both the L3-5 and parental CHO extracts were equally sensitive to dATP

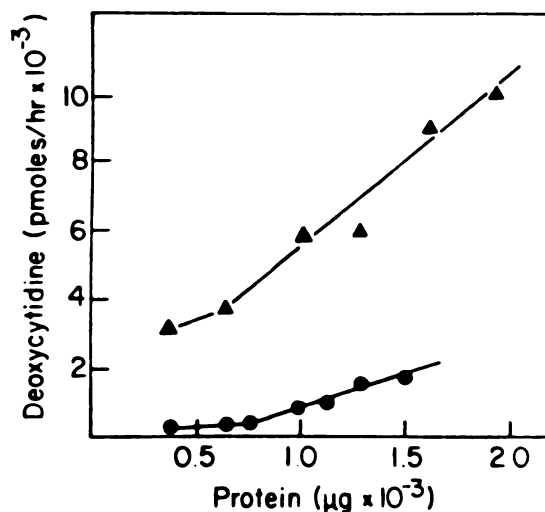


Fig. 2. CDP reductase activity in extracts of parental CHO (●) and L3-5 (▲). Varying amounts of each extract were assayed for CDP reductase activity as described in Materials and Methods. Each point represent the average of five separate experiments.

inhibition of CDP reduction with a 50% inhibition of 10 μM for both L3-5 and parental CHO extracts.

Since alterations in ribonucleotide reductase can apparently produce an aphidicolin-resistant phenotype, assays were run to determine if aphidicolin inhibits this enzyme. Even at 100 μM , fifteen times the value for 50% inhibition of the α -polymerase, the CDP reductase of the parental CHO showed no significant inhibition by aphidicolin (Table 5).

Polyacrylamide Gel Electrophoresis of Crude Extracts. From the results of the CDP reductase assays, it appeared that it may be possible to visualize the increased levels of ribonucleotide reductase by comparing L3-5 and parental CHO cell extracts on SDS polyacrylamide gel electrophoresis. The only significant differences in the gels were an increase in staining of the

Table 5. Effect of Aphidicolin on the CDP Reductase in Parental CHO

Addition to the assay	Specific activity ^a (pmol/h/mg protein)
100 μM aphidicolin in DMSO ^b	500
DMSO ^b	620
None	767

^aCDP reductase activity was determined as described in Materials and Methods.

^bThe final concentration of DMSO in the assay was 16.7%.

bands corresponding to polypeptides of molecular weight 55,000, 72,000, and 80,000 (Fig. 3). The values of 55,000 and 80,000 corresponded well with the subunit molecular weights of 55,000 and 84,000 for calf thymus ribonucleotide reductase (19). These two bands were found using several different preparations of the L3-5 cells. The 72,000-dalton polypeptide was not observed in all preparations, and may be a degradation product of a 76,000-dalton polypeptide. A densitometric scan of the gel indicated that the 76,000-dalton peak was smaller in the L3-5 extract than the corresponding peak in the parental CHO extract (Fig. 4). Integration of the 72,000 and 76,000 peaks in the L3-5 extract yielded a value equal to that of the 76,000 peak in parental CHO extract. The densitometric scan also showed that the 55,000 and 80,000 bands of the L3-5 extract were increased in approximately equal amounts, as would be expected for an increase in ribonucleotide reductase. These results suggest that the ribonucleotide reductase may be overproduced in the aphidicolin-resistant mutant.

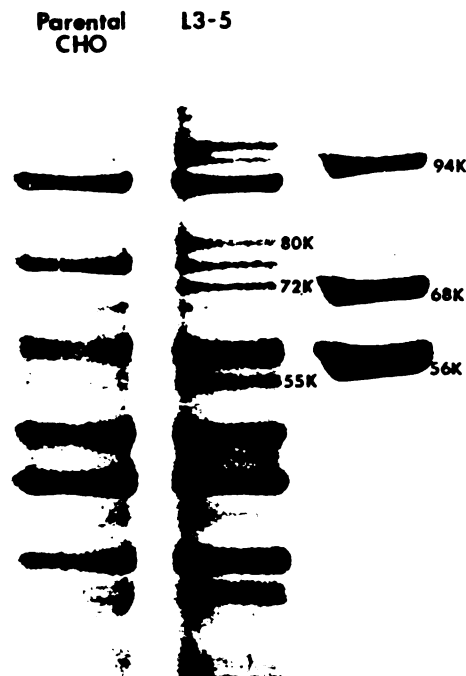


Fig. 3 SDS-polyacrylamide gel electrophoresis of parental CHO and L3-5 crude extracts prepared as described in Materials and Methods. The parental CHO and L3-5 lanes contain 30 μ g each of protein. Lane 3 contains 8 μ g each of the molecular weight standards phosphorylase A (94,000), bovine serum albumin (68,000), and catalase (56,000).

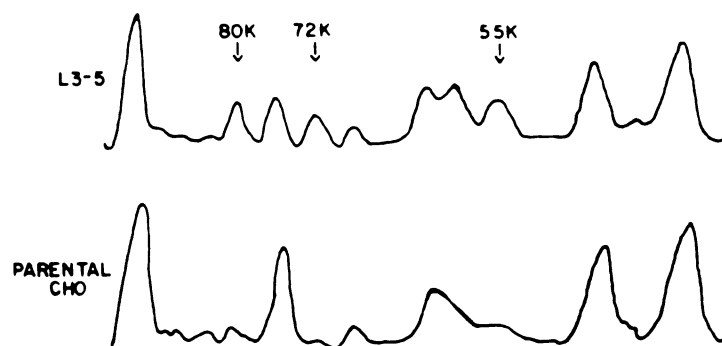


Fig. 4. Densitometric scan of a region of the SDS-polyacrylamide gel of parental CHO and L3-5 extracts.

DISCUSSION

This report describes the biological and biochemical characteristics of a series of CHO mutants which are resistant to aphidicolin. The resistance of these mutants does not appear to result from a structurally altered or overproduced α -polymerase. The resistance of the mutants seems instead to be due to an alteration in the levels of ribonucleotide reductase which results in the overproduction of the four dNTPs.

The aphidicolin-resistant mutants, when cultured in the presence of deoxythymidine, deoxyadenosine, and araC, showed considerably more resistance to these inhibitors than did the parental CHO. Deoxythymidine, deoxyadenosine, and araC in high concentrations all act to depress DNA synthesis. Deoxythymidine and deoxyadenosine are phosphorylated to the triphosphates which allosterically interact to inhibit the ribonucleotide reductase activity. The supply of deoxycytidine nucleotides then becomes inadequate for DNA synthesis. The large dCTP pool expansion which occurs in the mutants confers resistance to deoxythymidine and deoxyadenosine to the cell since the dCTP supply does not become a limiting factor for DNA synthesis. The additional dCTP produced also dilutes the araC nucleotide pool, thereby reducing the inhibition of the DNA polymerase by araCTP, the toxic form of araC.

The increased pools of the four dNTPs can be related to an increase in the ribonucleotide reductase activity. Two possible alterations in the ribonucleotide reductase which would account for the increased dNTP pools were investigated. Increasing the amount of ribonucleotide reductase in the cell is one mechanism by which the activity of the enzyme can increase. We found a 4- to 8-fold increase in the crude extract specific activity of the CDP reductase activity for L3-5 over the parental CHO. In addition, the electro-

phoretic analysis of the parental CHO and L3-5 extracts indicates that the two polypeptides of molecular weights similar to those reported for calf thymus ribonucleotide reductase are significantly overproduced in the L3-5 extracts.

Modification of the ribonucleotide reductase activity could also result if the allosteric control mechanism was disrupted. Desensitization of the enzyme to dATP, a general allosteric inhibitor of the ribonucleotide reductase, would cause increased activity and overproduction of all four dNTPs. If the mechanism for increased ribonucleotide reductase activity in the stepwise-selected mutants was a loss of allosteric inhibition by dATP, one would expect that L3-5, the mutant with the largest dNTP pools, would be the least sensitive to dATP compared to the parental CHO cells. We found no significant differences in the sensitivity to dATP of the L3-5 and parental CHO CDP reductases. Alternatively a structural alteration in the enzyme, not involving the allosteric sites, could also produce a more active enzyme. Such a structural alteration, affecting the intrinsic specific activity of the enzyme, would not be detected by the experiments conducted.

Meuth and Green (20) have demonstrated that mutants selected for resistance to araC also showed resistance to deoxythymidine and deoxyadenosine. These araC-resistant mutants were shown to have expanded dNTP pools (21). The alteration in ribonucleotide reductase activity showed a 4- to 10-fold higher level of enzyme specific activity but, unlike our aphidicolin-resistant mutants, also showed a partial desensitization of the enzyme to the allosteric negative effector dATP (20).

Enzyme levels of ribonucleotide reductase have been shown to be under strict cell cycle control (22). In mammalian cells the level of ribonucleotide reductase increases during the S phase of growth and falls during the G2 phase. If this cell cycle regulation were lost, the apparent effect in a nonsynchronized culture of cells would be an increase in the amount of ribonucleotide reductase present. Alternatively, the cells may retain cell cycle regulation of enzyme levels, but overproduce the enzyme during the S phase. Our studies do not distinguish between these two possible mechanisms for overproduction of ribonucleotide reductase.

The results for L3-5 in the CDP reductase assays coupled with the gel electrophoresis indicate that the aphidicolin-resistant mutants described overproduce ribonucleotide reductase. Each mutant isolated in a stepwise manner has increasingly higher resistance to aphidicolin, larger dNTP pools, and presumably has more reductase. The possibility of isolating mutants with an even greater overproduction of ribonucleotide reductase remains. Massive overproducers of this enzyme will be a great aid in studies characterizing the catalytic and structural properties of this enzyme. Our polyacrylamide gel

results suggest that two polypeptide chains are overproduced, lending credence to the hypothesis that the enzyme is composed of two polypeptide subunits. Further studies of these mutants may also answer questions concerning the control of biosynthesis of an enzyme composed of two different polypeptide chains.

The fact that mutants resistant to high concentrations of aphidicolin are only obtained by stepwise selection indicates that resistance to high aphidicolin concentrations could be due to overproduction of ribonucleotide reductase through gene amplification. Overproduction of ribonucleotide reductase could be linked to amplification of the genes programming the synthesis of two polypeptide chains. Amplification of genes coding for proteins with one polypeptide chain has been observed in methotrexate-resistant variants (23) and *N*-(phosphoacetyl)-L-aspartate-resistant variants (24). Highly resistant mutants could only be obtained in this stepwise manner. Mutants of mammalian cells that overproduce a protein composed of two different polypeptides could provide a system for answering many questions concerning the control of their synthesis.

ACKNOWLEDGMENTS

During the final phases of this work Dr. John A. Boezi died. John was both a dedicated scientist and educator. His enthusiasm and creativity were and still are an inspiration to those who knew him. This research was supported by grants from NIAID (AI-14357) to J.A.B., from NCI (CA-21104) to J.E.T., and from NIEHS (ES-01809) to C.C.C. Michigan Agricultural Experiment Station Journal Article Number 9625.

LITERATURE CITED

1. Ikegami, S., Taguchi, T., Ohashi, M., Oguro, M., Nagano, H., and Mano, Y. (1978). *Nature* **275**:458-460.
2. Longiaru, M., Ikeda, J.-E., Jarkovsky, Z., Horwitz, S.B., and Horwitz, M.S. (1979). *Nucleic Acids Res.* **6**:3369-3386.
3. Bucknall, R.A., Moores, H., Simms, R., and Hesp, B. (1973). *Antimicrob. Agents Chemother.* **4**:294-298.
4. Ohashi, M., Taguchi, T., and Ikegami, S. (1978). *Biochem. Biophys. Res. Commun.* **82**:1084-1090.
5. Pedrali-Noy, G., and Spadari, S. (1979). *Biochem. Biophys. Res. Commun.* **88**:1194-1202.
6. Wist, E., and Prydz, H. (1979). *Nucleic Acids Res.* **6**:1583-1590.
7. Oguro, M., Suzuki-Hori, C., Nagano, H., Mano, Y., and Ikegami, S. (1979). *Eur. J. Biochem.* **97**:603-607.
8. Habara, A., Kano, K., Nagano, H., Mano, Y., Ikegami, S., and Yamashita, J. (1980). *Biochem. Biophys. Res. Commun.* **92**:8-12.

9. Oguro, M., Shioda, M., Nagano, H., and Mano, Y. (1980). *Biochem. Biophys. Res Commun* **92**:13-19.
10. Sabourin, C.L.K., Reno, J.M., and Boczi, J.A. (1978). *Arch. Biochem. Biophys.* **187**:96-101.
11. Thompson, L.H., and Baker, R.M. (1975). In *Methods in Cell Biology*, Vol. 6, (ed.) Prescott, D.M., (Academic Press, New York), pp. 209-281.
12. Skoog, L. (1970). *Eur. J. Biochem.* **17**:202-208.
13. Lindberg, U., and Skoog, L. (1970). *Anal. Biochem.* **34**:152-160.
14. Giles, K.W., and Myers, A. (1965). *Nature* **206**:93.
15. Steeper, J.R., and Steuart, C.D. (1970). *Anal. Biochem.* **34**:123-130.
16. Laemmli, U.K. (1970). *Nature* **227**:680-685.
17. De Saint Vincent, B.R., and Buttin, G. (1979). *Somat. Cell Genet.* **5**:67-82.
18. Hopper, S. (1972). *J. Biol. Chem.* **247**:3336-3340.
19. Engstrom, Y., Eriksson, S., Thelander, L., and Akerman, M. (1979). *Biochemistry* **18**:2941-2948.
20. Meuth, M., and Green, H. (1974). *Cell* **3**:367-374.
21. Meuth, M., Aufreiter, E., and Reichard, P. (1976). *Eur. J. Biochem.* **71**:39-43.
22. Thelander, L., and Reichard, P. (1979). *Annu. Rev. Biochem.* **48**:133-158.
23. Alt, F.W., Kellems, R.E., Bertino, J.R., and Schimke, R.T. (1978). *J. Biol. Chem.* **253**:1357-1370.
24. Wahl, G.M., Padgett, R.A., and Stark, G.R. (1979). *J. Biol. Chem.* **254**:8679-8689.
25. Bradford, M. (1976). *Anal. Biochem.* **72**:248-254.

APPENDIX II

GENE 905

Double *cos* site vectors: simplified cosmid cloning *

(Genome libraries; chromosome walking; eukaryotic selectable marker; *Herpes virus*; thymidine kinase; HAT selection; retrovirus LTR)

Paul F. Bates and Robert A. Swift

Departments of Biochemistry and of Microbiology and Public Health, Michigan State University, East Lansing, MI 48824-1101 (U.S.A.) Tel. (517) 353-5024

(Received February 5th, 1983)

(Revision received August 11th, 1983)

(Accepted August 31st, 1983)

SUMMARY

A new vector for construction of cosmid libraries is described. Cosmid c2XB contains restriction sites for use in the insertion of foreign DNA and two λ *cos* sites separated by a blunt-end restriction site. The presence of two *cos* sites on a single plasmid eliminates the need to prepare two separate cosmid arms, and the internal blunt-end restriction site prevents cosmid concatemerization. Thus, a double restriction-enzyme digestion is sufficient to prepare the vector for subsequent ligation with DNA fragments which are dephosphorylated to prevent their self-ligation. The use of this vector system allows efficient cosmid cloning (1×10^5 colonies per μ g insert DNA) and eliminates background due to vector self-ligation. Furthermore, the procedure is so rapid as to eliminate the need to amplify cosmid libraries for storage and reuse. Also described is a cosmid vector for use in construction of cosmid libraries which are to be introduced into cultured eukaryotic cells. This vector contains the *Herpes simplex* virus thymidine kinase (HSV *tk*) gene as a selectable marker and a retroviral long terminal repeat (LTR) region as an enhancer sequence.

INTRODUCTION

Several recent studies of eukaryotic gene organization and/or expression demonstrate the need to clone large segments of DNA. In some cases the size

of a gene requires a large insert if the gene is to be isolated intact. Studies of gene families or of closely linked genes which are coordinately expressed may also require cloning of large DNA fragments. Furthermore, the study of extensive regions of a eukaryotic genome by serial isolation of a number of overlapping clones (i.e. genome walking) is facilitated by cloning fragments as large as possible.

Both plasmid and λ bacteriophage vectors are often inadequate for cloning large DNA fragments. The λ vectors available can only accommodate inserts up to about 20 kb. Theoretically, plasmids have the capacity to accommodate fragments of unlimited

* This is Journal Article No. 10755 from the Michigan Agricultural Experiment Station.

Abbreviations: bp, base pairs; c, cosmid; DTT, dithiothreitol; HAT, hypoxanthine-aminopterin-thymidine; HSV, *Herpes simplex* virus; kb, kilobase pairs; LB, Luria broth; LTR, long terminal repeat; tk, thymidine kinase; U, units.

size; however, the transformation efficiency of large plasmids is so low that their use in the construction of genomic libraries is not practical. Presently the primary system used to clone large DNA fragments efficiently is that of cosmid vectors (Collins and Hohn, 1978). These vectors are modified plasmids which contain a plasmid replicon, a selectable drug resistance marker, and the λ *cos* site. Cosmids can accept inserts of 30–45 kb, and the λ in vitro packaging system selects for large-size inserts and allows efficient introduction of the DNA into bacterial cells. However, because of several problems associated with cosmid cloning, this system has not yet been widely employed. Three main problems which have been encountered are vector concatemerization resulting in cosmids lacking inserted eukaryotic DNA, recombinational rearrangements caused by multiple inserts ligated into a single cosmid, and differential growth of cosmid clones causing misrepresentation of sequences in amplified cosmid libraries.

A cloning system which overcomes the first two of these problems has been described (Ish-Horowicz and Burke, 1981). Concatemerization of the vector was avoided by using a multiple enzyme digestion procedure involving four separate restriction enzyme digestion steps and two S1 nuclease treatments. The problem of insert rearrangement was minimized by treating the DNA to be cloned with phosphatase prior to its ligation to the vector; this prevents the insert DNAs from ligating to one another, and eliminates the need to purify insert DNA of the proper size (i.e. 30–45 kb). Because the Ish-Horowicz and Burke (1981) cloning scheme requires a considerable number of steps it becomes tedious when large numbers of cosmid libraries must be constructed. Furthermore, since differential growth of cosmids alters the representation of specific fragments in amplified cosmid libraries, it would be preferable to be able to rapidly construct a new cosmid library for each individual screening, thereby eliminating any amplification steps.

We describe here cosmid vectors which are easy to prepare and use, and avoid the problems of insert rearrangement and cosmid concatemers. This has been accomplished by constructing vectors containing two *cos* elements of λ separated by a blunt-end restriction enzyme site (see Figs. 1 and 3). The use of two *cos* sites on a single plasmid eliminates the need to separately prepare two cosmid arms, each of

which contains only a single cohesive end. After cleavage at the blunt-end restriction site, cosmid concatemerization does not occur during the subsequent ligation step when high ATP concentrations are employed (Ferretti and Sgaramella, 1981).

We have also constructed a cosmid vector for use in the transfection of cosmid libraries into eukaryotic cells. This vector uses the HSV *tk* gene as a selectable marker for cells which have incorporated cosmid DNA. A retroviral LTR is also contained on this vector as an enhancer sequence to aid in expression of the *tk* gene and adjacent cloned sequences.

MATERIALS AND METHODS

(a) Bacterial strains and enzymes

Escherichia coli HB101, $r_B^- m_B^- recA^-$ (Boyer and Roulland-Dussoix, 1969) was used for transformation and growth of most of the plasmids constructed. Those plasmids containing tandem repeats were grown in the strongly *recA^-* strain DH1 (Maniatis et al., 1982). Another *recA^-* strain, *E. coli* 1046 (Cami and Kourilsky, 1978) was used for the transduction and growth of the cosmid libraries. Restriction enzymes, T4 DNA ligase, T4 DNA polymerase, and Klenow fragment of *E. coli* DNA polymerase I were from New England Biolabs or Bethesda Research Labs. Calf intestinal phosphatase was from Boehringer. [α - ^{32} P]dCTP was from New England Nuclear.

(b) Nucleic acids

Preparative plasmid DNA extractions employed the alkaline lysis method (Birnboim and Doly, 1979). Analytical plasmid and cosmid DNA extractions (1.5 ml) followed the method of Ish-Horowicz and Burke (1981). Genomic DNA for use in construction of cosmid libraries was prepared from chicken erythrocytes by the method of Blin and Stafford (1976). DNA prepared by this method was generally greater than 150 kb in size as judged by electrophoresis in 0.2% agarose gels.

(c) Plasmid constructions

Details of the plasmid constructions are given in RESULTS AND DISCUSSION, sections a, d and e. In general, all restriction enzyme digestions were performed as recommended by the supplier. Cohesive ends were filled in using T4 DNA polymerase. DNA was incubated with T4 DNA polymerase (3 U/ μ g DNA) in the absence of added deoxynucleoside triphosphates for 5 min at 37°C. All four deoxynucleoside triphosphates were added to 100 μ M and the reaction was continued for 30 min at 37°C. A probe specific for the λ *cos* site was prepared by incubating Charon28 DNA (Rimm et al., 1980) with T4 polymerase for 70 min before adding [α -³²P]dCTP to 10 μ M and the other dNTPs to 100 μ M and incubating for an additional hour. The Charon28 DNA was then digested with *Bgl*II and *Ava*I to give a labeled λ end-specific probe. Partial end-filling reactions were performed with Klenow fragment of *E. coli* DNA polymerase I (0.1 U/ μ g DNA) and 200 μ M of the required deoxynucleotide used (e.g., 200 μ M dATP to partially fill *Eco*RI ends).

(d) Preparation of insert DNA for library construction

Insert DNA for construction of cosmid libraries was prepared by partial digestion with either *Mbo*I or *Eco*RI. 200 μ g of chicken erythrocyte DNA was digested at 37°C for 60 min in 0.5 ml with an amount of enzyme optimized to yield 30–50 kb fragments. The reaction was stopped by incubation at 70°C for 15 min. The pH was adjusted with 50 μ l of 2 M Tris·HCl pH 8.5 and calf intestinal phosphatase (2.5 U) was added. The DNA was dephosphorylated for 30 min at 37°C, and the phosphatase was inactivated at 70°C for 1.5 h. The DNA was then either precipitated with spermine (Hoopes and McClure, 1981) or used in ligations without further treatment.

(e) Preparation of vector DNA for cosmid library construction

c2XB DNA (20 μ g) was digested with *Sma*I (40 U) and *Bam*HI (20 U) or *Eco*RI (20 U) at 37°C for 4 h to ensure complete digestion. The enzymes were inactivated by incubation with 0.1% diethyl-

pyrocarbonate at 70°C for 15 min. The DNA was precipitated with spermine, washed with ethanol, and resuspended in 20 μ l of 10 mM Tris, pH 7.5, 0.1 mM EDTA.

(f) Ligations

Two different ligation conditions were used depending on whether blunt-end ligation was to be inhibited or not. To join cohesive and/or blunt ends, a buffer consisting of 66 mM Tris pH 7.5, 10 mM DTT, 5 mM MgCl₂ and 0.25 mM ATP was employed. Inhibition of blunt end ligation employed the same buffer except ATP was used at 5 mM.

(g) Packaging extracts, DNA packaging

Packaging extracts were prepared by protocol II of Maniatis et al. (1982) using strains BHB2688 and BHB2690 (Hohn, 1979). The efficiency of these extracts was 1×10^8 plaques/ μ g λ DNA. 5 μ l (1.5 μ g) of ligated DNA was packaged per reaction using a procedure detailed by Maniatis et al. (1982). Packaged cosmids were stored in phage buffer over CHCl₃. Cosmids were transduced into *E. coli* 1046. A saturated culture of 1046 grown in LB + 0.4% maltose was pelleted by centrifugation and resuspended in an equal volume of 10 mM MgSO₄. Up to 100 μ l of packaged cosmids was added to 100 μ l of the resuspended bacteria. Adsorption was allowed to take place for 20 min at 37°C, then 500 μ l of LB was added, and the incubation was continued for 30 min at 37°C. The mixture was spread on a 9-cm plate of L agar containing 75 μ g/ml ampicillin.

(h) Other methods

Standard published procedures were used for gel electrophoresis, transfer to nitrocellulose paper, nick translation, and hybridization (Maniatis et al., 1982). Elution of DNA fragments from agarose gels was by the method of Girvitz et al. (1980). Bacterial transformations were by a modification of the CaCl₂ procedure for *E. coli* χ 1776 (Maniatis et al., 1982; Hanahan, D., personal communication). Eukaryotic cell transformations were done by the calcium phosphate precipitation technique using *Ltk*⁻ *aprt*⁻ cells and the procedure described by Wigler et al. (1979), using HAT selection (Szybalska and Szybalski, 1962).

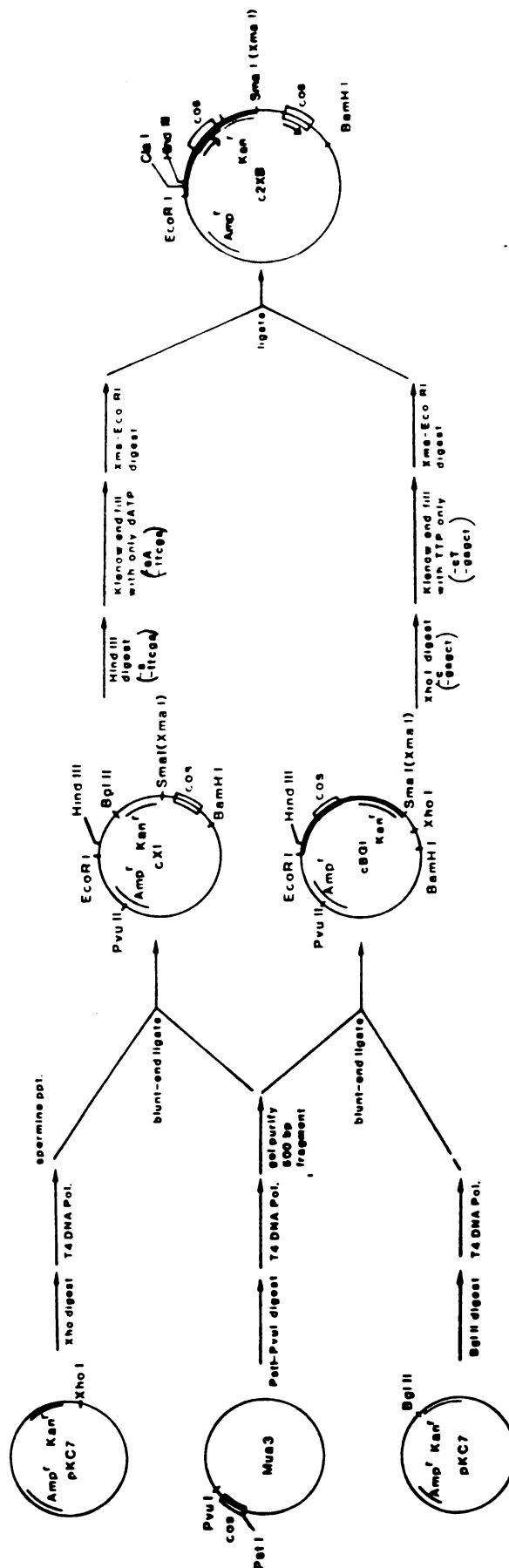


Fig. 1. Diagram illustrating steps used in construction of cosmid vector c2XB. Details of the construction are given in the RESULTS, section a. The heavy line indicates that portion of cBG1 which is cloned into the EcoRI and SmaI sites of cX1 to give c2XB. Arrows indicate the relative orientation of cos sites with respect to one another. The absolute orientation of the cos sites has not been determined.

RESULTS AND DISCUSSION

(a) Construction of vector c2XB

The construction of cosmid vector c2XB was performed as shown in Fig. 1. A 560-bp DNA fragment containing the λ *cos* site was isolated from the cosmid vector Mua3 (Meyerowitz et al., 1980) after digestion with *PvuI* + *PstI*. Ends of the fragment were filled in with T4 DNA polymerase, and it was inserted by blunt-end ligation into either the *XhoI* or *BglII* site of the plasmid pKC7 (Rao and Rogers, 1979) after these sites had also been blunt-ended with T4 polymerase. After transformation, colonies with plasmids containing the λ *cos* site were identified using ^{32} P-labeled Charon28 DNA (Rimm et al., 1980). DNA was prepared from several colonies with positive signals and digested with *EcoRI* + *SmaI* or *BamHI* + *SmaI*. In all the plasmids examined only a single *cos*-containing 560-bp fragment had been inserted (results not shown). These plasmids were named cX1 and cBG1 for those with the *cos* site at the *XhoI* or *BglII* site, respectively.

The cX1 and cBG1 plasmids were then combined to make the double *cos* site vector c2XB. A triple enzyme digestion scheme was designed to enrich for recombinants with two *cos* sites. cBG1 DNA was digested with *XhoI* and the sticky end was then partially filled with DNA polymerase I Klenow fragment using dTTP only. Partial end filling ensures that the *XhoI* cohesive ends will be unable to correctly base pair and will be unable to religate (see Fig. 1). This DNA was then digested with *XmaI* and *EcoRI*. This triple enzyme digest of cBG1 yielded three fragments but only the *XmaI*-*EcoRI* fragment containing the *cos* site had both cohesive ends intact and available for ligation. Similarly, cX1 DNA was digested with *HindIII*, partially filled with dATP, then digested with *EcoRI* and *XmaI*. The cX1 and cBG1 digestion products were mixed and ligated. After transformation of *E. coli* DH1, only plasmids with λ *cos* sites at both the *XhoI* and *BglII* sites were isolated.

For c2XB to be packaged *in vitro* the λ *cos* sites must be in the same orientation relative to one another, as shown by the arrows in Figs. 1 and 3. To test the relative orientation of the *cos* sites, plasmid DNA was isolated from ten colonies and tested for its utility for cosmid library construction. After

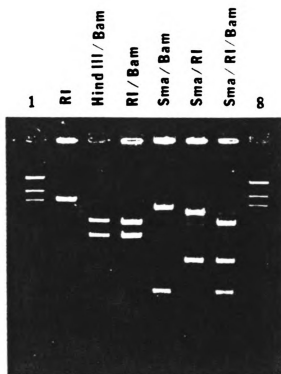


Fig. 2. Restriction enzyme mapping of c2XB cosmid vector. Electrophoresis of single, double, and triple restriction digests of c2XB are shown in lanes 2-7 (RI = *EcoRI*). Lanes 1 and 8 contain λ *HindIII* fragments as size markers. Electrophoresis was carried out on 6.5×9 cm gel of 0.7% agarose in Tris-acetate buffer for 1 h at 75 V.

digestion with *EcoRI* and *SmaI* the plasmid DNA was ligated to chicken DNA that had been partially digested with *EcoRI*, packaged *in vitro*, and transduced into *E. coli*. Four of the ten plasmids tested gave rise to ampicillin-resistant colonies, and so were judged to have the λ *cos* sites in the same relative orientation. DNA from one of these plasmids was prepared and the structure of c2XB shown was confirmed by restriction endonuclease mapping (Fig. 2). Single sites in c2XB, which can be used for cloning, are *BamHI*, *SalI*, *EcoRI*, and *HindIII*. The *BamHI* and *SalI* sites are the most useful sites for cosmid cloning, since these enzymes leave ends complementary to the 4-bp recognition sequence of enzymes *MboI* and *TaqI*, respectively.

(b) Construction of cosmid libraries with c2XB

A general method for construction of cosmid libraries with c2XB is shown in Fig. 3. A double

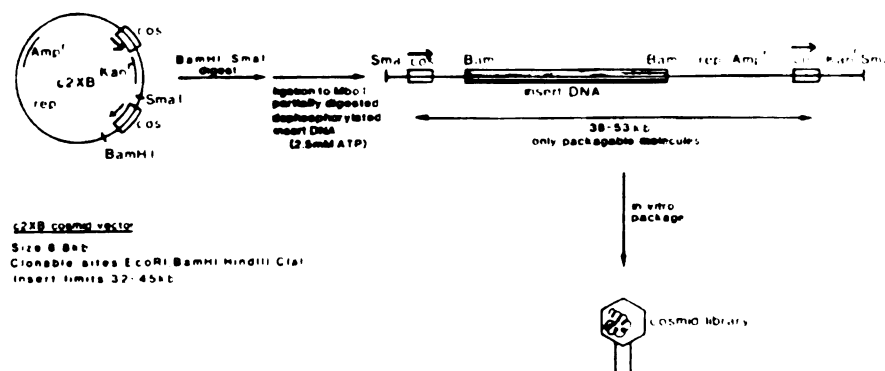


Fig. 3. Scheme for use of c2XB in construction of cosmid libraries. After digestion with *Bam*HI and *Sma*I, c2XB DNA is ligated to phosphatase-treated insert DNA partially digested with *Mbo*I. Note that *Bam*HI and *Mbo*I have the same cohesive GATC ends. A high ATP concentration is used during the ligation reaction to prevent the joining of the *Sma*I blunt ends (see RESULTS, section b)

digestion prepares the vector for cloning. Digestion of c2XB with *Sma*I creates a linearized molecule with blunt ends. Digestion with a second restriction enzyme, for example *Bam*HI, cleaves the vector into two fragments, each containing a λ *cos* site and each having one blunt and one sticky end. The vector DNA is then mixed in a 2:1 vector to insert mass ratio with partially digested (*Mbo*I), dephosphorylated insert DNA. The DNA is ligated at a DNA concentration of 300 μ g/ml with a high ATP concentration to inhibit concatemerization of the vector by blunt end ligation at the *Sma*I end. It has recently been demonstrated that an ATP concentration of 2.5 mM effectively inhibits blunt-end ligation but does not affect cohesive end ligation (Ferretti and Sgaramella, 1981). After ligation, the DNA is packaged *in vitro*, transduced into *E. coli* 1046, and ampicillin-resistant colonies are selected.

The cloning efficiency obtained with this vector is routinely 1×10^5 colonies per μ g insert DNA. This efficiency is slightly lower than that reported for some other cosmid cloning systems (Ish-Horowicz and Burke, 1981; Hohn and Collins, 1980), but this is probably due to the lower efficiency of our packaging extracts (1×10^8 plaques/ μ g λ DNA vs. 5×10^8 plaques/ μ g λ DNA obtained by others). With an average insert size of 37 kb (see below) only 1.2×10^5 cosmid clones would be required to have a 99% probability of cloning any sequence in the chicken genome. In this system only 1.2 μ g of partially digested insert DNA would be needed to generate a complete chicken library. Since the insert can be phosphatase treated, and need not be size

fractionated to prevent insertion of multiple fragments, both insert and vector DNA can be prepared for library construction in one day. Because libraries can be constructed so easily there is no need for amplification prior to screening. Thus, problems due to unequal growth of cosmid clones during amplification, causing unequal sequence representation in the library, are avoided.

(c) Characterization of cosmids

Independent colonies from a cosmid library constructed using c2XB and chicken DNA partially digested with *Eco*RI were picked randomly for further analysis. DNA prepared from these colonies was digested with *Eco*RI and analyzed by gel electrophoresis (Fig. 4). Insert sizes were determined by summing the bands. The average size was 37.6 ± 7.5 kb ($n = 12$). As expected, all five colonies examined showed different restriction patterns of their inserted chicken DNA (Fig. 4A), but only a single 5.1-kb vector band. The latter implies that a single copy of the vector is contained in each of the cosmids.

We noticed that the average insert size is smaller than expected. The packaged vector occupies only 5.1 kb of DNA (6.8-kb c2XB minus 1.7-kb fragment between *cos* sites lost during packaging; Fig. 3), and we expect that fragments up to 45 kb should be inserted. However, the average insert size obtained was only 37.6 kb. Similar results have been noted previously by others with the vectors pJB8 and Homer I, with average insert sizes of 38.8 kb and

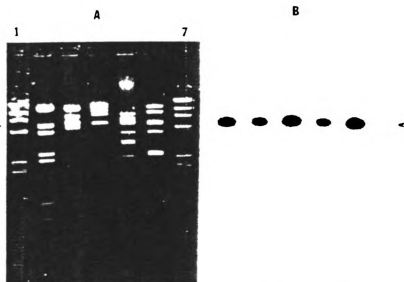


Fig. 4. *EcoRI* digests of cosmid DNA clones. A cosmid library of *EcoRI* partially digested chicken DNA was prepared. DNA was prepared from several randomly picked clones and digested with *EcoRI* and run on a 0.8% agarose gel. Lanes 1 and 7 contain λ *HindIII* fragments used as size markers. Panel A: Ethidium bromide-stained gel. Panel B: Autoradiograph of Southern transfer of gel, hybridized with ^{32} P-labeled pKC7 DNA (see Fig. 1). Arrowheads indicate 5.1-kb vector bands.

36 kb, respectively (Ish-Horowitz and Burke, 1981; Chia et al., 1982). Possibly the smaller insert sizes obtained reflect the dearth of larger fragments resulting from our partial digestion procedure, since Meyerowitz et al. (1980) have obtained inserts of 45.5 kb average size using sheared DNA. Alternatively, the *in vitro* packaging system used may be the cause of the smaller inserts. The same packaging system was employed in all three cases where small inserts were obtained, whereas Meyerowitz et al. (1980) used the system described by Blattner et al. (1978).

With other cosmid cloning systems, background colonies containing plasmids with insert DNA resulting from vector concatamerization are often a problem. The c2XB vector provides a simple system for checking vector background. As can be seen in Fig. 3, the *Kan'* gene should not be packaged. However, if the vector and recombinant have been ligated at the *SmaI* site, cosmids containing more than one vector molecule and, therefore, a *Kan'* gene, will be packaged. No kanamycin-resistant colonies have been detected in any of the libraries examined thus far. Other experiments in which no insert DNA is added to the ligation reaction also suggests that under the ligation conditions used vector concatamerization is not a problem. Ligation of the vector alone

produced no ampicillin- or kanamycin-resistant colonies after packaging and transduction (results not shown).

(d) Construction of c2RB

We have constructed a derivative of c2XB for use

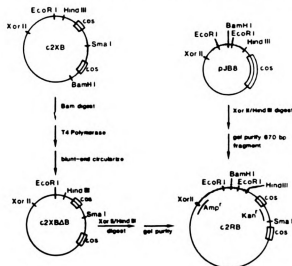
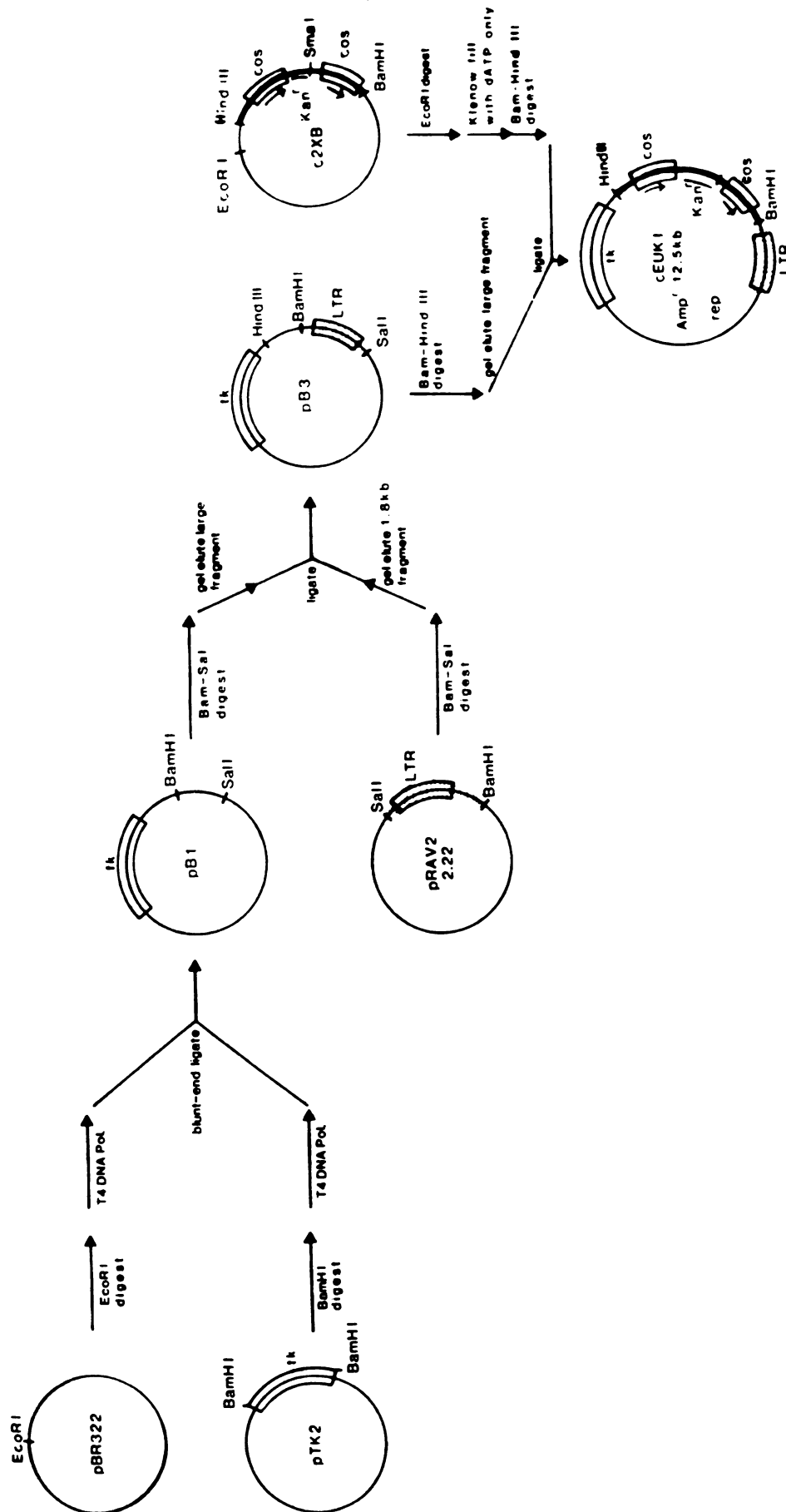


Fig. 5. Construction of the c2RB "walking" vector. The stippled segment indicates that portion of pJB8 cloned into c2XB Δ B to give c2RB. Details of the construction are given in RESULTS, section d.



in the preparation of libraries to be used for genomic walking. This vector, c2RB, is based upon the previously described cosmid vector pJB8 (Ish-Horowitz and Burke, 1981). In c2RB the *Bam*HI site has been moved so that it is flanked by *Eco*RI sites. *Eco*RI digestion cleanly excises DNA inserted at the *Bam*HI site, thus allowing rapid isolation of the inserted DNA for rescreening the cosmid library to isolate overlapping clones.

c2RB was constructed as shown in Fig. 5. Briefly, the *Bam*HI site of c2XB was eliminated by filling with T4 DNA polymerase followed by blunt-end ligation and transformation to isolate the recircularized plasmid. The resulting plasmid (c2XBΔB) was digested with *Xor*II + *Hind*III, and the large fragment was purified by gel electrophoresis. Similarly, pJB8 was digested with *Xor*II + *Hind*III, and the smaller fragment containing the *Bam*HI site was purified. After ligation and transformation of *E. coli* 1046, ampicillin- and kanamycin-resistant colonies were isolated, and the structure of c2RB was confirmed by restriction mapping (not shown). c2RB is the same size and contains the same restriction sites for library construction as c2XB. In parallel experiments with c2XB the same efficiency of 1×10^5 colonies/ μ g insert DNA was obtained with c2RB.

(e) Cosmid vector for use with eukaryotic cells

cEUK is a vector designed for use in construction of libraries of cosmids which will be introduced into eukaryotic cells. As shown in Fig. 6, it contains the HSV *tk* gene as a selectable marker in eukaryotic cells. Also, a retroviral LTR has been included as an "enhancer" sequence (Payne et al., 1982). The *tk* gene was isolated from the plasmid pTK2 (Hanahan et al., 1980) on a 3.6-kb *Bam*HI fragment, made blunt-ended with T4 DNA polymerase and inserted at the *Eco*RI site of pBR322 (after *Eco*RI cleavage and filling in the resultant cohesive ends). The resulting plasmid, pB1, with a single *Bam*HI site was digested with *Bam*HI and *Sal*I and the large fragment was isolated on a gel. This fragment was then ligated to a 1.8-kb LTR-containing *Bam*HI-*Sal*I fragment isolated from the Rous-associated-virus plasmid pRAV2 2.22 (Ju et al., 1980) creating the plasmid pB3. pB3 DNA was digested with *Bam*HI + *Hind*III, and the large fragment was iso-

lated by gel electrophoresis and ligated with the 2.6-kb *Bam*HI-*Hind*III fragment of c2XB. This 2.6-kb fragment contains the two λ *cos* sites and the Kan^r gene. The c2XB DNA was first digested with *Eco*RI and partially end-filled with dATP to enhance the recovery of the desired recombinant plasmid, cEUK. After transformation of *E. coli*, ampicillin- and kanamycin-resistant clones were selected and the structure shown for cEUK was confirmed by restriction enzyme mapping (not shown).

The ability of cEUK to transform tk⁻ eukaryotic cells to a tk⁺ phenotype was tested using the calcium phosphate precipitation technique and HAT selection (see MATERIALS AND METHODS, section h). Supercoiled cEUK DNA transformed mouse Ltk⁻aprt⁻ cells with a frequency of 1.5×10^2 colonies per 10^6 cells per μ g DNA. The enhancement effect of the LTR was evaluated by comparing the transformation efficiency of pB3 and pB1 DNA. Approximately a two-fold increase in efficiency was seen in the plasmid with the LTR (i.e., pB3) vs. the one without an LTR.

ACKNOWLEDGEMENTS

We wish to thank Drs. Hsing-Jien Kung and Jerry Dodgson for support for this project. We also thank Dr. Ed Fritsch for discussions of and interest in this work. We thank Dr. Gene Smith for the pRAV2 2.22 plasmid. This research was supported by the grants (to Dr. Dodgson) from the MSU Foundation, the Leukemia Research Foundation (ORD 28872), and a Biomedical Research Support Grant (No. 2-507 RR07049-15) from the Division of Research Resources, N.I.H.

REFERENCES

- Birnboim, H.C. and Doly, J.: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7 (1979) 1513-1523.
- Blattner, F.R., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Richards, J.E., Slightom, J.L., Tucker, P.W. and Smithies, O.: Cloning human fetal α -globin and mouse α -type globin DNA: preparation and screening of shotgun collections. *Science* 202 (1978) 1279-1284.

- Blin, N. and Stafford, D.W.: Isolation of high-molecular weight DNA. *Nucl. Acids Res.* 3 (1976) 2303-2310.
- Boyer, H.W. and Roulland-Dussoix, D.: A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* 41 (1969) 459-472.
- Cami, B. and Kourilsky, P.: Screening of cloned recombinant DNA in bacteria by in situ colony hybridization. *Nucl. Acids Res.* 5 (1978) 2381-2390.
- Chia, W., Scott, M.R.D. and Rigby, P.W.J.: The construction of cosmid libraries of eukaryotic DNA using the Homer series of vectors. *Nucl. Acids Res.* 10 (1982) 2503-2519.
- Collins, J. and Hohn, B.: Cosmids: A type of plasmid gene-cloning vector that is packagable in vitro in bacteriophage λ heads. *Proc. Natl. Acad. Sci. USA* 75 (1978) 4242-4246.
- Ferretti, L. and Sgarbetta, V.: Specific and reversible inhibition of the blunt end joining activity of the T4 DNA ligase. *Nucl. Acids Res.* 9 (1981) 3695-3705.
- Girvitz, S.C., Bacchetti, S., Rainbow, A.J. and Graham, F.L.: A rapid and efficient procedure for the purification of DNA from agarose gels. *Anal. Biochem.* 106 (1980) 492-496.
- Hanahan, D., Lance, D., Lipsich, L., Wigler, M. and Botchan, M.: Characteristics of an SV40-plasmid recombinant and its movement into and out of the genome of a murine cell. *Cell* 21 (1980) 127-139.
- Hohn, B.: In vitro packaging of λ and cosmid DNA, in Wu, R. (Ed.), *Recombinant DNA, Methods in Enzymology*, Vol. 68. Academic Press, New York, 1979, pp. 299-309.
- Hoopes, B.C. and McClure, W.R.: Studies on the selectivity of DNA precipitation by spermine. *Nucl. Acids Res.* 9 (1981) 5493-5504.
- Ish-Horowitz, D. and Burke, J.F.: Rapid and efficient cosmid cloning. *Nucl. Acids Res.* 9 (1981) 2989-2998.
- Ju, G., Boone, L. and Skalka, A.M.: Isolation and characterization of recombinant DNA clones of avian retroviruses: size heterogeneity and instability of the direct repeat. *J. Virol.* (1980) 1026-1033.
- Maniatis, T., Fritsch, E.F. and Sambrook, J.: *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1982.
- Meyerowitz, E.M., Guild, G.M., Prestidge, L.S. and Hogness, D.S.: A new cosmid vector and its use. *Gene* 11 (1980) 271-282.
- Payne, G.S., Bishop, J.M. and Varmus, H.E.: Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. *Nature* 295 (1982) 209-214.
- Rao, R.N. and Rogers, S.G.: Plasmid pKC7: A vector containing ten restriction endonuclease sites suitable for cloning DNA segments. *Gene* 7 (1979) 79-82.
- Rimm, D.L., Horness, D., Kucera, J. and Blattner, F.R.: Construction of coliphage lambda Charon vectors with *Bam*HI cloning sites. *Gene* 12 (1980) 301-309.
- Szybalska, E.H. and Szybalski, W.: Genetics of human cell lines. IV. DNA-mediated heritable transformation of a biochemical trait. *Proc. Natl. Acad. Sci. USA* 48 (1962) 2026-2034.
- Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G. and Chasin, L.: DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. *Proc. Natl. Acad. Sci. USA* 76 (1979) 1373-1376.

Communicated by M. Lieb.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03082 8416