# Structure and ~~expression of the~~ characterization of ~~rat~~ cytochrome $b_5$ genes

by Jacqueline Susan Rigby

A thesis submitted in partial
fulfilment of the requirements for
the degree of Doctor of Philosophy
in the University of London.

Department of Biochemistry
University College, London.

September 1990

ABSTRACT

Cytochrome $b_5$ is an amphipathic protein with a prosthetic haem group. It is found in various tissues such as liver, kidney, heart and brain as well as in erythrocytes. Liver microsomal cytochrome $b_5$ is 133 amino acids in length. The carboxy-terminal 36 amino acids constitute a hydrophobic tail and embed the protein in the membrane. Microsomal cytochrome $b_5$ functions in the cytochrome P450 mono-oxygenase system, acting to metabolise both exogenous and endogenous substances. Erythrocyte cytochrome $b_5$ is truncated after amino acid 97 and the protein is cytosolic. It functions in erythrocytes to reduce methaemoglobin. An absence of functional cytochrome $b_5$ in erythrocytes has been shown to lead to the medical disorder, methaemoglobinaemia.

Within a species, amino acids 1 to 96 are identical between microsomal and erythrocyte cytochrome $b_5$ - only amino acid 97 differs, in all species except bovine. It was not known whether there were one or two genes coding for the two forms of cytochrome $b_5$. If there was only one gene, the primary transcript may undergo differential RNA splicing to produce two different mature cytochrome $b_5$ mRNAs.

Rat cytochrome $b_5$ cDNA probes were used to screen a rat genomic library. Six different genomic clones were isolated which contained regions homologous to cytochrome $b_5$. These clones were analysed by restriction mapping, hybridization studies and DNA sequencing. Two of the genomic clones contained sequence for the same rat cytochrome $b_5$ processed, truncated pseudogene. From analysis of the other genomic clones, the functional microsomal cytochrome $b_5$ gene is more than 23 kb in length and consists of at least five exons. Only four exons from the functional rat cytochrome $b_5$ gene were isolated - the exon or exons containing amino acids 1-42 inclusive were not located. Due to this fact and the complicated pattern of hybridizing bands on a genomic Southern blot, it is still not clear whether there are one or two cytochrome $b_5$ genes.

To my mum, my sister and Nick

ACKNOWLEDGEMENTS

I would like to thank Liz Shephard for all her help, the time she spent on discussing my work over three years and reading this thesis. The encouragement and suggestions from Liz, Irvine Johnston and Ian Phillips were greatly appreciated. In my first few months practical training and guidance was, to a great extent, provided by Alan Ashworth. I am extremely grateful for his interest in my work. My friends and colleagues Lesley Forrest, Chris Griffith and Peter Bull provided both fun and invaluable support through difficult times, especially in the final year. Greatest thanks go to Nick Baker who has listened to my problems over the past two and a half years and allowed me the use of a word processor. He has worked extremely hard, typing some sections for me, as well as formatting the whole document. Thanks also go to Chris Taylorson and Peter Swann for the synthesis of oligonucleotides. Finally I would like to acknowledge the Medical Research Council for the provision of a grant to enable this work to be carried out.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

| | | |
|---|---|---|
| A | - | adenine or absorbance (as appropriate) |
| Amp | - | ampicillin |
| ATP | - | adenine triphosphate |
| $\alpha\alpha$ | - | amino acid |
| BCIG | - | 5-bromo-4-chloro-3-indoyl-beta galactoside |
| bp | - | base pairs |
| BSA | - | bovine serum albumin |
| C | - | cytosine |
| cDNA | - | complementary deoxyribonucleic acid |
| Ci | - | Curie |
| µCi | - | microCurie |
| cfu | - | colony forming units |
| cm | - | centimetres |
| CoA | - | coenzyme A |
| cpm | - | counts per minute |
| CTP | - | cytosine triphosphate |
| CGRP | - | calcitonin gene related peptide |
| C-terminal | - | carboxy-terminal |
| cytochrome $b_5$ reductase | - | NADH-cytochrome $b_5$ reductase |
| cytochrome P450 reductase | - | NADPH-cytochrome P450 reductase |
| °C | - | degrees Celsius |
| dATP | - | deoxyadenosinetriphosphate |
| dCTP | - | deoxycytosinetriphosphate |
| ddATP | - | dideoxyadenosinetriphosphate |
| ddCTP | - | dideoxycytosinetriphosphate |
| ddGTP | - | dideoxyguanosinetriphosphate |
| ddTTP | - | dideoxythymidinetriphosphate |
| DEAE | - | diethylaminoethyl |
| dGTP | - | deoxyguanosinetriphosphate |
| DMSO | - | dimethylsulphoxide |
| DNA | - | deoxyribonucleic acid |
| DNase | - | deoxyribonuclease |
| DTT | - | dithiothreitol |
| dTTP | - | deoxythymidinetriphosphate |
| E.coli | - | Escherichia coli |
| EDTA | - | ethylenediamine tetra-acetic acid (disodium salt) |
| FAD | - | flavin adenine dinucleotide |
| Fig. | - | figure |

| | | |
|---|---|---|
| FMN | - | flavin mononucleotide |
| g | - | gram or gravity (as appropriate) |
| µg | - | microgram |
| G | - | guanine |
| GTP | - | guanosine triphosphate |
| HEPES | - | N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid |
| Ig | - | immunoglobulin |
| IPTG | - | isopropyl-beta-D-thio-galactopyranoside |
| kb | - | kilobase |
| kD | - | kilodalton |
| µl | - | microlitre |
| M | - | molar |
| µm | - | micrometre |
| µM | - | micromolar |
| mA | - | milliamp |
| MES | - | 2-[N-Morpholino] ethanesulphonic acid |
| mg | - | milligram |
| ml | - | millilitre |
| mm | - | millimetre |
| mM | - | millimolar |
| mRNA | - | messenger ribonucleic acid |
| NAD | - | nicotinamide adenine dinucleotide |
| NADH | - | reduced nicotinamide adenine dinucleotide |
| NADPH | - | nicotinamide adenine dinucleotide phosphate |
| ng | - | nanogram |
| N-terminal | - | amino-terminal |
| OD | - | optical density |
| OLB | - | oligo labelling buffer |
| $^{32}P$ | - | phosphorus thirty-two |
| pfu | - | plaque forming unit |
| phage | - | bacteriophage |
| pmol | - | picomole |
| RF | - | replicative form |
| RNA | - | ribonucleic acid |
| RNase | - | ribonuclease |
| $^{35}S$ | - | sulphur thirty-five |
| SDS | - | sodium dodecyl sulphate |
| SSC | - | standard sodium citrate |
| SSPE | - | standard saline phosphate EDTA |
| T | - | thymine |

| | | |
|---|---|---|
| TBE | - | tris borate EDTA |
| $T_d$ | - | temperature of dissociation |
| TEMED | - | N,N,N',N'-tetramethylethylenediamine |
| TLC | - | thin layer chromatography |
| Tris-Cl | - | Tris (hydroxymethyl) aminoethane, pH with hydrochloric acid |
| TTP | - | thymidine triphosphate |
| uv | - | ultra violet |
| V | - | volt |
| W | - | watt |
| w/v | - | weight/volume |

# 1 INTRODUCTION

## 1.1 Forms of cytochrome $b_5$

Cytochrome $b_5$ is a type b cytochrome, containing the same prosthetic haem group, protoporphyrin IX, as haemoglobin and myoglobin (Strittmatter and Velick, 1956). It is a ubiquitous protein and is found in several subcellular organelles, such as the endoplasmic reticulum (Ito and Sato, 1968), Golgi bodies (Bergeron et al., 1973), nuclear membranes (Berezney and Crane, 1971; Franke et al., 1970) and the outer membrane of mitochondria (Sottocasa et al., 1967).

### 1.1.1 Endoplasmic reticulum cytochrome $b_5$

Microsomal cytochrome $b_5$ is an amphipathic protein (Spatz and Strittmatter, 1971) with a molecular weight of 18100 Daltons (Beck von Bodman et al., 1986). It is composed of a hydrophobic tail of 36 amino acid residues (Fleming et al., 1978), located at the carboxy-terminus (Ozols, 1972) which anchors the protein into membranes and a hydrophilic portion containing the haem binding domain (Ito and Sato, 1968).

Amino acid sequencing of microsomal cytochrome $b_5$ proteins from porcine (Ozols, 1974; Ozols and Gerard, 1977a; Abe et al., 1985), bovine (Ozols, 1975; Fleming et al., 1978), horse (Ozols et al., 1976; Ozols and Gerard, 1977b), rat (Ozols and Heinemann, 1982) and rabbit liver (Ozols, 1970; Kondo et al., 1979), revealed that the protein is 133 amino acids long and that the primary structure is highly conserved between species. Most of the interspecies heterogeneity lies at the amino- and carboxy- terminals and at the junction of the hydrophilic and hydrophobic domains of the liver cytochrome $b_5$.

### 1.1.2 Erythrocyte cytochrome $b_5$

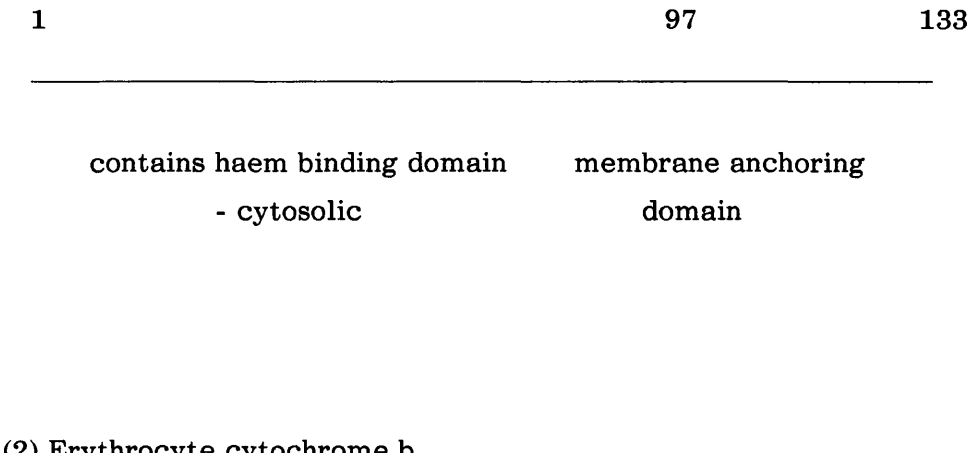
As shown by electron microscopy, the mature mammalian erythrocyte is without membranous structures, including endoplasmic reticulum (Beams and Kessel, 1966). It was therefore assumed that microsomal redox proteins would not be found in erythrocytes. However, a haemoprotein with properties similar to the cytochrome $b_5$ purified from solubilized microsomal membranes was found in the supernatant of red-cell haemolysates (Passon et al., 1972). A comparison of the cytochrome $b_5$ protein isolated from either trypsin or detergent solubilized microsomal membranes with the cytochrome $b_5$ located in erythrocytes, revealed that the spectral properties and molecular weights of the two proteins were similar (Passon et al., 1972). The prosthetic group of erythrocyte cytochrome $b_5$ was identified by paper chromatography to be

protohaem IX (Passon et al., 1972), which was the same as that reported for solubilized liver microsomal cytochrome $b_5$ (Strittmatter and Velick, 1956; Bois-Poltoratsky and Ehrenberg, 1967). The reactivity of the two proteins with various compounds was also compared. The reduced form of the solubilized liver cytochrome $b_5$ does not bind carbon monoxide (Petragnani et al., 1959), but is oxidized by oxygen, cytochrome c and methaemoglobin (Passon and Hultquist, 1972; Hultquist and Passon, 1971). The same properties were found for erythrocyte cytochrome $b_5$ (Passon et al., 1972). Both the erythrocyte and microsomal cytochromes $b_5$ oxidize cytochrome $b_5$ reductase [EC1.6.2.2] isolated from liver microsomal membranes and erythrocytes (Passon and Hultquist, 1972).

Immunochemical studies (Kuma et al., 1976; Goto-Tamura et al., 1976) also supported the similarity between the two forms of cytochrome $b_5$. An antibody raised to solubilized cytochrome $b_5$ from rat liver microsomes inhibited the reduction of erythrocyte cytochrome $b_5$ by NADH and cytochrome $b_5$ reductase.

The existence of an erythrocytic form of cytochrome $b_5$ was confirmed when the amino acid sequence of the protein isolated from rabbit (Schafer and Hultquist, 1983), human, porcine and bovine sources (Abe et al., 1985) was determined. In all cases the erythrocyte cytochrome $b_5$ was found to be 97 amino acids in length. Cytochrome $b_5$ in erythrocytes is 36 amino acids shorter than the intact microsomal membrane cytochrome $b_5$, but identical in length to the hydrophilic segment of microsomal membrane cytochrome $b_5$ (Douglas and Hultquist, 1978; Slaughter et al., 1982) (Fig. 1)

(1) Liver microsomal membrane cytochrome $b_5$




(2) Erythrocyte cytochrome $b_5$

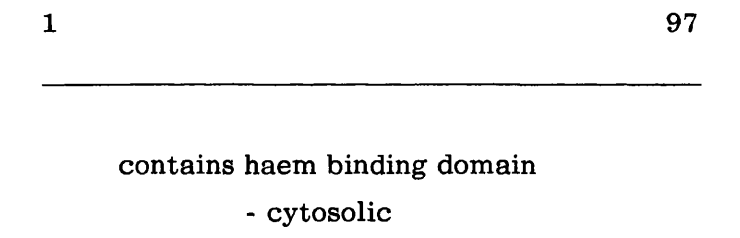



Figure 1. Structures of cytochrome $b_5$. Numbers indicate amino acid position.

### 1.1.3 Relationship between microsomal membrane-bound and cytosolic erythrocyte cytochrome $b_5$

The murine Friend virus-induced erythroleukaemic cell-line has many of the characteristics of immature erythroblasts and upon treatment with DMSO, differentiate as normal erythrocytes to polychromatophilic- and orthochromatophilic-like cells (Sato et al., 1971). These changes have been monitored by development of erythrocyte membrane antibody (Ikawa et al., 1973), accumulation of globin mRNAs (Ross et al., 1972) and synthesis of globin proteins (Boyer et al., 1972; Ostertag et al., 1972).

This cell line has been used as a model to study cytochrome $b_5$ during erythroid differentiation (Slaughter and Hultquist, 1979). The Friend erythroleukaemic cells were assayed for levels of membrane-bound and soluble forms of cytochrome $b_5$ at various stages of development. It was found that cytochrome $b_5$ is only present in the membrane fraction of the erythroleukaemic cell and that DMSO-induced differentiation of the cells to the polychromatophilic or orthochromatophilic erythroblast stages does not result in cytochrome $b_5$ in the soluble fraction. However, it is not known at which stage the `switch' between the membrane-bound and soluble forms of cytochrome $b_5$ occurs during erythroid maturation.

From the observation that during erythroid maturation the disappearance of endoplasmic reticulum results in the appearance of soluble cytochrome $b_5$, it was proposed that proteolysis of microsomal cytochrome $b_5$ occurs during erythrocyte development, producing the soluble form of the protein (Slaughter and Hultquist, 1979). This theory was supported by the amino acid analysis of tryptic peptides from bovine erythrocyte cytochrome $b_5$ (Slaughter et al., 1982). When this sequence was compared with the amino acid sequence of the bovine liver microsomal protein (Ozols, 1975; Fleming et al., 1978), the hydrophilic segment of microsomal cytochrome $b_5$ was identical to the erythrocytic protein.

However, comparison of the amino acid sequences of human, porcine (Kimura et al., 1984; Abe et al., 1985) and rabbit cytochromes $b_5$ (Schafer and Hultquist, 1983), indicates that, although within a species the liver and erythrocyte forms of cytochrome $b_5$ are identical for the first 96 amino acids, the carboxy-terminal amino acid of the erythrocyte protein, amino acid 97, is different, (Fig. 2).

| | amino acids 1-96 | | amino acid 97 | |
|---|---|---|---|---|
| | Liver | Erythrocyte | Liver | Erythrocyte |
| Bovine | Identical | | Ser | Ser |
| Human | Identical | | Thr | PRO |
| Rabbit | Identical | | Thr | PRO |
| Porcine | Identical | | Thr | SER |

Figure 2. Amino acid sequence comparison between liver and erythrocyte cytochrome $b_5$ in several species.

The changes in amino acid 97 correspond to substitution of the first base in the codon.

The amino acid sequence information suggests that there is one gene, rather than two, coding for cytochrome $b_5$. If there is just a single gene producing both the erythrocytic and microsomal membrane forms of cytochrome $b_5$ by means other than proteolysis, the two proteins are most likely to arise due to differential RNA splicing as occurs in the calcitonin/CGRP gene (Amara et al., 1982). The terminal amino acid, amino acid 97, in erythrocyte cytochrome $b_5$ can differ, in some species, from amino acid 97 in microsomal membrane cytochrome $b_5$. Therefore, there would have to exist two alternate exons, within the primary RNA transcript, containing sequence around the codon for amino acid 97. Depending on the tissue in which the cytochrome $b_5$ protein is to be expressed, one exon would be spliced out and the other remain in the mature mRNA. The codons for amino acid 97 in the two exons would not be the same in those species where amino acid 97 differed in the final proteins. The liver microsomal mRNA would also contain codons for amino acids 98 to 133, derived from one or more exons, whereas the erythrocyte mRNA would be truncated after amino acid 97 and its 3' non-coding region.

### 1.1.4 Mitochondrial cytochrome $b_5$

A cytochrome $b_5$-like protein, bound to the outer mitochondrial membrane, has been found in various tissues such as liver, kidney, heart, muscle and brain (Raw and Mahler, 1959; Ito, 1980a).

The protein from the outer mitochondrial membrane was shown, after solubilization with trypsin, to be different from the microsomal cytochrome $b_5$ in its immunological properties and by spectral and tryptic analysis (Fukushima and Sato, 1973; Ito, 1980b). The amino acid sequence of the haem-binding domain of the rat outer mitochondrial membrane cytochrome $b_5$ was determined by Lederer et al., (1983). A comparison of the first 91 amino acids between rat mitochondrial and microsomal cytochrome $b_5$ showed that the two proteins were 58% homologous. Conservation at the amino acid level is strongest in the central region of the protein (Lederer et al., 1983) (Fig. 3) around the haem crevice (Mathews et al., 1971, 1972) (Fig. 4). The area between the proximal His (amino acid 43) and the distal His (amino acid 67) of rat mitochondrial and microsomal cytochrome $b_5$ is greater than 76% homologous. Substitutions between the bovine microsomal and mitochondrial cytochromes $b_5$ are all found at the surface of the 3D structure as postulated by Mathews et al., (1971; 1972) except for the substitution of His-19 by Arg.

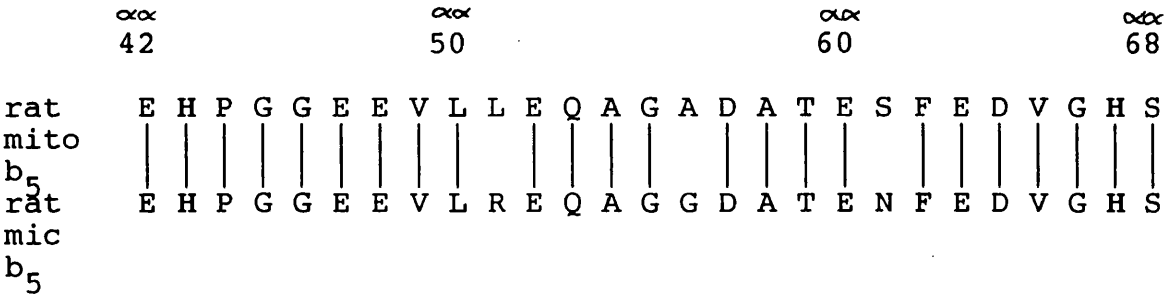



Figure 3. Comparison between rat mitochondrial cytochrome $b_5$ (rat mito $b_5$) and rat microsomal cytochrome $b_5$ (rat mic $b_5$) amino acid sequence around the proximal (amino acid 43) and distal (amino acid 67) histidines. The haem binding histidine residues are highlighted in bold. Homology between the mitochondrial and microsomal sequences is shown by vertical lines.

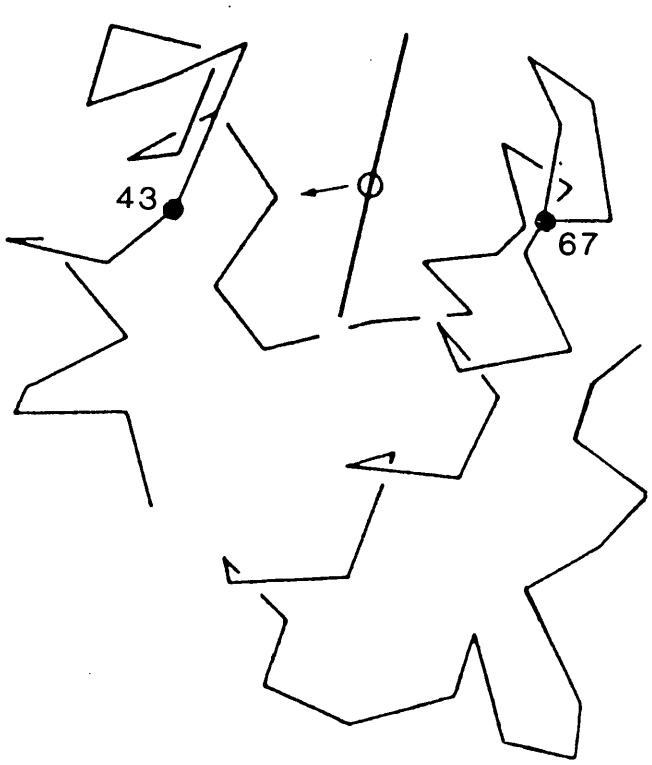



Figure 4. Schematic view of calf liver cytochrome $b_5$. The haem is held between His-43 and His-67, with the proximal histidine arrowed. The haem plane is at right angles to the page.

## 1.2 Cytochrome $b_5$ reductase

Cytochrome $b_5$ reductase [EC1.6.2.2] is a flavoprotein of molecular weight 33kD (Yubisui and Takeshita, 1980). It contains one mole of FAD per mole of enzyme (Kuma and Inomata, 1972).

Cytochrome $b_5$ reductase exists, like cytochrome $b_5$, in both soluble and membrane-bound forms. The protein is membrane-bound on the endoplasmic reticulum, the outer mitochondrial membrane (Sottocasa et al., 1967; Borgese and Mendolesi, 1980), Golgi membranes (Borgese and Mendolesi, 1980; Hino et al., 1978), and on outer nuclear membranes (Sagara et al., 1978) as well as existing as a soluble protein in erythrocytes (Passon and Hultquist, 1972; Kuma and Inomata, 1972).

It functions in the liver with cytochrome $b_5$ (section 1.3) in fatty acid metabolism (Holloway, 1971; Oshino et al., 1971) and the cytochrome P450 monooxygenase system (Hildebrandt and Estabrook, 1971; Ortiz de Montillado, 1986). In erythrocytes, cytochrome $b_5$ reductase is involved in the reduction of methaemoglobin via the soluble form of cytochrome $b_5$ (Hultquist and Passon, 1971).

Cytochrome $b_5$ reductase is an amphipathic protein with a large hydrophilic catalytic domain and a smaller hydrophobic membrane-binding section (Spatz and Strittmatter, 1973). In the liver the protein is 300 amino acids in length (Ozols et al., 1985) and in erythrocytes it is 275 amino acids in length (Yubisui et al., 1986).

The membrane-binding domain of cytochrome $b_5$ reductase has been reported to be at the carboxy-terminal, when cytochrome $b_5$ reductase from rabbit liver microsomal membranes was digested with carboxypeptidase (Mihari et al., 1978) and at the amino-terminal end when cytochrome $b_5$ reductase from steer liver was digested with carboxypeptidase (Kensil et al., 1983). Analysis of the amino-terminal domain of cytochrome $b_5$ reductase from steer liver (Ozols et al., 1984) showed that it comprises 25 amino acids and that the amino-terminal glycine is modified by myristic acid. A comparison of the amino acid sequence of cytochrome $b_5$ reductase isolated from steer (Ozols et al., 1985) and porcine liver (Crabb et al., 1980) with the amino acid sequence of human erythrocyte cytochrome $b_5$ reductase (Yubisui et al., 1986) shows that the erythrocyte protein begins at amino acid 26 of the steer liver sequence. This suggests that the hydrophobic domain of cytochrome $b_5$ reductase is located at the amino-terminus. The sequence of a human liver cytochrome $b_5$ reductase cDNA clone (Yubisui et al., 1987), 57 bases short at the 5' end and a full-length human liver microsomal cDNA clone (Bull, 1990) confirms the evidence.

Amino acid sequencing (Yubisui et al., 1986) has shown that only one residue in the first 23 amino acids in the human erythrocyte cytochrome $b_5$ reductase, Ser-12, is different from the corresponding amino acids in steer or porcine liver. Therefore, there is clearly a close link between erythrocytic and membrane-bound cytochrome $b_5$ reductase. The soluble erythrocyte form of cytochrome $b_5$ reductase has been shown to have similar properties to the membrane-bound form of cytochrome $b_5$ reductase. They are flavoproteins having similar molecular weights and both catalyze the reduction of cytochrome $b_5$ reductase isolated from human erythrocytes (Passon and Hultquist, 1972; Kuma and Inomata, 1972). Immunological studies have also supported the suggestion that there is a similarity between the membrane-bound and soluble forms of cytochrome $b_5$ reductase. An antibody raised against solubilized NADH-cytochrome $b_5$ reductase from rat liver microsomes inhibits the rate of cytochrome $b_5$ reduction by NADH (Kuma et al., 1976; Borgese et al., 1982).

From studies on methaemoglobinaemic patients it appears that the two forms of the protein are the products of one gene (Leroux et al., 1975; Lostanlen et al., 1981). The generalized deficiency of cytochrome $b_5$ reductase seen in about 10% of congenital methaemoglobinaemic patients is due to a deficiency in both the membrane-bound and soluble forms of the protein (section 1.3.3.1). The hydrophilic domain of human liver membrane-bound cytochrome $b_5$ reductase (Yubisui et al., 1987) is identical at the amino acid level with the soluble human erythrocyte cytochrome $b_5$ reductase (Yubisui et al., 1986). This indicates that both forms of the protein are encoded by the same gene and that the soluble form arises either by proteolysis of the membrane-bound cytochrome $b_5$ reductase during erythroid maturation, the use of an alternate promoter or by differential splicing of one primary RNA transcript giving rise to two mRNAs.

A cDNA clone coding for human cytochrome $b_5$ reductase has been used to analyze human-rodent somatic cell hybrids by Southern blot hybridization. These findings indicate that cytochrome $b_5$ reductase is encoded by a single gene located on human chromosome 22 (Bull et al., 1988).

### 1.2.1 Interaction between cytochromes $b_5$ and $b_5$ reductase

Modifications of 7 lysine residues on cytochrome $b_5$ reductase leads to the loss of the association with cytochrome $b_5$ (Loverde and Strittmatter, 1968) as observed by analyzing catalytic activity. Analysis of the tertiary structure of cytochrome $b_5$ shows that the surface of the protein has a large area of negatively charged side-chain carboxyl groups and a negatively charged exposed haem propionate group (Mathews et al., 1971). Between species these

carboxyl groups are highly conserved. Modification of the carboxyl groups of cytochrome $b_5$ indicates that the cytochrome $b_5$ carboxyl side chains of Glu-47,-48 and -52 and the haem propionate are involved in the interaction with cytochrome $b_5$ reductase (Dailey and Strittmatter, 1979).

## 1.3 Functions of cytochrome $b_5$

### 1.3.1 Cytochrome P450 monooxygenase system

Cytochromes P450 are the terminal oxidases of the hepatic microsomal mixed-function monooxygenase system. Cytochromes P450 play a central role in the metabolism of various compounds. Some are endogenous like steroids and fatty acids (section 1.3.2), whereas others are foreign compounds such as drugs, carcinogens and environmental pollutants (Ortiz de Montillado, 1986).

The preferred source of the two reducing equivalents required for the reduction of various compounds oxidized by the cytochrome P450 system is generally NADPH. However, NADH can also act as an electron donor to cytochromes P450, although the reactions tend to proceed at a slower rate (Fig. 5). Addition of NADH to a reaction being supported by NADPH sometimes enhances the rate of reaction (Nilsson and Johnson, 1963).

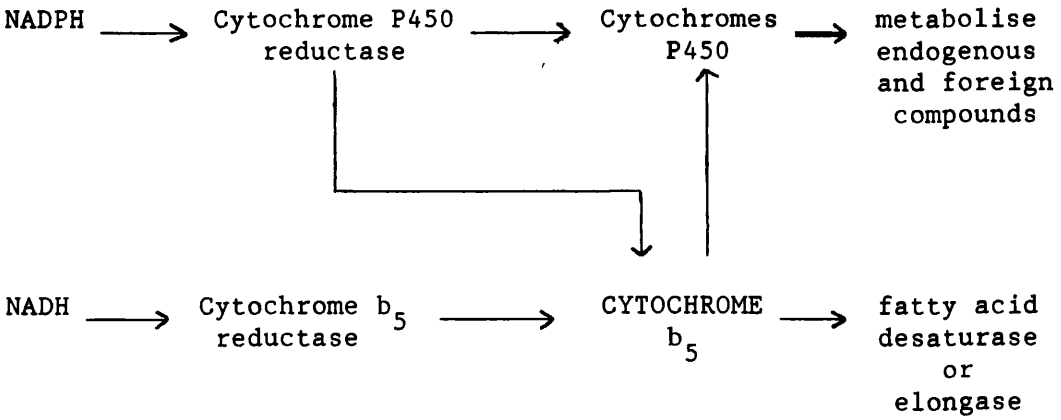



Figure 5. Function of microsomal cytochrome $b_5$.

⟶ shows passage of electrons.

NADPH-cytochrome P450 reductase (NADPH-cytochrome c reductase) [EC1.6.2.4] is an FAD and FMN containing enzyme (Iyanagi and Mason, 1973) and has been shown to participate in the cytochrome P450 monooxygenase system. The requirement of electron transfer from cytochrome P450 reductase to cytochromes P450 (Fig. 5) was shown by studies using antibodies specific for cytochrome P450 reductase (Raftell and Orrenius, 1970; Omura, 1969; Masters et al., 1971; Prough and Burke, 1975; Noshiro and Omura, 1978). At least one electron must be transferred by cytochrome P450 reductase to cytochromes P450.

Cohen and Estabrook (1971a; 1971b; 1971c) observed that NADH actually promotes the NADPH driven cytochrome P450 dependent N-demethylation of aminopyrine to a greater extent than could be accounted for by their combined effects. They demonstrated that cytochrome $b_5$ was functional in the supply of electrons to cytochromes P450 and postulated that the synergistic effect was due to more efficient utilization of the second electron from NADH via cytochrome $b_5$. This conclusion was supported by Correia and Mannering, (1973a; 1973b) who named it `NADH synergism'. The synergistic effect was later studied using specific inhibitory antibodies which confirmed the involvement of cytochrome $b_5$ in certain P450 catalyzed reactions. A cytochrome $b_5$ antibody was used to study the demethylation of ethylmorphine (Mannering et al., 1974). The antibody inhibited ethylmorphine N-demethylase in the presence of both NADH and NADPH together, whereas in the presence of NADPH alone, the cytochrome $b_5$ antibody had little effect on the reaction. If the cytochrome $b_5$ antibody was pre-incubated with pure cytochrome $b_5$, the inhibitory effect of the antibody on ethylmorphine metabolism was neutralized. These results indicate that cytochrome $b_5$ can act to donate the second electron to cytochromes P450 in the presence of NADH. However, when NADPH alone is present, cytochrome $b_5$ has little, if any, effect as an electron donor to cytochromes P450. In this situation, some other compound, such as P450 reductase, must provide the second electron. Any transfer of electrons from NADPH to cytochrome $b_5$ which does take place has been shown to be via P450 reductase (Enoch and Strittmatter, 1979) (Fig. 5).

From studies with reconstituted rabbit liver microsomes, cytochrome $b_5$ seems to be obligatory, as is P450 reductase, in the p-nitroanisole-o-demethylation reaction (Sugiyama et al., 1979). This process is catalyzed by a specific cytochrome P450 which has a high affinity for cytochrome $b_5$.

The overall conclusion from the above data is that the first electron donated to reduce cytochromes P450 is provided by NADPH-dependent cytochrome P450 reductase. In the presence of NADH, the second electron can be donated to certain cytochromes P450 by cytochrome $b_5$. Indeed, in some instances, cytochrome $b_5$ may even be obligatory, as in the reconstituted p-nitroanisole O-demethylation system (Sugiyama et al., 1979).

Administration of certain drugs is known to enhance the hydroxylation of those drugs, and sometimes other compounds as well, by selectively inducing cytochromes P450 and cytochrome P450 reductase (Remmer, 1972; Nebert and Gelboin, 1968). This effect was first observed on administration of phenobarbital to animals (Orrenius and Ernster, 1964).

Selective induction of cytochrome $b_5$ has also been observed (Denk et al., 1977). Griseofulvin treated mice had higher levels of mixed-function oxygenase activities, despite a decreased cytochrome P450 level. The induction of cytochrome $b_5$ by certain compounds stimulates the supply of the second electron to cytochromes P450. This results in an increase in the rate of breakdown of the applied compound.

1.3.1.1 Interaction of cytochrome $b_5$ and cytochromes P450

Phospholipid has been found to greatly stimulate cytochrome $b_5$-mediated cytochrome P450 hydroxylation reactions (Chiang, 1981).

As studied by electronparamagnetic resonance, the haem plane of cytochromes P450 is in the same plane as the membrane surface. The cytochrome $b_5$ haem, however, has a random orientation (Rich et al., 1979). These facts suggest that there is no tight complex between cytochromes P450 and cytochrome $b_5$. They appear to interact by random collisions or alternatively via other intermediate species.

1.3.2 Fatty acid and steroid metabolism

1.3.2.1 Fatty acid desaturation

At least three different fatty acid desaturation reactions are known to occur in rat liver microsomal membranes. These are catalyzed by the Δ9, Δ6 and Δ 5-desaturases:-

(1) Δ9-desaturase converts stearyl CoA to oleyl CoA and converts palmitic acid to palmiteoleic acid (Marsh and James, 1962).

(2) Δ6-desaturase introduces double bonds between C6 and C7 of certain fatty acids such as oleic acid and linoleic acid (Brenner and Peluffo, 1966).

(3) Δ5-desaturase desaturates fatty acids containing 20 carbon atoms (Castuma et al., 1972).

All three fatty acid desaturase-catalyzed reactions appear to require either NADPH or NADH and molecular oxygen (Fig. 5)

Stearyl-CoA desaturase, a Δ9 desaturase, is the most studied of these enzymes. In the presence of NADPH, the desaturation of stearyl CoA by rat liver microsomal membranes has been shown to involve the reduction of cytochrome $b_5$ by cytochrome P450 reductase (Enoch and Strittmatter, 1979). Cytochromes P450 are not functional components of the desaturase system (Oshino et al., 1966).

In vitro studies showed that NADH-supported stearyl CoA desaturation requires cytochrome $b_5$ (Holloway and Katz, 1972) and cytochrome $b_5$ reductase (Holloway and Wakil, 1970). A cytochrome $b_5$ antibody inhibits NADH- and NADPH-stimulated stearyl CoA desaturation. Both electron donors are equally sensitive to the antibody inhibition (Oshino and Omura, 1973). Therefore, cytochrome $b_5$ is the common factor in the two routes (Fig. 5).

Experiments in vitro with rat liver microsomal membranes indicate the importance of the hydrophobic membrane-binding section of cytochrome $b_5$ in desaturation reactions. Substitution of the hydrophilic fragment of cytochrome $b_5$ for the complete amphipathic protein prevents detectable NADH-dependent stearyl CoA desaturase activity (Shimakata et al., 1972; Strittmatter et al., 1974). Antibodies raised against the hydrophilic portion of cytochrome $b_5$ have also been used to show that cytochrome $b_5$ is a component of the Δ6-desaturation of linoleic acid and oleic acid in rat liver microsomes (Okayasu et al., 1976; Lee et al., 1977).

1.3.2.2 Fatty acid elongation

Fatty acid elongation in microsomal membranes requires either NADH or NADPH to provide the reducing equivalents (Stoffel and Arch, 1964; Nugteren, 1965). With the use of antibodies, cytochrome $b_5$ has been shown to participate in fatty acid chain elongation initiated by NADPH and involving cytochrome $b_5$ reductase (Ilan et al., 1981; Nagao et al., 1983) (Fig. 5). Cytochrome $b_5$ is also necessary for the transfer of electrons from NADH via cytochrome $b_5$ reductase to an elongase enzyme system (Keyes et al., 1979) (Fig. 5). When malonyl CoA, the precursor for fatty acid elongation in liver microsomal membranes, is added to a rat liver microsomal membrane system in the presence of ATP it stimulates the reoxidation of cytochrome $b_5$ two- to four-fold (Alfano and Cinti, 1977; Keyes et al., 1979). Correspondingly, malonyl CoA is incorporated into existing microsomal fatty acids. The involvement of cytochrome $b_5$ was confirmed by the use of an antibody against cytochrome $b_5$. In the presence of the antibody the incorporation of

malonyl CoA into fatty acids was reduced by 60% in a reconstituted system (Keyes et al., 1979).

Brain microsomal membranes have also been found to involve the cytochrome $b_5$/$b_5$reductase system in fatty acid elongation. Cytochrome $b_5$ reductase antibodies inhibit NADH-cytochrome $b_5$ reductase activity and palmitoyl CoA elongation (Takeshita et al., 1985).

1.3.2.3 Cholesterol biosynthesis

Lanosterol is converted to cholesterol by a series of processes in the rat liver. Immunological evidence indicates that microsomal cytochrome $b_5$ is involved in the introduction of the C5 double bond in cholesterol biosynthesis (Reddy et al., 1976). It is also required as part of an NADH-dependent electron transport system to demethylate lanosterol at C4. This was demonstrated by treating microsomes with trypsin which destroys cytochrome $b_5$ and leads to the loss of demethylation activity. The activity is restored with the addition of pure cytochrome $b_5$ (Fukushima et al., 1981).

1.3.2.4 Prostaglandin synthesis

Prostaglandin synthetase is an enzyme system bound to the endoplasmic reticulum (Rollins and Smith, 1980). The activity has been reconstituted in phospholipid vesicles (Strittmatter et al., 1982). Binding of cytochromes $b_5$ and $b_5$ reductase to these vesicles in the presence of NADH indicates that reduced cytochrome $b_5$ is used as an electron donor in prostaglandin formation.

1.3.2.5 Plasmalogen biosynthesis

Plasmalogen synthesis in porcine spleen microsomes has been shown by using a specific cytochrome $b_5$ antibody to require cytochrome $b_5$ (Paltauf et al., 1974). The cytochrome $b_5$ antibody inhibited plasmalogen synthesis to the same extent whether NADH or NADPH was acting as the electron donor. The mechanism for plasmalogen synthesis, therefore, appears to be similar to that of the stearyl CoA desaturase system in rat liver microsomes (Oshino et al., 1966), (section 1.3.2.1).

1.3.3 Reduction of methaemoglobin in erythrocytes

Haemoglobin is maintained in its active state in erythrocytes by reduced pyridine nucleotides. These ultimately reduce the ferric haem in methaemoglobin to counteract the autooxidation of haemoglobin (Gutmann et al., 1947; Jaffé and Neumann, 1968). NADH is responsible for most of the methaemoglobin reducing capacity in erythrocytes (Gibson, 1948; Scott et al.,

1965; Hegesh and Avron, 1967a; Scott and McGraw, 1962). NADPH-dehydrogenase activity is not neutralized by antibody raised against methaemoglobin reductase (Leroux and Kaplan, 1972) and is present at normal levels in the erythrocytes of methaemoglobinaemic patients.

The major erythrocytic NADH-diaphorase was isolated from normal human erythrocytes and shown to reduce methaemoglobin (Scott and McGraw, 1962). The enzyme was purified and its molecular properties were studied (Kuma and Inomata, 1972). It was determined to have a molecular weight of approximately 33kD with one mole of FAD, as the prosthetic group, per mole of enzyme. This identified the protein as being similar to NADH-dependent cytochrome $b_5$ reductase (Strittmatter and Velick, 1957; Passon and Hultquist, 1972) and NADH-methaemoglobin reductase (Hultquist and Passon, 1971). Methaemoglobinaemic patients (section 1.3.3.1) with NADH-diaphorase deficiency (Scott, 1960) were also shown to have a deficiency in cytochrome $b_5$ reductase (Kitao et al., 1974; Leroux et al., 1975).

These facts indicated that a single protein was responsible for the activities observed and it is now classified as cytochrome $b_5$ reductase [EC1.6.2.2] (section 1.2).

In both haemolysates and reconstituted systems, cytochrome $b_5$ reductase reduces methaemoglobin much more slowly than in vivo (Hegesh and Avron, 1967b; Abe and Sugita, 1979). Therefore, it was realised that another component was involved in the methaemoglobin reduction system.

Cytochrome $b_5$ was isolated from human erythrocytes (Passon et al., 1972) and found to be reduced by cytochrome $b_5$ reductase in vitro (Passon and Hultquist, 1972). When cytochrome $b_5$ is added to an in vitro system containing NADH and cytochrome $b_5$ reductase, it stimulates the reduction of methaemoglobin by cytochrome $b_5$ reductase (Hultquist and Passon, 1971; Passon and Hultquist, 1972).

Cytochrome $b_5$ and NADH-cytochrome $b_5$ reductase together were, therefore, postulated to be responsible for methaemoglobin reduction and a model for the transfer of electrons was proposed (Hultquist and Passon, 1971) (Fig. 6).

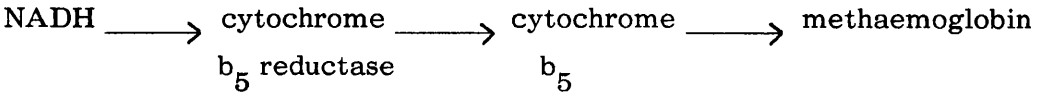



Figure 6. Function of erythrocyte cytochrome $b_5$

$\longrightarrow$ indicates passage of electrons

This model is also supported by the observation that deficiencies in cytochrome $b_5$ reductase, cytochrome $b_5$ or in the generation of NADH all lead to symptoms of methaemoglobinaemia (section 1.3.3.1).

1.3.3.1 Methaemoglobinaemia

In normal erythrocytes, less than 1% of the total haemoglobin is in the methaemoglobin form (Jaffé, 1964). This level is maintained by a methaemoglobin reductase system comprising cytochrome $b_5$ and cytochrome $b_5$ reductase (section 1.3.3).

Any deficiency in this system results in the accumulation of the methaemoglobin in erythrocytes, as the autooxidation of haemoglobin exceeds the capacity of the reduction system.

Hereditary methaemoglobinaemia has several possible causes:-
(1) methaemoglobin M diseases which are dominantly inherited and are due to an alteration, in one of several ways, to the amino acid sequence of the haem pocket in adult globin chains.
(2) methaemoglobinaemia with deficient generation of NADH due to decreased glutathione synthesis (Townes and Morrison, 1962).
(3) methaemoglobinaemia with cytochrome $b_5$ reductase deficiency (Gibson, 1948; Scott and Griffith, 1959; Scott, 1960; Leroux et al., 1975; Kaplan and Beutler, 1967; Schwartz et al., 1972; Schwartz et al., 1983) which is recessively inherited and has been reported in over 150 patients worldwide. There are, infact, three types of methaemoglobinaemia due to deficiency in cytochrome $b_5$ reductase.

(a) Type I - the deficiency is restricted to erythrocytes. The disease is benign and the only clinical symptom is cyanosis, which rarely needs treating (Gibson, 1948; Scott and Griffith, 1959). This uncomplicated methaemoglobinaemia results from a mutation in paired alleles which affects the stability, function or possibly solubilization of the hydrophilic segment of the protein.

(b) Type II - this more serious form of the disease is found in about 10% of patients. In addition to high methaemoglobin levels in erythrocytes they show deficient cytochrome $b_5$ reductase activity in microsomal membranes of many tissues including the brain. Patients with type II cytochrome $b_5$ reductase deficiency are mentally retarded and die prematurely (Leroux et al., 1975). As myelin phospholipid is usually rich in oleyl CoA, it is thought that reduced desaturation of fatty acids due to cytochrome $b_5$ reductase deficiency may be partly responsible for the effects on the central nervous system. A patient with the severe form of methaemoglobinaemia associated with mental retardation was assayed for the ability to elongate fatty acids in

platelets and leukocytes. Compared to normal subjects only 60% activity was found in the methaemoglobinaemic patient (Takeshita et al., 1987). Reduction in the capacity to elongate fatty acids may also, therefore, result in neurological damage.

Type II methaemoglobinaemia is thought to be due to either gene deletions or a mutation in paired alleles affecting both the hydrophobic and hydrophilic sections of cytochrome $b_5$ reductase. This could affect the activity, stability or attachment of the protein to the endoplasmic reticulum.

(c) Type III - this form of methaemoglobinaemia is defined by deficiency of cytochrome $b_5$ reductase in erythrocytes, leukocytes and platelets. Patients do not show generalized deficiency or the neurological symptoms associated with type II methaemoglobinaemia (Arnold et al., 1978; Tanishama et al., 1985). The cytochrome $b_5$ reductase deficiency is, in these patients, restricted to haemopoietic cells. This type of methaemoglobinaemia cannot be explained by the simple theory of different mutations at the same locus leading to erythrocytic or generalized cytochrome $b_5$ reductase deficiency. A transcription factor could possibly be absent in the haemopioetic cell-line of these patients.

(4) Methaemoglobinaemia with cytochrome $b_5$ deficiency in erythrocytes has been reported in one patient (Hegesh et al., 1986). This individual had normal cytochrome $b_5$ reductase levels and this protein, when purified from the patient's blood, had all the expected properties. However, the level of cytochrome $b_5$ protein isolated from the patient's erythrocytes was approximately 25% of the normal value. The haemolysate from the erythrocytes of the patient showed a very low rate of methaemoglobin reduction in the presence of excess NADH and added pure cytochrome $b_5$ reductase. Addition of pure cytochrome $b_5$ to the haemolysate enhanced the rate of methaemoglobin reduction. These results indicate that methaemoglobinaemia can be caused by cytochrome $b_5$ deficiency. It also confirms that cytochrome $b_5$ is required in the methaemoglobin reduction system.

### 1.3.4 Function of mitochondrial cytochrome $b_5$

It is thought that both microsomal and mitochondrial liver membranes possess a rotenone-insensitive NADH-cytochrome c reductase system involving cytochrome $b_5$ and cytochrome $b_5$ reductase (Strittmatter, 1963; Raw et al., 1958; Mahler et al., 1958; Raw and Mahler, 1959; Raw et al., 1960).

Mitochondrial cytochrome $b_5$ has been shown to play a role in the NADH-cytochrome c reductase system associated with the outer mitochondrial membrane (Fig. 7) by analysis with specific antibodies against microsomal and mitochondrial cytochrome $b_5$ (Ito, 1980b).

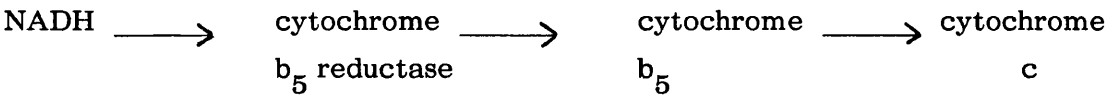



Figure 7. Function of mitochondrial cytochrome $b_5$ in the NADH-cytochrome c reductase system in mitochondria.

⟶ indicates passage of electrons

## 1.4 Aim

The aim of this project was to isolate and characterize genomic clones coding for rat cytochrome $b_5$. These clones were to be partially sequenced to allow the intron/exon borders to be defined. This information would provide evidence concerning the mode of origin of the truncated form of cytochrome $b_5$ in erythrocytes. Firstly, it was of interest to determine whether there was an intron/exon boundary near the codon for amino acid 97. If so, this would provide support for the existence of one cytochrome $b_5$ gene. If an intron was not located at this point, the theory that two genes code for cytochrome $b_5$ would have more credence. Secondly, the cytochrome $b_5$ exons within the genomic clones would be mapped in relation to several restriction endonuclease sites. The sizes of restriction fragments of rat genomic DNA on Southern blots probed with a rat liver cytochrome $b_5$ cDNA clone could then be compared to the expected sizes from the analysis of the genomic clones. In this way, genomic DNA restriction fragments could be shown to contain liver cytochrome $b_5$ exons. If all the cytochrome $b_5$ exons were isolated and accounted for, any remaining genomic DNA fragments which hybridized to the rat liver cDNA clone must be due to another cytochrome $b_5$ gene, although not necessarily a functional one.

# 2 MATERIALS AND METHODS

## 2.1 Bacterial Growth Media

### 2.1.1 Liquid Media

Bactotryptone, Bactoyeast extract and Bactoagar were from Difco. Agarose was from Sigma.

LB Medium: Bactotryptone (10g), Bactoyeast extract (5g) and NaCl (10g) were dissolved in water, the pH adjusted to 7.5, the volume made up to 1 litre and autoclaved.

2xTY Medium: Bactotryptone (16g), Bactoyeast extract (10g) and NaCl (5g) were made up to 1 litre with water and autoclaved.

M9 Medium: $Na_2HPO_4$ (6g), $KH_2PO_4$ (3g), NaCl (0.5g) and $NH_4Cl$ (1g) were dissolved in water, the pH adjusted to 7.4 and the volume made up to 1 litre. After autoclaving, the following components (each sterilized separately) were added:- 1M $MgSO_4$ (2ml), 20% glucose (10ml) and 1M $CaCl_2$ (0.1ml).

SOB Medium: Bactotryptone (20g), Bactoyeast extract (5g), NaCl (0.58g) and KCl (0.19g) were dissolved in water and made up to 1 litre (pH6.8-7.0). This was autoclaved for 30 minutes and used within 2-3 weeks.

SOB/$Mg^{2+}$: a 2M stock of $Mg^{2+}$ (1M $MgCl_2$, 1M $MgSO_4$) was made in water and sterilized by filtration through a 0.45µm filter (Millipore). This was added to autoclaved SOB just before use, to a final concentration of 20mM $Mg^{2+}$.

SOC: a 2M stock of glucose was made in water and sterilized by filtration through a 0.45µm filter (Millipore). This was added to autoclaved SOB just before use, to a final concentration of 20mM.

### 2.1.2 Media containing agar

Liquid media were prepared according to the recipes given above, except that before autoclaving the appropriate amount of Bactoagar was added: for plates, 15g/litre and for top agar 7g/litre.

2.1.3 Media containing agarose

Liquid media were prepared according to the recipes above except that before autoclaving the appropriate amount of agarose was added: for plates 15g/litre and for top agarose 7g/litre.

2.1.4 Antibiotics

Media were cooled to 55°C before addition of antibiotic.

Ampicillin (Amp): a stock solution of the sodium salt of Amp (Beecham) was made up to 50mg/ml in water , sterilized by filtration through a 0.45µm nitrocellulose filter (Millipore) and stored in aliquots at -20°C. Amp was added to the appropriate sterile media at a concentration of 50µg/ml. Liquid media containing Amp was used immediately. Plates containing Amp were stored at 4°C and used within 1-2 weeks.

2.2 Bacterial Strains

Three strains of Escherichia coli (E. coli) were used in these studies:-
JM101: supE, thi, Δ(lac-proAB), {F' traD36, proAB, $lacI^qZ$ M15}.

JM109: recA1, endA1, gyrA96, thi, hsd1r17, supE44, rel1A1, Δ(lac-proAB), {F', traD36, proAB, $lacI^qZ$ M15}.

The bacteria were streaked on an M9 agar plate and incubated at 37°C for 1-2 days. The plates were then sealed and stored at 4°C. The bacteria were restreaked every 4 weeks. These two strains (Messing, 1979) were used for the growth of recombinant M13 bacteriophage.

VCS 257: this strain (obtained from Stratagene) is a subclone of DP50 supF. It was used to plate out the lambda Charon 4A library as it has a high lambda phage DNA plating efficincy. The bacteria were streaked on an LB agar plate and incubated overnight at 37°C to obtain single colonies. The plate was sealed and stored at 4°C. The bacteria were replated every 1-2 months.

## 2.3 Restriction Endonuclease Digestion

### 2.3.1 Digestion buffers

10x Low salt: 100mM Tris-Cl(pH7.5)/100mM $MgCl_2$ - used for SacI.
10x Medium salt: 100mM Tris-Cl(pH7.5)/100mM $MgCl_2$/500mM NaCl - used for EcoRI, HindIII, PstI and AluI.
10x High salt: 100mM Tris-Cl(pH7.5)/100mM $MgCl_2$/1M NaCl - used for BamHI and SalI.
10x SmaI buffer: 200mM KCl/100mM Tris-Cl(pH7.5)/100mM $MgCl_2$ - used for SmaI only.
The volume of the digest varied according to the amount of DNA and the concentration of the restriction endonuclease. Digestions generally contained final concentrations of 10-100µg/ml DNA, 100µg/ml BSA (Pentax fraction V, Miles Scientific, nuclease free)/1mM dithiothreitol/4mM spermidine (Sigma)/ 10µg/ml RNase A/1x salt buffer (see above) and 2-10 units of restriction endonuclease per µg of DNA. Samples were incubated in water baths at 37°C for all restriction endonucleases used, other than SmaI which was incubated at 25°C, for 1-3 hours or for 16 hours in the case of genomic DNA. Reaction mixtures were either loaded directly onto agarose gels (section 2.4.2) stored at -20°C for later analysis or digested with a second restriction endonuclease.

## 2.4 Electrophoresis of DNA

### 2.4.1 Stock solutions

10x TBE: 108g Tris base, 55g Boric acid and 9.3g EDTA made up to 1 litre with water (pH approximately 8.3).
Ethidium Bromide: made up as a 10mg/ml or 1mg/ml solution in water and stored in the dark at 4°C.
10x Load Buffer: 25% Ficoll (w/v), 0.025% bromophenol blue, 0.025% xylene cyanol.

### 2.4.2 Agarose Gels

DNA was electrophoresed in horizontal agarose gels containing ethidium bromide. Varying percentages were used (0.3-1.8%) depending on the size of the DNA to be analyzed and the resolution required. Agarose (Sigma) was dissolved in 1x TBE by heating, cooled to approximately 55°C and made to a final concentration of 0.5µg/ml with respect to ethidium bromide. The solution was poured into a template containing a comb to form wells. The gel was allowed to solidify for 45-60 minutes at room temperature or at 4°C if

the percentage agarose was below 0.6%. One-tenth volume of 10x load buffer was added to the DNA samples and DNA molecular weight markers before loading onto gels. The most commonly used marker was a 1kb ladder (BRL) with 23 different sized fragments ranging from 75bp-12.2kb. Electrophoresis was carried out in 1x TBE buffer containing 0.5µg/ml ethidium bromide, at 10V/cm for approximately 1.5 hours or at 1.5V/cm for 16 hours. DNA was visualised by photography under ultra-violet light using Polaroid 55 film in a Polaroid Land camera (f4.5 for 45 seconds) fitted with a red filter. The sizes of the DNA restriction fragments were determined by first drawing a standard curve from the DNA molecular weight markers (Southern, 1975).

## 2.5 Isolation of Plasmid DNA

### 2.5.1 Growth and Harvesting of Bacteria (on a large-scale)

LB medium (100ml) containing the appropriate antibiotic (50µg/ml Amp) was inoculated with 0.1ml of bacteria from a glycerol stock culture. The culture was incubated overnight at 37°C in a 250ml flask with shaking. The bacterial cells were harvested by centrifugation at 4000g for 10 minutes at 4°C. The pellet was washed on ice by resuspending it in 15ml of ice-cold 10mM Tris-Cl(pH8.0)/100mM NaCl/1mM EDTA. This was then respun at 4000g for 10 minutes at 4°C.

### 2.5.2 Isolation of Plasmid DNA (on a large-scale)

Plasmid DNA was isolated as described by Birnboim and Doly (1979) and modified by Ish-Horowicz and Burke (1981).

Stock Solutions

Solution I: 50mM glucose/25mM Tris-Cl(pH8.0)/10mM EDTA. This solution was autoclaved for 15 minutes at 10lb/in$^2$ and stored at 4°C.
Solution II: 0.2N NaOH/1% SDS. This was made up from stock solutions (10N NaOH and 10% SDS) in water just before use.
Solution III: potassium acetate (pH4.8) was made up as follows:- to 60ml 5M potassium acetate was added 11.5ml glacial acetic acid and 28.5ml water. This solution was stored at 4°C.

The washed bacterial pellet was resuspended in 4ml of cold solution I, left at room temperature for 5 minutes, then transferred to a 50ml capped polypropylene centrifuge tube. 8ml of freshly prepared solution II was added, the tube capped and the contents mixed by gently inverting the tube several

times. After 10 minutes on ice, 4ml of solution III was added and the contents mixed by inversion several times. The tube was incubated on ice for a further 10 minutes, then centrifuged at 10000g for 10 minutes at 4°C. The supernatant was transferred to a 30ml Corex tube. Isopropanol (0.6 volumes) was added to the supernatant, the tube contents mixed well and incubated at room temperature for 15 minutes. The tube was then spun at 9000g for 30 minutes at room temperature. The pellet was washed in 70% ethanol, dried and resuspended in 1.25ml 10mM Tris-Cl(pH8.0)/1mM EDTA. The solution was transferred to sterile microfuge tubes and 39µl of 5M NaCl plus 25µl of 10mg/ml RNase A was added. This was incubated at 37°C for 90 minutes. The sample was first extracted with an equal volume of buffered phenol [the phenol was buffered once with 1M Tris-Cl(pH8.0)/1mM EDTA and twice with 100mM Tris-Cl(pH8.0)/1mM EDTA], then with an equal volume of buffered phenol/chloroform/isoamylalchohol (25:24:1) and finally with an equal volume of chloroform/isoamylalchohol (24:1). Lengthy vortexing (1 minute) was carried out at each stage of extraction. The phases were separated by centrifugation in a microfuge for 2 minutes. To the final aqueous phase, two volumes of absolute ethanol were added and mixed well. The tubes were stored at -70°C for 20 minutes. The plasmid DNA was pelleted by centrifugation in a microfuge for 15 minutes at room temperature. The pellets were rinsed with 70% ethanol, dried briefly and resuspended in 0.1ml 10mM Tris-Cl(pH8.0)/1mM EDTA. The resuspended DNA was heated at 65°C for 5 minutes to destroy endogenous nucleases and stored at -20°C.

2.5.3 Small-scale Rapid Plasmid Preparation

5ml of LB medium (containing the appropriate antibiotic) was inoculated with 5µl of a bacterial glycerol stock and incubated with shaking at 37°C overnight. Stock solutions were as described in section 2.5.2. 1.5ml of the overnight culture was transferred to a 1.5ml microfuge tube and centrifuged for 1 minute at room temperature in a microfuge. The medium was removed and the dry pellet resuspended by vortexing in 100µl of ice-cold solution I. The tube was stored at room temperature for 5 minutes, then 200µl of freshly prepared solution II was added. The contents of the tube were mixed by inverting the tube rapidly 2-3 times and the sample placed on ice for 5 minutes. 150µl of ice-cold solution III was then added, the tube vortexed gently in an inverted position for 10 seconds and placed on ice for 5 minutes. The tube was then centrifuged for 10 minutes at room temperature in a microfuge. The supernatant was extracted with an equal volume of buffered phenol/chloroform/isoamylalchohol (25:24:1) with thorough vortexing (1 minute). Two volumes of ethanol were added to the aqueous phase and the

contents of the tube mixed. The tube was left to stand at room temperature for 2 minutes, then centrifuged for 5 minutes in a microfuge at room temperature. The supernatant was removed, the pellet washed with 1ml of 70% ethanol then dried briefly. The pellet was finally resuspended in 10µl of 10mM Tris-Cl(pH8.0)/1mM EDTA. The sample was heated at 65°C for 5 minutes to destroy endogenous nucleases and stored at -20°C.

2.6 Preparation of total cDNA insert and smaller restriction fragments of cDNA insert from plasmid DNA

2.6.1 Restriction Endonuclease Digestion and Gel Electrophoresis

Closed circular plasmid DNA was digested with the appropriate restriction endonuclease(s), as described in section 2.3. The digest was then electrophoresed on a horizontal agarose gel containing ethidium bromide, as described in section 2.4. The gel was then visualised under uv light in order to locate and excise the required band.

2.6.2 Isolation of DNA fragment from gel

The fragments were isolated in one of two ways:-

2.6.2.1 DEAE Affinity

DEAE paper (Schleicher and Schuell NA45 Membrane filter, 0.45µm) was cut to a size slightly wider than the gel tracks and slightly deeper than the gel. Two slits were cut in the agarose gel using a scalpel blade, one just in front and one behind the DNA band to be isolated. Two pieces of NA45 paper were placed in the slits and the gel electrophoresed at 200V for 3 minutes to run the DNA fragment onto the NA45 paper in front. The piece of NA45 paper behind serves to prevent unwanted DNA being eluted. Ultra violet light was used to ensure that the DNA of interest had transferred to the NA45 paper and was now absent in the agarose gel. The NA45 paper with the DNA fragment of interest bound was rinsed briefly in 10mM Tris-Cl(pH8.0)/1mM EDTA, then placed in a microfuge tube containing 400µl of 1M NaCl/50mM Arginine (Sigma) (sterile filtered). The tube was incubated at 70°C for 1-2 hours to elute the DNA from the NA45 paper. The liquid was then removed to a fresh tube and 1ml of absolute ethanol added. The tube contents were mixed and placed on dry ice for 15 minutes, then brought to room temperature and centrifuged for 15 minutes in a microfuge at room temperature. The pellet was washed in 70% ethanol, dried and resuspended in 40µl of 10mM Tris-Cl(pH8.0)/1mM EDTA. This was extracted with an equal volume of buffered phenol/chloroform/isoamylalchohol (25:24:1) and then the organic fraction

was back-extracted with 10µl of 10mM Tris-Cl(pH8.0)/1mM EDTA. The aqueous layers were combined and one volume of 4M ammonium acetate/four volumes of absolute ethanol were added. The contents of the tube were mixed and placed on dry ice for 15 minutes. After the tube had reached room temperature it was centrifuged for 15 minutes at room temperature in a microfuge. The pellet was rinsed with 70% ethanol, dried briefly and resuspended in 10µl of 10mM Tris-Cl(pH8.0)/1mM EDTA.

2.6.2.2 'Glassmilk'

The Geneclean kit (BIO101 Inc.) was used to isolate DNA fragments from agarose gels. The band of DNA required was cut out of the gel using a scalpel blade and weighed in a microfuge tube. 2-3 volumes of saturated sodium iodide solution were added and the tube incubated at 45-55°C for 5 minutes to dissolve the agarose. For solutions containing up to 5µg DNA, 5µl of the 'Glassmilk' (an insoluble silica matrix) was added. An additional 1µl of 'Glassmilk' is added for each 0.5µg DNA above 5µg. The tube contents were mixed and placed on ice for 5 minutes. The silica matrix with the DNA bound was centrifuged for 5 seconds at room temperature in a microfuge. The sodium iodide supernatant was discarded and the pellet washed three times with 10-50 volumes of ice-cold 'NEW wash' (NaCl/Tris-Cl/EDTA/48% ethanol pH7-8.5 depending on the temperature). Washing was carried out by resuspending the pellet gently using a pipette, then centrifuging for 5 seconds at room temperature in a microfuge. After the supernatant from the third wash had been removed, the tube was centrifuged again for 2-3 seconds to ensure all the liquid had been removed. The DNA was eluted from the 'Glassmilk' by resuspending the pellet in 5µl or more of 10mM Tris-Cl(pH8.0)/ 1mM EDTA then incubating at 45-55°C for 2-3 minutes. The tube was centrifuged for 30 seconds at room temperature in a microfuge and the supernatant containing the eluted DNA placed in a new tube. A second elution of DNA from the 'Glassmilk' was carried out as described above to increase recovery.

2.6.2.3 Quantitation of the DNA fragment isolated

A sample of the DNA fragment isolated by one of the methods described above was loaded onto an agarose gel with known amounts of a DNA fragment of the same size in adjacent tracks. These standards were derived from restriction endonuclease digests of caesium chloride plasmid preparations. Comparison of the intensity of the bands under uv light gave a good indication of the amount of DNA in the sample eluted from the gel.

## 2.7 [$^{32}$P] radioactive labelling of DNA

### 2.7.1 Random Primer Labelling

The $pb_5$(1)D cDNA insert and the 280bp SacI/EcoRI restriction fragment from $pb_5$(1)D (section 3.1.2) were prepared as described in section 2.6. Just before either probe was required for labelling, a sample was boiled for 3 minutes to denature the DNA, then placed immediately on ice to quick-cool. The random primer labelling was carried out as described by Feinberg and Vogelstein, (1984).

Solutions

Solution O: 1.25M Tris-Cl(pH8.0)/0.125M $MgCl_2$,
Solution A: 1ml solution O/18µl 2-mercaptoethanol/5µl dATP/5µl dTTP/5µl dGTP (each triphosphate previously dissolved in 3mM Tris-Cl(pH7.0)/0.2mM EDTA at a concentration of 0.1mM),
Solution B: 2M HEPES (pH6.6 with 4M NaOH),
Solution C; Hexadeoxyribonucleotides evenly suspended in 3mM Tris-Cl(pH7.0)/0.2mM EDTA at 90 OD units/ml,
Oligo labelling buffer (OLB): solutions A:B:C were mixed in a ratio of 100:250:150.

The labelling reaction was carried out by addition of the following components in the stated order: $H_2O$ (to a final volume of 50µl), 10µl OLB, 2µl BSA (10mg/ml), DNA (up to 32.5µl), 2.5µl [$\alpha^{32}$P]-dCTP (New England Nuclear; Neg013A, 800Ci/mmol.10µCi/µl), 5 units of Klenow fragment of Escherichia coli DNA polymerase I.
The reaction was left at room temperature for 3-16 hours. The efficiency of incorporation of radioactivity was determined using Whatman DE81 paper (Maniatis et al., 1982). 1µl of the reaction mixture was spotted in duplicate onto 1cm square pieces of Whatman DE81 paper. One piece was used to determine total counts, the second to determine incorporated counts. The incorporated counts were assayed by washing the filter paper 6 times for 5 minutes each in 0.5M disodium hydrogen orthophosphate, twice for one minute each in distilled water and once for 1 minute in 95% ethanol. Both filters were dried under a heat lamp. Radioactivity on the filters was determined by liquid scintillation spectrophotometry (Phillips model 4700 scintillation counter) in 4ml of Aquasol (New England Nuclear). Incorporation of radioactivity was generally in the range 50-80%. DNA samples were labelled to a specific activity of $3x10^8$-$1x10^9$ cpm/µg. Unincorporated deoxyribonucleotides were separated from the labelled DNA by

chromatography on a column (0.6x8cm) of Sephadex G-50 (Pharmacia) suspended in 5xSSPE (20xSSPE is 3.6M NaCl, 0.2M $NaH_2PO_4$, 0.02M EDTA, adjusted to pH7.4). The column was washed with 3 volumes of 5xSSPE, saturated with denatured salmon sperm DNA (Sigma) (100µg), then washed with 1 volume of 5xSSPE. The sample was loaded onto the column and the labelled DNA eluted with 5xSSPE. 100µl fractions containing radioactively labelled DNA were combined (total volume 0.4-0.5ml) and the DNA concentration determined on Whatman DE81 paper by liquid scintillation spectrophotometry.

2.7.2 Radioactive Labelling of cDNA insert using M13 phage

The recombinant bacteriophage mE4.3II single-stranded DNA was kindly supplied by Dr.A.Ashworth from this laboratory (section 3.1.2). 2µg of mE4.3II was primed by incubating 8µl of DNA with 2µl (15ng) of the 17mer M13 universal primer, 1µl 100mM Tris-Cl(pH8.5) and 100mM $MgCl_2$ in a microfuge tube at 55°C for 1 hour followed by slow cooling to room temperature for 30 minutes. This primed M13 was then labelled by incubating with 1µl 100mM Tris-Cl(pH8.5)/100mM $MgCl_2$, 1µl 0.5mM dATP, 1µl 0.5mM dGTP, 1µl 0.5mM dTTP, 1µl 50mM dCTP, 5µl (50µCi) [$\alpha^{32}$P]-dCTP and 5 units of the Klenow fragment of E. coli polymerase I, at room temperature for 15 minutes. This was followed by addition of 1µl 0.5mM dCTP as a cold chase and incubation at room temperature for a further 15 minutes. The percentage incorporation and specific activity were determined as described in section 2.7.1. The labelled M13 clone mE4.3II was then digested with the restriction endonuclease EcoRI (section 2.3) to excise the rat cytochrome $b_5$ cDNA insert subcloned into this M13 bacteriophage vector. SDS was added to the digested mE4.3II to a final concentration of 0.1%. The sample was then extracted with an equal volume of buffered phenol/chloroform/isoamylalchohol (25:24:1) and the organic phase re-extracted with 10µl of 10mM Tris-Cl(pH8.0)/1mM EDTA. An equal volume of 4M ammonium acetate and 4 volumes of absolute ethanol were added to the combined aqueous phases. The tube was placed on dry ice for 15 minutes, brought to room temperature and centrifuged for 10 minutes in a microfuge at room temperature. The pellet was washed in 70% ethanol, dried briefly and resuspended in 15µl 10mM Tris-Cl(pH8.0)/1mM EDTA. The DNA was electrophoresed on a 1.5% agarose gel and the EcoRI cDNA fragment isolated from the gel using NA45 paper as described in section 2.6.2.1 except that once the DNA had been eluted from the NA45 paper in 1M NaCl/50mM Arginine, it was used directly.

2.7.3 [$\gamma^{32}$P] Radioactive Labelling of DNA by end-labelling

Oligonucleotides were end-labelled using T4 polynucleotide kinase (Maniatis et al., 1982). The reaction mix contained 2.5µl 4x kinase buffer [0.2M Tris-Cl(pH7.6), 40mM $MgCl_2$, 20mM dithiothreitol, 400µM spermidine, 400µM EDTA], 1µl (10µCi [$\gamma^{32}$P]-ATP (New England Nuclear, 800Ci/mmole) and 1µl (10units) T4 polynucleotide kinase. The reaction was incubated at 37°C for 1 hour. The efficiency of incorporation of radioactivity and the specific activity was assayed using Whatman DE81 paper, as described in section 2.7.1. Incorporation of radioactivity was generally in the range 70-90% and the specific activity $7x10^7$-$1x10^8$. Unincorporated deoxyribonucleotides were separated from the labelled DNA by ion-exchange chromatography on a column (0.7x4cm) containing DE52 cellulose (Whatman) equilibrated with 0.1M NaCl/10mM Tris-Cl(pH8.0)/1mM EDTA. Following application of the sample, the unincorporated nucleotides were washed from the column with 0.1M NaCl/10mM Tris-Cl(pH8.0)/1mM EDTA. The labelled oligonucleotide was eluted with 1M NaCl/10mM Tris-Cl(pH8.0)/1mM EDTA. 100µl fractions containing radioactively labelled oligonucleotide were combined (total volume 0.4-0.7ml) and the concentration of the DNA determined by liquid scintillation spectrophotometry.

2.8 Genomic Library

A rat genomic library in the lambda Charon 4A vector was kindly supplied by Sargent et al., 1979.

2.8.1 Preparation of Plating Bacteria

Single colonies from streaked VCS 257 plates were inoculated into 5ml LB medium supplemented with 10mM $MgSO_4$ (1M $MgSO_4$ stock) and 0.2% maltose (20% maltose stock) after autoclaving the medium. The 1M $MgSO_4$ and 20% maltose stocks were sterilized through a 0.45µm filter. The culture of VCS 257 was grown overnight at 30°C, to prevent overgrowth, with shaking. The following day, the bacteria were pelleted at 2500g for 10 minutes and resuspended in 0.4 times the original volume of 10mM $MgSO_4$. The bacterial suspension was stored at 4°C and used within 1-2 days. Alternatively, the VCS 257 culture was set up as above, incubated for 7 hours at 37°C with shaking and used without further treatment the same day.

2.8.2 Adsorption of Phage

For plating on 82mm petri dishes, 200µl (7 hour culture) or 100µl (overnight culture resuspended in 0.4x original volume of 10mM $MgSO_4$) of plating cells were mixed with the required dilution of bacteriophage. For plating on 150mm petri dishes, 300µl (7 hour culture) or 150µl (overnight culture resuspended in 0.4x original volume of 10mM $MgSO_4$) of plating cells were mixed with the required dilution of bacteriophage. In both cases the bacteria and phage were mixed gently and incubated at 37°C for 20 minutes to allow adsorption.

2.8.3 Plating out the phage

3ml of 0.7% LB top agar, or 0.7% top agarose if screening a library (to prevent filters peeling off the top layer) containing 10mM $MgSO_4$ (sterile filtered and added after agar/agarose is autoclaved) was added to the adsorbed phage. This mixture was vortexed gently and poured onto dry, prewarmed 1.5% LB bottom agar plates. These plates were incubated and inverted at 37°C for 14-16 hours.

2.8.4 Titring of phage library

A series of dilutions of the lambda Charon 4A genomic libraries were prepared in lambda diluent (100mM NaCl/50mM Tris-Cl(pH7.5)/10mM $MgSO_4$) and plated out as described in sections 2.8.2 and 2.8.3.

2.8.5 Screening of Genomic Library

2.8.5.1 Immobilization of phage DNA onto nitrocellulose or nylon filters

Approximately $4x10^4$ pfu were plated out in 0.7% LB agarose onto 1.5% LB bottom agar in 150mm petri dishes as described in sections 2.8.2 and 2.8.3 and incubated at 37°C for 14 hours. The plates were chilled at 4°C for 1 hour to allow the agarose to harden. Nitrocellulose filters (Millipore, type HA, 0.45µm) or Hybond-N nylon filters (Amersham International, 0.45µm) were laid on the plate and marked asymmetrically with Indian ink using a 25 gauge needle. The filters were removed and the phage DNA denatured by floating the filters on puddles of 0.5M NaOH/1.5M NaCl, DNA-side up for 45-60 seconds, then neutralized by floating on puddles of 0.5M Tris-Cl(pH8.0)/1.5M NaCl for 5 minutes, again DNA-side up. The filters were then briefly rinsed in 2xSSPE and air dried prior to fixing the DNA. This was carried out by baking at 80°C for 1.5-2 hours under vacuum for the nitrocellulose filters or exposure

to uv light wrapped in Saran wrap, DNA-side down on a transilluminator for 3 minutes for the nylon filters. Duplicate filters were prepared by a second round of replica plating as above, except that the filters were left in contact with the plates for 2 minutes.

2.8.5.2 Prewash

These baked/uv-fixed filters were then prewashed at 42°C for 1-2 hours in 50mM Tris-Cl(pH8.0)/1M NaCl/1mM EDTA/0.1% SDS with shaking to remove bacteria and other debris. The filters were then prehybridized and hybridized with the appropriate probe.

2.8.5.3 Prehybridization procedure for [$\alpha^{32}$P]-labelled probes

Solutions

Deionised formamide: formamide (BDH) was deionised batchwise by stirring at room temperature for 30 minutes with Amberlite ion exchange resin and stored at -20°C.

Denatured salmon sperm DNA: this was dissolved in water (10mg/ml) then sheared several times through a 25 gauge needle, denatured by boiling for 10 minutes and stored at -20°C. The DNA was reboiled for 5 minutes prior to use and then quickly cooled on ice.

Denatured Escherichia coli DNA: this was dissolved in water (1mg/ml), denatured by boiling for 10 minutes and stored at -20°C. The DNA was reboiled for 5 minutes prior to use and then quickly cooled on ice.

50x Denhardt's solution with BSA; ficoll (5g), polyvinylpyrolydine (5g) and BSA (5g) (Pentax fraction V, Miles Scientific) were dissolved in water and the volume made up to 500ml. The solution was filter sterilized through a 0.45µm nitrocellulose filter (Millipore) (Denhardt, 1966).

20xSSPE: 3.6M NaCl, 0.2M $NaH_2PO_4$, 0.02M EDTA (pH7.4).

Prehybridization buffer

This consisted of either:

(1) formamide buffer: 50% formamide/5x Denhardt's with BSA/5xSSPE/0.1% SDS/200µg/ml salmon sperm DNA/10µg/ml E. coli DNA,

or,

(2) aqueous buffer: 5x Denhardt's with BSA/5xSSPE/0.1% SDS/200µg/ml salmon sperm DNA/10µg/ml E. coli DNA.

The filters were heat sealed in a plastic bag with prehybridization buffer (100µl/$cm^2$ filter) and incubated at 42°C with formamide buffer/65°C with aqueous buffer for 4-16 hours.

2.8.5.4 Hybridization procedure for [$\alpha^{32}$P]-labelled probes

Solutions

These were as described for prehybridization (section 2.8.5.3.).

Hybridization buffer

This consisted of either:

(1) formamide buffer: 50% formamide/2x Denhardt's with BSA/5xSSPE/0.1% SDS/100µg/ml salmon sperm DNA/10µg/ml E. coli DNA/6% (w/v) polyethylene glycol 6000,

or,

(2) aqueous buffer: 2x Denhardt's solution with BSA/5xSSPE/0.1% SDS/100µg/ml salmon sperm DNA/10µg/ml E. coli DNA/6% polyethylene (w/v) glycol 6000.

The prehybridization buffer was removed from the bag, hybridization buffer added (50µl/$cm^2$) and incubated at 42°C with formamide buffer/65°C with aqueous buffer for 1 hour.

2.8.5.5 Addition of probe

mE4.3II insert labelled with [$\alpha^{32}$P]-dCTP (section 2.7.2), $pb_5$(1)D insert or the 280bp SacI/EcoRI restriction fragment from $pb_5$(1)D insert both labelled with [$\alpha^{32}$P]-dCTP (section 2.7.1) were boiled for 3 minutes, cooled on ice and added at an equivalent of 2ng/ml of hybridization buffer. Hybridization was allowed to proceed for 16-20 hours.

2.8.5.6 Filter washes

The filters were washed twice with 2xSSPE/0.1% SDS for 15 minutes each at room temperature and twice with 0.1x SSPE/0.1% SDS for 15 minutes each at room temperature. The filters were blotted dry, wrapped in Saran wrap and autoradiographed at -70°C.

2.8.5.7 Location of positive signal

Plaques giving positive signals on the duplicate filters were located by alignment of the filters on the plates. Agar plugs were picked using pipette tips and the phage eluted in 1ml of lambda diluent (100mM NaCl/50mM Tris-Cl(pH7.5)/10mM $MgSO_4$) at room temperature for 2 hours, then stored at 4°C with 50µl of chloroform.

2.8.5.8 Further screens

Positives from the first screen were titred, plated out on 82mm petri dishes at a density of 1-2x$10^3$ pfu and screened as above. This was continued until the phage were plaque-pure - usually by the third screen when the phage are plated out at a density of 1x$10^2$ pfu.

2.8.6 Rapid small-scale isolation of lambda phage DNA

2.8.6.1 Amplification of phage

VCS 257 plating bacteria (section 2.8.1) were added to LB top agar without bacteriophage and poured onto LB bottom agar plates (section 2.8.3). When the agar had set, 2-12µl of bacteriophage in lambda diluent were pipetted in a concentrated spot onto the top agar. When the bacteriophage solution had dried, the plates were incubated at 37°C for 14 hours. The large plaques were picked into 1ml of lambda diluent, left at room temperature for 2 hours and stored at 4°C with 50µl of chloroform. This amplification step was repeated as above using the primary amplification of bacteriophage to obtain a highly amplified phage stock.

2.8.6.2 Plate Lysis

100-200µl of the secondary amplification of bacteriophage suspension (>1x$10^9$ pfu/ml) (section 2.8.6.1) from the positive plaque was plated out in top agarose (sections 2.8.2, 2.8.3) and poured onto a wet, prewarmed bottom agarose plate. The plate was incubated upright wrapped in wet tissues in a box, to encourage lysis, at 37°C for 14 hours. 5ml of lambda diluent was pipetted onto the plate which was then rocked gently at room temperature for 1-2 hours to elute the phage. The lambda diluent, containing the bacteriophage, was then transferred to a Corex centrifuge tube and centrifuged at 8000g for 10 minutes at 4°C. The supernatant was incubated at 37°C for 30 minutes with RNase A and DNase I, both at a concentration of 1µg/ml. An equal volume of an ice-cold solution containing 20% (w/v) polyethylene glycol 6000 and 2M NaCl in lambda diluent was added and incubated at 0°C for 1 hour. This was then centrifuged at 10000g for 20 minutes at 4°C, the precipitated phage resuspended by vortexing in 0.5ml of lambda diluent and transferred to a microfuge tube. The tube was centrifuged for 5 seconds at room temperature in a microfuge to remove debris. 5µl of 10% SDS and 5µl of 0.5M EDTA(pH8.0) were added prior to incubation at 68°C for 15 minutes. This mixture was extracted once with an equal volume of buffered phenol/chloroform/isoamylalcohol (25:24:1) and once with chloroform/isoamylalcohol (24:1). An equal volume of isopropanol was added

to the final aqueous phase and incubated at -70°C for 20 minutes to precipitate the DNA. The tube was centrifuged for 15 minutes at room temperature in a microfuge, the pellet washed with 70% ethanol and dried. The DNA was resuspended in 50µl of 10mM Tris-Cl(pH8.0)/1mM EDTA.

2.8.7 Rapid large-scale Isolation of lambda phage DNA

After the amplified phage had been titred (sections 2.8.6.1 and 2.8.4 respectively), $1.25x10^8$ bacteriophage were mixed with 5ml of fresh VCS 257 plating cells from an overnight culture (sections 2.8.1) in a 1 litre flask. This was incubated at room temperature for 5 minutes. 250ml of prewarmed LB media containing 10mM $MgSO_4$ was added to the flask and the contents incubated at 37°C with shaking to achieve lysis (5-8 hours). Once this had occurred, solid NaCl was added to a concentration of 0.5M with 1ml of chloroform and the flask incubated at 37°C for 5 minutes with shaking. The lysate was centrifuged at 6000g for 10 minutes to remove debris and the supernatant stored at 4°C with 1ml of chloroform overnight. The supernatant was taken without chloroform and incubated at 37°C for 1 hour with DNase I and RNase A, both at a final concentration of 1µg/ml. Polyethylene glycol 6000 solid was added to a final concentration of 10% (w/v) and the mixture stirred at room temperature until the polyethylene glycol had dissolved. This was incubated for 1 hour at 0°C, then centrifuged at 6000g for 10 minutes to precipitate the phage. The pellets were resuspended in 5ml of lambda diluent with 10µl of silicone antifoaming agent (BDH) added. The suspension was centrifuged at 6000g for 2 minutes to remove debris, the supernatant extracted once with an equal volume of chloroform/isoamylalcohol (24:1) and then the organic layer re-extracted with 1ml of lambda diluent. 50µl of 10% SDS and 50µl of 0.5M EDTA(pH8.0) were added to the combined aqueous phases prior to incubation at 68°C for 15 minutes. This mixture was extracted once with an equal volume of buffered phenol, once with phenol/chloroform/isoamylalcohol (25:24:1) and once with chloroform/isoamylalcohol (24:1). An equal volume of isopropanol was added to the final aqueous phase and incubated at -70°C for 20 minutes. The thawed solution was centrifuged at 9000g for 30 minutes at room temperature. The pellet was washed in 70% ethanol, dried and resuspended in 300µl of 10mM Tris-Cl(pH8.0)/1mM EDTA.

## 2.9 Southern Blot Analysis

This method was devised by Southern, 1975.

### 2.9.1 Preparation of the filter

In order to analyze the DNA on an agarose gel by hybridization, after the gel had been photographed it was treated as described by Southern, 1975. The gel was soaked in 0.5M NaOH/1.5M NaCl (2ml/ml of gel) twice for 15 minutes each and then twice in 0.5M Tris-Cl(pH7.5)/1.5M NaCl for 15 minutes each with gentle shaking at room temperature. The DNA was transferred either to nitrocellulose or nylon (Hybond-N, Amersham International). A glass plate was covered with Whatman 3MM paper 'wicks' and placed so that the 'wicks' were in a reservoir of 20xSSC [3M NaCl/0.3M sodium citrate (pH7.0)]. The gel was placed face downwards on top of the saturated 'wick'. A piece of membrane was cut to the same size as the gel, soaked in distilled water for 5 minutes in the case of nitrocellulose and placed on top of the gel. Two sheets of Whatman 3MM paper cut to the same size as the membrane and soaked in 2xSSC were put on top of the membrane. Two layers of disposable nappy were placed on to the Whatman 3MM paper on top of the membrane to draw the transfer buffer up through the gel and membrane. The side of the nappy that would be next to the baby's bottom is placed downwards in direct contact with the pieces of Whatman paper. Finally, a second glass plate and a 500g weight were placed on top of the disposable nappy. The gel was blotted for 4-16 hours, depending on the size of the DNA fragments being transferred. After transfer had taken place, the position of the wells were marked on the membane with a pencil and the membrane either baked at 80°C for 1.5-2 hours under vacuum if it was nitrocellulose or fixed under uv light for 3 minutes if it was a nylon membrane.

## 2.10 Southern Blot Hybridization Procedure for [$\alpha^{32}$P]-labelled Probe

### 2.10.1 Prehybridization of the Filter

The prehybridization buffer was either a formamide based or an aqueous buffer.

#### Solutions

The stock solutions used were as described in section 2.8.5.3 for screening a genomic library.

Prehybridization buffer:

(1) formamide buffer: 50% formamide/5x Denhardt's with BSA/5xSSPE/0.1% SDS/200μg/ml salmon sperm DNA,

or,

(2) aqueous buffer: 5x Denhardt's with BSA/5xSSPE/0.1% SDS/200μg/ml salmon sperm DNA.

Prehybridization:-

The filter was heat sealed in a plastic bag with prehybridization fluid (100μl/$cm^2$ filter) and incubated at 42°C with formamide buffer/65°C with aqueous buffer for 4-16 hours.

2.10.2 Hybridization of the Filter

The hybridization buffer was either a formamide based or aqueous based buffer.

Solutions

These were as described in section 2.8.5.4 for screening a genomic library.

Hybridization buffer:

(1) formamide buffer: 50% formamide/2x Denhardt's with BSA/5xSSPE/0.1% SDS/100μg/ml salmon sperm DNA/6% (w/v) polyethylene glycol 6000,

or,

(2) aqueous buffer: 2x Denhardt's with BSA/5xSSPE/0.1% SDS/100μg/ml salmon sperm DNA/6% (w/v) polyethylene glycol 6000.

The prehybridization buffer was removed from the bag, hybridization buffer added (50μl/$cm^2$ filter) and the filter incubated at 42°C with formamide buffer/65°C with aqueous buffer for 1 hour.

2.10.3 Addition of Probe

Probe, labelled using the random primer method (section 2.7.1) was boiled for 3 minutes, cooled on ice and added at an equivalent of 2ng/ml of hybridization buffer for plasmid Southerns or 10ng/ml for genomic Southerns. The filter was hybridized for 18-20 hours at 42°C with a formamide buffer/65°C with an aqueous buffer.

2.10.4 Filter washes

To remove the unbound [$\alpha^{32}$P]-labelled probe the filter was washed twice with 2xSSPE/0.1% SDS for 15 minutes each at room temperature, once with 0.1xSSPE/0.1% SDS for 15 minutes at room temperature and once with

0.1xSSPE/0.1% SDS for 15 minutes at 50°C. The filter was blotted dry, covered in Saran wrap and autoradiographed at -70°C using Fuji or Kodak X-Omat X-ray film. If necessary, the filter was washed again for 15 minutes in 0.1xSSPE/0.1% SDS at temperatures ranging from 55-65°C and the blot re-exposed to autoradiographic film.

2.11 Hybridization procedure for [$\gamma^{32}$P]-labelled oligonucleotide probes

Solutions

These were as for [$\alpha^{32}$P] hybridizations (section 2.10), except for the use of 6xSSC instead of 5xSSPE.

20xSSC: NaCl (175.3g) and sodium citrate (88.2g) were dissolved in water and the volume made up to 1 litre and autoclaved.

2.11.1 Prehybridization

Filters were prehybridized at 5°C below the melting temperature of the duplex ($T_d$) as calculated by the equation:

$$T_d = 94 - \frac{820}{l} - 1.2(100-h) + 8$$

where l=length of the oligonucleotide, h=% homology; for nucleotides > 24 bases (Lathe, 1985)

or using the equation:

$T_d$ = 4°C(G.C) + 2°C(T.A); for nucleotides < 24 bases (Itakura et al., 1984)

where $T_d$ is the recommended stringent wash temperature in degrees C under salt conditions of 6xSSC.

Prehybridization buffer

6xSSC/10x Denhardt's solution with BSA/0.1% SDS/50µg/ml salmon sperm DNA.

The filter was heat sealed in a plastic bag with prehybridization buffer (100µl/$cm^2$ filter) and incubated for 2 hours at 5°C below the calculated $T_d$.

2.11.2 Hybridization

Hybridization buffer

This had the same constitution and volume as the prehybridization buffer (section 2.11.1).

The prehybridization buffer was removed and hybridization fluid added to the bag (50µl/$cm^2$ filter). Hybridization was carried out at the same temperature as the prehybridization step.

2.11.3 Addition of Probe

[$\gamma^{32}$P] end-labelled oligonucleotide probe (section 2.7.3) was added with the hybridization buffer at a concentration of 1-2ng/ml of hybridization fluid. The filter was hybridized for 3-4 hours at 5°C below the $T_d$.

2.11.4 Filter washes

The filter was washed in 6xSSC/0.1% SDS/0.05% sodium pyrophosphate three times for 5 minutes each at room temperature, then in the same solution for 2 minutes at the $T_d$ of the oligonucleotide probe. Filters were covered in Saran wrap and autoradiographed at -70°C.

2.12 Complete removal of probe from filters

Radiolabelled DNA was removed from Southern blots and phage lifts by incubating in boiling 0.1% SDS with gentle shaking until the solution reached room temperature - approximately 1 hour. The filter was exposed overnight to X-ray film to ensure that all the probe had been removed.

2.13 Hybridization procedure for Dried Gels

2.13.1 Preparation of Agarose Gel

The genomic DNA, digested (section 2.3) and run out on an agarose gel (section 2.4), was denatured by soaking the gel in 0.5M NaOH/0.15M NaCl for 30 minutes at room temperature with gentle shaking. The DNA was then neutralized by soaking the gel in 0.5M Tris-Cl(pH8.0)/0.15M NaCl for 30 minutes at 4°C with shaking. The gel was then placed on two sheets of Whatman 3MM paper and covered with Saran wrap. It was dried with only the vacuum until the gel was nearly flat (approximately 30 minutes), then with heating at 60°C under vacuum until the gel was completely dry (0.5-1 hour). The gel was stored on the paper, covered in Saran wrap until use.

2.13.2 Hybridization

The dried gel was separated from the 3MM paper by soaking in a shallow dish of distilled water. The gel with it's supporting paper, was then placed in a plastic hybridization bag, pressed against the plastic and separated from the paper. The gel was prehybridized in the heat sealed plastic bag with a solution of 6xSSC/10x Denhardt's/0.1% SDS/10µg/ml denatured salmon sperm DNA using 50µl of fluid/$cm^2$ of gel. This was carried out at 12°C below the $T_d$ of the oligonucleotide. After 2 hours, $2x10^6$cpm of end-labelled oligonucleotide (section 2.7.3) was added and hybridization allowed to proceed for 16 hours.

2.13.3 Dried Gel Washes

The gel was first washed in the plastic hybridization bag for 15 minutes at room temperature in 6xSSC. It was then washed in the same way on a piece of Saran wrap using a pump to remove the liquid. The filter was then stringently washed in 6xSSC for 2.5 minutes at 5°C below the $T_d$ of the oligonucleotide on Saran wrap as before. The gel was then covered in Saran wrap and exposed to Kodak X-Omat AR film between two intensifying screens at -70°C. A further wash for 1 minute in 6xSSC at 5°C below the $T_d$ of the oligonucleotide probe was used if necessary.

2.14 Subcloning of DNA into plasmid vectors pSPT19 and pUC19

2.14.1 Ligation of DNA fragments into plasmid vectors pSPT19 or pUC19

DNA from a lambda genomic clone was digested with the appropriate restriction endonuclease (section 2.3), as was the plasmid DNA. Samples of these digests were then run on agarose gels (section 2.4) to ensure that the enzyme had cut the DNA. The volume of the rest of the digest was made up to 50µl, with 10mM Tris-Cl(pH8.0)/1mM EDTA, SDS added to a final concentration of 0.1% and EDTA added to a final concentration of 0.01M. These digests were then extracted once with an equal volume of phenol/ chloroform/isoamylalcohol (25:24:1) and the DNA precipitated with 1 volume of 4M ammonium acetate/4 volumes absolute ethanol on dry ice for 15 minutes. The DNA was pelleted by centrifugation for 15 minutes in a microfuge at room temperature. The pellet was washed in 70% ethanol, dried and resuspended in water. Alternatively, if there were many fragments generated by the restriction endonuclease digest of the lambda DNA and only one was required in the plasmid vector, the fragment was selected prior to ligation. This was achieved by digesting the DNA with one or more suitable endonucleases, running the whole digest on an agarose gel and eluting the required DNA fragment onto NA45 paper (section 2.6.2.1) or extracting using

the Geneclean kit (section 2.6.2.2). 200ng of digested lambda clone DNA was mixed with 10ng of digested plasmid vector in a volume of 10µl, containing 60mM Tris-Cl(pH7.5)/6mM $MgCl_2$/10mM DTT/1mM ATP and 10 units of T4 DNA ligase. The mixture was incubated at 12°C for 16 hours or at room temperature for 1-16 hours.

2.14.2 Simple transformation of E.coli JM101 and JM109 by pSPT19 and pUC19 recombinant plasmid

Solutions

SB2 transformation buffer: KCl (7.4g), $MnCl_2.4H_2O$ (8.9g), $CaCl_2.2H_2O$ (1.5g) were dissolved in distilled deionized water. 20ml of a 0.5M K-MES(pH6.3) stock was added and the volume made up to 1 litre. The final pH was 6.2. This was filtered through a 0.45µm filter and stored at 4°C.

For pUC19 transformations only: 5-bromo-4-chloro-3-indoyl-beta galactoside (BCIG) was made up as a 20mg/ml stock solution in dimethylformamide and stored in aliquots at -20°C, isopropyl-beta-D-thio-galactopyranoside (IPTG) was made up as a 0.5M solution in water and stored in aliquots at -20°C.

Competent cells were prepared by the method of Hanahan (1985).

A single JM101/JM109 colony from an M9 plate was inoculated into 1ml of SOB and vortexed to distribute the bacteria. The bacterial suspension was added to 9ml of SOB in a well-rinsed sterile 250ml flask and incubated at 37°C with gentle shaking until an optical density of 0.35-0.6 OD units/ml at a wavelength of 550nm was reached for JM109 and 0.2-0.4 OD units/ml at a wavelength of 550nm was reached for JM101. This takes 1.5-3.5 hours. The culture was transferred to 50ml polypropylene Falcon tubes and chilled on ice for 10-60 minutes. The cells were pelleted by centrifugation at 900g for 15 minutes at 4°C. The pellets were then resuspended in a volume of SB2 that was $^1/_3$rd of the original culture volume by gentle pipetting. The tubes were placed on ice for 10-60 minutes then recentrifuged as above. The cells were resuspended in a volume of SB2 that was $^1/_{12.5}$th of the original culture volume i.e. 2.5ml of cells were resuspended in 200µl. 200µl aliquots of competent cells were put into polypropylene tubes and the ligated mixture added. As controls, 1ng (10µl) of uncut plasmid vector and a religated vector control were also added to a 200µl aliquot of competent cells. The tube contents were swirled gently and placed on ice for 30-60 minutes. Following this, the cells were heat-shocked by incubation in a 42°C water bath for exactly 90 seconds and then immediately chilled on ice. To each tube was added 0.2ml of SOC and these were then incubated at 37°C for 45-60 minutes with gentle shaking.

2.14.2.1 Plating out of pSPT19 transformants

100µl and 20µl aliquots of the ligation mix/competent cells were spread onto LB agar plates, containing 100µg/ml Amp, using a glass spreader. The smaller volume was diluted on the plate with SOB to a volume of 100µl. For the uncut vector, 10µl and 50µl aliquots (diluted in SOB on the plate as above) were taken and spread on LB agar plates, containing 100µg/ml Amp. Plates were inverted and incubated overnight at 37°C.

2.14.2.2 Plating out of pUC19 transformants

Just before use, LB plates containing 100µg/ml Amp were taken and 40µl of 20mg/ml BCIG and 40µl of 0.5M IPTG were spread onto the plates with a glass spreader. The competent cell/ligation mixture was then spread as for the pSPT19 transformation. Plates were incubated inverted at 37°C overnight. Efficiency of transformation for both vectors was between $1x10^5$ and $1x10^7$ cfu/µg.

2.14.3 Preparation of recombinant plasmid DNA

Transformed colonies were picked using toothpicks into 3ml SOB with Amp (50µg/ml) and incubated at 37°C overnight with shaking. 1.5ml of this culture was poured into a microfuge tube and centrifuged for 1 minute at room temperature in a microfuge. A small-scale rapid plasmid preparation was carried out on the pellet (section 2.5.3). This plasmid DNA was digested with the appropriate restriction endonuclease and electrophoresed on an agarose gel to size the inserts in the plasmid vector. From the remaining culture, 0.5ml aliquots were diluted 1:1 with 30% glycerol/LB and stored at -70°C to form a stock.

2.15 DNA sequence determination

The sequence of the cloned DNA fragments was determined by the M13 dideoxynucleotide chain termination method (Sanger et al., 1977; 1980) using the vectors of Messing and Vieira (1982).

2.15.1 Cloning of DNA fragments into M13 vectors

DNA restriction fragments from the inserts in the plasmid vectors pSPT19 and pUC19 which hybridized to the [$\alpha^{32}$P]-dCTP labelled rat cytochrome $b_5$ cDNA and/or to [$\gamma^{32}$P]-ATP labelled oligonucleotides specific for rat cytochrome $b_5$, were subcloned into M13mp18 or M13mp19 vectors. The M13 vector was digested with the appropriate restriction endonuclease(s) and a sample of the digest run on an agarose gel to ensure it had cut to completion.

The remainder of the digest was made up to a volume of 50µl with 10mM Tris-Cl(pH8.0)/1mM EDTA, SDS added to a final concentration of 0.1% and EDTA added to a final concentration of 0.01M. The digest was then extracted once with an equal volume of phenol/chloroform/isoamylalcohol (25:24:1) and the DNA precipitated with 1 volume of 4M ammonium acetate/4 volumes absolute ethanol on dry ice for 15 minutes. The DNA was pelleted by centrifugation for 15 minutes in a microfuge at room temperature. The pellet was washed in 70% ethanol, dried and resuspended in water.

The DNA fragments from the plasmid clones were prepared in the same way as described for the M13 vectors above unless there were several fragments generated by the restriction endonuclease digest. The single fragment required was then isolated on an agarose gel as described in section 2.6.2.1 or section 2.6.2.2.

2.15.1.1 Ligation of DNA fragments to the M13 replicative form (RF) DNA

DNA fragments (100ng) from the plasmids were mixed with 10ng of M13mp18 or M13mp19 RF form DNA in a final volume of 10µl containing 66mM Tris-Cl(pH7.5)/6mM $MgCl_2$/10mM DTT/1mM ATP and 10 units of T4 DNA ligase. The mixture was incubated at 12°C for 16 hours or at room temperature for 1-16 hours. A control ligation of cut vector re-ligated under the same conditions was also performed.

2.15.1.2 Transformation of E.coli JM101 and JM109 by recombinant M13 phage

To prepare competent cells, several single colonies from a freshly streaked culture of JM101 or JM109 (on M9 minimal media) were inoculated into 20ml 2xTY medium in a 250ml flask and incubated at 37°C with shaking until an optical density of 0.35-0.6 OD units/ml at a wavelength of 550nm was reached for JM109 and 0.2-0.4 OD units/ml at a wavelength of 550nm was reached for JM101. (1.5-3.5 hours). Alternatively, a single colony of E.coli JM101/JM109 was inoculated into 10ml 2xTY medium in a Universal tube and incubated at 37°C overnight. Then 1ml of this culture was added to 20ml 2xTY medium in a 250ml flask and incubated at 37°C with shaking until the optimum $OD_{550}$ of the culture was reached, as above (1-2 hours). At the same time as initiating the culture to prepare the competent cells, two drops of an overnight culture of JM101/JM109 in 2xTY media were added to 10ml of 2xTY media. This was incubated at 37°C with shaking until the $OD_{550}$ is approximately 0.3 (3-4 hours) (exponential growth phase) and was stored on ice until the competent cells had been prepared. The JM101/JM109 competent cells were prepared as described in section 2.14.2 up to the stage of chilling

on ice after heat shock. Then each ligation/competent cell mixture was split in half into pre-chilled polypropylene tubes. An uncut M13 vector control (1ng) was also included as for the plasmid transformations (section 2.14.2). This was aliquoted out at this stage into 10µl and 100µl volumes.

2.15.1.3 Plating out of transformation

To 3ml of 1xTY top agar (at 45°C) 40µl of 20mg/ml BCIG, 40µl of 0.5M IPTG (section 2.14.2) and 200µl of the exponential culture of JM101/JM109 were added. This was vortexed briefly, poured into the tube containing the transformed cells, mixed by inversion and poured onto the surface of a pre-warmed, dry 1xTY agar plate. After the top agar had solidified, the plates were inverted and incubated at 37°C overnight.
Efficiency of transformation was $5x10^5$-$1x10^6$pfu/µg.

2.15.2 Preparation of single-stranded M13 recombinant bacteriophage DNA

A single colony of E.coli JM101/JM109 was inoculated into 5ml of 2xTY medium and incubated overnight at 37°C with shaking. This culture was then diluted 100-fold in 2xTY media. Recombinant (white) plaques were picked using toothpicks into 2ml of the diluted culture and incubated at 37°C with shaking for 5 hours. 1.5ml of this culture was poured into a microfuge tube and centrifuged for 5 minutes at room temperature in a microfuge. 800µl of the supernatant was transferred to a fresh microfuge tube and mixed with 200µl of 2.5M NaCl/20% (w/v) polyethylene glycol 6000 and incubated at room temperature for 30 minutes. The solution was then centrifuged for 10 minutes at room temperature in a microfuge and all the polyethylene glycol supernatant removed. The pellet was resuspended in 100µl 10mM Tris-Cl(pH8.0)/1mM EDTA by gentle vortexing. 50µl of buffered phenol was added, the mixture vortexed for 20 seconds, left at room temperature for 15 minutes then re-vortexed for 20 seconds. The tubes were centrifuged for 2 minutes in a microfuge at room temperature. To the aqueous phase was added 10µl of 3M sodium acetate (pH6.0) and 300µl of absolute ethanol. The tube was placed on dry ice for 20 minutes and the DNA recovered by centrifugation for 15 minutes at room temperature in a microfuge. The pellet was washed with 200µl of absolute ethanol, dried and resuspended in 35µl of 10mM Tris-Cl(pH8.0)/0.1mM EDTA. 5µl of this M13 single-stranded DNA preparation was run on a 0.6% agarose gel alongside M13 single-stranded DNA containing no insert, to select recombinant clones. The remainder of the DNA was used for sequencing reactions if they contained the insert of interest. The rest of the phage supernatant and pellet were stored at -20°C.

2.15.3 Preparation of RF DNA from M13 recombinant bacteriophage

The recombinant (white plaques) were picked using toothpicks and grown up in 2xTY media containing JM101/JM109 and grown for 5 hours at 37°C as for single-stranded M13 DNA preparation (section 2.15.2). 1.5ml of the culture was then spun down in a microfuge tube for 5 minutes at room temperature. The supernatant was removed and the pellet resuspended in 100µl solution I for plasmid preparations (section 2.5.2). The RF M13 DNA was prepared on a small-scale as for plasmid DNA described in section 2.5.3.

2.16 DNA sequencing reactions

2.16.1 Klenow fragment of DNA polymerase I catalysed reactions

DNA sequencing reactions were performed with [$\alpha^{35}$S]-dATP (Biggin et al., 1983).

2.16.1.1 Priming

7µl of recombinant M13 single-stranded phage DNA (section 2.15.2) was combined with 1µl 100mM Tris-Cl(pH8.5)/100mM $MgCl_2$ and 1µl (1.25ng) Universal M13 17mer sequencing primer (5'GTAAAACGACGGCCAGT3'). This was heated at 55-60°C for 30-60 minutes in an oven and allowed to cool to room temperature. If not used immediately, this was stored at -20°C.

2.16.1.2 Sequencing reactions

2.16.1.2.1 Non-commercially prepared sequencing mixes

Deoxynucleotides were stored as 0.5mM stocks in water. Working solutions (made up in 10mM Tris-Cl(pH8.0)/0.1mM EDTA) were:

T°: 10µM dTTP/220µM dCTP/220µM dGTP

C°: 220µM dTTP/10µM dCTP/220µM dGTP

G°: 220µM dTTP/220µM dCTP/10µM dGTP

A°: 220µM dTTP/220µM dCTP/220µM dGTP

Dideoxynucleotides were stored as 10mM stocks in water. Working stocks (made up in 10mM Tris-Cl(pH8.0)/0.1mM EDTA) were 0.5mM ddTTP, 0.2mM ddCTP, 0.3mM ddGTP, 0.1mM ddATP.

Reaction mixes for T, C, G and A were made by mixing equal volumes of the appropriate deoxy- and dideoxy-nucleotide working stock solutions.

2.16.1.2.2 'Stratagene' prepared sequencing mixes

DNA sequencing mixes were also used as supplied by 'Stratagene Cloning Systems'. The DNA sequencing reactions had the following final concentrations:

| | T | C | G | A |
|---|---|---|---|---|
| Tris-Cl(pH7.5) | 50mM | 50mM | 50mM | 50mM |
| $MgCl_2$ | 5mM | 5mM | 5mM | 5mM |
| DTT | 5mM | 5mM | 5mM | 5mM |
| ddATP | 0 | 0 | 0 | 112µM |
| ddCTP | 0 | 36µM | 0 | 0 |
| ddGTP | 0 | 0 | 60µM | 0 |
| ddTTP | 100µM | 0 | 0 | 0 |
| dATP(cold) | 1µM | 1µM | 1µM | 1µM |
| dATP[$^{35}$S] | 1µM | 1µM | 1µM | 1µM |
| dCTP | 20µM | 2µM | 20µM | 20µM |
| dGTP | 20µM | 20µM | 3µM | 20µM |
| dTTP | 2µM | 20µM | 20µM | 20µM |

2.16.1.2.3 Incorporation of [$\alpha^{35}$S] for Klenow based reactions

To the primed mix was added 15µCi [$\alpha^{35}$S]-dATP (New England Nuclear; 600Ci/mmole) and 5 units of the Klenow fragment of DNA polymerase I. 2.5µl aliquots were dispensed into four microfuge tubes labelled T, C, G and A. To each tube 2µl of the relevant nucleotide reaction mix was added and the sample incubated at room temperature for 20 minutes. 1µl of 0.5mM dATP containing 0.1 units of Klenow polymerase was added to each of the four tubes and the incubation continued for a further 20 minutes at room temperature. The samples were then split into two or three tubes depending on the size of the combs to be used for the vertical polyacrylamide/urea gels and either loaded directly, or stored at -20°C for up to one week.

2.16.2 'Sequenase' catalysed reactions

'Sequenase' (United States Biochemical Corporation) is a modification of bacteriophage T7 DNA polymerase (Tabor and Richardson, 1987). The method of sequencing with 'Sequenase' is a modification of that used with the Klenow fragment of DNA polymerase I as outlined above.

2.16.2.1 Priming

7µl of recombinant M13 single-stranded phage DNA (section 2.15.2) was combined with 2µl Sequenase buffer (200mM Tris-Cl(pH7.5)/100mM $MgCl_2$/ 250mM NaCl) and 1µl (0.5pmol) 17mer Universal Sequencing primer. This mixture was heated at 55-60°C for 30-60 minutes in an oven, then allowed to slow cool to room temperature. If not used immediately the primed recombinant DNA was stored at -20°C.

2.16.2.2 Sequenase sequencing reaction mixes

Stock solutions:

dGTP labelling mix (5x concentrate):7.5µM dGTP/7.5µM dCTP/7.5µM dTTP

ddG termination mix: 80µM dGTP, 80µM dATP, 80µM dCTP, 80µM dTTP, 8µM ddGTP, 50mM NaCl

ddA termination mix: 80µM dGTP, 80µM dATP, 80µM dCTP, 80µM dTTP, 8µM ddATP, 50mM NaCl

ddT termination mix: 80µM dGTP, 80µM dATP, 80µM dCTP, 80µM dTTP, 8µM ddTTP, 50mM NaCl

ddC termination mix: 80µM dGTP, 80µM dATP, 80µM dCTP, 80µM dTTP, 8µM ddCTP, 50mM NaCl

2.16.2.3 Incorporation of [$\alpha^{35}$S]-dATP for Sequenase based reactions

2.16.2.3.1 For standard reactions (1-500 bases from the primer)

To the primed mix was added 1µl 0.1M DTT, 2µl 1xdGTP labelling mix (dilute 5x concentrate with water), 5µCi [$\alpha^{35}$S]-dATP (600Ci/mmol) and 2µl of Sequenase enzyme diluted 1:8 in ice-cold 10mM Tris-Cl(pH8.0)/1mM EDTA. This was incubated at room temperature for 5 minutes. Meanwhile, 2.5µl of the ddT/ddC/ddG/ddA termination mixes were added to tubes labelled T, C, G, A respectively and these were prewarmed at 37°C for 1-5 minutes. When the labelling reaction was complete, 3.5µl of the labelled mix was added to each tube containing a dideoxynucleotide termination mix. Incubation was allowed to proceed at 37°C for 5 minutes. The reactions were split into two or three tubes depending on the size of the combs to be used for the polyacrylamide/urea sequencing gel and either loaded directly onto the gel or stored at -20°C for up to one week.

2.16.2.3.2 For extended reactions (beyond 500 nucleotides)

The same protocol as above was used except that concentrations of dNTPs in the labelling reaction was increased 5-fold, the amount of [$\alpha^{35}$S]-dATP was increased to 10μCi and the labelling reaction lengthened to 10 minutes.

2.16.3 Gel electrophoresis of sequencing reactions

Stock solutions

Bis-acrylamide (40%): acrylamide (BDH Electran grade) (38g) and N,N'-methylenebisacrylamide (2g) were dissolved in water and made up to 100ml. The solution was deionised by stirring for 30 minutes with 5g of mixed-bed resin (Sigma, MB1), filtered and stored in the dark at 4°C.

Formamide dyes: 98% deionised formamide/25mM EDTA(pH8.0)/0.03% bromophenol blue/0.03% xylene cyanol.

10xTBE: as detailed in section 2.4.1.

6% polyacrylamide/7M urea gels (50ml) were made up by mixing 5ml of 10xTBE, 7.5ml of stock 40% bis-acrylamide, 21g urea (BRL, ultrapure) and 10ml of water. This was stirred at room temperature until the urea had dissolved. The volume was made up to 50ml with water. Freshly made ammonium persulphate (0.4ml of a 10% solution) and TEMED (40μl) were added, the solution swirled to mix and the gel poured immediately into a sequencing cassette lying at 45°. The cassette was 40cm long by 20cm wide by 0.35mm thick. Shark's-tooth combs (BRL) were clamped into position with the teeth outermost to form the top of the gel. The gel was allowed to polymerise for at least 1 hour and was used within 24 hours. The combs were removed and then replaced with the teeth just touching the surface of the gel to form the wells. Sequencing reaction products (2.5μl with the wider combs (24 wells) or 2μl with the narrow combs (48 wells) were mixed with formamide dyes (2μl), the samples boiled for 3 minutes and loaded immediately onto the gel. The electrophoresis buffer was 1xTBE. Electrophoresis was at 30mA (approximately 1500V) for times between 1.5 and 6 hours. Gels were transferred to Whatman 3MM paper, covered with Saran wrap and dried at 80°C for 40 minutes under vacuum. The Saran wrap was removed and the gel exposed to either Fuji RX or Kodak X-Omat X-ray film overnight at room temperature without an intensifying screen.

The Cruachem oligonucleotide synthesizer machines used the phosphoamidite method of synthesis.

2.17 Purification of oligonucleotides

Oligonucleotides were synthesised on oligonucleotide synthesiser automated machines.

2.17.1 Deprotection of bases and cleavage

The frit at one end of the column was pierced. 1ml of ammonium hydroxide was taken up in a syringe and inserted into the other end of the column. The open end of the column was placed inside a screw cap tube and the ammonium hydroxide gently flushed through the column into the tube ensuring that the beads were deposited in the tube. The capped tube was then placed in a waterbath at 55°C for 5 hours followed by cooling to room temperature by placing at -20°C (for approximately 30 minutes). The oligonucleotide was dried down in a freeze-drier overnight then resuspended in water (100µl). This was then butanol extracted until the volume was about 50µl (two or three extractions).

2.17.2 Gel purification

The oligonucleotides were then purified by electrophoresis on 20% polyacrylamide/7M urea gels. Stock solutions used were as for sequencing gels (section 2.16.3). The gel was made up by mixing 5ml 10xTBE, 25ml 40% bis-acrylamide stock and 21g urea then making the volume up to 50ml. Ammonium persulphate (0.4ml of fresh 10% solution) and TEMED (40µl) were added to the solution and the gel poured immediately into a sequencing cassette (40cm long by 20cm wide by 0.35mm thick) lying at 45°. Shark's-tooth combs (BRL) were clamped into position and the gel allowed to polymerise for 1 hour. Formamide dyes (2µl) were added to aliquots of the oligonucleotide (3µl) and loaded onto the gel. The electrophoresis buffer was 1xTBE and proceeded at 37W for 1½-3 hours depending on the size of the oligonucleotide being purified. As a guide, bromophenol blue runs at 10 bases and xylene cyanol at 28 bases.

After electrophoresis the gel was removed from the plate onto Saran wrap. The gel was covered with another piece of Saran wrap and the oligonucleotide visualized by uv shadowing. This involves placing the wrapped gel on a TLC plate and viewing under long-wave uv light. The position of the oligonucleotide was marked and the gel slice cut out. The oligonucleotide was eluted from the gel slice by placing in a microfuge tube, grinding with a toothpick, then adding approximately 0.75ml of 0.1M ammonium bicarbonate, regrinding, and incubating at 50°C for 10 minutes. The tube was then placed on a slowly rotating daisywheel overnight at room temperature. The tube was centrifuged briefly and the supernatant transferred to a fresh microfuge tube.

The acrylamide was washed with 0.1M ammonium bicarbonate (100µl), centrifuged briefly and the supernatants combined. The volume was made up to exactly 1ml with 0.1M ammonium bicarbonate. The oligonucleotide was purified on a precalibrated Sephadex G25 column (Pharmacia, NAP-10 column, 1.3 x 2.7cm gel bed). The NAP-10 column was equilibrated with 20% ethanol (approximately 20ml). The 1ml sample was then added and eluted with 1.5ml 20% ethanol into a glass Corex tube. The sample was lyophilised overnight in a freeze-drier, the DNA resuspended in sterile water (100µl) and the $A_{260}$ read in a spectrophotometer to assess the concentration of the oligonucleotide, knowing that 1 OD unit represents the equivalent of 20µg/ml of DNA.

2.18 Isolation of Genomic DNA from Frozen Tissue

Solutions

Extraction solution:

A 0.5M EDTA solution (pH8.0) was autoclaved. To this was added solid proteinase K to a concentration of 100µg/ml and sarcosyl to a concentration of 0.5%.

Frozen liver from Sprague-Dawley rats was ground with dry ice pellets in a coffee grinder. 1g of tissue was weighed out and added to 5ml of extraction solution. This mixture was incubated overnight at 50°C. Ensuring that the solution was homogeneously viscous, the DNA was extracted three times with an equal volume of buffered phenol. The phenol extraction was carried out by gentle shaking on a roller platform to ensure that the DNA remained intact. The final aqueous solution was treated with 100µg/ml of DNase-free RNase A at 37°C for 3 hours. The sample was gently extracted twice with phenol/ chloroform/isoamylalcohol (25:24:1) and once with chloroform/isoamylalcohol (24:1). To the aqueous solution, two volumes of ethanol were added, the solution mixed gently and the DNA spooled out with a hooked pasteur pipette into 400µl of 10mM Tris-Cl/1mM EDTA. The DNA was left to dissolve at 4°C then the concentration determined. The DNA was stored at 4°C.

The λgt11 library was screened using a polyclonal antibody raised in rabbit.

# 3 RESULTS

## 3.1 Isolation of genomic clones

### 3.1.1 Libraries

Two rat liver genomic libraries, kindly supplied by Sargent et al. (1979), were screened with partial and full-length rat cytochrome $b_5$ cDNA clones. These libraries were prepared by partial digestion of rat (Sprague-Dawley) genomic DNA with either EcoRI or HaeIII restriction endonucleases. The EcoRI fragments were inserted directly into the EcoRI site of lambda Charon 4A (Fig. 8) and the HaeIII fragments were ligated to EcoRI linkers before insertion into Charon 4A at the same site.

### 3.1.2 cDNA probes

The rat liver cytochrome $b_5$ cDNA clones used in this investigation were isolated from a rat liver $\lambda$gt11 library. The cDNA clones were sequenced by Dr. Ashworth in Dr. Shephard's laboratory.

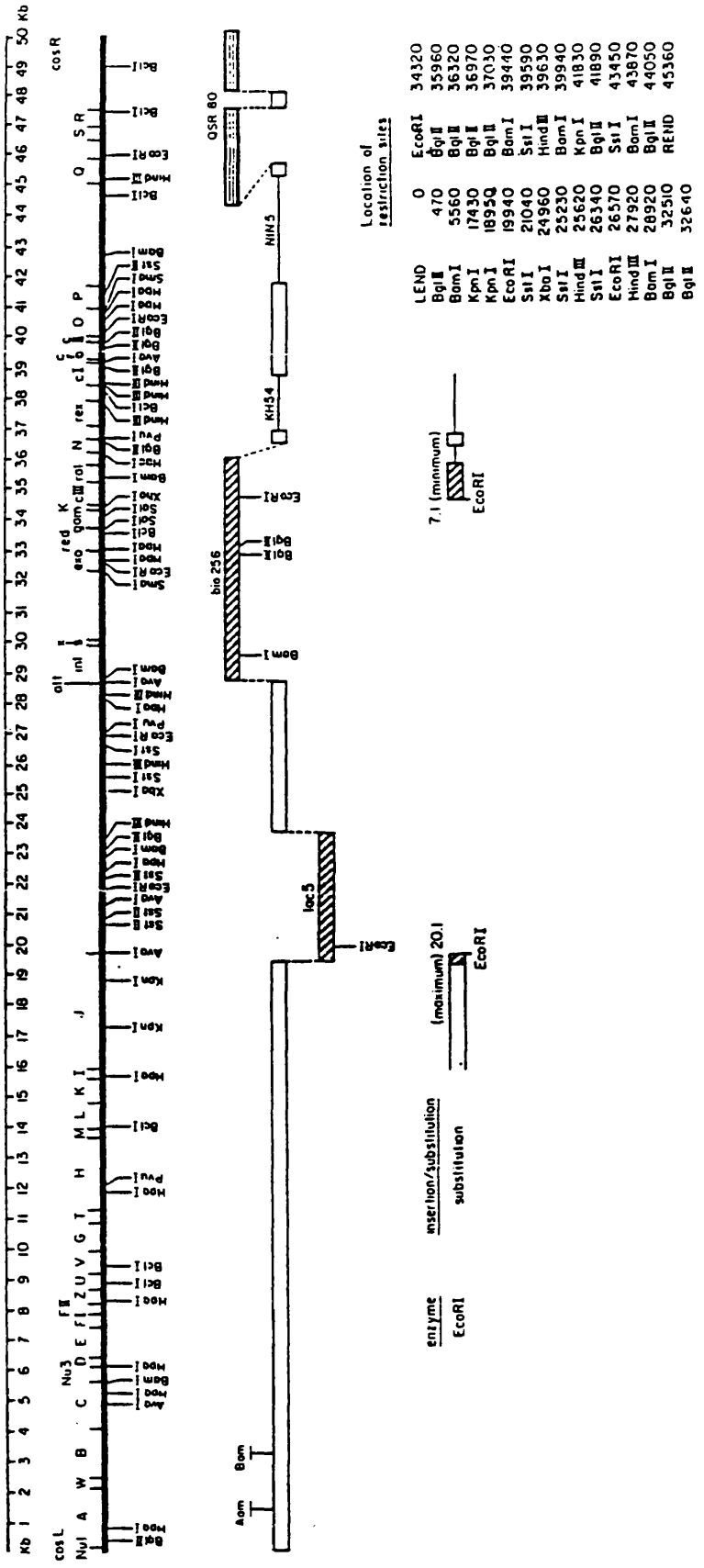



Figure 8. Map of the lambda vector Charon 4A from Maniatis et al. (1982). Total length 45.4 kb. Wild type lambda is shown above Charon 4A.

mE4.3II

An initial screen of the HaeIII rat liver genomic library was carried out using a partial length cDNA clone coding for rat cytochrome $b_5$, as a full-length clone had not been isolated at this stage. This clone was 560bp in length and contained the sequence from the first poly[A] tail to the codon for amino acid 40 (Fig. 9). The cDNA insert had previously been subcloned into the M13 bacteriophage vector and called mE4.3II. When screening the genomic libraries it was necessary to prevent cross-hybridization between the sections of the lac gene present in both the M13 and lambda Charon 4A vectors. Therefore, the cytochrome $b_5$ cDNA insert was isolated from the M13 vector. The single-stranded DNA form of mE4.3II was primed using the universal sequencing primer, then radioactively labelled as described in section 2.7.2, to obtain a probe with high specific activity. The double stranded clone was then digested with the restriction endonuclease EcoRI and the cDNA insert isolated on a 1.5% agarose gel (section 2.7.2).

$pb_5$(1)D

Once a full-length cytochrome $b_5$ cDNA clone became available (sequence shown in Fig. 9), this was then used to carry out further rescreening of the libraries. The cDNA insert of this clone was inserted into the EcoRI site of the plasmid vector pSPT19 and called pb5(1)D. For screening the genomic libraries, the total cDNA insert was separated from the vector by digestion with the restriction endonuclease EcoRI. The digest was electrophoresed on a 1% agarose gel and the 820bp cytochrome $b_5$ cDNA insert isolated (section 2.6) and radioactively labelled (section 2.7.1).

```
                                                  1
                                                  |------
                                          Met Ala Glu Gln Ser Asp
GGCTGTGTTGCAGGGCCCGGAAGCCTCACTGTTCCGAA ATG GCC GAG CAG TCA GAC


                10                                          20

------36mer----------------------------|
Lys Asp Val Lys Tyr Tyr Thr Leu Glu Glu Ile Gln Lys His Lys Asp
AAG GAT GTG AAG TAC TAC ACT CTG GAG GAG ATT CAG AAG CAC AAA GAC


                                    30

Ser Lys Ser Thr Trp Val Ile Leu His His Lys Val Tyr Asp Leu Thr
AGC AAG AGC ACC TGG GTG ATC CTA CAT CAT AAG GTG TAC GAT CTG ACC


        40                                          50

Lys Phe Leu Glu Glu His Pro Gly Gly Glu Glu Val Leu Arg Glu Gln
AAG TTT CTC GAA GAG CAT CCT GGT GGG GAA GAA GTC CTA AGA GAG CAA


                        60

                |--15mer----------|
Ala Gly Gly Asp Ala Thr Glu Asn Phe Glu Asp Val Gly His Ser Thr
GCT GGG GGT GAT GCT ACT GAG AAC TTT GAG GAC GTC GGG CAC TCT ACG


70                                      80
                                                |-SacI-|
Asp Ala Arg Glu Leu Ser Lys Thr Tyr Ile Ile Gly Glu Leu His Pro
GAT GCA CGA GAA CTG TCC AAA ACA TAC ATC ATC GGG GAG CTC CAT CCA


                90                                          100
                                                |--------------
    |------------------27mer----------|         |---18mer------
Asp Asp Arg Ser Lys Ile Ala Lys Pro Ser Glu Thr Leu Ile Thr Thr
GAT GAC AGA TCA AAG ATA GCC AAG CCT TCG GAA ACC CTT ATC ACT ACT


                                    110
---------26mer---|
-------|
Val Glu Ser Asn Ser Ser Trp Trp Thr Asn Trp Val Ile Pro Ala Ile
GTC GAG TCT AAT TCC AGT TGG TGG ACC AAC TGG GTG ATC CCA GCC ATC


        120                                         130

Ser Ala Leu Val Val Ala Leu Met Tyr Arg Leu Tyr Met Ala Glu Asp
TCA GCC CTG GTG GTA GCT CTG ATG TAT CGC CTC TAC ATG GCA GAA GAT
```

Ter
TAACCTGTCTGTCCGAAGCCAAGGAAGGAAAAGACTGCCCCAGAGAGGGGAGAAAAGAAGCCAG
TGTTAATCACTTCCACTGACAGAAACCCTCCCCCTGAGAATGTAATTGTAATATATCTGTCTCC
CTCTCCTCCTATGCTAGGAGAACAAACATGGGACTCTTTGTACTCTTAAACTTTCAAATGTGCC
TTTTTACTCAACTTCATTTTGACATTTCTTCACTACGTAATTTACTTATTGTAAACATGATCTT
TTTAAAATATATCTGGCTTGTAAAGTACA* CCAGGTGTGCCTGTTTGTGTGGTATTTTATATTT
AGTATTTAGTGTTTTGGAGTTGTTTAACAGAATTGCTTAACGTACTAACTGTTTGGACCCAAAA
AAAA

Figure 9. Amino acid and cDNA sequence of rat cytochrome $b_5$. The first poly[A] addition site is indicated (*). The SacI site used to create a 5' fragment for screening the genomic libraries is shown (SacI) and the positions of the five oligonucleotides used are marked. The proximal and distal histidines are highlighted. Possible poly[A] addition sites are underlined. The numbering relates to the amino acid position.

5' SacI/EcoRI fragment of $pb_5(1)D$

A third rat cytochrome $b_5$ cDNA derived probe was later used to rescreen both the genomic libraries. Characterization of all the original genomic clones showed that the 5' coding sequence of cytochrome $b_5$ had not been obtained. This cDNA probe was designed to prevent the isolation of clones that only contained 3' coding sequence. Use was made of a single SacI restriction endonuclease site centred at amino acid 82 of the rat cytochrome $b_5$ protein. The $pb_5(1)D$ recombinant plasmid construct was digested with the restriction endonucleases EcoRI and SacI. The digest was electrophoresed on a 1.2% agarose gel, the fragment containing the 5' end of the cDNA - 282bp in length - eluted from the gel (sections 2.6) and radioactively labelled (section 2.7.1). This SacI/EcoRI fragment consisted of a short section of the 5' non-coding region and the coding region from amino acids 1-81 inclusive of the rat cytochrome $b_5$ cDNA (Fig. 9).

3.1.3 Screening of rat genomic libraries

The first time the libraries were screened, approximately $1x10^6$ plaques from the HaeIII library were plated in top agarose (sections 2.8.1, 2.8.2, 2.8.3). These plaques were probed with the truncated rat cytochrome $b_5$ cDNA insert from mE4.3II (section 3.1.2). Bacteriophage DNA was transferred to nitrocellulose or nylon membranes in duplicate and fixed by baking (nitrocellulose) or uv light (nylon) (section 2.8.5.1). The filters were then prewashed (section 2.8.5.2) to remove the agarose and bacterial debris. The bound DNA was hybridized to [$\alpha^{32}$P]-dCTP labelled mE4.3II insert (section 3.1.2) as described in sections 2.8.5.3, 2.8.5.4 and 2.8.5.5 under hybridization conditions of 50% formamide/5xSSPE at 42°C. Two final washes at room temperature for 15 minutes each in 0.1xSSPE (section 2.8.5.6) were carried out prior to exposure to X-ray film. Plaques giving a positive signal with the mE4.3II cDNA insert (Fig. 10a) were located on the agarose plates and large areas picked (approximately $7mm^2$). The phage isolated were eluted from the agarose plugs (section 2.8.5.7) and plated out in top agarose at a lower density (section 2.8.5.8). The DNA was transferred to filters as before and hybridized with the same probe, mE4.3II, under the same conditions (Fig. 10b). This screening was carried out a third time, after which the positive plaques could be isolated from the plates individually (Fig. 10c). Two different plaque-pure clones, $Rgb_5 1$ and $Rgb_5 4$, were isolated from the HaeIII genomic library at this stage.

Before complete analysis of these two genomic clones was carried out, a full-length rat cytochrome $b_5$ cDNA clone became available. The libraries were then both screened with the full-length rat cytochrome $b_5$ cDNA insert from $pb_5(1)D$, generated as described in section 3.1.2, in case $Rgb_5 1$ and $Rgb_5 4$ together did not contain the whole cytochrome $b_5$ gene. The only clones isolated from plating out another $5x10^5$ plaques from each library using the full-length cytochrome $b_5$ cDNA insert were identical to $Rgb_5 1$ or $Rgb_5 4$.


Figure 10a. Primary screen of genomic library. Filters in duplicate - A and B.

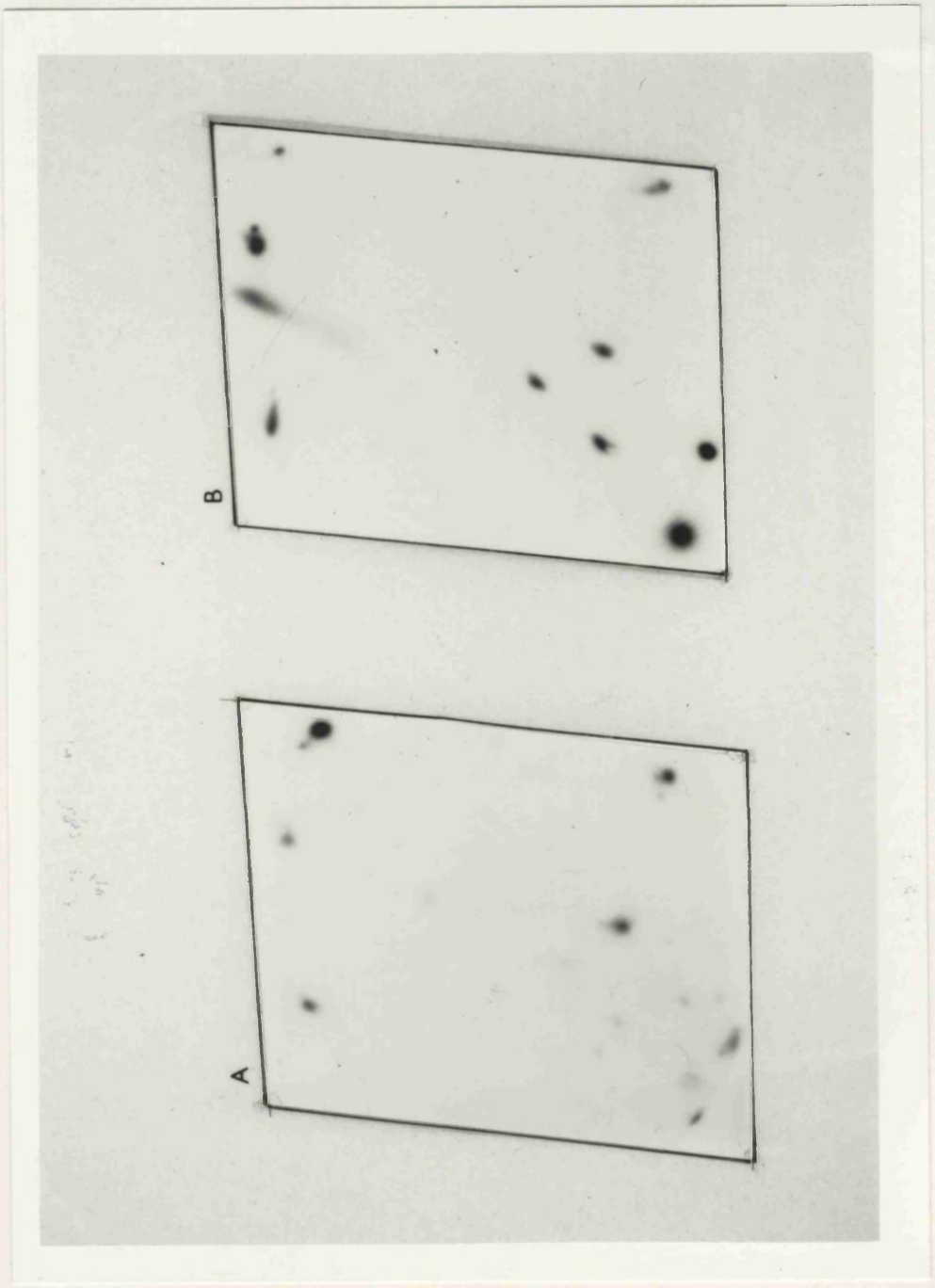



Figure 10b. Secondary screen of genomic library.
Filters in duplicate - A and B.

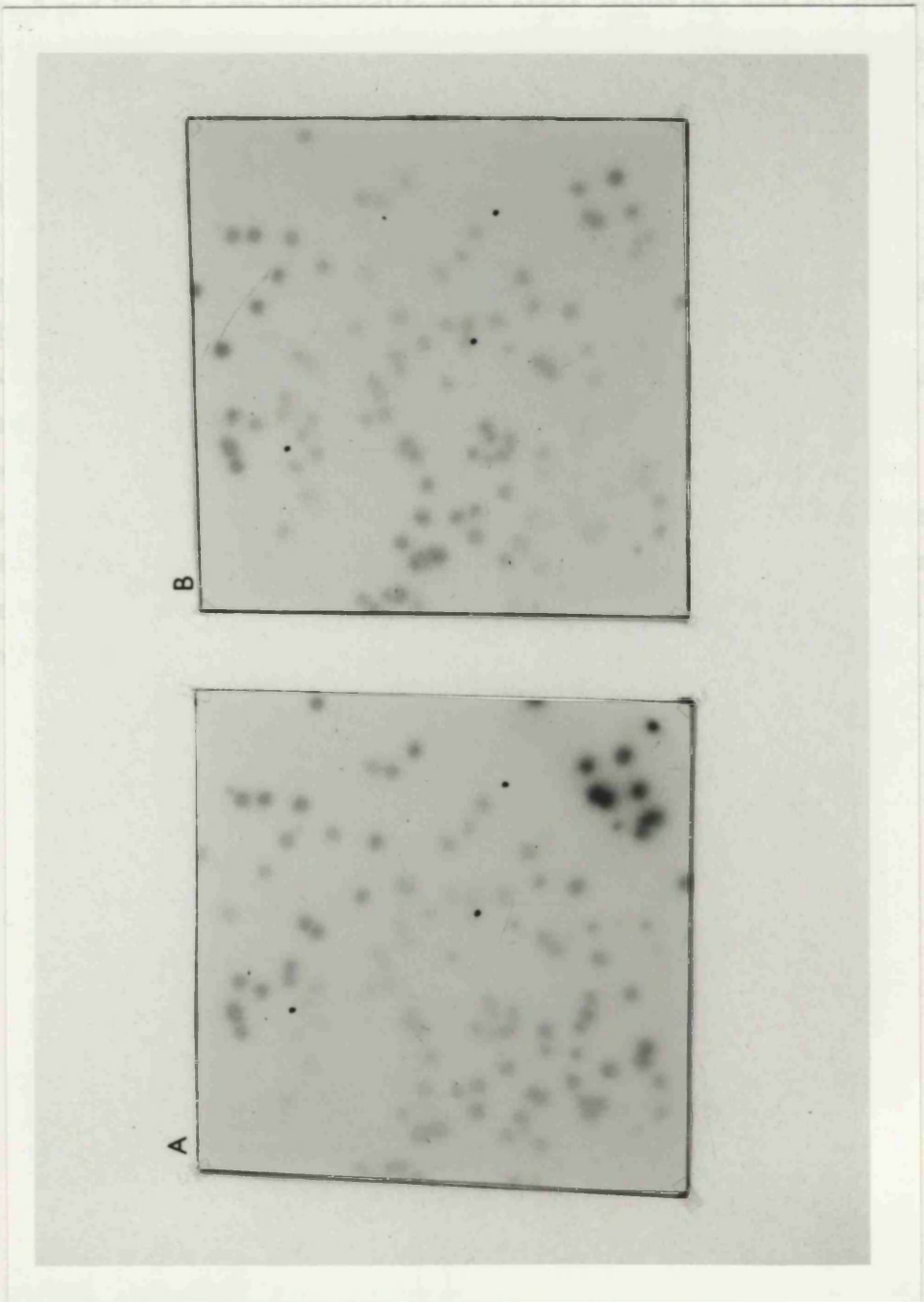



Figure 10c. Tertiary screen of genomic library. Filters in duplicate - A and B.

Once it had been established by hybridization and sequencing that neither $Rgb_5$1 nor $Rgb_5$4 contained the 5' end of the cytochrome $b_5$ gene, the third cDNA derived probe was used to rescreen both the rat genomic libraries. This was the 5' SacI/EcoRI fragment of $pb_5$(1)D, obtained as described in section 3.1.2. Approximately $5x10^5$ bacteriophage from each library were plated out (sections 2.8.1, 2.8.2, 2.8.3). Five additional clones were isolated and designated $Rgb_5$2, $Rgb_5$3, $Rgb_5$5, $Rgb_5$6 and $Rgb_5$7. The genomic clones $Rgb_5$3 and $Rgb_5$7 were identical to each other, whilst the rest all differed. The genomic clone $Rgb_5$6 was isolated from the HaeIII library, whilst the $Rgb_5$2, $Rgb_5$3, $Rgb_5$5 and $Rgb_5$7 clones were all from the EcoRI library.

## 3.2 Characterization of genomic clones

### 3.2.1 Restriction mapping of the genomic clones with EcoRI and Southern blot hybridization with a full-length rat cytochrome $b_5$ cDNA clone

DNA from plaque-pure positives at third screen was isolated - on a small-scale initially (section 2.8.6). The phage DNA was digested with the restriction endonuclease EcoRI, which separates the lambda arms from the insert and cuts the insert internally. The digests were then electrophoresed on a 20cm 0.6% agarose gel to ensure good separation (section 2.4). The sizes of the EcoRI fragments from each genomic clone are shown in table 1. All the lambda clones contained total insert sizes of between 10.1 and 14.2kb.

Table 1 Fragments from all rat cytochrome $b_5$ genomic clones digested with EcoRI.

| Rat $b_5$ clone | $Rgb_5$1 (H) | $Rgb_5$2 (E) | $Rgb_5$3/7 (E) | $Rgb_5$4 (H) | $Rgb_5$5 (E) | $Rgb_5$6 (H) |
|---|---|---|---|---|---|---|
| Sizes of frags. cut with EcoRI (bp) | 6150 | 3800 | 6500 | 6300 | 6500 | 6500 |
| | 4700 | 3200 | 2400 | 3000 | 3400 | 2400 |
| | 1200 | 2450 | 2200 | 870 | 2200 | 1600 |
| | 200 | 2200 | 1300 | | 400 | 1550 |
| | | 1450 | 850 | | | 1300 |
| | | 700 | | | | 500 |
| | | 400 | | | | |
| Total size of insert | 12250 | 14200 | 13250 | 10170 | 12500 | 13750 |

The underlined fragments indicate those which hybridize to the full-length rat cytochrome $b_5$ cDNA clone. The 1300bp fragment (from $Rgb_5$3/7 genomic clones) only hybridized under reduced stringency (section 3.2.4) - hybridization at 60°C/final wash at 50°C in 0.1xSSPE. Restriction endonuclease sites, E - EcoRI and H - HaeIII, indicate from which library the clone was isolated. Sizes given are in base pairs.

The EcoRI digested genomic clones were analyzed by Southern blot hybridization (section 2.9 and 2.10) using as a probe the full-length rat cytochrome $b_5$ cDNA insert from $pb_5$(1)D. The conditions of hybridization were in aqueous solution/5xSSPE at 65°C (sections 2.10.1, 2.10.2 and 2.10.3). A final wash was carried out for 15 minutes in 0.1xSSPE at 50°C (section 2.10.4).

$\underline{Rgb_5 1}$

This genomic clone, isolated from the HaeIII library, has a 6.15kb EcoRI fragment which hybridizes to the rat cytochrome $b_5$ cDNA (Fig. 11).

$\underline{Rgb_5 4}$

Also isolated from the HaeIII library, $Rgb_5$4 has an 870bp EcoRI fragment which hybridizes to the rat cytochrome $b_5$ cDNA (Fig. 11).

$\underline{Rgb_5 2}$

Isolated from the EcoRI library, $Rgb_5$2 has a 2.2kb fragment which hybridizes to the rat cytochrome $b_5$ cDNA (Fig. 12).

$\underline{Rgb_5 3/Rgb_5 7}$

These genomic clones, isolated from the EcoRI library, are identical in terms of the location of the EcoRI sites. They have a 2.2kb EcoRI fragment which hybridizes to the rat cytochrome $b_5$ cDNA (Fig. 12). The 1300bp EcoRI fragment only hybridizes under reduced stringency conditions (hybridize 60°C/ final wash 50°C in 0.1xSSPE) (section 3.2.4).

$\underline{Rgb_5 5}$

This clone, isolated from the EcoRI library has a 6.5kb EcoRI fragment which hybridizes to the rat cytochrome $b_5$ cDNA (Fig. 12).

$\underline{Rgb_5 6}$

The final clone, isolated from the HaeIII library has a 1.6kb EcoRI fragment which hybridizes to the rat cytochrome $b_5$ cDNA (Fig. 12).

None of the other EcoRI fragments in any of the genomic clones hybridized to the rat cytochrome $b_5$ cDNA under these stringent conditions.


Figure 11A. Photograph of lambda genomic clones $Rgb_5 1$ and $Rgb_5 4$ digested with EcoRI and run on a 0.6% agarose gel. Lane a - 1kb ladder (BRL marker), lane b - $Rgb_5 1$ digested with EcoRI, lane c - $Rgb_5 4$ digested with EcoRI. Sizes given in kilobase pairs.

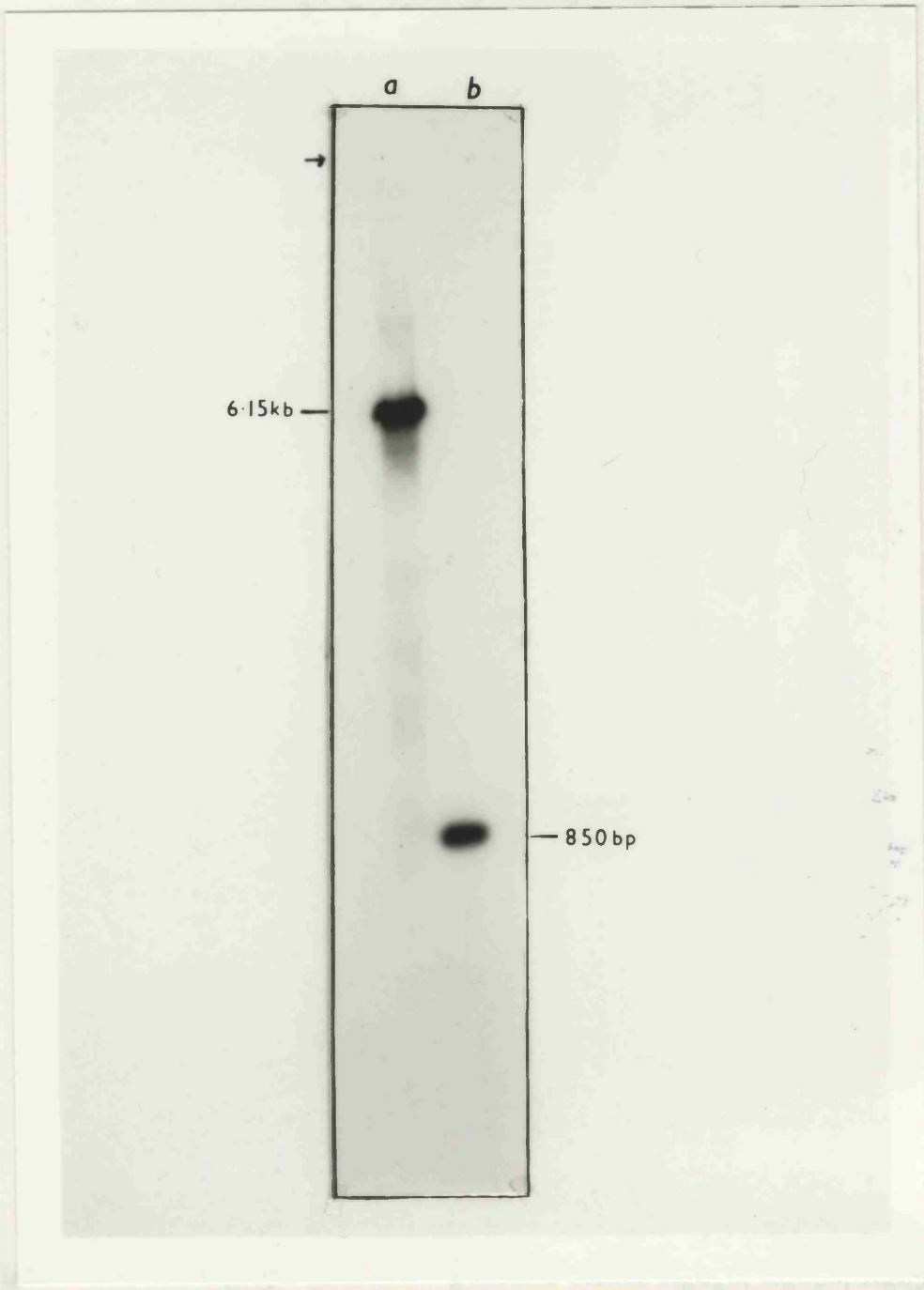



Figure 11B. Southern blot of agarose gel (Figure 11A) probed with the full-length rat cytochrome $b_5$ cDNA. Final wash at 50°C. Lane a - $Rgb_5$1, lane b - $Rgb_5$4. Sizes given in kilobase pairs. ⟶ marks the wells.

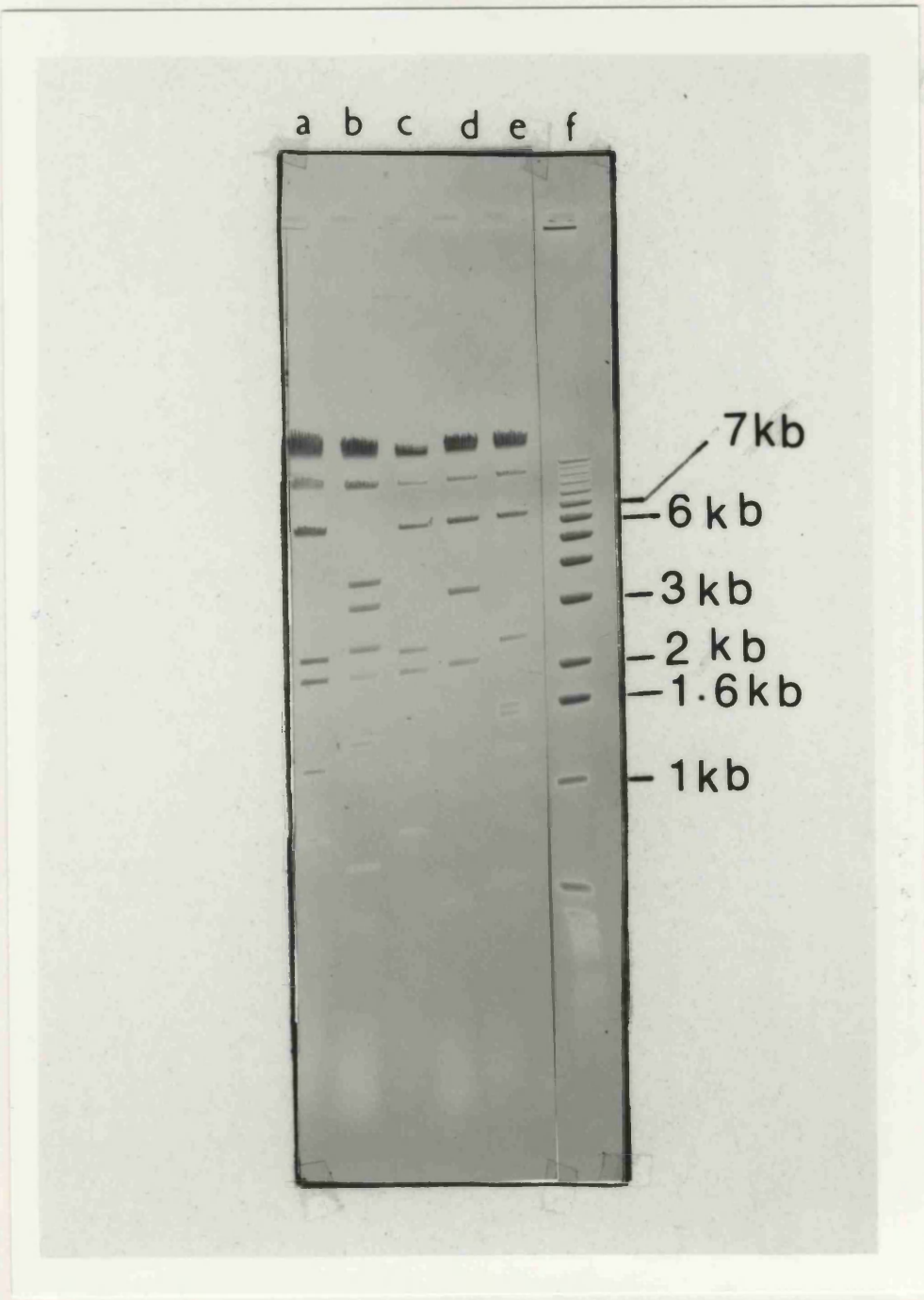



Figure 12A. Photograph of lambda genomic clones $Rgb_5$2, $Rgb_5$3, $Rgb_5$7, $Rgb_5$5 and $Rgb_5$6 digested with EcoRI and run on a 0.6% agarose gel. Lane a - $Rgb_5$7 digested with EcoRI, lane b - $Rgb_5$2 digested with EcoRI, lane c - $Rgb_5$3 digested with EcoRI, lane d - $Rgb_5$5 digested with EcoRI, lane e - $Rgb_5$6 digested with EcoRI, lane f - 1kb ladder (BRL marker). Sizes given in kilobase pairs.

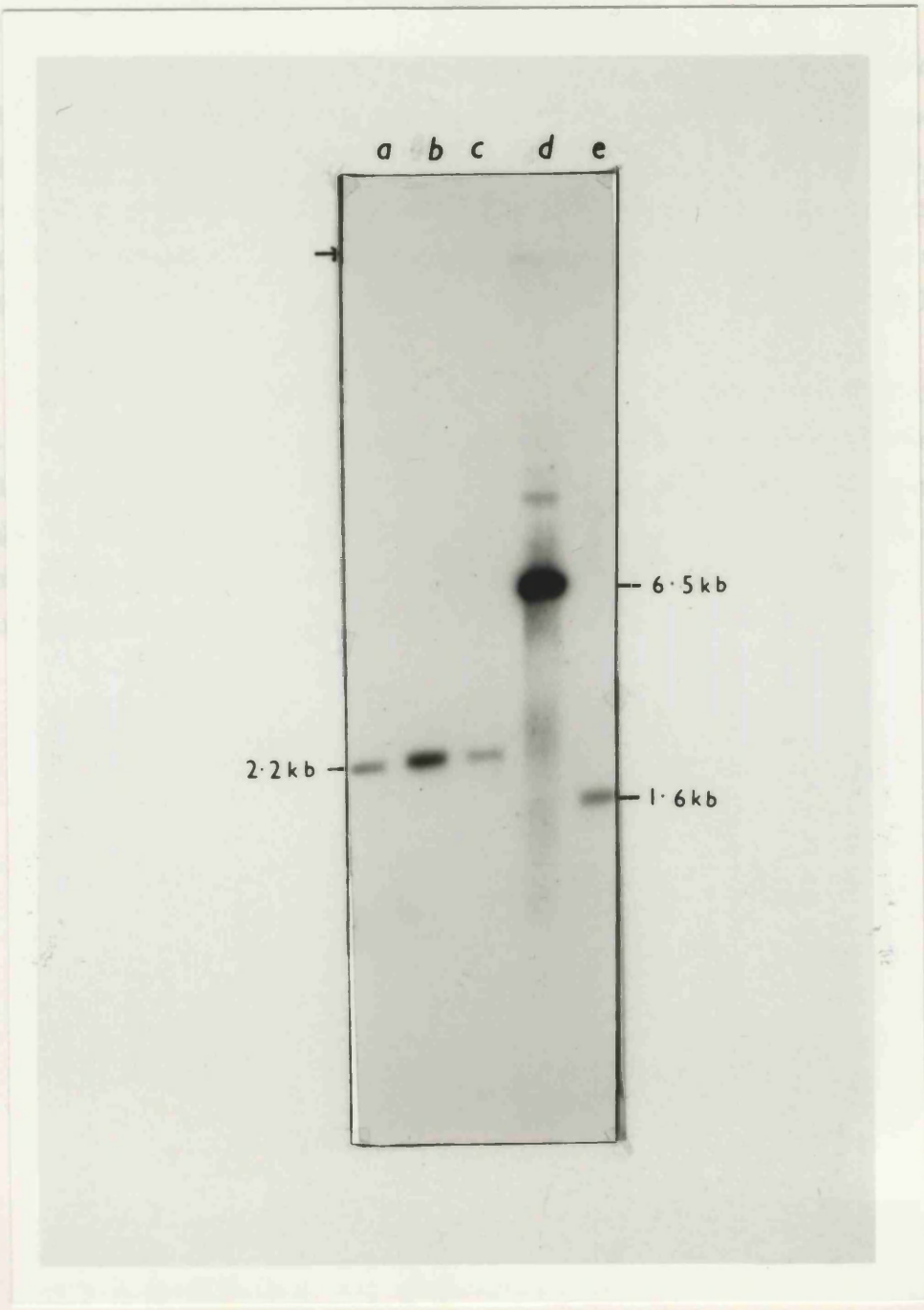



Figure 12B. Southern blot of agarose gel probed with the full-length labelled rat cytochrome $b_5$ cDNA. Final wash at 50°C. Lane a - $Rgb_5$7, lane b - $Rgb_5$2, lane c - $Rgb_5$3, lane d - $Rgb_5$5, lane e - $Rgb_5$6. Sizes given in kilobase pairs. ⟶ marks the wells.

3.2.2 Subcloning of lambda EcoRI fragments into plasmid vectors, restriction mapping and Southern blot hybridization.

To enable the genomic clones to be mapped, they were digested with the restriction endonuclease EcoRI and the fragments subcloned into either of the plasmid vectors pSPT19 (Pharmacia) (Fig. 13a) or pUC19 (Fig. 13b) cut with EcoRI. The nomenclature of the EcoRI fragments in pSPT19 and pUC19 is shown in table 2.

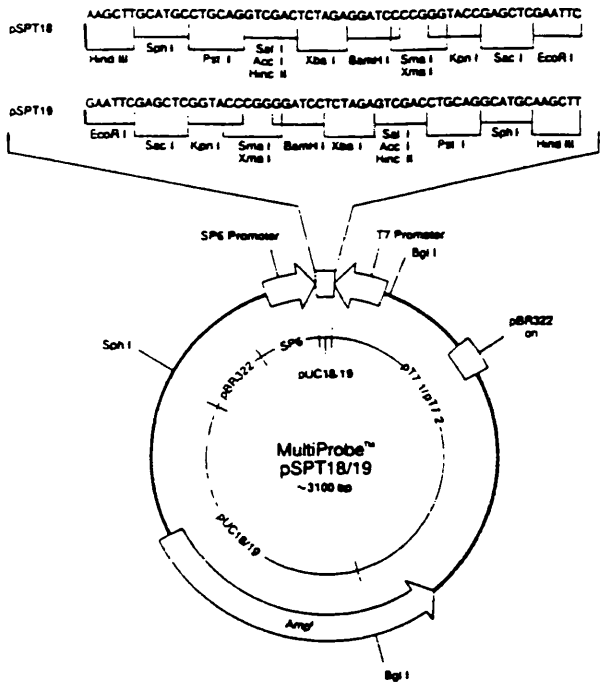
The larger EcoRI fragments in pSPT19 and pUC19 (6150bp and 6500bp) which hybridized to the rat cytochrome $b_5$ cDNA were restriction mapped by single and double enzyme digests. These were 1.3, derived from the $Rgb_5 1$ genomic clone (Fig. 14A(1)) and 6.5E, derived from the $Rgb_5 5$ genomic clone (Fig. 15) respectively. The similarity of the restriction map of subclone 6.5E (Fig. 15) to subclone 1.2 (Fig. 14A(2)) is discussed in section 3.2.3.2 and Fig. 26.

Analysis of 1.3 and 6.5E by Southern blot hybridization was then carried out to find smaller restriction fragments which contained coding regions. The full-length rat cytochrome $b_5$ cDNA insert from $pb_5(1)D$ was used as a probe (Figs. 16 and 17 respectively). Hybridization was under aqueous conditions containing 5xSSPE at 65°C with a final wash at 50°C for 15 minutes in 0.1xSSPE (section 2.10).

Table 2 Nomenclature of plasmid subclones.

| Rat cytochrome $b_5$ genomic clone | Size of EcoRI fragment (bp) | Nomenclature of fragment in plasmid pSPT19 |
|---|---|---|
| $Rgb_5 1$ | 6150 | 1.3 |
| $Rgb_5 1$ | 4700 | 1.2 |
| $Rgb_5 1$ | 200 | 1.24 |
| $Rgb_5 4$ | 6300 | 4.14 |
| $Rgb_5 4$ | 3000 | 4.20 |
| $Rgb_5 4$ | 870 | 4.10 |
| | | |
| Rat cytochrome $b_5$ genomic clone | Size of EcoRI fragment (bp) | Nomenclature of fragment in plasmid pUC19 |
| $Rgb_5 5$ | 6500 | 6.5E |
| $Rgb_5 2$ | 2200 | E1.2B |
| $Rgb_5 3/7$ | 2200 | E3.1o |
| $Rgb_5 3/7$ | 1300 | E3.1v |
| $Rgb_5 6$ | 1600 | H4.1T |

(a)




(b)

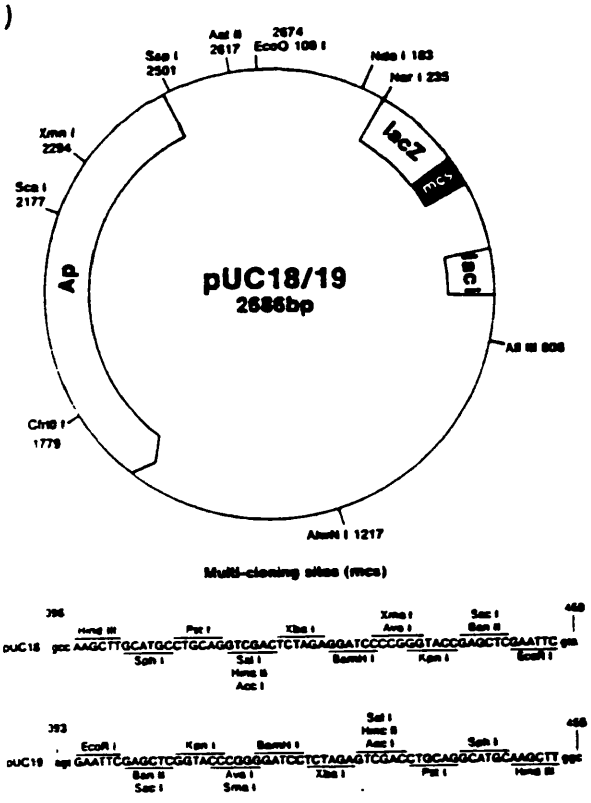
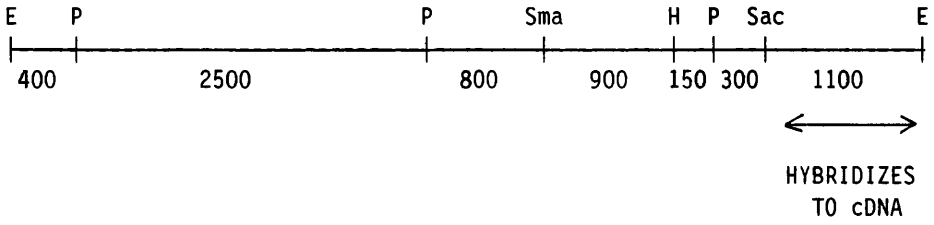


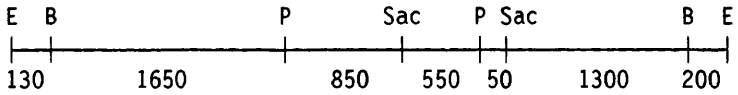

Figure 13(a). Map of the plasmid cloning vector pSPT19 (Pharmacia). The multiple cloning site is shown separately in detail.

Figure 13(b). Map of the plasmid cloning vector pUC19. The multiple cloning site is shown separately in detail.

(1) 1.3 (6150bp)

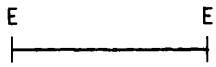



(2) 1.2 (4700bp)




does not hybridize to cDNA

(3) 1.24 (200bp)




does not cleave with Hind III

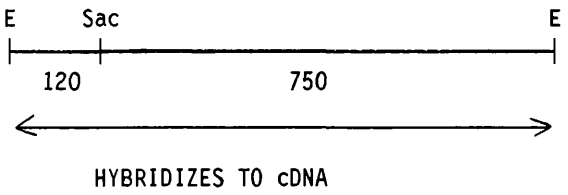
does not hybridize to cDNA

Figure 14A

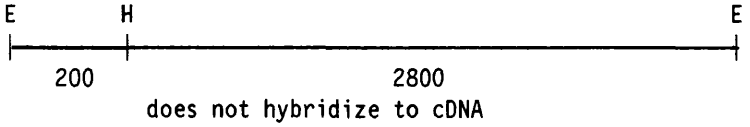
Restriction maps of $Rgb_5 1$ subclones in pSPT19 including fragments hybridizing to rat cytochrome $b_5$ cDNA. Sizes given are in base pairs.

E = EcoRI, P = PstI, Sma = SmaI, H = HindIII, Sac = SacI, B = BamHI

(1) 4.10 (870 bp)




(2) 4.20 (3000bp)




(3) 4.14 (6300bp)

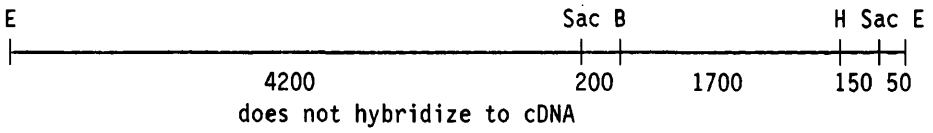



Figure 14B

Restriction maps of $Rgb_5$4 subclones in pSPT19 including fragments hybridizing to rat cytochrome $b_5$ cDNA . Sizes given are in base pairs.

E = EcoRI, Sac = SacI, H = HindIII, B = BamHI.

(1) 6.5E (6500bp)

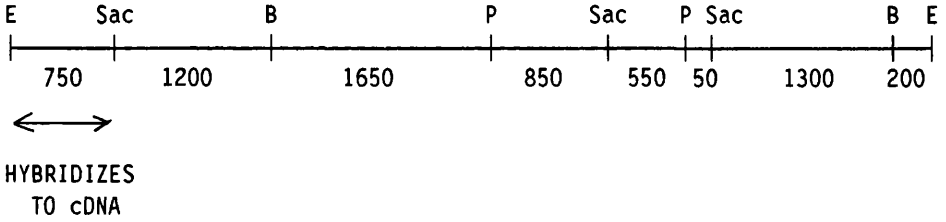



Figure 15

Restriction map of $Rgb_5$5 subclone in pUC19 including fragments hybridizing to rat cytochrome $b_5$ cDNA. Sizes given are in base pairs.

E = EcoRI, Sac = SacI, B = BamHI, P = PstI

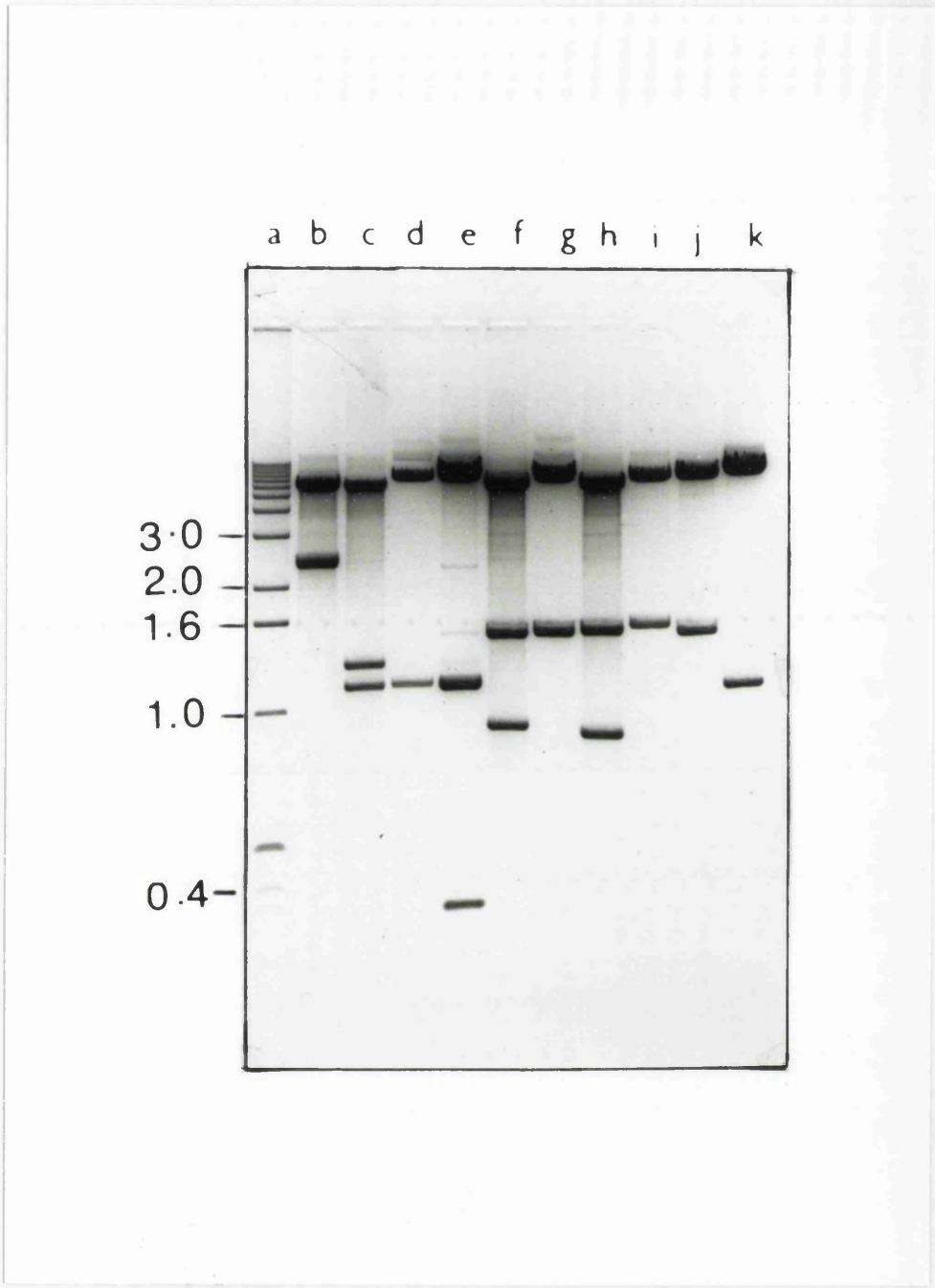



Figure 16A. Photograph of subclone 1.3, from the lambda genomic clone $Rgb_5$1, digested with various enzymes (single and double digests) run out on a 1.2% agarose gel. Lane a - 1kb (BRL marker), lane b - 1.3 digested with SmaI, lane c - 1.3 digested with SacI and SmaI, lane d - 1.3 digested with SacI, lane e - 1.3 digested with SacI and SalI, lane f - 1.3 digested with SmaI and SalI, lane g - 1.3 digested with SalI, lane h - 1.3 digested with SmaI and HindIII, lane i - 1.3 digested with HindIII, lane j - 1.3 digested with HindIII and SalI, lane k - 1.3 digested with SacI and HindIII. Sizes given in kilobase pairs.

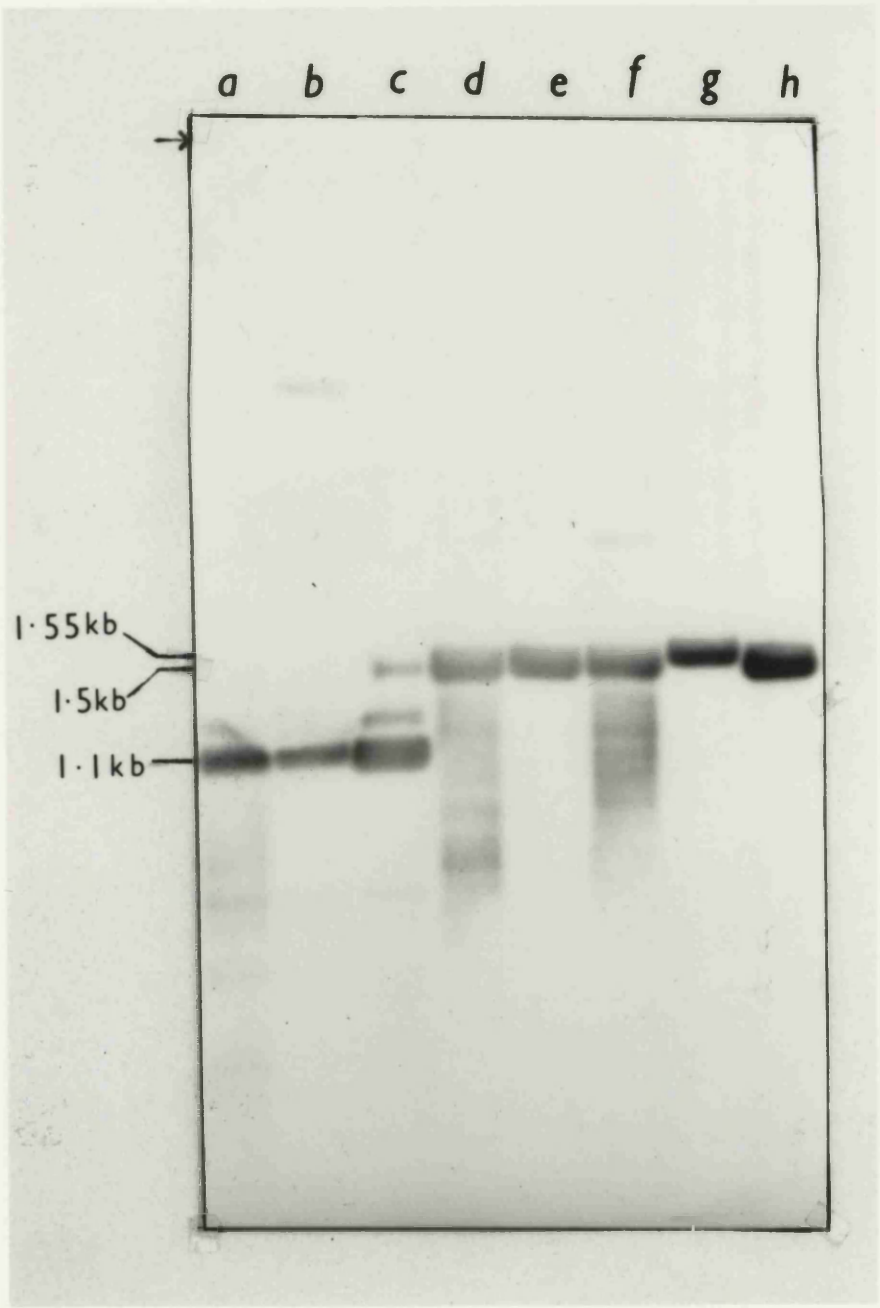



Figure 16B. Southern blot of the gel from Figure 16A probed with the full-length rat cytochrome $b_5$ cDNA. Final wash at 50°C. Lane a - 1.3 digested with SacI and SmaI, lane b - 1.3 digested with SacI, lane c - 1.3 digested with SacI and SalI, lane d - 1.3 digested with SmaI and SalI, lane e - 1.3 digested with SalI, lane f - 1.3 digested with SmaI and HindIII, lane g - 1.3 digested with HindIII, lane h - 1.3 digested with HindIII and SalI. Sizes shown are in kilobase pairs (kb). ⟶ marks the wells.

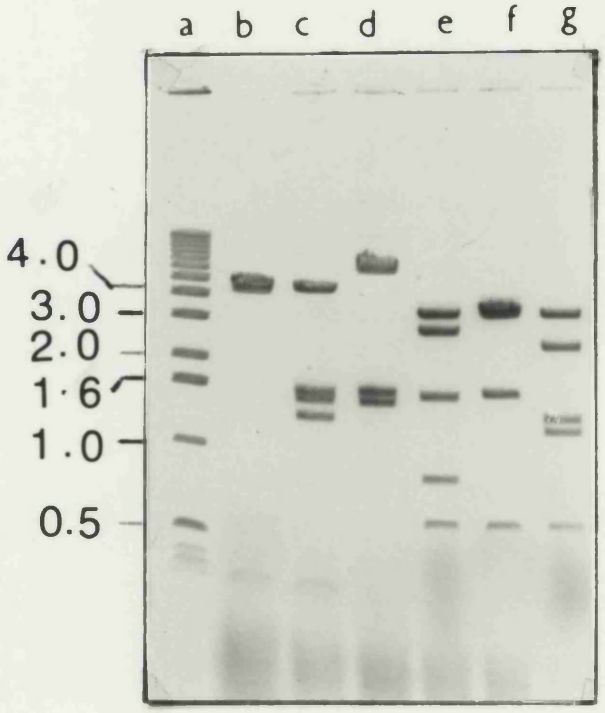



Figure 17A. Photograph of subclone 6.5E, from the lambda genomic clone $Rgb_5 5$, digested with various restriction endonucleases and run out on a 0.8% agarose gel. Lane a - 1kb (BRL marker), lane b - 6.5E digested with BamHI, lane c - 6.5E digested with PstI and BamHI, lane d - 6.5E digested with PstI, lane e - 6.5E digested with SacI and PstI, lane f - 6.5E digested with SacI, lane g - 6.5E digested with SacI and BamHI. Sizes shown are in kilobase pairs (kb).

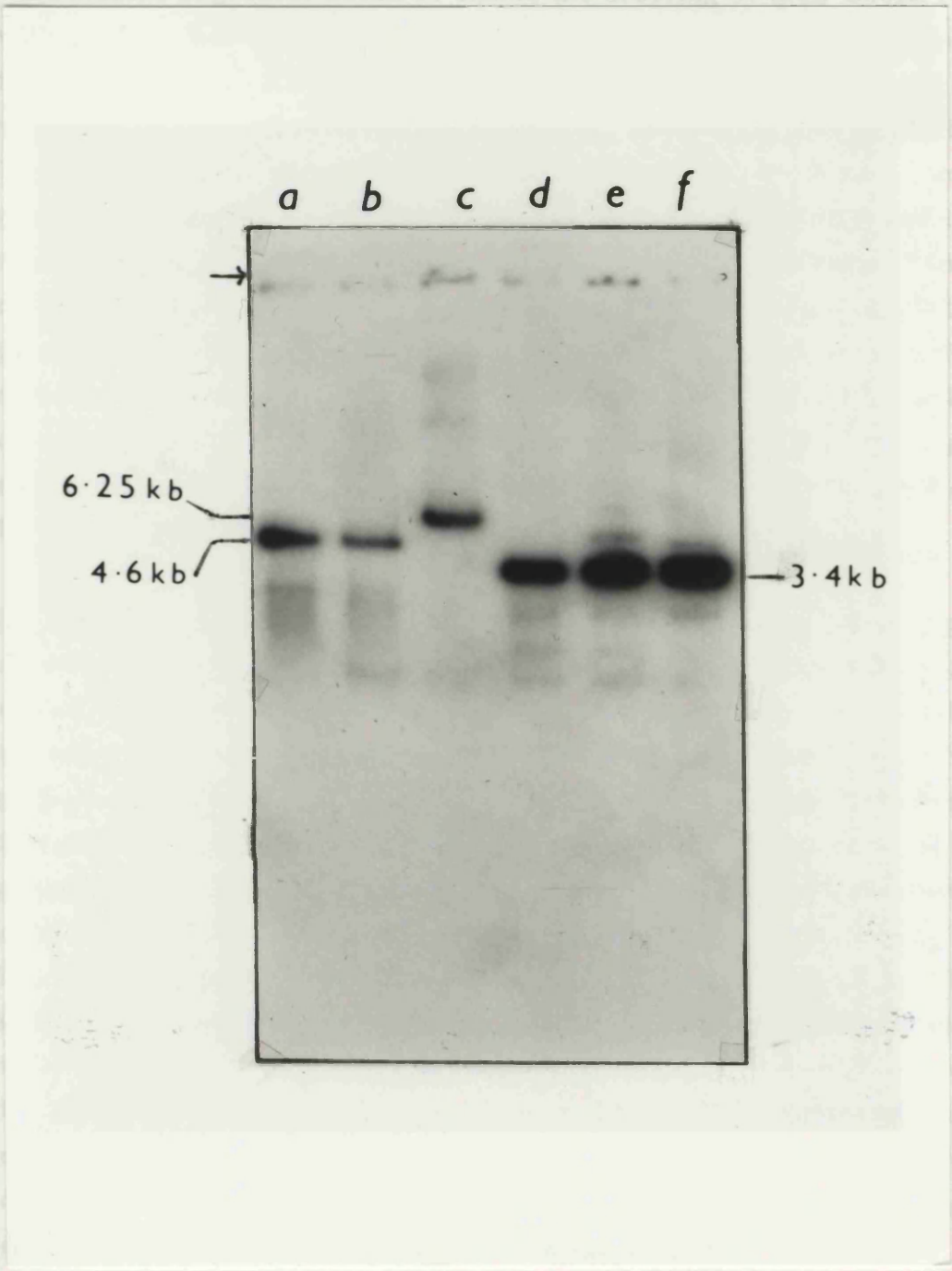



Figure 17B. Southern blot of the gel from Figure 17A probed with the full-length rat cytochrome $b_5$ cDNA. Final wash at 50°C. Lane a - 6.5E digested with BamHI, lane b - 6.5E digested with PstI and BamHI, lane c - 6.5E digested with PstI, lane d - 6.5E digested with SacI and PstI, lane e - 6.5E digested with SacI, lane f - 6.5E digested with SacI and BamHI. Sizes shown are in kilobase pairs (kb). ——→ marks the wells.

$Rgb_5$1 and $Rgb_5$4

Before genomic clones other than $Rgb_5$1 and $Rgb_5$4 were isolated and whilst the libraries were being rescreened, the genomic clones $Rgb_5$1 and $Rgb_5$4 were studied in greater detail to enable the ordering of their EcoRI fragments within the lambda arms. This was carried out by digesting all the EcoRI fragments subcloned into pSPT19 - 1.3, 1.2, 1.24 (Fig. 14A) and 4.10, 4.20, 4.14 (Fig. 14B) - with the restriction endonucleases SacI, BamHI and HindIII. The genomic clones in lambda Charon 4A were also digested with these enzymes and Southern blots of the agarose gels probed with the full-length cytochrome $b_5$ cDNA clone (data not shown). From a knowledge of the restriction endonuclease sites present in the lambda arms and the individual plasmid subclone EcoRI genomic fragments, the restriction maps of $Rgb_5$1 and $Rgb_5$4 genomic clones were produced (Figs. 18 and 19 respectively). These detailed restriction maps show that the genomic clones $Rgb_5$1 and $Rgb_5$4 do not overlap. The possibility that the genomic clone $Rgb_5$5 overlaps $Rgb_5$1 and $Rgb_5$4 is discussed in section 3.2.3.2 and Fig. 26.

3.2.3 Southern blot hybridization of clones using oligonucleotides as probes.

As there were such small areas of coding region within large lambda genomic clones, it was a possibility that more exons existed within them which were too short to hybridize to the cDNA probe. The genomic clones and/or their plasmid subclones were screened with each one of four oligonucleotides - the 26mer, the 27mer, the 15mer and the 36mer (Fig. 20). The subclones from the genomic clones $Rgb_5$1 and $Rgb_5$4 were analysed with the oligonucleotides after the restriction fragments hybridizing to the rat cytochrome $b_5$ cDNA had been sequenced (the 1.1kb SacI fragment from subclone 1.3 and the 870bp EcoRI fragment from subclone 4.10; section 3.2.4). The genomic clones isolated later on, $Rgb_5$2, $Rgb_5$3/7, $Rgb_5$5 and $Rgb_5$6, were digested with EcoRI and hybridized with all the oligonucleotides. This allowed the location of various regions of rat cytochrome $b_5$ coding sequence within the genomic clones to be quickly ascertained. The subclones were then studied in greater detail.

The clones or subclones which hybridized to any of these oligonucleotides are discussed below.

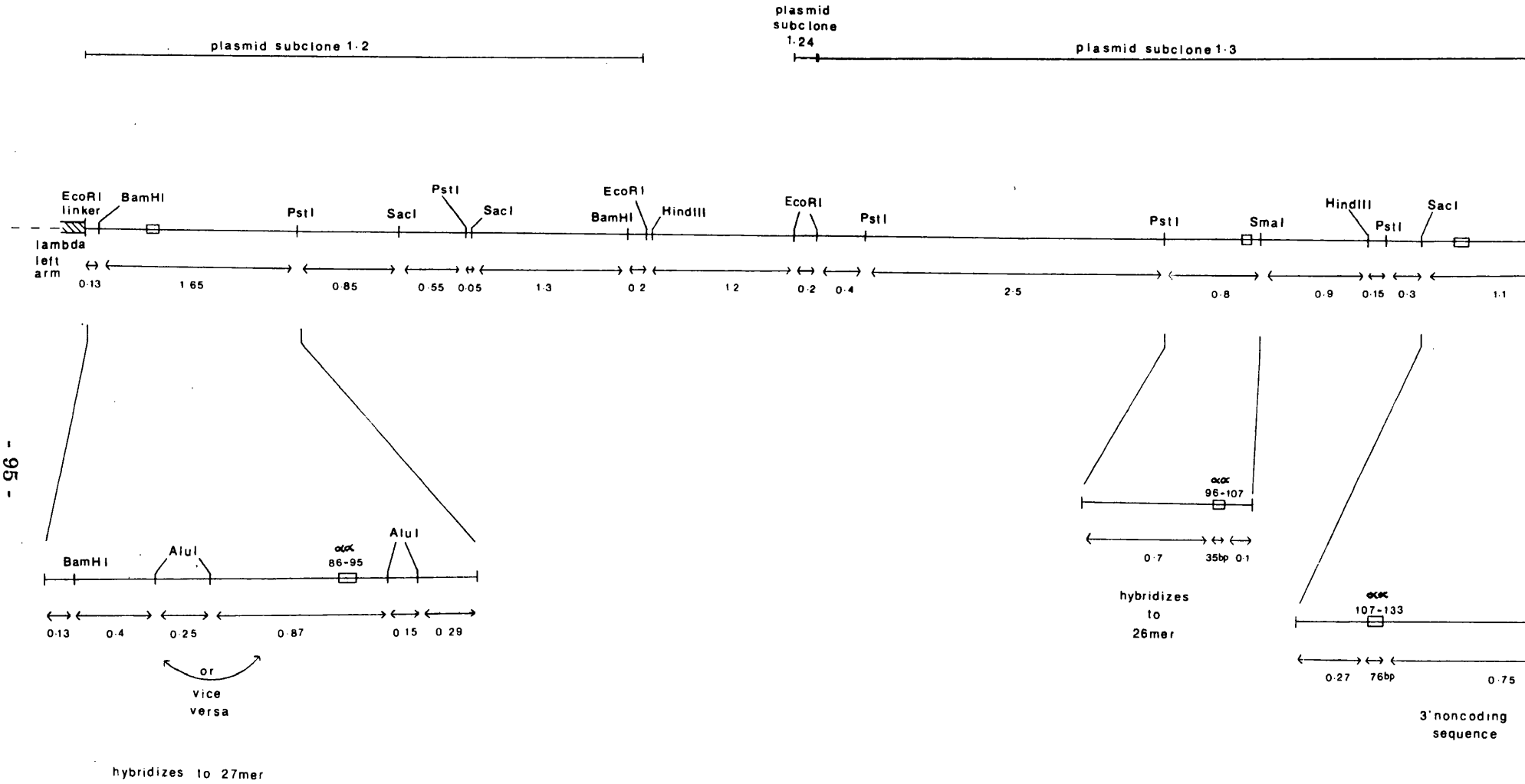



Figure 18. Restriction map of genomic clone, $Rgb_5 1$, in lambda Charon 4A. Sizes given in kilobases unless otherwise stated. Approximately to scale, except for exons (open boxes). Numbers above exons indicate amino acid position. Shaded boxes indicate lambda DNA.

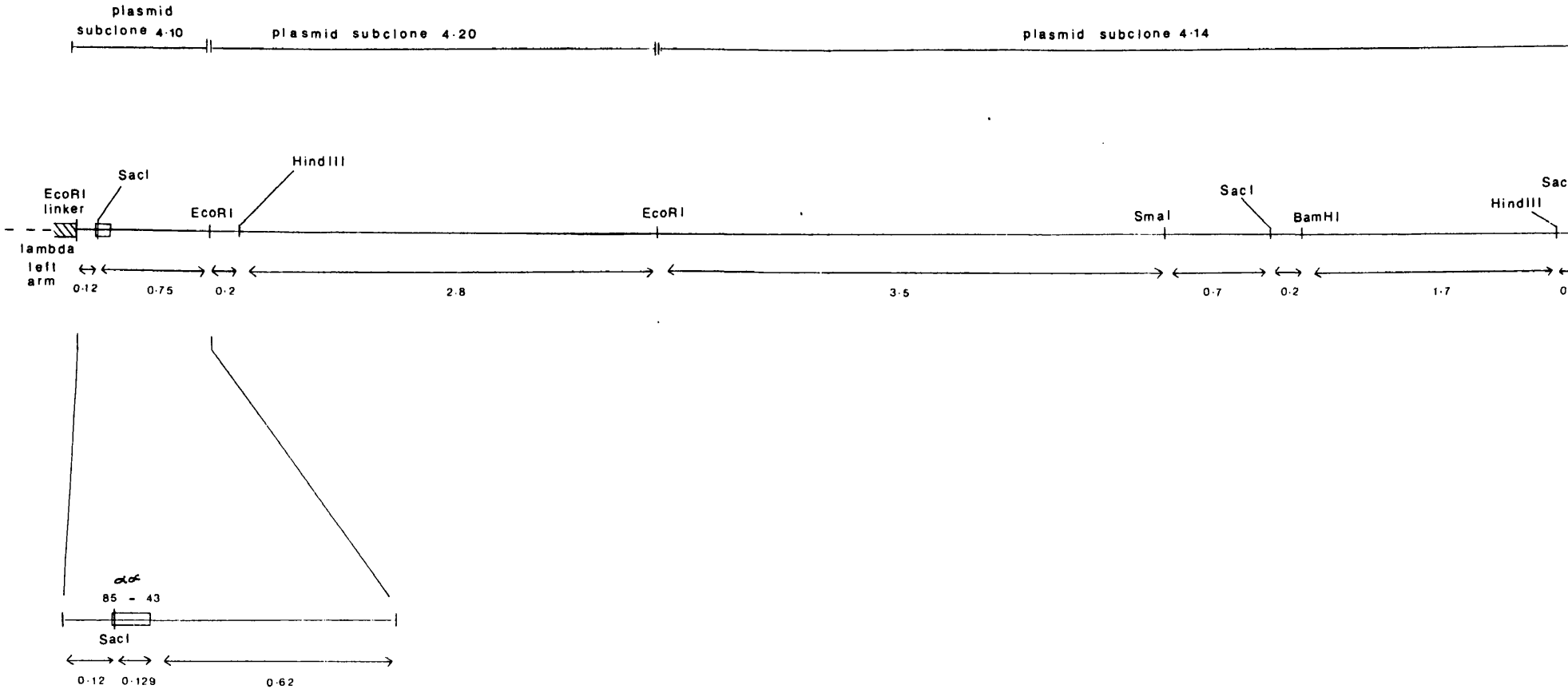



Figure 19. Restriction map of genomic clone, $Rgb_5$4, in lambda Charon 4A. Sizes given in kilobases unless otherwise stated. Approximately to scale, except for exons (open boxes). Numbers above exons indicate amino acid position. Shaded boxes indicate lambda DNA.

3.2.3.1 Oligonucleotide probes

Five specifically designed oligonucleotides, end-labelled with [$\gamma^{32}$P]-ATP (section 2.11), were used. The sequences of the oligonucleotides are shown in Fig. 20 and their positions within the rat cytochrome $b_5$ cDNA sequence in Fig. 9.

The 27mer (Fig. 20), running from amino acids 87 to 95 inclusive, is a specific probe for the hydrophilic segment of cytochrome $b_5$.

A 26mer oligonucleotide (Fig. 20), running from amino acid 98 to within amino acid 106, provides a probe specific for the 3' hydrophobic segment of cytochrome $b_5$.

An 18mer oligonucleotide (Fig. 20) running from amino acid 98 to amino acid 103 inclusive, is a shorter probe specific for the hydrophobic segment of cytochrome $b_5$.

The fourth oligonucleotide is a 36mer (Fig. 20) which codes for amino acids 4 to 15 inclusive. Clones were screened for the presence or absence of the 5' end of the coding region with the 36mer.

The final oligonucleotide, a 15mer (Fig. 20) coding for amino acids 58 to 62 inclusive of the rat cytochrome $b_5$ protein was also used to identify exons in genomic clones.

(1) 15mer - antisense strand. $T_d$ = 36°C.

```
        62              58
5'      AAA GTT CTC AGT AGC     3'
```

(2) 18mer - sense strand. $T_d$ = 47°C.

```
      98                  103
5'    CTT ATC ACT ACT GTC GAG    3'
```

(3) 26mer - complementary strand. $T_d$ = 70°C.

```
      98                                  106
5'    GAA TTA GAC TCG ACA GTA GTG ATA AG    3'
```

(4) 27mer - antisense strand. $T_d$ = 71°C.

```
      95                                  87
5'    CGA AGG CTT GGC TAT CTT TGA TCT GTC   3'
```

(5) 36mer - sense strand. $T_d$ = 79°C.

```
       4                                                 15
5'    TCA GAC AAG GAT GTG AAG TAC TAC ACT CTG GAG GAG    3'
```

Figure 20. Sequences of the oligonucleotides used with their dissociation temperatures ($T_d$). The numbers indicate amino acid position.

3.2.3.2 Southern blot hybridization with oligonucleotide probes

Restriction endonuclease digests of the genomic clones in lambda Charon 4A or of the subclones in plasmid vectors were blotted by the method of Southern (1975) and probed with the oligonucleotides described above (section 3.2.3.1). Stringent hybridization and washing conditions were used after calculation of the temperature of dissociation of the duplex ($T_d$) (section 2.11.1), unless otherwise stated.

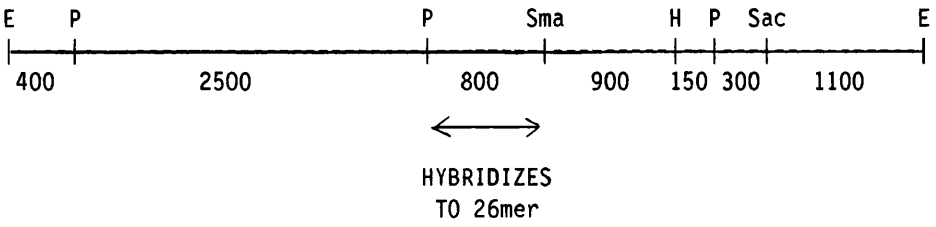
$Rgb_5$1 derived subclones

Southern blot hybridization using the 26mer (amino acids 98-106) and 27mer (amino acids 87-95) oligonucleotides revealed additional information on the $Rgb_5$1 derived subclones 1.2 and 1.3 in the plasmid vector pSPT19 (Fig. 21).

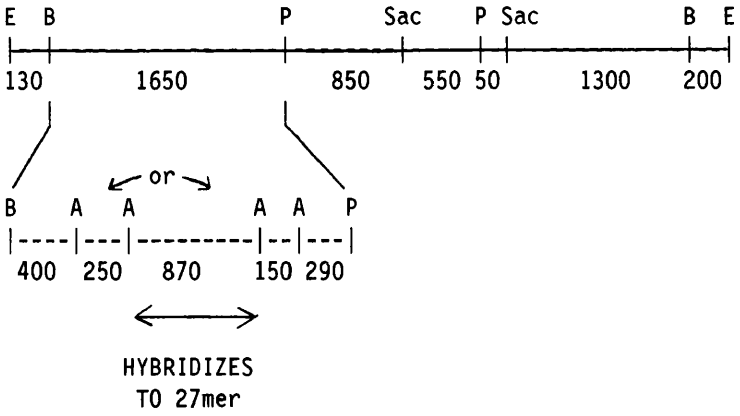
26mer - the 26mer oligonucleotide (amino acids 98-106), hybridized to the 800bp PstI/SmaI fragment of the 6.15kb EcoRI subclone in pSPT19 (subclone 1.3 - Fig. 21(1)) at a hybridization temperature of 65°C ($T_d$ -5°C) and with a final stringent wash of 2 minutes in 6xSSC at 70°C ($T_d$) (Fig. 22A).

27mer - the 27mer oligonucleotide (amino acids 87-95) hybridized to the 1800bp PstI fragment of the 4.7kb EcoRI subclone in pSPT19 (subclone 1.2 - Fig. 21(2)) at a hybridization temperature of 66°C ($T_d$ -5°C) and with a final stringent wash of 2 minutes in 6xSSC at 71°C ($T_d$) (Fig. 22B).

(1) 1.3 (6150bp)




(2) 1.2 (4700bp)




(3) 4.10 (870bp)

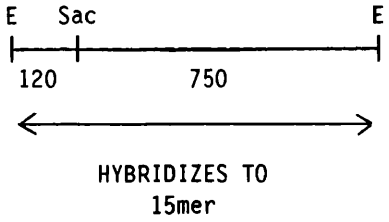



Figure 21

Restriction maps of $Rgb_5$1 and $Rgb_5$4 subclones in pSPT19 including fragments hybridizing to the 26mer, 27mer and 15mer oligonucleotides. Sizes given are in base pairs.

E = EcoRI, P = PstI, Sma = SmaI, H = HindIII, Sac = SacI, B = BamHI, A = AluI

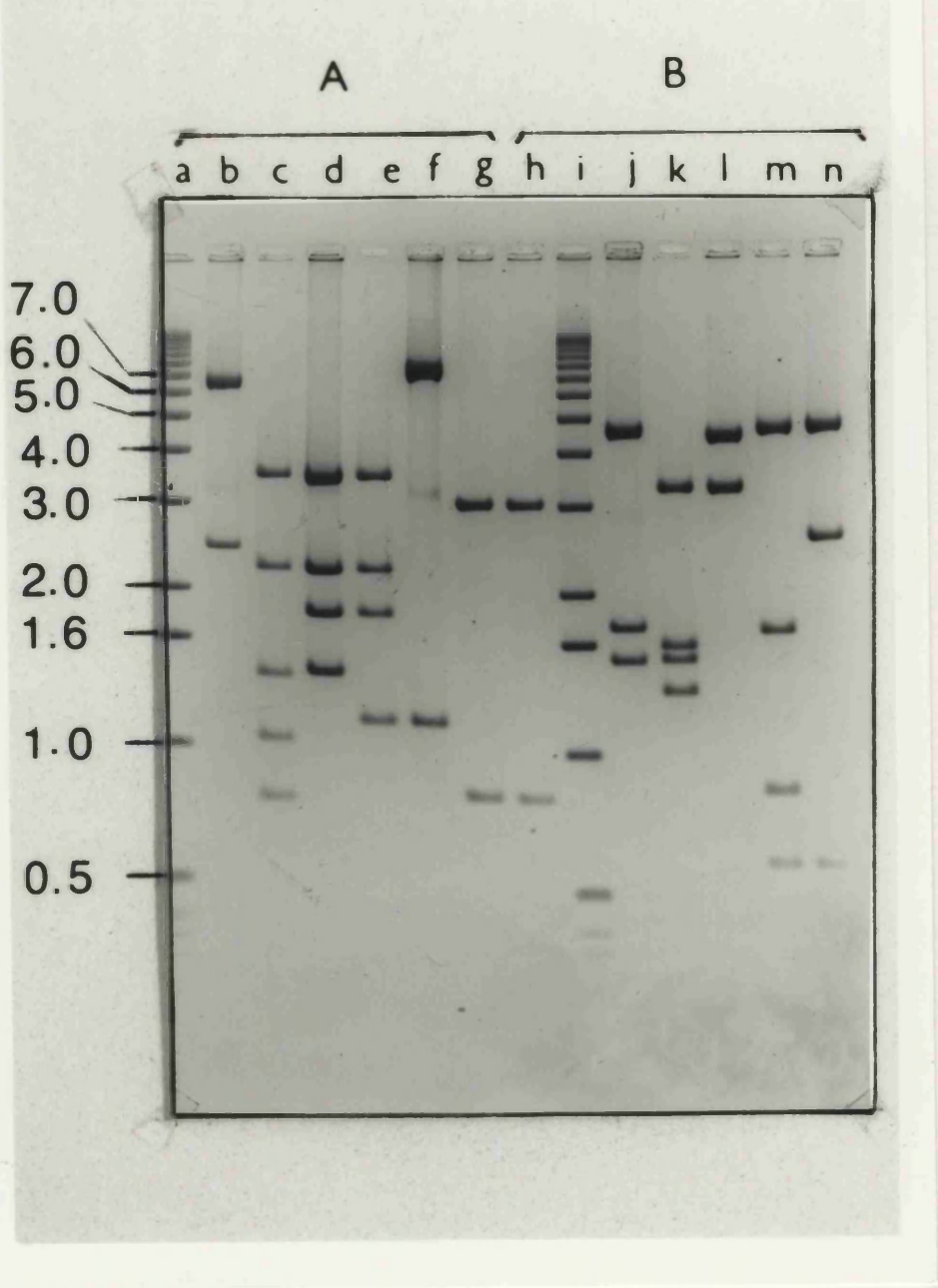



Figure 22(1). Photograph of the subclones 1.3(A) and 1.2(B), from the lambda genomic clone $Rgb_5 1$, digested with various restriction endonucleases and run out on a 0.8% agarose gel. A positive control of the rat cytochrome $b_5$ cDNA is included ($pb_5(1)D$).

A) Lane a - 1kb ladder (BRL marker), lane b - 1.3 digested with SmaI, lane c - 1.3 digested with SmaI and PstI, lane d - 1.3 digested with PstI, lane e - 1.3 digested with SacI and PstI, lane f - 1.3 digested with SacI, lane g - $pb_5(1)D$ digested with EcoRI.

B) Lane h - $pb_5(1)D$ digested with EcoRI, lane i - 1kb ladder (BRL marker), lane j - 1.2 digested with PstI, lane k - 1.2 digested with PstI and BamHI, lane l - 1.2 digested with BamHI, lane m - 1.2 digested with SacI and PstI, lane n - 1.2 digested with SacI.

Sizes are in kilobase pairs (kb).

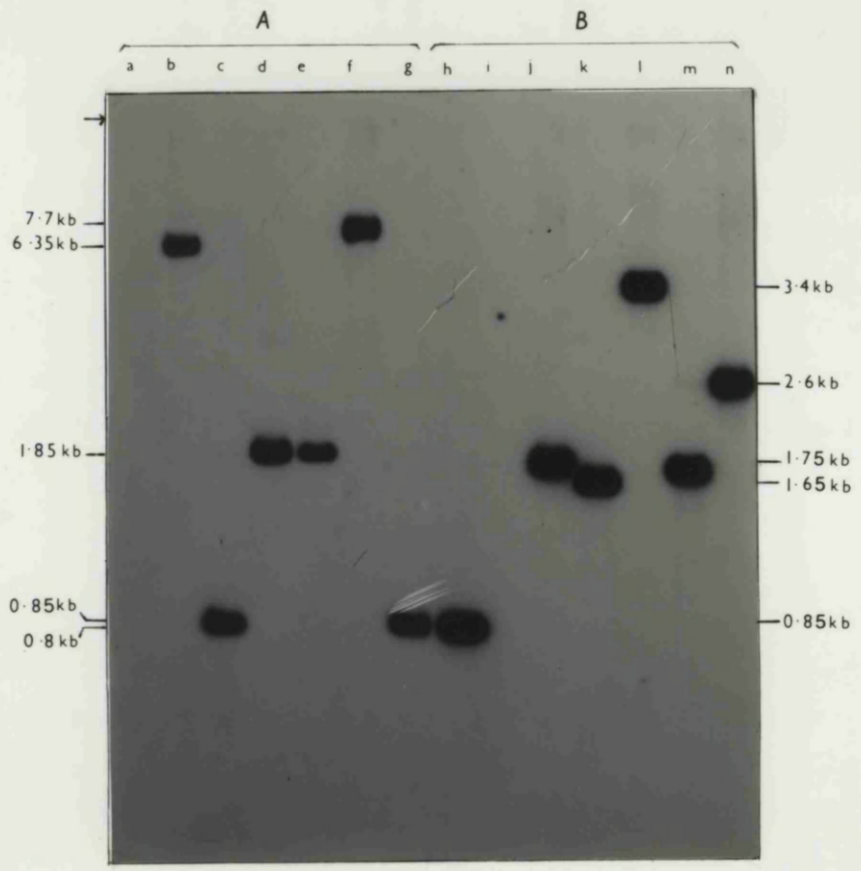



Figure 22(2)A. Southern blot of the gel from Figure 22(1)A, probed with the end-labelled 26mer oligonucleotide. Hybridized at $T_d$ -5°C. Lane a - 1kb ladder (BRL marker), lane b - 1.3 digested with SmaI, lane c - 1.3 digested with SmaI and PstI, lane d - 1.3 digested with PstI, lane e - 1.3 digested with SacI and PstI, lane f - 1.3 digested with SacI, lane g - $pb_5(1)D$ digested with EcoRI.

Figure 22(2)B. Southern blot of the gel from Figure 22(1)B, probed with the end-labelled 27mer oligonucleotide. Hybridized at $T_d$ -5°C. Lane h - $pb_5(1)D$ digested with EcoRI, lane i - 1kb ladder (BRL marker), lane j - 1.2 digested with PstI, lane k - 1.2 digested with PstI and BamHI, lane l - 1.2 digested with BamHI, lane m - 1.2 digested with SacI and PstI, lane n - 1.2 digested with SacI. Sizes are in kilobase pairs (kb). ⟶ marks the wells.

Therefore, together, these two $Rgb_5$1 derived EcoRI fragments - the 800bp PstI/SmaI fragment from 1.3 and the 1800bp PstI fragment from 1.2 - must contain at least amino acids 87 to 106 in two exons separated by an intron greater than 7kb in length (Fig. 18) The full restriction maps of subclones 1.3 and 1.2 indicating the sites of oligonucleotide hybridization are as shown in Fig. 21.

The rat cytochrome $b_5$ cDNA clone did not hybridize to the 800bp PstI/SmaI fragment of subclone 1.3 nor to the 1800bp PstI fragment of subclone 1.2 under the usual stringent hybridization conditions of 65°C in aqueous solution/5xSSPE with a final stringent wash of 15 minutes at 50°C in 0.1xSSPE (section 2.10). This was due to the fact that the two exons are very short. The lengths of the exons as determined by DNA sequencing were 35 and 30 bp respectively (section 3.2.4).

$Rgb_5$4 derived subclones

15mer - the 15mer oligonucleotide (amino acids 58-62) hybridized to the 870bp EcoRI fragment in pSPT19 (subclone 4.10, Fig. 21(3)) under hybridization conditions of 31°C ($T_d$ -5°C) and with a final wash for 2 minutes in 6xSSC at 36°C ($T_d$) (data not shown).

$Rgb_5$5 derived subclones

27mer - the 27mer oligonucleotide (amino acids 87-95) hybridized to the 6.5kb EcoRI fragment of the $Rgb_5$5 genomic clone (Fig. 23). This oligonucleotide was used to probe a Southern blot of restriction digests of this fragment subcloned into pUC19 (6.5E), run out on an agarose gel (Fig. 24). The restriction map of the $Rgb_5$5 subclone 6.5E in relation to the 27mer oligonucleotide hybridization, can be seen in Fig. 25(1).

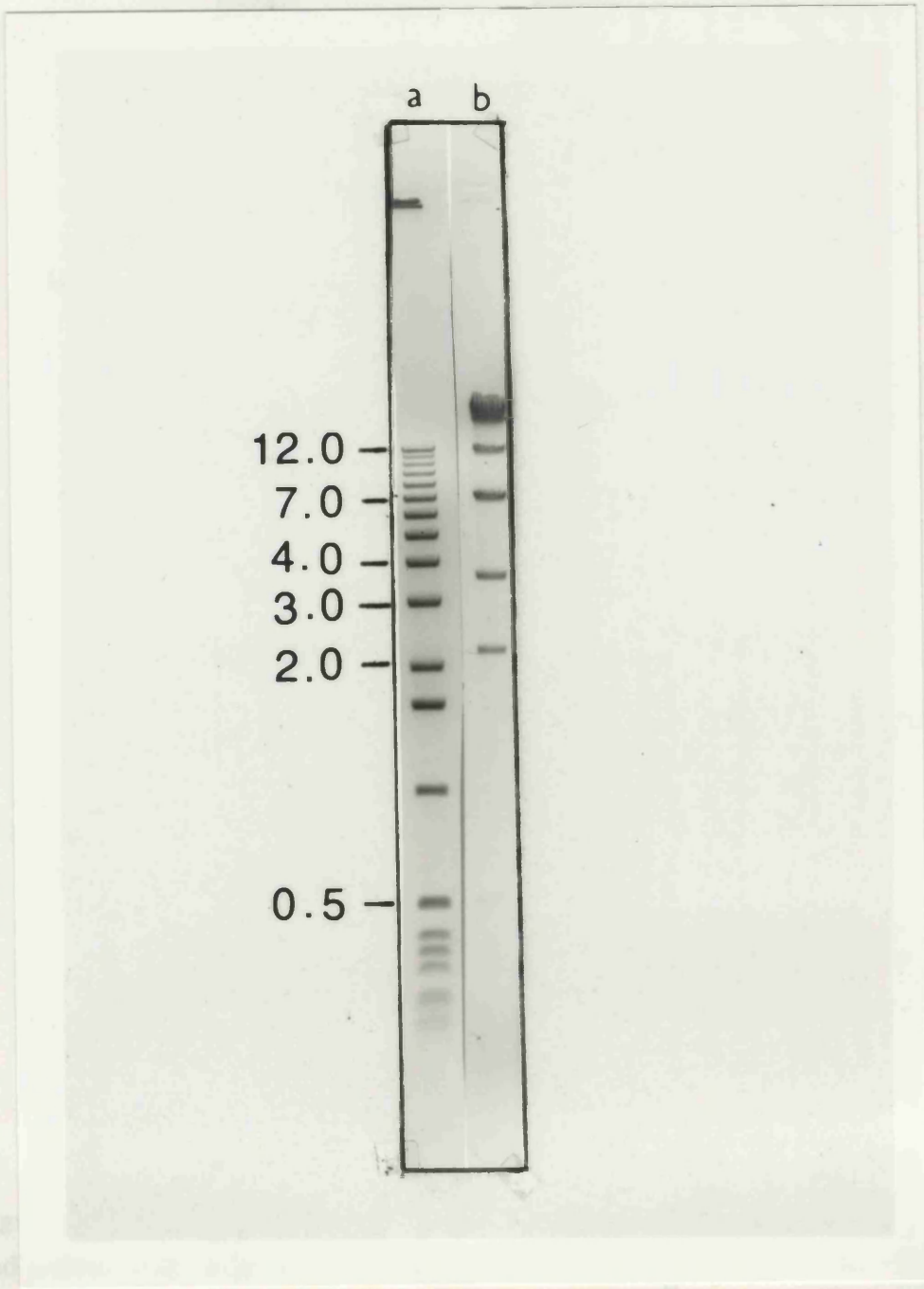



Figure 23(1). Photograph of the lambda genomic clone, $Rgb_5 5$, digested with EcoRI and run out on a 0.6% agarose gel. Lane a - 1kb (BRL marker), lane b - $Rgb_5 5$ digested with EcoRI. Sizes are in kilobase pairs (kb).

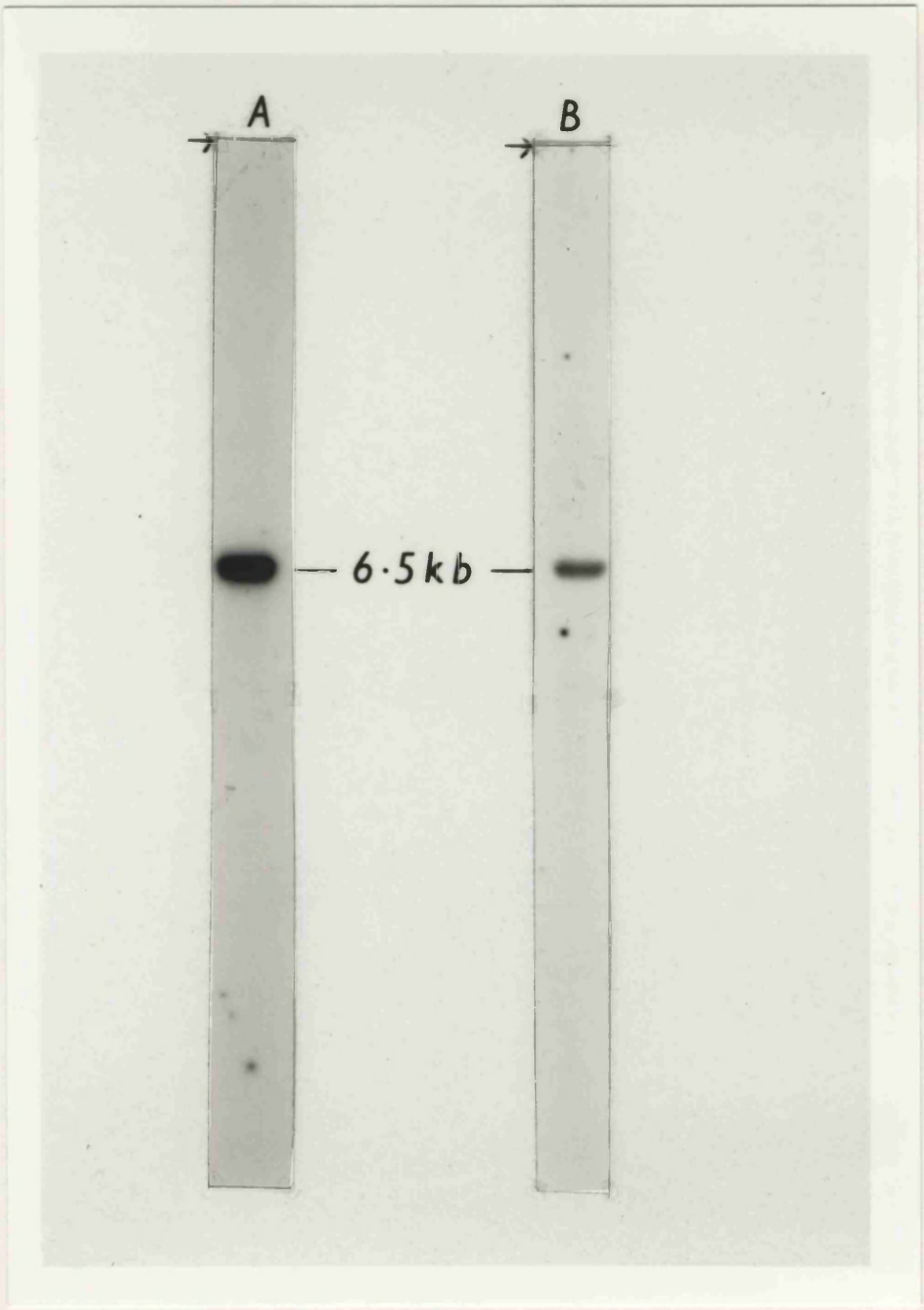



Figure 23(2)A. Southern blot of the gel from Figure 23(1), probed with the end-labelled 27mer oligonucleotide. Hybridized at $T_d$ -5°C. Lane A - $Rgb_5 5$ digested with EcoRI.

Figure 23(2)B. Southern blot of the gel from Figure 23(1), probed with the end-labelled 15mer oligonucleotide. Final wash at $T_d$ -5°C. Lane B - $Rgb_5 5$ digested with EcoRI.

Sizes are in kilobase pairs (kb). ⟶ marks the wells.

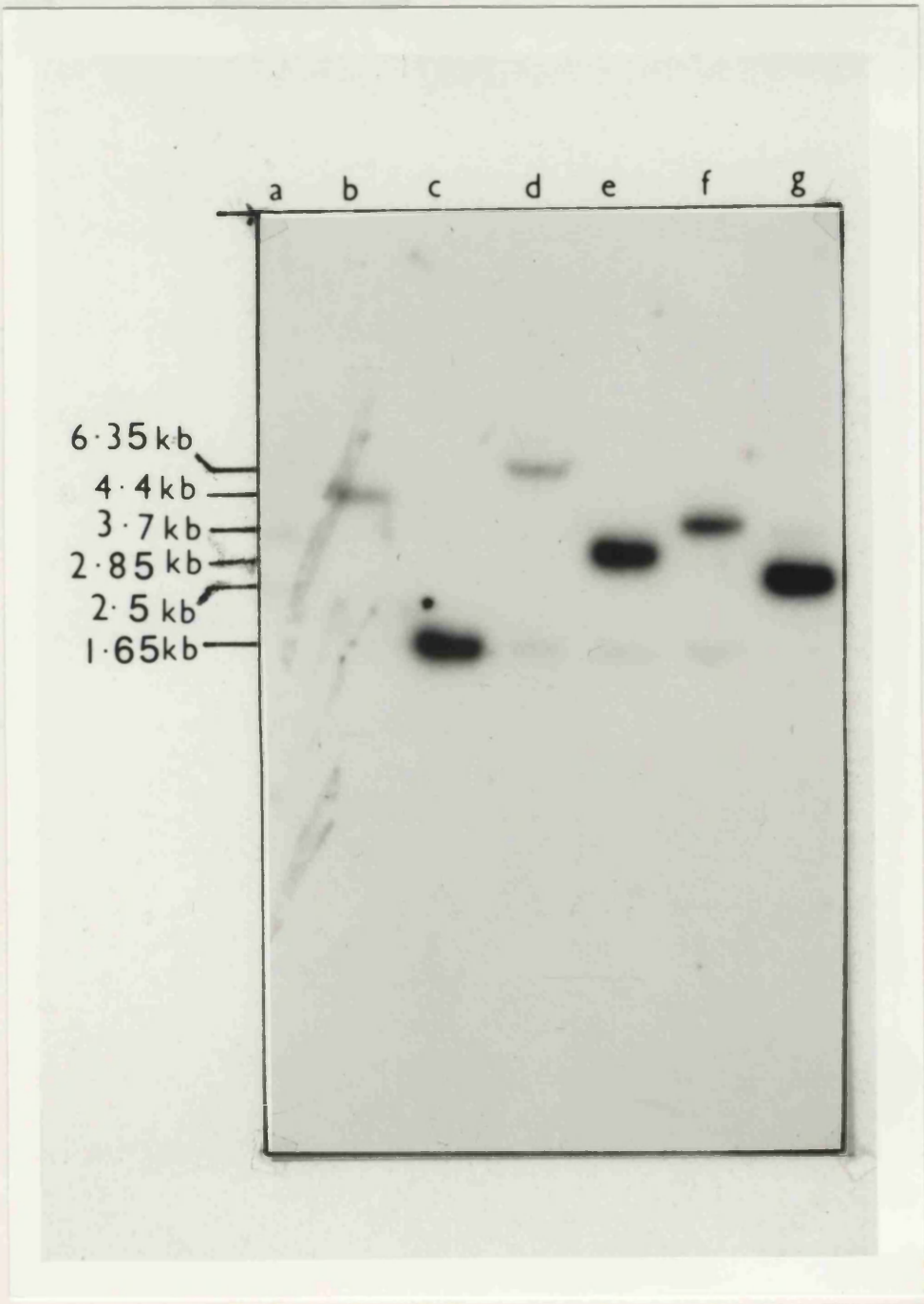
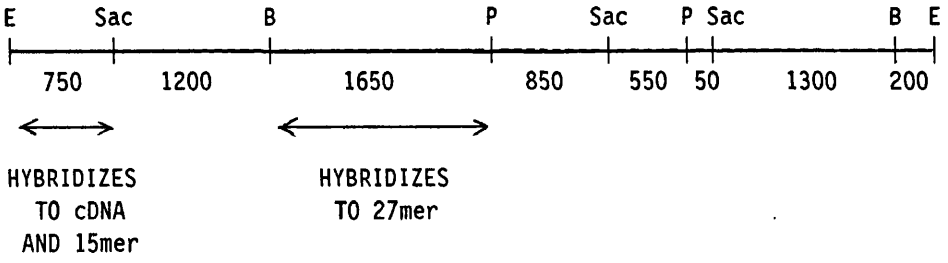
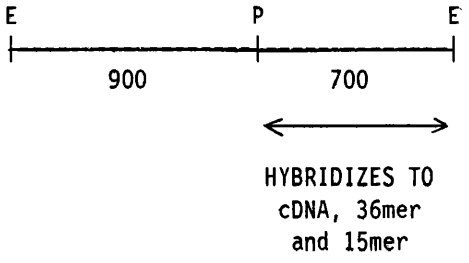



Figure 24. Southern blot of the agarose gel from Figure 17A, probed with the end-labelled 27mer oligonucleotide. Hybridized at $T_d$ -5°C. Lane a - 1kb (BRL marker), lane b - 6.5E digested with BamHI, lane c - 6.5E digested with PstI and BamHI, lane d - 6.5E digested with PstI, lane e - 6.5E digested with SacI and PstI, lane f - 6.5E digested with SacI, lane g - 6.5E digested with SacI and BamHI. Sizes are in kilobase pairs (kb). ⟶ marks the wells.

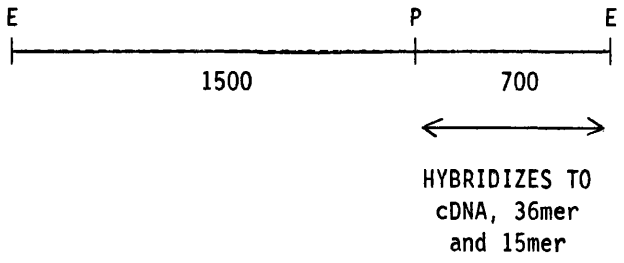
(1) 6.5E from $Rgb_5$5 genomic clone (6500bp)



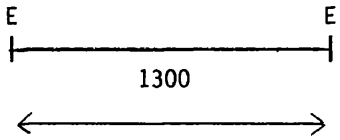

(2) H4.1T from $Rgb_5$6 genomic clone (1600bp)




(3) E3.1o from $Rgb_5$3/7 genomic clones (2200bp)

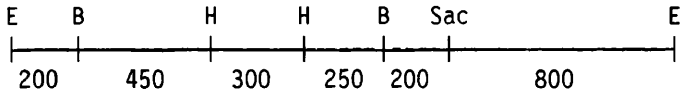


(4) E3.1v from $Rgb_5$3/7 genomic clones (1300bp)




HYBRIDIZES TO cDNA (under reduced stringency -
hybridize 60°C/final wash 50°C in 0.1xSSPE)

(5) E1.2B from $Rgb_5$2 genomic clone (2200bp)




An EcoRI fragment of this size hybridizes to the rat cytochrome $b_5$ cDNA and 36mer only when in the lambda clone. On subcloning into a plasmid vector, this fragment (E1.2B) does not hybridize to these two probes.

Figure 25.

Restriction maps of EcoRI fragments from $Rgb_5$ clones in pUC19 including fragments hybridizing to rat cytochrome $b_5$ cDNA and oligonucleotides. Sizes given are in base pairs.

E = EcoRI, Sac = SacI, B = BamHI, P = PstI, H = HindIII

15mer - the 15mer oligonucleotide (amino acids 58-62) hybridized to the 6.5kb EcoRI fragment of the $Rgb_5$5 genomic clone (Fig. 23) under hybridization conditions of 31°C ($T_d$ -5°C) and with a final wash of 2 minutes in 6xSSC at 36°C ($T_d$). The restriction map of the $Rgb_5$5 subclone 6.5E in relation to the 15mer oligonucleotide hybridization can be seen in Fig. 25(1).

Therefore, the $Rgb_5$5 genomic clone contains a 6.5kb EcoRI fragment to which both the 15mer and 27mer hybridize.

Comparing the genomic EcoRI subclone 6.5E (from $Rgb_5$5) with the two subclones 4.10 (from $Rgb_5$4) and 1.2 (from $Rgb_5$1) shows that 6.5E is identical in terms of a restriction map and hybridization pattern to 4.10 and 1.2 together (Fig. 26). The full restriction maps of the lambda clones $Rgb_5$1 (Fig. 18) and $Rgb_5$4 (Fig. 19) show that as these clones are isolated from the HaeIII library, they have HaeIII restriction sites, rather than EcoRI sites at one end of both of the subclones 4.10 and 1.2 (Fig. 26). From this it appears that there is a 950bp HaeIII fragment missing between the subclones 4.10 and 1.2 and that subclone 6.5E from the lambda clone $Rgb_5$5 (isolated from the EcoRI library) represents an overlapping fragment. Sequence data (section 3.2.4) also supports this idea, as 200bp of DNA sequence is identical between the subclones 6.5E and 1.2. However, on looking at the sizes of the other EcoRI fragments observed in $Rgb_5$1 and $Rgb_5$4 (Table 1 and Figs. 11 and 12) it can be seen that there are no other common EcoRI fragments. The EcoRI fragments on either side of the 6.5E subclone are not the same size as those expected from the $Rgb_5$1 (Fig. 18) or $Rgb_5$4 (Fig. 19) restriction maps. Therefore, from this it would appear that the lambda clone $Rgb_5$5 is not in fact an overlapping clone of $Rgb_5$1 and $Rgb_5$4.

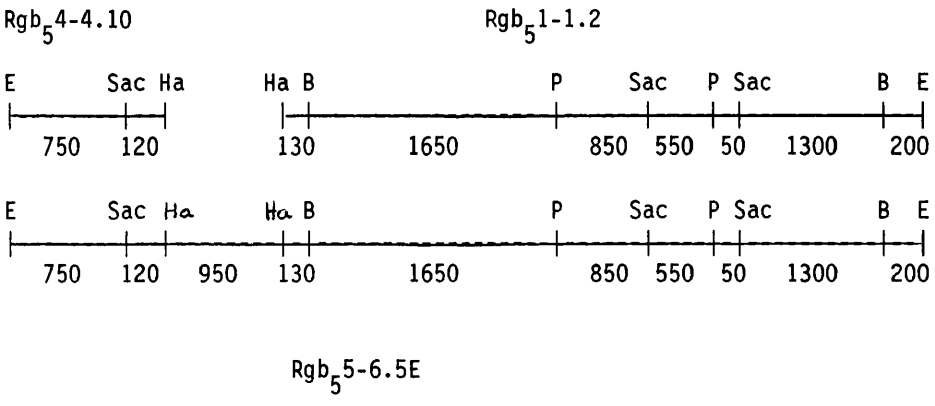



Figure 26

Overlap of the $Rgb_5$5 subclone 6.5E with the $Rgb_5$1 subclone 1.2 and the $Rgb_5$4 subclone 4.10. Sizes are in base pairs.

E = EcoRI, Sac = Sac I, B = BamHI, P = Pst I, Ha = HaeIII.

This could be explained in one of two ways. The lambda genomic clone $Rgb_5$5 could represent a different gene to that of the $Rgb_5$1 and $Rgb_5$4 clones and there are two genes coding for cytochrome $b_5$. This seems unlikely because the restriction maps over 5.5kb of DNA are identical (Fig. 26) and the sequence data shows that they are the same over 200bp. The other possibility is that one of the libraries has rearranged to result in the juxtaposition of EcoRI fragments which do not lie side by side in the rat genome.

$Rgb_5$3/$Rgb_5$7 and $Rgb_5$6 derived subclones

The genomic clones $Rgb_5$3 and $Rgb_5$7 are identical, as seen by the sizes of the EcoRI fragments (Fig. 12).

36mer - the 36mer oligonucleotide (amino acids 4-15) hybridized weakly to the 2.2kb EcoRI fragment of the genomic clones $Rgb_5$3 or $Rgb_5$7 and to the 1.6kb EcoRI fragment of $Rgb_5$6 under normal conditions (data not shown).

The 700bp PstI/EcoRI fragment of the 2.2kb EcoRI subclone from $Rgb_5$3/7 in pUC19, E3.1o (Fig. 25(3)), hybridized to the 36mer under slightly reduced stringency conditions (Fig. 27). Hybridization was at 65°C ($T_d$ -14°C) with the final wash in 6xSSC for 2 minutes at 70°C ($T_d$ -9°C). This was carried out because there was initially no signal with the 36mer at 74°C ($T_d$ -5°C). There was the possibility that the oligonucleotide was spanning an intron so that a 100% match was not being achieved.


Figure 27A. Photograph of the subclone E3.1o (from the lambda genomic clone $Rgb_5 3$) digested with PstI and EcoRI and run out on a 0.8% agarose gel. Lane a - 1kb ladder (BRL marker), lane b - E3.1o digested with PstI and EcoRI, lane c - $pb_5(1)D$ digested with EcoRI. $pb_5(1)D$ is the full-length cDNA Sizes shown are in kilobase pairs (kb).

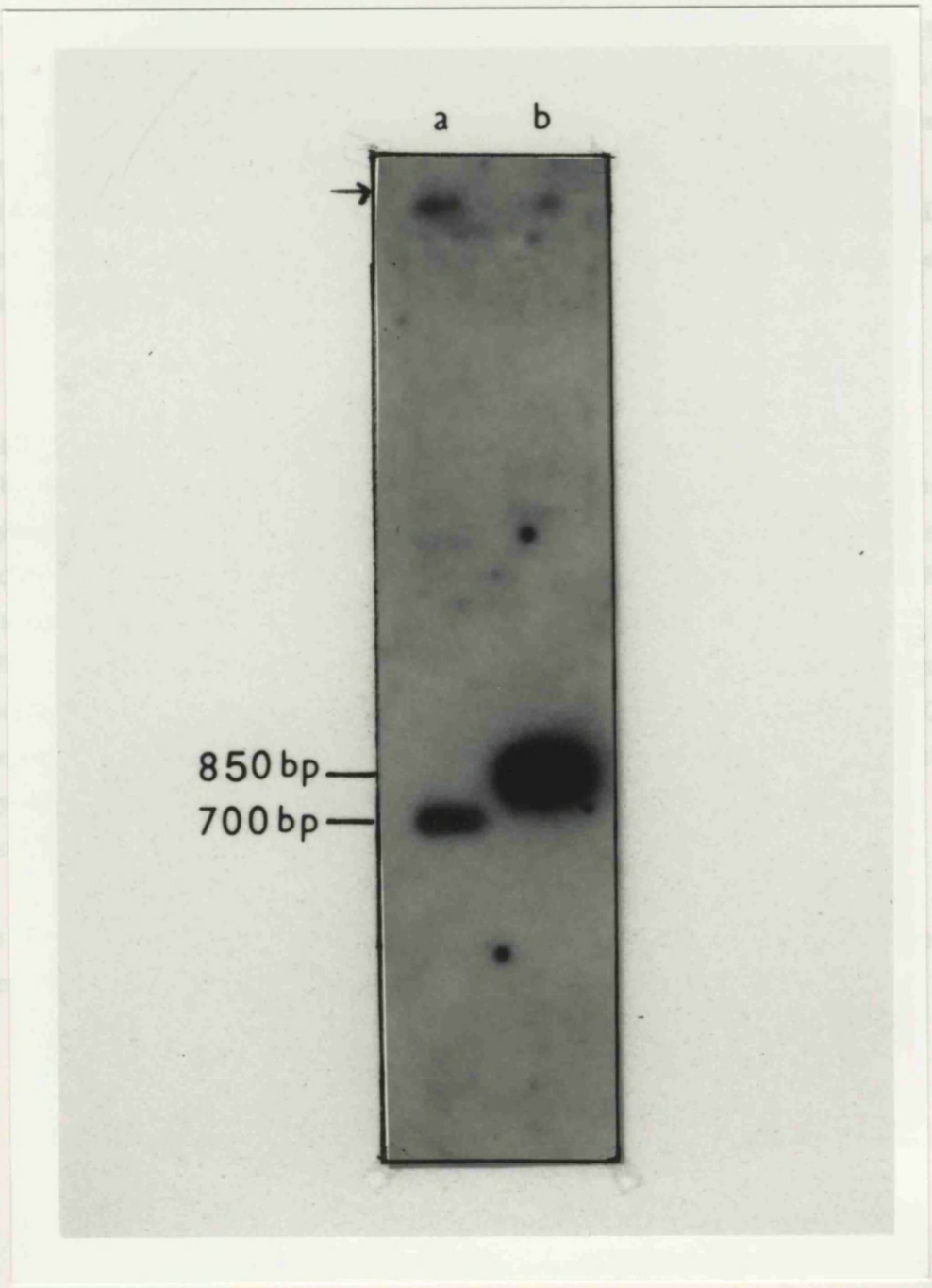



Figure 27B. Southern blot of the agarose gel from Figure 27A, probed with the end-labelled 36mer oligonucleotide. Hybridization at $T_d$ -14°C / final wash at $T_d$ -9°C. Lane a - E3.1o digested with PstI and EcoRI, lane b - $pb_5$(1)D digested with EcoRI. Sizes shown are in base pairs (bp).

⟶ marks the wells.

The restriction maps of the hybridizing EcoRI fragments from $Rgb_5$3/7 (from the EcoRI genomic library) and $Rgb_5$6 (from the HaeIII genomic library) - subclones E3.1o and H4.1T in pUC19 respectively, suggest that they are identical over the first 1.6kb. It then appears that the E3.1o insert extends for 600bp beyond the H4.1T insert (Figs. 25(2) and 25(3)). Only the restriction endonuclease PstI, of six common enzymes tried, cut either insert. If H4.1T and E3.1o are cut with the restriction endonucleases EcoRI (cleaves within the multiple cloning site of pUC19) and PstI and the digests run on an agarose gel, both give a 700bp PstI/EcoRI fragment to which the 36mer oligonucleotide (amino acids 4-15) hybridizes (Fig. 27). From the EcoRI restriction fragments of $Rgb_5$3/7 and $Rgb_5$6 (table 1) it can be seen that these clones overlap (Fig. 28).

$\underline{Rgb_5}$2 derived subclones

The $Rgb_5$2 genomic clone contains a 2.2kb EcoRI fragment to which the rat cytochrome $b_5$ cDNA hybridizes (Fig. 12). The 36mer oligonucleotide (amino acids 4-15) also hybridized to this 2.2kb EcoRI fragment under the slightly reduced stringency conditions used for the 36mer with all the genomic clones (Fig. 29). However, even when this 2.2kb EcoRI fragment was specifically isolated from an agarose gel and subcloned into the plasmid vector pUC19, named E1.2B (Fig. 25(5)), the fragment no longer hybridized to either the rat cytochrome $b_5$ cDNA or to the 36mer oligonucleotide under the same conditions.

The restriction map (Fig. 25(5)) shows that the pUC19 subclone, E1.2B is not the same as the pUC19 subclone E3.1o (Fig. 25(3)) - the 2.2kb EcoRI fragment derived from $Rgb_5$3.

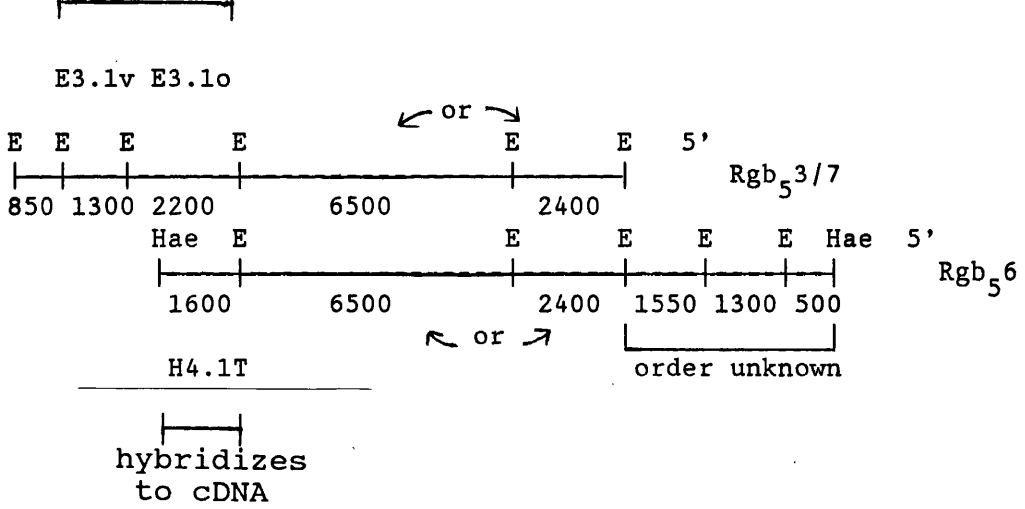



Figure 28.

Diagram of the overlap of $Rgb_5$3/7 and $Rgb_5$6 genomic clones. Sizes given are in base pairs. The names of the subclones are indicated above the relevant EcoRI fragments - E3.1v, E3.1o and H4.1T.

E = EcoRI, Hae = HaeIII

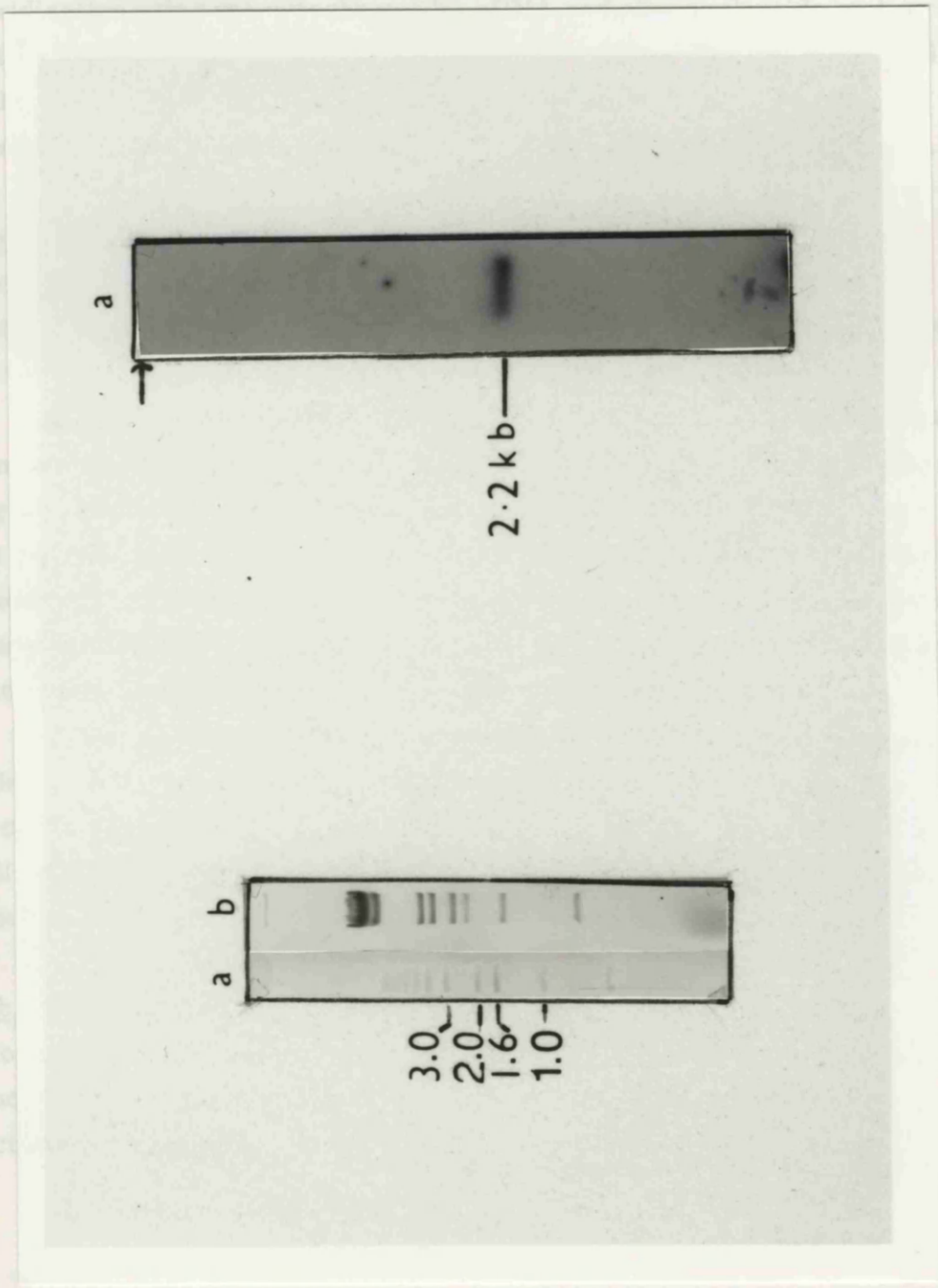



Figure 29A. Photograph of the lambda genomic clone $Rgb_5 2$ digested with EcoRI and run out on a 0.6% agarose gel. Lane a - 1kb ladder (BRL marker), lane b - $Rgb_5 2$ digested with EcoRI. Sizes shown are in kilobase pairs (kb).

Figure 29B. Southern blot of the agarose gel from Figure 29A, probed with the end-labelled 36mer oligonucleotide. Hybridization at $T_d$ -14°C / final wash at $T_d$ -9°C. Lane a - $Rgb_5 2$ digested with EcoRI. Sizes shown are in kilobase pairs (kb). ⟶ marks the wells.

### 3.2.4 Sequence analysis

From the information obtained by restriction mapping and Southern blot hybridization using rat cytochrome $b_5$ cDNA and oligonucleotide probes, certain restriction fragments were chosen to subclone into M13mp18 and M13mp19 vectors (section 2.15). Single-stranded DNA from the M13 constructs was prepared (section 2.15.2) for sequence analysis (section 2.16).

#### $Rgb_5$4 - subclone 4.10

The 870bp EcoRI fragment (subclone 4.10 in pSPT19, Fig. 14B) from the $Rgb_5$4 genomic clone, which hybridized to the rat cytochrome $b_5$ cDNA (Fig. 11) was subcloned into EcoRI cut M13mp19 in both orientations.

Sequencing, catalyzed by the Klenow fragment of DNA polymerase I, from the EcoRI site at the extreme left-hand end of the $Rgb_5$4 genomic clone (Fig. 19) located one exon after approximately 110 bases of intron (Figs. 30 and 31). This exon contained bases coding for amino acid 85 to amino acid 43 inclusive of the rat cytochrome $b_5$ protein and was followed by more intron sequence. This 870bp EcoRI fragment contained no further coding sequence as determined by sequencing from both ends of the EcoRI fragment.

The remainder of the $Rgb_5$4 genomic clone (9.9kb) also contained no coding sequence, as deduced by hybridization studies with the rat cytochrome $b_5$ cDNA and the 36mer oligonucleotide (amino acids 4-15). However, it is still possible that a small exon does exist within the $Rgb_5$4 genomic clone which is too short to hybridize to either, or both, of these probes.

#### $Rgb_5$1 - subclone 1.3

Two restriction fragments from the 6.15kb EcoRI fragment of the $Rgb_5$1 genomic clone (subclone 1.3 in pSPT19, Fig. 14A) were subcloned into M13 vectors for sequencing.

```
                                                Met Ala Glu Gln Ser Asp
GGCTGTGTTGCAGGGCCCGGAAGCCTCACTGTTCCGAA ATG GCC GAG CAG TCA GAC

                10                                                  20

Lys Asp Val Lys Tyr Tyr Thr Leu Glu Glu Ile Gln Lys His Lys Asp
AAG GAT GTG AAG TAC TAC ACT CTG GAG GAG ATT CAG AAG CAC AAA GAC

                                    30

Ser Lys Ser Thr Trp Val Ile Leu His His Lys Val Tyr Asp Leu Thr
AGC AAG AGC ACC TGG GTG ATC CTA CAT CAT AAG GTG TAC GAT CTG ACC

        40          INTRON                          50
                      |
Lys Phe Leu Glu Glu His Pro Gly Gly Glu Glu Val Leu Arg Glu Gln
AAG TTT CTC GAA GAG CAT CCT GGT GGG GAA GAA GTC CTA AGA GAG CAA

                60

Ala Gly Gly Asp Ala Thr Glu Asn Phe Glu Asp Val Gly His Ser Thr
GCT GGG GGT GAT GCT ACT GAG AAC TTT GAG GAC GTC GGG CAC TCT ACG

70                                      80                  INTRON
                                                              |
Asp Ala Arg Glu Leu Ser Lys Thr Tyr Ile Ile Gly Glu Leu His Pro
GAT GCA CGA GAA CTG TCC AAA ACA TAC ATC ATC GGG GAG CTC CAT CCA

                90                  INTRON              100
                                      |
Asp Asp Arg Ser Lys Ile Ala Lys Pro Ser Glu Thr Leu Ile Thr Thr
GAT GAC AGA TCA AAG ATA GCC AAG CCT TCG GAA ACC CTT ATC ACT ACT

                    INTRON      110
                      |
Val Glu Ser Asn Ser Ser Trp Trp Thr Asn Trp Val Ile Pro Ala Ile
GTC GAG TCT AAT TCC AGT TGG TGG ACC AAC TGG GTG ATC CCA GCC ATC

        120                                     130

Ser Ala Leu Val Val Ala Leu Met Tyr Arg Leu Tyr Met Ala Glu Asp
TCA GCC CTG GTG GTA GCT CTG ATG TAT CGC CTC TAC ATG GCA GAA GAT
```

Ter
TAACCTGTCTGTCCGAAGCCAAGGAAGGAAAAGACTGCCCCAGAGAGGGGAGAAAAGAAGCCAG
TGTTAATCACTTCCACTGACAGAAACCCTCCCCCTGAGAATGTAATTGTAATATATCTGTCTCC
CTCTCCTCCTATGCTAGGAGAACAAACATGGGACTCTTTGTACTCTTAAACTTTCAAATGTGCC
TTTTTACTCAACTTCATTTTGACATTTCTTCACTACGTAATTTACTTATTGTAAACATGATCTT
TTTAAAATATATCTGGCTTGTAAAGTACA*CCAGGTGTGCCTGTTTGTGTGGTATTTTATATTT
AGTATTTAGTGTTTTGGAGTTGTTTAACAGAATTGCTTAACGTACTAACTGTTTGGACCCAAAA
AAAA

Figure 30. Sequence of cytochrome $b_5$ cDNA with known intron sites marked. Numbering indicates amino acid position. The proximal and distal histidines are highlighted. Possible poly[A] addition sites are underlined. The first poly[A] addition site is indicated (*).

The 1.1kb SacI fragment from 1.3 (Fig. 14A) which hybridized to the rat cytochrome $b_5$ cDNA (Fig. 11) was subcloned into SacI cut M13mp19 in both orientations. This fragment lies at the extreme right-hand end of the $Rgb_5 1$ genomic clone (Fig. 18). Sequencing, catalysed by the Klenow fragment of DNA polymerase I, from the internal SacI site located an exon which ran from within amino acid 107 of rat cytochrome $b_5$ through to amino acid 133 and into the 3' non-coding sequence. The coding sequence began after 270bp of intron from the internal SacI site (Figs. 30 and 32).

The second restriction fragment subcloned into M13 vectors from the plasmid subclone 1.3 was the 800bp PstI/SmaI fragment (Fig. 14A) which hybridized to the 26mer oligonucleotide (amino acids 98-106) (Fig. 21(1)). This fragment was subcloned into both M13mp18 and M13mp19 cut with the restriction endonucleases SmaI and PstI. After sequencing 100bp of intron from the SmaI site in M13mp19, using the `Sequenase' enzyme (section 2.16.2), a small exon of 35 base pairs was found which ran from within amino acid 107 to amino acid 96. This was followed by intron sequence (Figs. 30 and 33).

$Rgb_5 1$ - subclone 1.2

Within the 4.7kb EcoRI fragment of the $Rgb_5 1$ genomic clone (subclone 1.2 in pSPT19) a 1650bp BamHI/PstI fragment (Fig. 14A) hybridized to the 27mer oligonucleotide (amino acids 87-95) (Fig. 21(2)). This BamHI/PstI fragment was subcloned into both M13mp18 and M13mp19 cut with the restriction endonucleases BamHI and PstI. Sequencing from either end using the `Sequenase' enzyme (section 2.16.2) did not reveal any coding region. This indicated that the exon lay beyond 490bp from the PstI site and beyond 370bp from the BamHI site.

Therefore, a restriction endonuclease with a recognition site of four bases that produces blunt ends on DNA cleavage was chosen to digest the 1650bp BamHI/PstI fragment of the subclone 1.2. This would allow smaller sections to be ligated into the SmaI site of M13mp19 - SmaI being a restriction endonuclease producing blunt ends. The restriction endonuclease AluI was chosen because it did not cut too often nor too little within the sequence of the 1650bp BamHI/PstI fragment read so far. For ease of digestion, the plasmid subclone 1.2, was cut just with the restriction enzyme PstI. As there is a single PstI site in the multiple cloning site of pSPT19, this digest produces an 1800bp fragment, within which is the 1650bp BamHI/PstI fragment that hybridized to the 27mer oligonucleotide (amino acids 87-95) (Fig. 21(2)). The digest was then electrophoresed on a 0.8% agarose gel, the 1800bp PstI fragment was sliced out and the DNA purified using `Glassmilk'

(section 2.6.2.2). This 1800bp fragment was then digested with the restriction endonuclease AluI which cut within the fragment four times. This meant that there were three possible AluI/AluI blunt-ended fragments able to insert into the SmaI cut M13mp19 vector. These AluI fragments were ligated into the M13mp19 vector in both orientations and transformed into JM101 (section 2.15.1). Small-scale plasmid preparations of the replicative form (RF) of the M13 recombinants were carried out (section 2.15.3) and the DNA digested with the restriction endonucleases EcoRI and PstI, which cut either side of the insert. (The SmaI sites are not regenerated with AluI cut inserts). The digests were then electrophoresed on a 1.2% agarose gel (section 2.4). This gel was blotted by the method of Southern (1975) and probed with the 27mer oligonucleotide (amino acids 87-95) to determine which of the recombinant clones contained the rat cytochrome $b_5$ exon. Two M13mp19 AluI clones containing an 870bp insert which hybridized to the 27mer (amino acids 87-95) were sequenced using the `Sequenase' enzyme (section 2.16.2). The clones were found to be in opposite orientations. Within one of these was found an exon of 30 base pairs, which ran from amino acid 86 to amino acid 95 of the rat cytochrome $b_5$ protein. Intron sequence lay either side (Figs. 30 and 34).

Rgb$_5$5 - subclone 6.5E

The 6.5kb EcoRI fragment (6.5E in pUC19, Fig. 25(1)) was characterized by partial sequencing. The 6.5E subclone was digested with the restriction endonucleases PstI and BamHI and the fragments cloned into M13mp18. One of these M13 subclones analyzed was sequenced from the PstI site and found to have identical sequence over 200 bases to the BamHI/PstI insert from the plasmid subclone 1.2 (Fig. 14A) of the Rgb$_5$1 genomic clone.

Rgb$_5$6 and Rgb$_5$3/7

To analyze the EcoRI fragment from Rgb$_5$6 which hybridized to the 36mer oligonucleotide (amino acids 4-15), the H4.1T subclone in pUC19 (Fig. 25(2)) was digested with the restriction endonucleases EcoRI and PstI. This gave two possible PstI/EcoRI restriction fragments to ligate into M13mp19 cut with these enzymes. The 700bp fragment in M13mp19, which hybridizes to the 36mer oligonucleotide, was called H4.1A/mp19 Pst/Eco and sequenced. It was expected that the 5' end of the rat cytochrome $b_5$ gene would be found. However, reading from the EcoRI site revealed sequence that was very similar, although not identical to the rat cytochrome $b_5$ cDNA. The EcoRI site was itself part of this sequence - one base alteration around amino acid 105 produces an EcoRI site (Fig. 35). Longer runs of the 700bp fragment, H4.1A/mp19 Pst/Eco, on sequencing gels combined to give sequence similar to

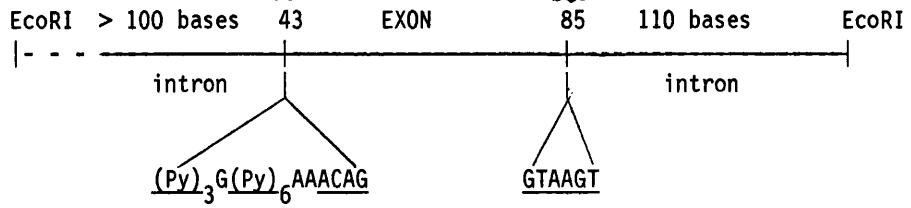



Figure 31.
$Rgb_5$4 4.10 EcoRI subclone showing intron sequence data on either end of the exon.
Underlined bases show homology to consensus intron/exon border sequence.

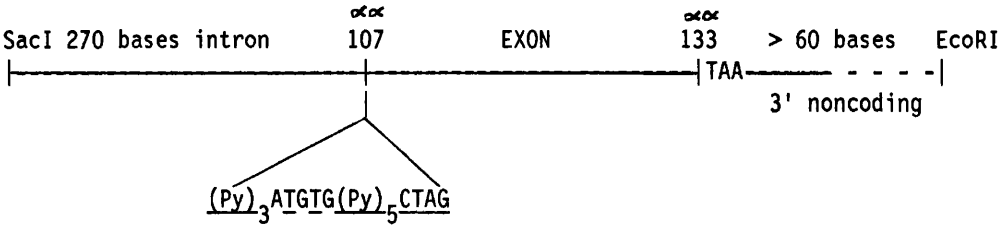



Figure 32.
$Rgb_5$1 1.3 EcoRI/SacI fragment (from 6.15kb EcoRI subclone) showing intron/exon border.
Underlined bases show homology to consensus intron/exon border sequence.

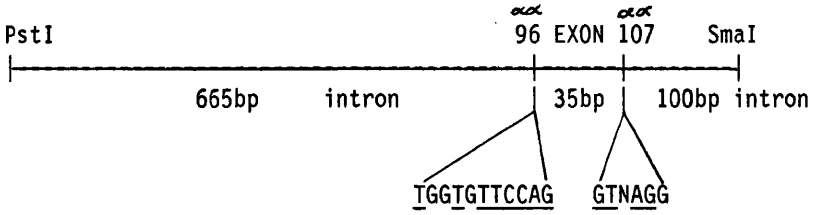



Figure 33.
Location of exon within 800bp SmaI/PstI fragment from $Rgb_5$1 EcoRI subclone 1.3.
Underlined bases show homology to consensus intron/exon border sequence.

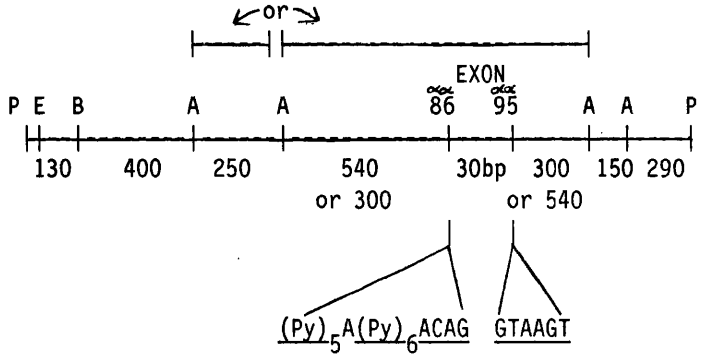



Figure 34.
Location of exon within 1800bp PstI fragment from $Rgb_5$1 EcoRI subclone 1.2.
P = PstI, E = EcoRI, A = AluI, B = BamHI.
Underlined bases indicate homology to consensus intron/exon border sequence.
Sizes given are in base pairs.

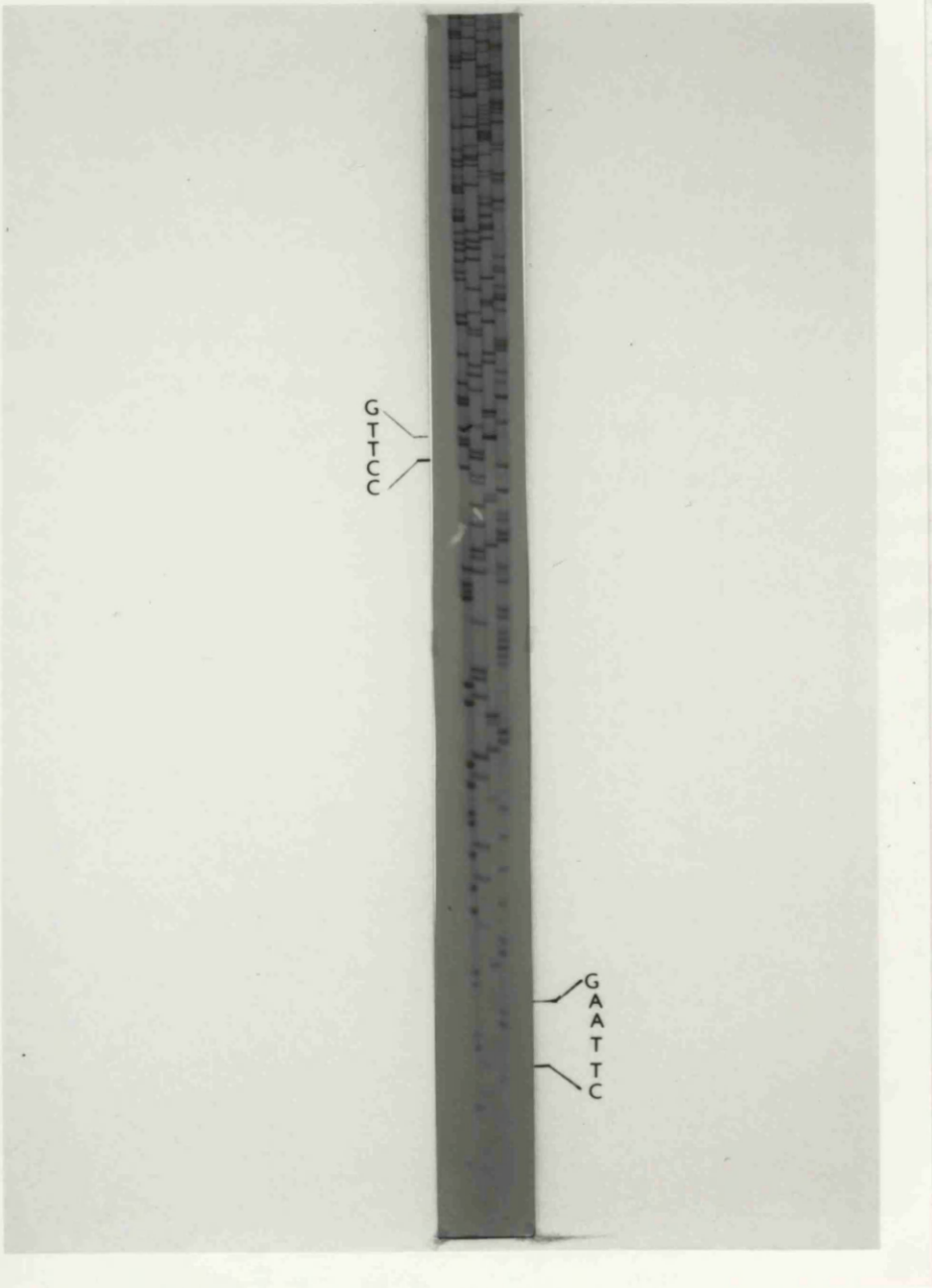



Figure 35. Sequencing gel of the rat cytochrome $b_5$ pseudogene clone, H4.1/mp19 Pst/Eco. This is the 700bp PstI/EcoRI fragment from the plasmid subclone H4.1T (Fig. 25(2)), from the genomic clone $Rgb_5 6$. The gel was electrophoresed for 1½ hours. The 5' end is at the top of the gel. The EcoRI site (CTTAAG) is labelled. A reference sequence of CCTTG is also labelled to indicate where to line up the gel with that in figure 36. The tracks are T, C, G, A from left to right.

rat cytochrome $b_5$ cDNA from amino acid 105 (the EcoRI site) to within the 5' non-coding region without any intervening "non-coding" sequence (Figs. 36 and 37). As this sequence similarity begins within the rat cytochrome $b_5$ cDNA sequence, a Southern blot of $Rgb_5$6 and $Rgb_5$3 clones digested with the restriction endonuclease EcoRI was hybridized with the rat cytochrome $b_5$ cDNA under reduced stringency conditions (hybridize at 60°C not 65°C, final wash at 50°C for 15 minutes in 0.1xSSPE). This would ascertain whether the EcoRI fragment adjacent to that already subcloned and partially sequenced, or any other within the genomic clones, continued the similarity to the cytochrome $b_5$ cDNA sequence beyond amino acid 105. Under these reduced stringency conditions the 1.3kb EcoRI fragment in $Rgb_5$7 and $Rgb_5$3 appeared to be hybridizing to the rat cytochrome $b_5$ cDNA (data not shown). This was confirmed by sequencing the 1.3kb fragment from $Rgb_5$3, subclone E3.1v, inserted into M13mp19 at the EcoRI site. The subclone in M13mp19, E3.1v2/mp19 Eco, reading from the EcoRI site began immediately with sequence similar to the rat cytochrome $b_5$ cDNA at amino acid 106 and continued through to amino acid 119 or 120 without any intervening sequence. (Fig. 38). Beyond this, no similarity to the coding region of rat cytochrome $b_5$ cDNA could be found.

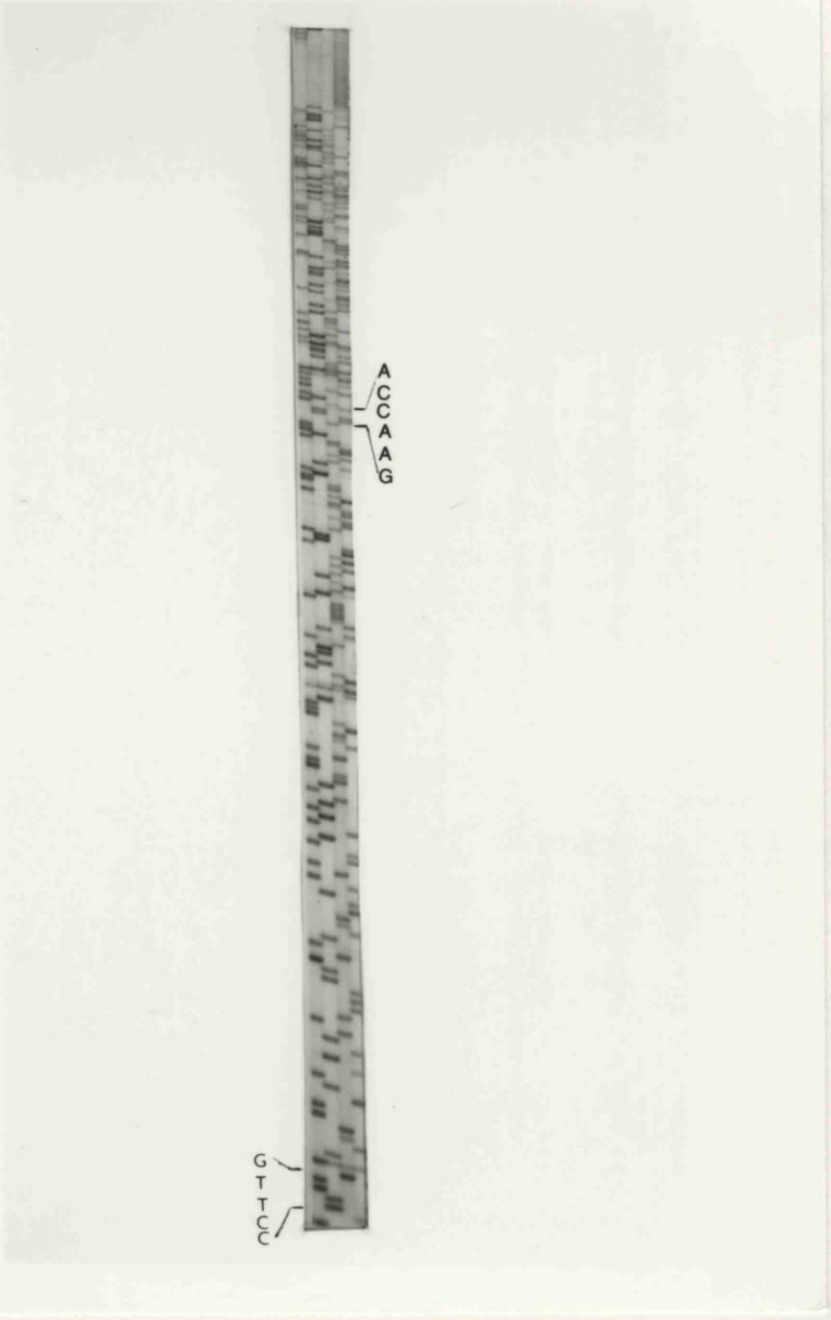



Figure 36. Sequencing gel of the rat cytochrome $b_5$ pseudogene clone, H4.1/mp19 Pst/Eco. This is the 700bp PstI/EcoRI fragment from the plasmid subclone H4.1T (Fig. 25(2)), from the genomic clone $Rgb_5 6$. The gel was electrophoresed for 3½ hours. The 5' end is at the top of the gel. The reference sequence CCTTG indicates where to line up the gel with that in Figure 35. The reference sequence GAACCA indicates where to line up the gel with that in Figure 37. The tracks are T, C, G, A from left to right.

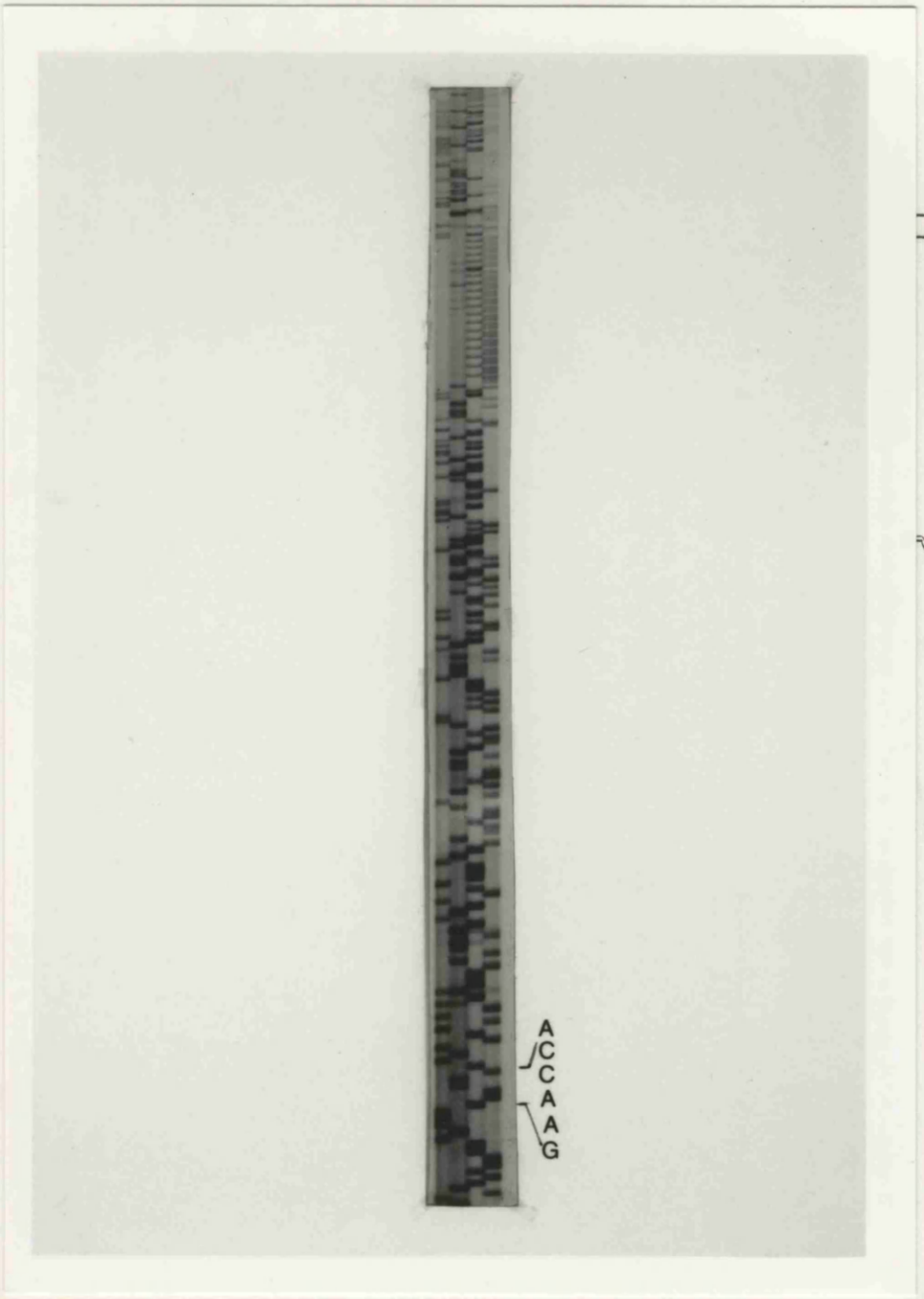



Figure 37. Sequencing gel of the rat cytochrome $b_5$ pseudogene clone, H4.1/mp19 Pst/Eco. This is the 700bp PstI/EcoRI fragment from the plasmid subclone H4.1T (Fig. 25(2)), from the genomic clone $Rgb_5 6$. The gel was electrophoresed for 5½ hours. The 5' end is at the top of the gel. The reference sequence GAACCA indicates where to line up the gel with that in Figure 36. The direct repeat sequence at the 5' end of the pseudogene is shown (5' NANAAAATTGAAATGTA 3'). The altered Met codon (AGG) is shown. The tracks are T, C, G, A from left to right.

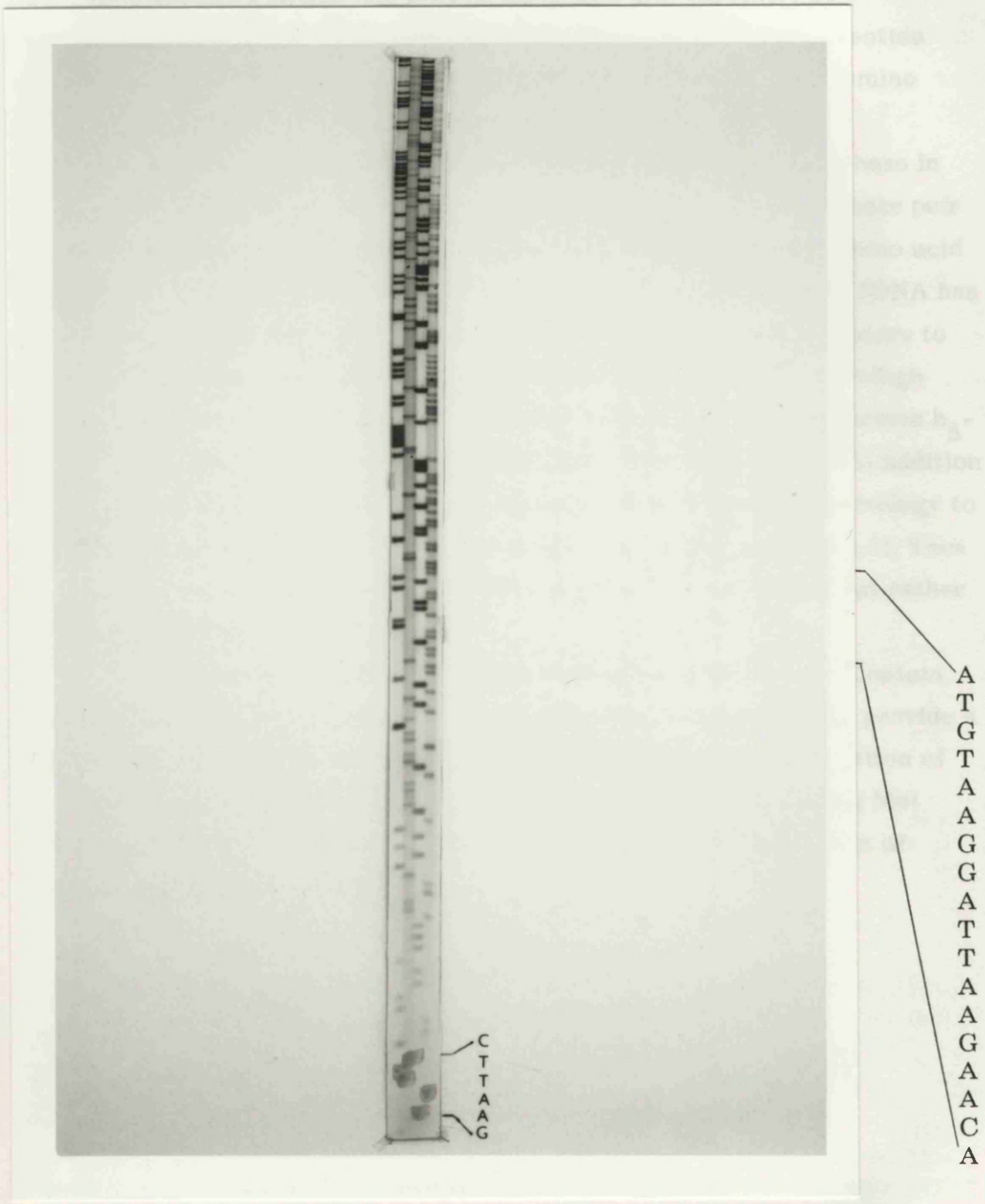



Figure 38. Sequencing gel of the rat cytochrome $b_5$ pseudogene clone, E3.1v2/mp19 Eco. This is the 1300bp fragment from the plasmid subclone E3.1v (Fig. 25(4)) from the genomic clone $Rgb_5$3. The gel was electrophoresed for 1½ hours. The 5' end is at the bottom of the gel. The EcoRI site (GAATTC) is labelled, which is the same site as that in Figure 35 reading in the opposite direction. The direct repeat sequence at the 3' end of the pseudogene is shown (5' ACAAGAATTAGGAATGTA 3'). The tracks are T, C, G, A from left to right.

The overall comparison of the rat cytochrome $b_5$ cDNA with the sequences of H4.1A/mp19 Pst/Eco and E3.1v2/mp19 Eco is shown in Fig. 39.

H4.1A/mp19 Pst/Eco and E3.1v2/mp19 Eco are, together, 64% homologous at the amino acid level and 78% homologous at the nucleotide level to the rat cytochrome $b_5$ cDNA over the coding section from amino acids 1-119 inclusive, excluding deletions and insertions.

In the cytochrome $b_5$-like sequence there is a deletion of one base in amino acid 6, a two base pair insertion around amino acid 34, a one base pair insertion after amino acid 98 and a one base pair insertion around amino acid 109 (Fig. 39). The Met (ATG) at the 5' end of the rat cytochrome b5 cDNA has also been altered to AGG (Figs. 37 and 39). The $b_5$-like sequence appears to be truncated at the 3' end, finishing after amino acid 119 or 120, although there is no obvious poly[A] tail. Beyond amino acid 120 of the cytochrome $b_5$-like sequence there are three regions which show homology to poly[A] addition consensus sequences (underlined in Fig. 39) and a short stretch of homology to the 3' noncoding sequence of the rat cytochrome $b_5$ cDNA (section 4.1). Thus the sequence beyond amino acid 119 does not seem to be an intron but rather a 3' non-coding region.

It therefore appears that $Rgb_5$3/7 and $Rgb_5$6 genomic clones contain sequence for a rat cytochrome $b_5$ pseudogene. This sequence cannot provide a functional protein due to the base deletion and insertions, the truncation of the 3' end of the sequence and alteration of the translation initiating Met codon. As there are no intervening sequences, it appears to be a form of processed pseudogene.

$b_5$ p/gene NAN**AAAATTGAAATGTA**GAAGAAGAAGAAGGAAGAAAGAAAGAAAGAAAGA

$b_5$ p/gene AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAACAGTGTACCACCCTGAAT

```
                                                        1
                                                  Met Ala Glu Gln Ser Asp
b5 GGCTGTGTTGCAGGGCCCGGAAGCCTCACTGTTCCGAA ATG GCC GAG CAG TCA GAC
b5 p/gene     T    T   G  ****GT     T  G  G    T  GA     G
                                          Arg     Gly     Ala

                10                                          20
Lys Asp Val Lys Tyr Tyr Thr Leu Glu Glu Ile Gln Lys His Lys Asp
AAG GAT GTG AAG TAC TAC ACT CTG GAG GAG ATT CAG AAG CAC AAA GAC
*                G        C      GA                      C
                Cys             Gly                     Gln

                                30
Ser Lys Ser Thr Trp Val Ile Leu His His Lys Val Tyr Asp Leu Thr
AGC AAG AGC ACC TGG GTG ATC CTA CAT CAT AAG GTG TAC GAT CTG ACC
 T       A  G           G   G   C   C  G       | T
Ile     Asn Ser         Met             Arg    /\
                                               TA

        40                                          50
Lys Phe Leu Glu Glu His Pro Gly Gly Glu Glu Val Leu Arg Glu Gln
AAG TTT CTC GAA GAG CAT CCT GGT GGG GAA GAA GTC CTA AGA GAG CAA
          G             T       A                       G
                       Leu     Glu                     Arg

                        60
Ala Gly Gly Asp Ala Thr Glu Asn Phe Glu Asp Val Gly His Ser Thr
GCT GGG GGT GAT GCT ACT GAG AAC TTT GAG GAC GTC GGG CAC TCT ACG
          A C     C T                   T   T   C TG        T
            His     Ser                           Cys

70                                          80
Asp Ala Arg Glu Leu Ser Lys Thr Tyr Ile Ile Gly Glu Leu His Pro
GAT GCA CGA GAA CTG TCC AAA ACA TAC ATC ATC GGG GAG CTC CAT CCA
A   T   A   G             GTG C         T   A CT  T
Asn Val Gln Gly           Val His             Leu Phe

                90                                          100
Asp Asp Arg Ser Lys Ile Ala Lys Pro Ser Glu Thr Leu Ile Thr Thr
GAT GAC AGA TCA AAG ATA GCC AAG CCT TCG GAA ACC CTT ATC ACT ACT
 GC A    AG C   C T T   A A   A A C   T  G   G  TC |  T
Gly Asn Lys Pro His Leu Thr     Thr     Gly Ser Ser/
                                                  G
```

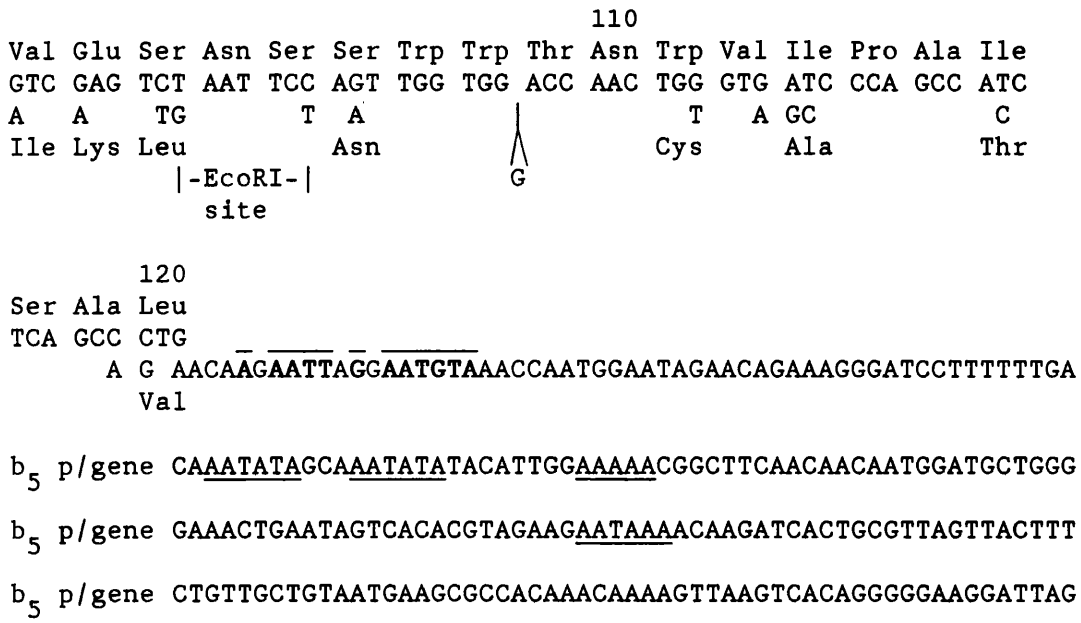



Figure 39. Amino acid and nucleotide comparison of rat cytochrome $b_5$ ($b_5$) (upper lines) and its pseudogene ($b_5$ p/gene) (lower lines). Alterations in the pseudogene are indicated - both at nucleotide and amino acid level. Deletions in the pseudogene sequence are indicated (*) and insertions are beneath the sequence. The proximal histidine is conserved (highlighted) and the distal histidine is altered to a cystine (Cys) (highlighted). The possible poly[A] addition signals are underlined, as is the only likely poly[A] tail found. Numbers indicate amino acid position. The internal EcoRI site in the pseudogene is shown. The direct repeats at the start and end of the pseudogene are indicated.

3.3 Genomic Southern

Genomic DNA from Sprague-Dawley rats was prepared (section 2.18) and digested with the restriction endonuclease EcoRI (section 2.3). The digest was loaded onto a 0.8% agarose gel 20cm in length (section 2.4). The gel was treated as described in section 2.9 and blotted by the method of Southern (1975) onto nylon membranes. The fixed blot was probed with the full-length [$\alpha^{32}$P]-dCTP labelled rat cytochrome $b_5$ cDNA insert prepared from the $pb_5$(1)D construct (section 2.7.1) Aqueous solutions were used for the prehybridization and hybridization of the filter at 65°C (section 2.10). Enough $pb_5$(1)D cDNA labelled probe was added to the hybridization fluid to give a concentration of $2x10^6$ cpm/ml of fluid. The final stringent wash was in 0.1xSSPE for 15 minutes at 60°C. After exposure of the filter to X-ray film, up to seven bands from EcoRI cut rat genomic DNA hybridized to the full-length rat cytochrome $b_5$ cDNA probe. There may be a doublet at 7.5kb (the band is more intense and wider than expected from a single band), single bands at 6.3kb, 5.2kb and 2.1kb as well as two very faint bands at 3.4kb and 2.8kb (Fig. 40). The Southern blot gave the same banding pattern with a final wash in 0.1xSSPE for 15 minutes at 50°C (data not shown).

It is not clear from this whether there are one or two functional genes coding for cytochrome $b_5$, as discussed in section 4.2.

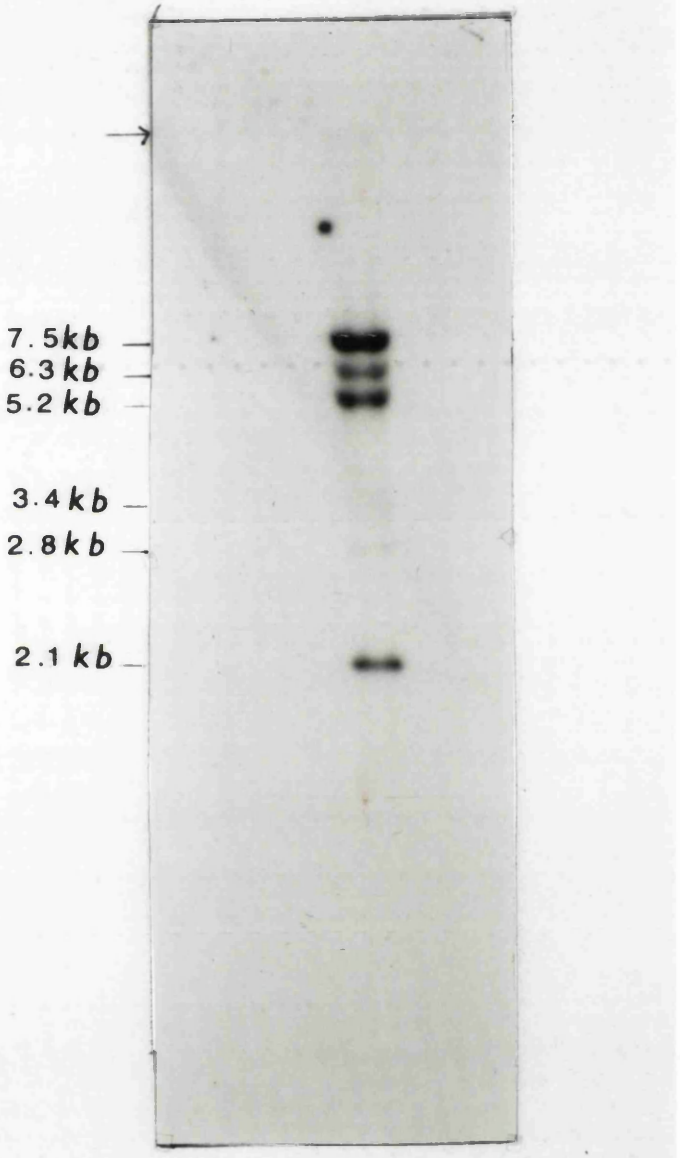



Figure 40. Southern blot of rat genomic DNA digested with EcoRI, run out on a 0.8% agarose gel and probed with the full-length rat cytochrome $b_5$ cDNA. Final wash at 60°C in 0.1xSSPE. Sizes shown are in kilobase pairs (kb).

⟶ marks the wells.

## 4.1 Pseudogenes

The term pseudogene is used to describe a DNA sequence which is related to a functional gene, but is itself defective and unable to produce an active product.

The first stretch of DNA classified as a pseudogene was the 5s rRNA-related gene from Xenopus laevis (Jacq et al., 1977). This pseudogene is truncated at the 5' end by 16 base pairs.

The sequence analyzed from the H4.1T and E3.1v2 genomic DNA clones in M13 vectors shows 64% homology to the rat cytochrome $b_5$ cDNA sequence at the amino acid level and 78% at the nucleotide level (excluding deletions/insertions). This DNA sequence could not produce a functional polypeptide due to the presence of various mutations (Fig. 39):

(1) deletion of one base in amino acid 6.
(2) insertion of two bases - TA or AT - before or after amino acid 34 respectively.
(3) 5' ATG translation initiation codon converted to AAG.
(4) distal His-67 not conserved (although proximal His-43 is conserved) (by analogy with globins, Mathews, 1980).
(5) insertion of one base after amino acid 109.
(6) insertion of one base after amino acid 98.
(7) sequence truncated at 3' end - ceases after amino acid 119 or 120.

The deletion and insertions lead to the production of several stop codons. Therefore, this cytochrome $b_5$-like sequence can be seen to represent a rat cytochrome $b_5$ pseudogene.

There are two categories that pseudogenes can fall into:-

(1) those that retain the introns found in the functional gene
(2) those that lack the intervening sequences of the functional gene and are known as processed pseudogenes.

As there are no introns in the H4.1T and E3.1v2 subclones isolated from the rat genomic library, it appears that this rat cytochrome $b_5$ pseudogene is a processed one.

The absence of introns in processed pseudogenes is precise, such that the DNA sequence 5' and 3' of the introns is juxtaposed. This applies to the rat cytochrome $b_5$ pseudogene.

The sequence homology between a pseudogene and the functional counterpart usually ceases at the points corresponding to the 5' and 3' ends of the transcript from the functional gene. However, the homology between the cytochrome $b_5$ pseudogene and the functional cytochrome $b_5$ gene ceases after amino acid 119 or possibly amino acid 120. It does not appear that this represents the site of an intron. Firstly, the rat cytochrome $b_5$ gene does not have an intron at this point. Secondly, no partially processed pseudogene has been reported to date. Finally, there are three (or four) potential poly[A] addition sites shortly after amino acid 119/120 of the pseudogene. Therefore, it seems that the cytochrome $b_5$ pseudogene is prematurely terminated at amino acids 119/120.

There have been other processed pseudogenes reported to be shorter than the corresponding functional mRNA, as will be discussed later (section 4.1.1).

Most processed pseudogenes also possess a poly[A] tail immediately 3' to the point at which the homology between the pseudogene and its functional counterpart ceases to exist. The mouse pseudo $\alpha_3$ globin gene (Nishioka et al., 1980; Vanin et al., 1980) is an exception to this general rule. The point of sequence divergence is 17-20 nucleotides prior to the site of polyadenylation in the functional gene and there is no poly[A] tail present. The rat cytochrome $b_5$ gene appears to also lack a poly[A] tail. From long sequencing gels, the only possible poly[A] tail is 5 A's in length. This starts 8 nucleotides downstream from one AATATA sequence and 17 nucleotides downstream from a second AATATA sequence. The distance between the poly[A] addition signal and a poly[A] tail is usually between 10 and 30 nucleotides (Proudfoot and Brownlee, 1976).

Studies on the chromosomal localization indicate that most pseudogenes that retain introns are linked to their functional counterparts. However, processed pseudogenes are not found on the same chromosome as the functional gene (Battey et al., 1982), so do not appear to have arisen by gene duplication and intron removal.

4.1.1 The evolution of processed pseudogenes

Processed pseudogenes have only been found in the genomes of mammals, except for a single calmodulin processed pseudogene found in the chicken (Stein et al., 1983).

Processed pseudogenes are usually colinear with the functional mRNA up to the position of the CAP nucleotide and most also possess a poly[A] tail at the 3' end. It seems probable, therefore, that an RNA intermediate is the source of the pseudogenes. It is believed that a cDNA copy of the mRNA is integrated into the genome to form the pseudogene. On insertion it is thought that the gene is inactivated immediately, as the point of integration is random. Therefore, the necessary promoter sites would not be positioned correctly at the 5' end.

Direct repeats are often found flanking processed pseudogenes (Lemischka and Sharp, 1982; Karin and Richards, 1982; Chen et al., 1982). They are generally 7-17bp in length and occur immediately 3' to the poly[A] tail and immediately 5' to the point where the homology between the pseudogene and its functional counterpart ceases.

The rat cytochrome $b_5$ pseudogene shows direct repeats at both ends of the gene which are 15 base pairs in length (Fig. 41).

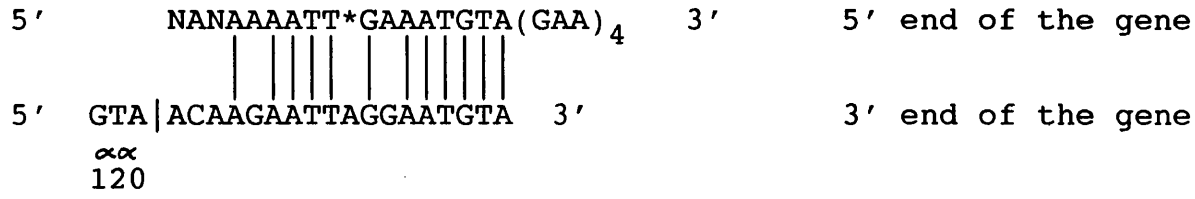



Fig. 41. Direct repeats in rat cytochrome $b_5$ pseudogene. Homology between bases at 5′ and 3′ ends of the cytochrome $b_5$ pseudogene is indicated by vertical lines.

The repeat sequence at the 5' end of the rat cytochrome $b_5$ pseudogene occurs immediately 5' to the runs of (GAA) and (GAAA). The direct repeat at the 3' end of the pseudogene lies almost immediately 3' to amino acid 120.

Processed pseudogenes with flanking direct repeats as in the case of the rat cytochrome $b_5$ pseudogene are postulated to have arisen in several ways. Insertion could have occurred at a staggered break in the DNA, generated by two nearby but independent topoisomerase I molecules that have made single-stranded breaks on opposite strands of the DNA. The variable distance possible between the topoisomerase molecules results in different lengths of stagger which leads to various lengths of direct repeats.

It has been observed that the direct repeats flanking pseudogenes are A-rich. If the staggered break were to occur so that the A-rich sequence was part of a 5' overhang, the poly[T] sequences at the 5' end of the cDNA could have formed a hybrid with these A-rich sequences (Moos and Gallwitz, 1983). The complementary strand would then have been synthesized and repair of the break would produce direct repeats flanking the pseudogene. The other alternative is that the staggered break occured so that the A-rich sequence was part of a 3' overhang. The poly[A] tail of the mRNA could have formed a hybrid with the exposed T-rich sequence (Vanin, 1985), the mRNA then being copied by reverse transcriptase or by DNA polymerase Y (Weissbach, 1977). The mRNA would then have been replaced in the next round of replication and the DNA repaired to again give rise to flanking direct repeats. The extremely A-rich sequence in the long run of (GAAA) repeats at the 5' end of the rat cytochrome $b_5$ pseudogene could have served in this respect.

The rat cytochrome $b_5$ processed pseudogene is not a direct DNA copy of the mRNA. The homology between the pseudogene and its functional counterpart appears to cease after amino acid 119 or 120. Five processed pseudogenes have been reported which are not direct DNA copies of their respective mRNAs. These are the pseudogenes for the mouse $\alpha_3$ globin (Nishioka et al., 1980; Vanin et al., 1980), the human Ig$\varepsilon$ (Battey et al., 1982; Ueda et al., 1982), the human Ig$\lambda$1 (Hollis et al., 1982), the mouse corticotrophin-$\beta$-lipoprotein precursor pseudogene (Notake et al., 1983) and the human cytochrome $b_5$ gene (Yoo and Steggles, 1989).

Two cytochrome $b_5$ pseudogenes have been isolated from human liver genomic libraries (Yoo and Steggles, 1989). One of these contains the complete nucleotide sequence of the human liver cytochrome $b_5$ cDNA but the other is truncated at the 5' end, starting at amino acid 43. Both human pseudogenes continue into the 3' non-coding region but only the truncated pseudogene is shown to possess a poly[A] tail.

The λ Ig light chain is encoded by an mRNA that has V, J and C regions joined together. In the human pseudogene no V region sequences are present although the J and C regions are found, as in the functional gene.

The functional Ig $\varepsilon$ gene has V, D, J and C regions joined together. The human Ig $\varepsilon$ pseudogene contains sequence homologous to the four exons found in the constant (C) region but has no V, D or J sequences 5' to the C region sequences.

The mouse corticotrophin-$\beta$-lipoprotein precursor pseudogene is only homologous to the functional gene from within the second exon down to the polyadenylation site.

The mouse pseudo $\alpha_3$ gene is unusual in that the homology between it and its functional counterpart extends for at least 350 nucleotides 5' to the initiation site of $\alpha$-globin mRNA (Paulakis et al., 1980). At the 3' end the homology ceases 17-20 nucleotides prior to the point of polyadenylation and unlike other processed pseudogenes there is no poly[A] tail present (Nishioka et al., 1980; Vanin et al., 1980). This fact can be explained in the $\alpha$-globin gene by observing that the sequence of the functional gene at the point where the homology between it and the pseudogene ceases is GAGTAGG which could have been used as a 5' splice site (Sharp, 1981). Therefore, an extra splicing event could have removed the poly[A] tail. Truncation of the gene at the 3' end and the probable absence of a poly[A] tail is also seen in the rat cytochrome $b_5$ gene.

Studying the rat cytochrome $b_5$ cDNA sequence shows that if the homology between this sequence and the pseudogene finishes after amino acid 120, the next two bases in the functional gene are GT which could serve as a 5' donor splice site (Sharp, 1981), (Fig. 42).

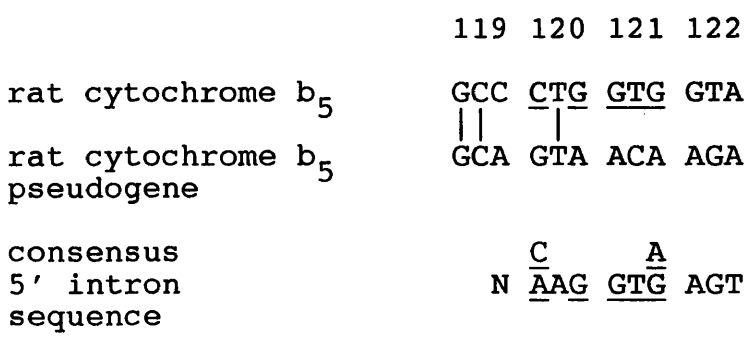



Figure 42. Possibility of extra splicing event giving rise to a truncated rat cytochrome $b_5$ pseudogene without a poly[A] tail. The underlined bases show homology between the consensus 5′ intron sequence and the rat cytochrome $b_5$ sequence. Numbers indicate amino acid position.

Comparing the rat cytochrome $b_5$ cDNA sequence to the consensus intron splicing sequence at this point shows good homology.

If there has been an extra splicing event in the formation of the rat cytochrome $b_5$ pseudogene, this would account for the fact that the 3' end of the protein and the poly[A] tail is absent in the pseudogene. Clearly if the 3' end of this aberrant intron was beyond the poly[A] tail then the rat cytochrome $b_5$ processed pseudogene would not possess one.

However, it was observed that there are possibly three (or four) poly[A] addition signals present within 160bp 3' to amino acid 119 of the rat cytochrome $b_5$ pseudogene:-

(1) AATATA

(2) AATATA

(3) AATAAA

(4) AACA(A)GA

which are marked in Fig. 39.

The only poly[A] sequence within 250 bases of amino acid 119 of the rat cytochrome $b_5$ pseudogene is five A's in length. This run is 18bp from one of the AATATA sequences so could represent a very short poly[A] tail. The sequence of 43 bases immediately prior to this shows some similarity to the sequence in the functional rat cytochrome $b_5$ cDNA just before its first poly[A] tail (Fig. 43).

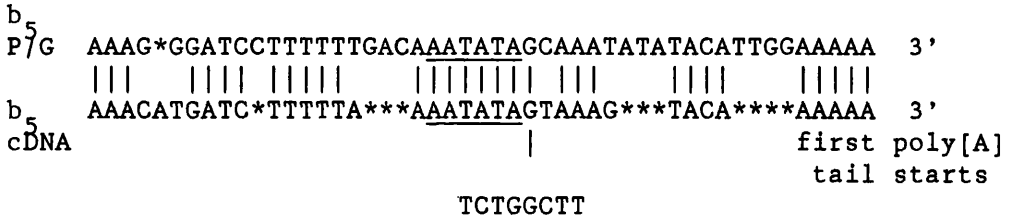



Figure 43. Comparison of the 3' noncoding region of rat cytochrome $b_5$ cDNA ($b_5$ cDNA) and its pseudogene ($b_5$ P/G). Homology is indicated by vertical lines. The poly[A] addition sites are underlined.

However, there is no obvious homology between the rat cytochrome $b_5$ pseudogene and the functional rat cytochrome $b_5$ noncoding section from where amino acid 119 ceases and the section of the rat cytochrome $b_5$ pseudogene shown above (Fig. 43) begins. This covers a region of 37bp in the pseudogene. It is possible that this 3' section of the pseudogene (before the sequence of five A's) is derived from the functional cytochrome $b_5$ cDNA. However, the sequence of 37bp between this homologous region and amino acid 119 of the cytochrome $b_5$ pseudogene seems to bear no resemblance to the functional cytochrome $b_5$ cDNA. The stretch of five A's in the pseudogene is shorter than usually found in a poly[A] tail and the region beyond this is not A-rich. If this did represent a poly[A] tail, it is unclear from where the sequence of 37 bases, which is not homologous to the rat cytochrome $b_5$ cDNA, could have been derived. The homology between the functional rat cytochrome $b_5$ gene 3' non-coding region over 43 bases to the 3' sequence in the rat cytochrome $b_5$ pseudogene could be due to a functional region of "noncoding" DNA sequence. Alternatively, it could be a chance occurrence of homologous A/T rich noncoding regions as it is also possible to line up the same stretch of cytochrome $b_5$ pseudogene sequence with a separate section of the cytochrome $b_5$ 3' noncoding region. Analyzing the area immediately 3' to the postulated poly[A] tail of the pseudogene does not show any sequence which could represent a direct repeat of that at the 5' end of the pseudogene, whereas there is a direct repeat pairing between a region almost immediately 3' to amino acid 120 and a region upstream of the 5' end of the gene.

In conclusion, it seems likely that the cytochrome $b_5$ mRNA which gave rise to the processed pseudogene, underwent an extra splicing event. The 5' end of this intron would be directly after amino acid 120 and the 3' end would be beyond the poly[A] tail of the rat cytochrome $b_5$ cDNA.

4.2 Origins of liver and erythrocyte cytochromes $b_5$

Amino acid sequencing of cytochrome $b_5$ from liver and erythrocytes in several species has gone some way towards determining the mode of origin of its two forms. Within a species, amino acids 1-96 are identical in liver and erythrocyte cytochrome $b_5$, whereas amino acid 97 can differ in the two tissues (section 1.1.3).

Genomic Southern blots were carried out to determine whether there were one or two genes responsible for the two different cytochrome $b_5$ proteins (section 3.3). The sizes of the DNA restriction fragments which hybridized to a full-length rat cytochrome $b_5$ cDNA probe on a genomic Southern blot were to be compared with the sizes of known restriction fragments from the lambda genomic clones containing coding sequences. Both genomic libraries were screened four times each, (after the initial screen of the HaeIII library), using as probes the full-length rat cytochrome $b_5$ cDNA, two overlapping 5' oligonucleotides to radioactively label by "fill-in", the 36mer oligonucleotide (Fig. 20) and the 5' EcoRI/SacI fragment from $pb_5(1)D$ (section 3.1.2). Unfortunately, the genomic clones isolated did not include coding sequence 5' to amino acid 43. Therefore, it is not known how many exons there are at the 5' end of the gene. There is also the added confusion that at least one cytochrome $b_5$ pseudogene exists in rat and there may be more. Analysis of the sizes of the EcoRI fragments in the genomic clones which contain coding sequence for cytochrome $b_5$ and its pseudogene would lead to the expectation of one of the following series of bands on a genomic Southern, I or II. The first possibility (I) would arise if there was one gene coding for cytochrome $b_5$, the second (II) if there were two genes coding for cytochrome $b_5$. The fragments are labelled either as strong, indicating that the band should be intense, or as weak, indicating that the band will only be faint if visible at all after the final stringent wash (60°C/0.1xSSPE) used for the genomic Southern (Fig. 40).

| | |
|---|---|
| [I] (1) > 6150 bp (strong) | amino acids 133-96 - cytochrome $b_5$ |
| (2) 6500 bp (strong) | amino acids 95-43 - cytochrome $b_5$ |
| (3) 2200 bp (strong) | cytochrome $b_5$ pseudogene (amino acids 1-105) |
| (4) 1300 bp (weak) | cytochrome $b_5$ pseudogene (amino acids 106-119) |
| (5) At least one other fragment of unknown size | amino acids 42-1 - cytochrome $b_5$ |

[II] (1) > 6150 bp (strong) amino acids 133-96 - cytochrome $b_5$

(2) 6500 bp (strong) amino acids 95-43 - cytochrome $b_5$

(3) > 4730 bp (weak) amino acids 95-86 - cytochrome $b_5$

(4) > 870 bp (strong) amino acids 85-43 - cytochrome $b_5$

(5) 2200 bp (strong) cytochrome $b_5$ pseudogene (amino acids 1-105)

(6) 1300 bp (weak) cytochrome $b_5$ pseudogene (amino acids 106-119)

(7) At least two, (8) other fragments of unknown sizes — both for amino acids 42-1 - cytochrome $b_5$

If there is one gene coding for cytochrome $b_5$ (possibility [I]):

| Genomic Southern | Lambda Clone (EcoRI fragment) |
|---|---|
| 7.5 kb (one of two?) | > 6150 bp - amino acids 133-96 |
| 6.3 kb | 6500 bp - amino acids 95-43 |
| 2.1 kb | 2200 bp - pseudogene (amino acids 1-103) |
| 5.2 kb | [? 5' fragment cytochrome $b_5$ - amino acids 42-1] |
| ? 7.5 kb (one of two?) | [? 5' fragment cytochrome $b_5$ - amino acids 42-1] |
| not visible | 1300 bp - pseudogene (amino acids 106-119) |
| 3.4 kb | ? unknown pseudogene(s) |
| 2.8 kb | |

If there are two genes coding for cytochrome $b_5$ (possibility [II]):

| Genomic Southern | Lambda Clone (EcoRI fragment) |
|---|---|
| 7.5 kb | > 6150 bp - amino acids 133-96 |
| 6.3 kb | 6500 bp - amino acids 95-43 |
| 2.1 kb | 2200 bp - pseudogene (amino acids 1-103) |
| 5.2 kb | ? > 870 bp - amino acids 85-43 and one intense fragment - amino acids 42-1 |
| 7.5 kb | |
| 3.4 kb | two 5' fragments (faint) cytochrome $b_5$ - amino acids 42-1 |
| 2.8 kb | |
| not visible | 1300 bp - pseudogene (amino acids 106-119) |
| not visible | > 4730 bp - amino acids 95-86 |

From this, it is not clear whether there are one or two genes coding for cytochrome $b_5$. There are too many unknown factors. Isolation of the 5' exon(s) (amino acids 1-42) may clarify the banding pattern on the genomic Southern.

Chicken liver cytochrome $b_5$ has also been studied to determine whether there are one or two genes coding for the protein (Zhang and Somerville, 1988). Chicken genomic DNA digested with EcoRI, HindIII and BglII was electrophoresed on an agarose gel and blotted by the method of Southern (1975). The DNA was hybridized with either the full-length cDNA probe or a 3' HindIII fragment (amino acids 91-132) encoding the hydrophobic domain. The results obtained could be explained by there being two genes coding for cytochrome $b_5$, one of which does not possess the coding sequence for the hydrophobic domain. Alternatively, the results may be equally well explained by the existence of only one gene coding for cytochrome $b_5$.

Similar experiments were attempted during this PhD project, to determine whether there was one or two hydrophilic cytochrome $b_5$ domains in the rat genome. Genomic DNA was digested separately with the restriction enzymes EcoRI, BamHI and HindIII and electrophoresed on agarose gels. The gels were dried down and probed directly (section 2.13) rather than being blotted onto nitrocellulose or nylon membranes. This is reported to increase the sensitivity when using oligonucleotides as probes (Miyada and Wallace, 1987). The gels were probed with end-labelled (section 2.11) domain-specific oligonucleotides. These were the 27mer (amino acids 87-95) which hybridizes to the hydrophilic domain and as a control, the 18mer (amino acids 98-103) which hybridizes specifically to the hydrophobic domain. If the 27mer had hybridized to one genomic fragment this would have indicated that there was one gene coding for cytochrome $b_5$. The presence of two hybridizing genomic fragments would have shown that there were two genes coding for cytochrome $b_5$. These experiments, however, proved unsuccessful - no hybridizing fragments could be seen, even on long exposure of the gels to fast Kodak X-Omat X-ray film.

4.2.1 Two genes coding for cytochrome $b_5$?

If there are two genes coding for cytochrome $b_5$, then they have remained almost 99% conserved at the amino acid level in rat - only amino acid 97 differs between liver and erythrocyte cytochrome $b_5$ over the first 97 amino acids. The homology between species over the first 96 amino acids of liver cytochrome $b_5$ is high - human and pig cytochromes $b_5$ show 93% homology at the amino acid level over this distance. Other more distantly related species show slightly lower homology - cow and rabbit are 83% homologous, for example. However, from this it is clear that complete conservation of the cytochrome $b_5$ protein is not necessary for function. Therefore, if there are two genes, it seems unusual that the liver and erythrocyte forms of cytochrome $b_5$, in all species analyzed to date, remain identical over the first 96 amino acids within a species. As both liver and erythrocyte forms of cytochrome $b_5$ have been found in distantly related species, possible duplication giving rise to two genes could not be a recent event. Homology between two cytochrome $b_5$ genes could have been maintained by gene conversion as is seen in both the human $\alpha$-globin ($\alpha 1, \alpha 2$) (Hess et al., 1983) and the foetal globin ($^{A}\gamma$, $^{G}\gamma$) genes (Slightom et al., 1980). The duplicated foetal globin genes produce identical polypeptides except at one amino acid (136) which is either glycine ($^{G}\gamma$) or alanine ($^{A}\gamma$), (Slightom et al., 1980). The two adult globin genes $\alpha 1$ and $\alpha 2$ encode identical polypeptides.

Therefore, it is possible that one cytochrome $b_5$ gene partially duplicated early in evolution and the two genes have remained homologous in the same way as the human $\alpha$-globin and $\gamma$-globin genes.

If two cytochrome $b_5$ genes do exist, then there must be a mechanism or mechanisms to determine which gene is active in different tissues. As discussed in section 1.1.3, from experiments with the murine Friend virus-induced erythroleukaemic cell-line (Slaughter and Hultquist, 1979), it appears that differentiating erythrocytes produce membrane-bound and not soluble cytochrome $b_5$ up to and possibly beyond the polychromatophilic and orthochromatophilic erythroblast stages. Therefore, there seems to be a developmental 'switch' during erythrocyte maturation to produce the soluble form, which does not occur in other cell types. The mechanism of this 'switch' is completely unknown for cytochrome $b_5$. If there are two genes coding for cytochrome $b_5$, the liver protein producing gene would have to be inactivated and the erythrocyte protein producing gene would need to be activated only in developing erythrocytes. For this to occur there would need to be at least one tissue-specific trans-acting molecule that could bind to gene regulatory regions. These may be at the 5' or 3' ends of the gene or even within an exon or intron and could include an enhancer element. Primary, secondary and

tertiary structure of regulatory regions as well as the degree and sites of methylation and on a larger scale, chromatin structure, could all be involved in modulating gene expression in either a positive or negative sense.

4.2.2 One gene coding for cytochrome $b_5$?

Liver and erythrocyte cytochromes $b_5$ could arise from a single gene. Different proteins can be generated from one gene by several means:-

(1) DNA rearrangement, as occurs in immunoglobulin genes to provide multiple proteins (Early et al., 1979; Early et al., 1980a; Davis et al., 1980).

(2) One gene may be preceded by two promoter regions, each with its own tissue-specific regulatory elements. This is seen in the mouse $\alpha$-amylase gene which gives rise to mRNAs specific for the liver and salivary gland (Hagenbuchle et al., 1981; Young et al., 1981).

(3) A single gene may result in tissue-specific expression of one of two different mRNAs varying only in the length of their 3' untranslated regions. This has been observed for the vimentin gene in chicken (Capetanaki et al., 1983) and is thought to be due to specific transcription termination at one or the other poly[A] sites or to post-transcriptional processing of the 3' region.

(4) Protein splicing could occur during development.

(5) Co- or post-transcriptional change of one or more bases giving rise to two proteins from a single primary transcript. This is seen in apolipoprotein-B (Powell et al., 1987).

(6) Alternative splicing of a single primary transcript to give two alternative proteins. There are many examples of differential RNA processing including the calcitonin/CGRP proteins (Amara et al., 1982).

Taking each of these possibilities, they will be discussed in relation to cytochrome $b_5$.

The huge variety of immunoglobulins arises out of the rearrangement of large sections of DNA before transcription, to give many combinations of fragments. The cytochrome $b_5$ gene, however, is not functioning to provide variation to anywhere near the extent of the immunoglobulins. The rearrangement of what would be a small extra fragment seems highly unlikely. This could be confirmed by Southern blot analysis of genomic DNA isolated from cells at various stages of development.

If a single gene is preceded by two promoter regions, it is only the 5' untranslated region which differs between the two mRNAs. The amino acid content is not affected. This is clearly not the case for the cytochrome $b_5$ gene.

Tissue-specific expression arising from a difference in the 3' untranslated region also results in a protein being identical in all tissues or at all stages of development. Again this would not explain the structures of the cytochrome $b_5$ proteins.

Proteolytic processing was originally thought to be responsible for the production of a truncated erythrocyte cytochrome $b_5$ protein due to the bovine amino acid sequence being identical over the first 97 amino acids in liver and erythrocytes (Slaughter et al., 1982) (section 1.1.3). However, as discussed previously (section 1.1.3), the primary structures of cytochrome $b_5$ have shown that in all other known cases, amino acid 97 differs in the two tissues. It is difficult to imagine a mechanism involving protein splicing that would alter just the terminal amino acid of the erythrocytic cytochrome $b_5$ protein (or an internal amino acid in the liver cytochrome $b_5$ protein).

An unusual form of tissue-specific RNA processing is found to produce two forms of apolipoprotein-B (Powell et al., 1987). Intestinal apolipoprotein-B48 (apo-B48) is colinear with the amino-terminal section of hepatic apolipoprotein-B100 (apo-B100). Isolation of cDNA clones from the small intestine showed that apo-B48 is identical to apo-B100 up to amino acid 2152. There is then an alteration of one base in the codon for amino acid 2153, from a C to a T, which produces a stop codon in apo-B48. Amino acid 2153 is reported to be glutamine (CAA) in five published sequences of hepatic cDNA and also in three published genomic sequences. It was shown that apo-B is a single copy gene, that differential splicing was not involved in the production of apo-B48 and that the stop codon is not found in the genome. Therefore, two different proteins are being produced from the same primary transcript due to a single mRNA nucleotide change.

A similar situation could be responsible for the diversity in cytochrome $b_5$ proteins. Whichever codon is used for threonine at amino acid position 97 in the liver form of cytochrome $b_5$ - either ACU, ACC, ACA, or ACG, a single alteration of the first base of the triplet, A, to C, produces proline (for human and rabbit erythrocytes) or to U gives serine (for porcine erythrocytes) (section 1.1.3). To produce a stop codon from leucine (CTT) at amino acid 98 in the rat, chicken, horse, rabbit and porcine cytochrome $b_5$, three base alterations are required. Although these alterations could take place, it would require several base changes to produce the truncated erythrocyte cytochrome $b_5$, whereas in the case of the apolipoprotein-B proteins, only a single base change is necessary to produce the same effect. The other alternative for cytochrome $b_5$ is that a single nucleotide alteration takes place to change amino acid 97 and three additional nucleotides are inserted into the transcript to produce a stop codon. This is seen in trypanosomes,

where nucleotides are inserted by an RNA editing mechanism (Benne et al., 1986; Feagin et al., 1987). However, either of these possibilities seems unnecessarily complicated for an organism to adopt in the case of cytochrome $b_5$.

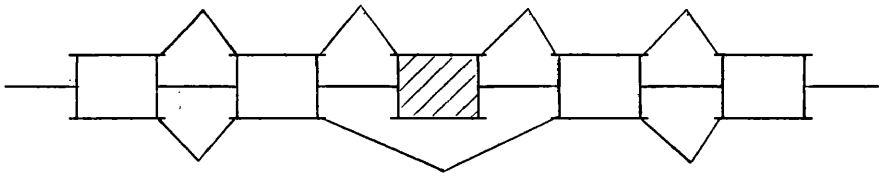
The final explanation for the production of two cytochrome $b_5$ proteins is that alternate splicing of a single primary RNA transcript gives rise to two mature mRNAs.

A primary RNA transcript is said to undergo alternate splicing if at least one pair of donor and acceptor sites that are joined together in the formation of one mRNA, do not do so in the formation of a second mRNA.

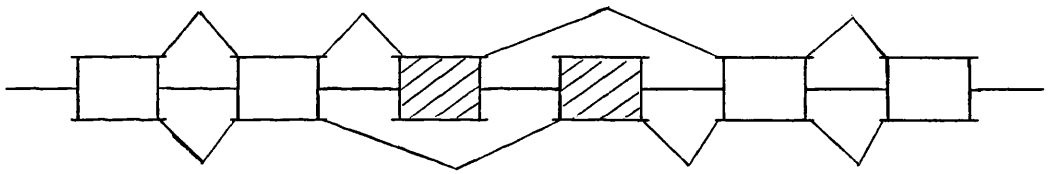
Differential RNA processing is being found increasingly to be responsible for the production of more than one protein from a single gene. It provides a strategy for generating diversity in gene expression and is used as a means of generating tissue-specific expression and gene regulation during development. Organisms ranging from Drosophila to humans, including their RNA and DNA viruses, use alternative RNA splicing to generate protein diversity.

The general patterns observed in alternate splicing are shown below in diagrammatic form with examples (Fig. 44).

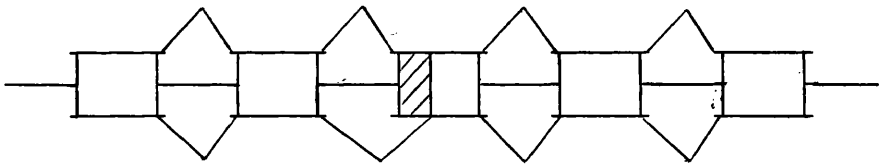
1) cassette - eg. mouse $\alpha$ A-crystallin gene (King and Piatigorsky, 1983)



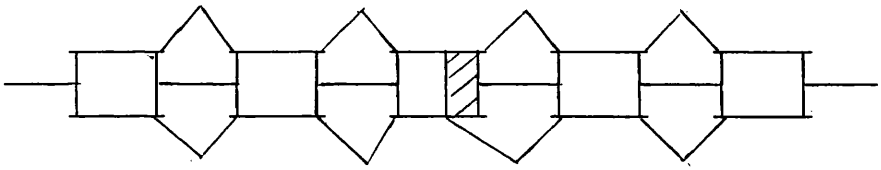
2) mutually exclusive exons - eg. myosin light chain 1/3, found in chicken, rat and mouse (Periasamy et al., 1984; Nabeshima et al., 1984).



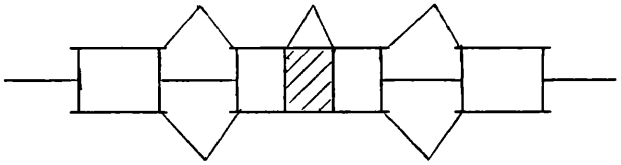
3) internal acceptor site - eg. fibronectin in rat and human (Kornblihtt et al., 1984a; 1984b; Schwarzbauer et al., 1983; Tamkun et al., 1984).



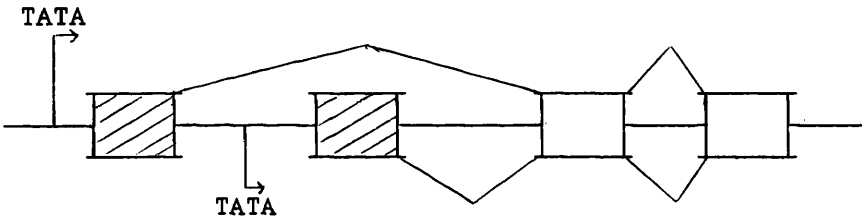
4) internal donor site - eg. human fibronectin (Kornblihtt et al., 1984a; 1984b).



5) retained intron - eg. Drosophila P transposable element (Laski et al., 1986).



6) alternative promoters giving alternative 5' exons - eg. myosin light chain 1/3 (Periasamy et al., 1984; Nabeshima et al., 1984).




7) alternative poly[A] sites giving alternative 3' exons - eg. calcitonin (Amara et al., 1982).

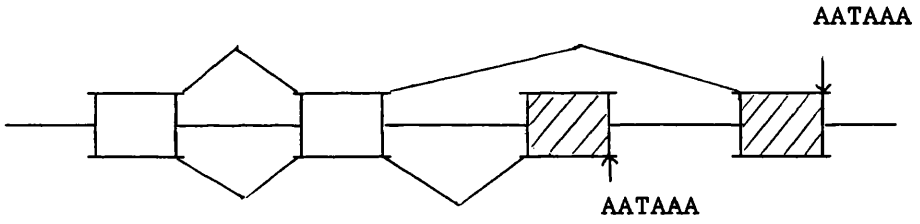



Figure 44. Classification of types of differentially spliced genes. Boxes indicate exons, horizontal lines indicate introns. Alternate splicing patterns are indicated. The AATAAA sequence is the poly[A] addition signal. TATA is the promoter region.

There are, of course, differentially spliced primary transcripts which fit into more than one of these general classification groups.

Differential splicing could explain the existence of the two cytochrome $b_5$ proteins. The point of divergence between erythrocyte and liver cytochrome $b_5$ mRNAs does not correspond to a genomic splice junction as might have been expected if the gene was alternately spliced. Analysis of the intron/exon structure from sequencing the rat cytochrome $b_5$ genomic clones shows that there is an intron between amino acids 95 and 96 rather than after amino acid 97 (Fig. 30) - the terminal amino acid of erythrocyte cytochrome $b_5$. The next exon noted in the genomic clones contains codons for amino acids 96-107. Therefore, if the cytochrome $b_5$ proteins are produced by differential splicing, there must be an alternate exon to this which contains codons for amino acids 96 and 97. Amino acid 97 would differ in the two exons in most species. There would probably also be two separate stop codons and 3' untranslated regions - one stop codon would lie immediately after amino acid 97 in the undetected exon to truncate the erythrocyte protein, followed by its 3' untranslated region. The other stop codon has been identified in the rat cytochrome $b_5$ genomic clone and occurs after amino acid 133 followed by 3' untranslated sequence. This theoretical situation is shown in diagrammatic form below (Fig. 45). The exon containing codons for amino acids 96 and 97 only plus a 3' untranslated region could also theoretically lie downstream of both other 3' exons or between the exons for amino acids 96-107 and 108-133 as well as upstream of the exon for amino acids 96-107 as shown. The exons containing sequence for amino acids 1-95 would be shared by both liver and erythrocyte mRNAs.

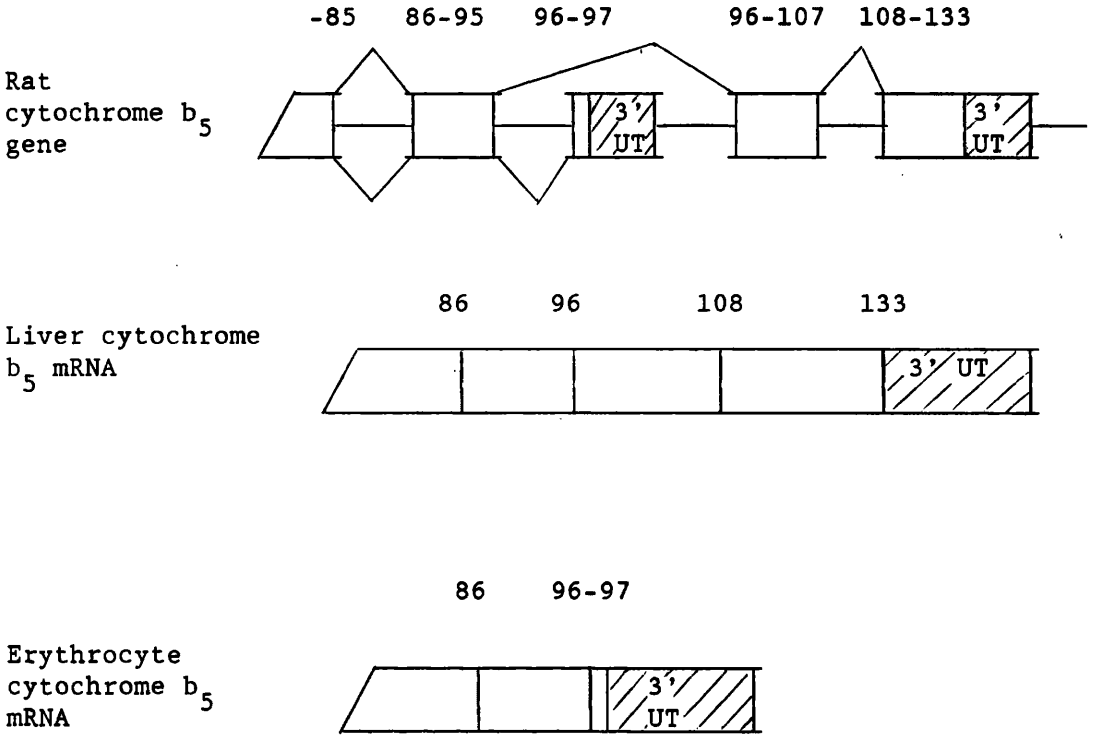


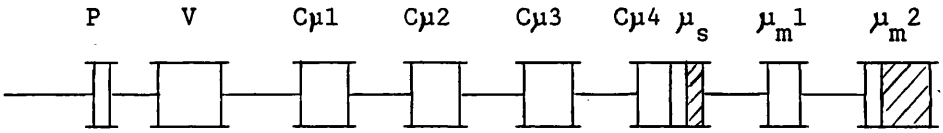

Figure 45. Showing possible alternate splicing of a single rat cytochrome $b_5$ gene using alternative exons and two alternative poly[A] addition sites to produce a liver protein of 133 amino acids and a truncated erythrocyte protein of 97 amino acids. Amino acid 97 would differ in the two exons and thus also the two mRNAs in most species. Open blocks indicate coding sequence, shaded blocks indicate 3' untranslated regions (3' UT ) and single horizontal lines indicate introns. The numbers indicate amino acid positions.

Therefore, the cytochrome $b_5$ gene in this instance would utilize a combination of cassette and alternative polyadenylation site mechanisms to produce two different mRNAs.

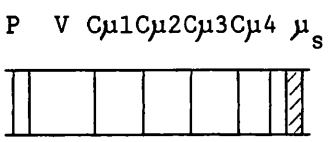
A similar situation is seen in the immunoglobulin $\mu$ gene which produces membrane-bound ( $\mu_m$) and secreted ( $\mu_s$) forms of IgM in early B cells and mature lymphocytes respectively. There are two $\mu_m$ specific exons, the most 3' of which consists of just two amino acids plus a stop codon, then its own 3' non-coding region and poly[A] addition site (Early et al., 1980b) as seen below in Figure 46.

It is now known that two pre-mRNAs are produced from the IgM gene to give rise to the membrane-bound ($\mu_m$) and secreted ($\mu_s$) forms of the protein. Regulation of expression of the IgM gene is, therefore, at the level of transcription termination. If the postulated mechanism of alternate splicing of a single rat cytochrome $b_5$ gene (Fig. 45) is correct, it would seem likely that cytochrome $b_5$ expression is also regulated in this manner.

IgM gene:




secreted ( $\mu_s$ ) mRNA:




membrane-bound ( $\mu_m$ ) mRNA:

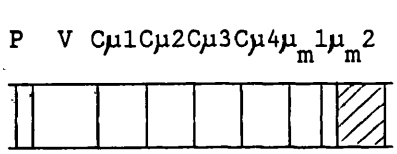



Figure 46. Splicing patterns for IgM $\mu_m$ and $\mu_s$ mRNAs using different 3' exons and polyadenylation sites. P is the signal peptide exon, V is the rearranged $V_H$ exon and Cμ1 - Cμ4 are the constant regions. The alternate exons are $\mu_s$ with its 3' non-coding region and $\mu_m1$ and $\mu_m2$ with their 3' non-coding region. The $\mu_m2$ exon only contains six bases coding for amino acids. Open blocks indicate coding sequence, shaded blocks indicate 3' untranslated regions and single horizontal lines indicate introns.

The mechanisms responsible for alternate splicing remain unknown, although general theories have been postulated. When the differential splicing arises from the use of two polyadenylation sites, the choice of one over the other has been hypothesized to be due to trans-acting factors.

Studies of the immunoglobulin $\mu$ gene shows correlation between the length of the C$\mu$4 to $\mu_m$1 intron and the ratio of $\mu_m$ to $\mu_s$ mRNAs. The shorter the intron, the more the $\mu_m$ poly[A] addition site is used (Tsurushita and Korn, 1987; Galli et al., 1988). Insertion of homologous sequences restores the correct expression showing that the spacing is critical (Galli et al., 1988). It was the distance between the poly[A] addition sites rather than the distance between splice sites which mattered (Galli et al., 1988) which indicates that the choice of poly[A] site is crucial to splicing. This is supported by experiments which impair $\mu_m$ splicing (Galli et al., 1988). Removal of the $\mu_m$1 splice acceptor site, or the $\mu_m$ splice donor site which prevents normal C$\mu$4 to $\mu_m$1 splicing, does not increase $\mu_s$ RNA production. Deletion of the $\mu_m$1 exon which prevents C$\mu$4/ $\mu_m$1 and $\mu_m$1 /$\mu_m$2 splicing also does not result in increased $\mu_s$ RNA levels. Therefore, there is no apparent competition between a C$\mu$4/ $\mu_m$1 or $\mu_m$1/ $\mu_m$2 splice event and selection of the $\mu_s$ poly[A] site, as was suggested by Tsurushita and Korn, 1987. The dominant factor in transcription seems to be poly[A] site selection with splicing occuring after this selection. This is consistent with the fact that generally, polyadenylation is thought to preceed splicing (Darnell, 1982; Nevins and Darnell, 1978; Lai et al., 1978). The choice of poly[A] site would probably involve a trans-acting poly[A] site factor.

The calcitonin/CGRP gene, however, seems to show that exon splicing dictates expression rather than polyadenylation. This gene is developmentally regulated and the protein products cell-specific, resulting in the CGRP protein in neurons and the calcitonin protein in thyroid 'C' cells (Fig. 47).

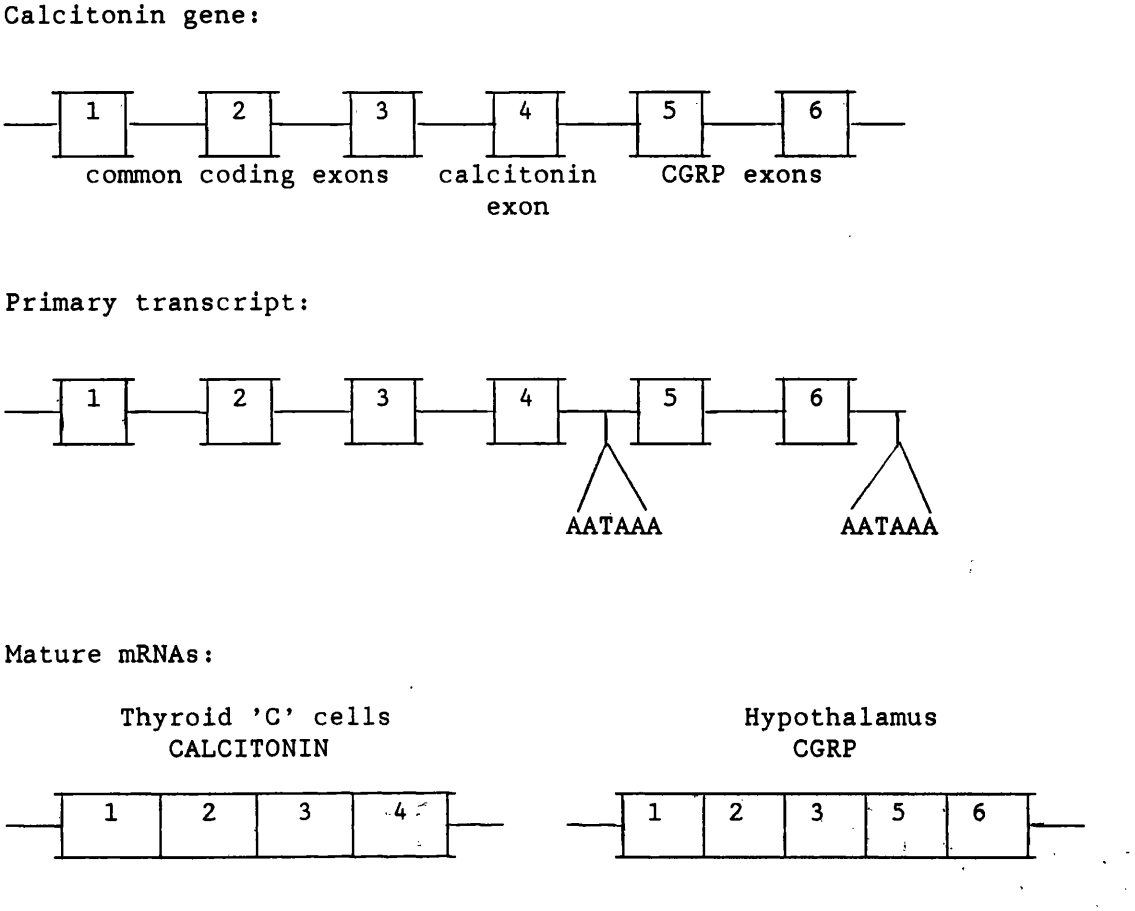



Figure 47. Splicing patterns for calcitonin and CGRP mRNAs using different 3' exons and polyadenylation sites. The alternate exons are exon 4 or exons 5 and 6 with their 3' non-coding regions. Open blocks indicate exons and horizontal lines indicate introns.

The poly[A] site selection does not seem to be regulated in itself. Expression of exons 4, 5 and 6 in different cell-lines does not show the poly[A] site specificity seen with the whole gene (Leff et al., 1987). The theory postulated is that there is a neuron-specific factor which commits the splicing of exons 3 to 5 by inhibiting the use of the calcitonin poly[A] site. B-lymphocytes, which have no "splice commitment machinery", cause polyadenylation at the calcitonin site (Leff et al., 1987). A gene with a mutant calcitonin poly[A] signal transfected into cell lines which normally favour calcitonin production do not show splicing of exons 3 to 4 (for calcitonin) or exons 3 to 5 (for CGRP) (Leff et al., 1987). Removal of sequence from the calcitonin-specific 3'-splice junction and its replacement with a heterologous splice acceptor showed that there could be a cis-acting element near the calcitonin acceptor site which inhibits the formation of calcitonin RNA in CGRP producing cells (Emerson et al., 1989). Therefore, there appears to be regulation by cis-acting sequences to which trans-acting factors may bind. It seems that both splice site and polyadenylation site selection can be responsible for this kind of RNA processing.

If the cytochrome $b_5$ proteins arise from a single gene by differential processing as suggested, either model of mRNA expression could explain the outcome.

It is necessary in the future to determine whether there are one or two genes coding for cytochrome $b_5$ from genomic Southern blots and the complete intron/exon structure of the cytochrome $b_5$ gene(s) from genomic clones. Alternatively, if cytochrome $b_5$ cDNA clones can be isolated from an erythroid cell library, this may answer the question.

If there are two genes coding for cytochrome $b_5$ the mechanism of the developmental "switch" needs to be elucidated. Equally, if there is a single gene coding for cytochrome $b_5$, the possibilities of co- or post-transcriptional change and differential splicing will need to be investigated.

BIBLIOGRAPHY

Abe, K., Kimura, S., Kizawa, R., Anan, F.K., Sugita, Y. (1985) J. Biochem. 97, 1659-1668.

Abe, K., Sugita, Y. (1979) J. Biochem. 101, 423-428.

Alfano, J., Cinti, D. (1977) Pharmacologist 19, 233.

Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S., Evans, R.M. (1982) Nature 298, 240-244.

Arnold, H., Botcher, H.W., Hufnagel, D., Lohr, G.W. (1978) XVIII Congress of International Society of Haematology, Abstracts (II) Paris 752 (abstr.).

Battey, J., Max, E.E., McBride, W.O., Swan, D., Leder, P. (1982) Proc. Natl. Acad. Sci USA 79, 5956-5960.

Beams, H.W., Kessel, R.G. (1966) Am. J. Anat. 118, 471.

Beck von Bodman, S., Schuler, M.A., Jollie, D.R., Sligar, S.G. (1986) 83, 9443-9447.

Benne, R., Van Den Burg, J., Brakenhoff, J.P.J., Sloof, P., Van Boom, J.H., Tromp, M.C. (1986) Cell 46, 819-826.

Berezney, R., Crane, F.L. (1971) Biochem. Biophys. Res. Commun. 43, 1017-1023.

Bergeron, J.J.M., Ehrenreich, J.H., Siekevitz, P., Palade, G.E. (1973) J. Cell. Biol. 59, 73-88.

Biggin, M.D., Gibson, T.J., Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.

Birnboim, H.C., Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523.

Bois-Poltoratsky, R., Ehrenberg, A. (1967) Eur. J. Biochem. 2, 361-365

Borgese, N., Macconi, D., Parola, L., Pietrini, G. (1982) J. Biol. Chem. 257, 13854-13861.

Borgese, N., Mendelosi, J. (1980) J. Cell. Biol. 85, 501-515.

Boyer, S.H., Wuu, K.D., Noyes, A.N., Young, R., Scher, W., Friend, C., Preisler, H.D., Bank, A. (1972) Blood 40, 823-835.

Brenner, R.R., Peluffo, R.O. (1966) J. Biol. Chem. 241, 5213-5219.

Bull, P.C. (1990) PhD thesis, University of London.

Bull, P.C., Shephard, E.A., Povey, S., Santisteban, I., Phillips, I. R. (1988) Ann. Hum. Genet. 52, 263-268.

Capetanaki, Y.G., Ngai, J., Flytzanis, C.N., Lazandes, E. (1983) Cell 35, 411-420.

Castuma, J.C., Catala, A., Brenner, R.R. (1972) J. Lipid Res. 13, 783-789.

Chen, M-J., Shimada, T., Moulton, A.D., Harrison, M., Nienhaus, A.W. (1982) Proc. Natl. Acad. Sci. USA 79, 7435-7439.

Chiang, J.Y.L. (1981) Arch. Biochem. Biophys. 211, 662-673.

Cohen, B.S., Estabrook, R.W. (1971)a Arch. Biochem. Biophys. 143, 37-45.

Cohen, B.S., Estabrook, R.W. (1971)b Arch. Biochem. Biophys. 143, 46-53.

Cohen, B.S., Estabrook, R.W. (1971)c Arch. Biochem. Biophys. 143, 54-65.

Correia, M.A., Mannering, G.J. (1973)a Mol. Pharmacol. 9, 455-469.

Correia, M.A., Mannering, G.J. (1973)b Mol. Pharmacol. 9, 470-485.

Crabb, J.W., Tarr, G.E., Yasunobu, K.T., Iyanagi, T., Coon, M.J. (1980) Biochem. Biophys. Res. Commun. 95, 1650-1655.

Dailey, H.A., Strittmatter, P. (1979) J. Biol. Chem. 254, 5388-5396.

Darnell, J.E.Jr (1982) Nature 297, 365-371.

Davis, M.M., Calame, K., Early, P.W., Livant, D.L., Joho, R., Weissman, I.L., Hood, L. (1980) Nature 283, 733-739.

Denhardt, D. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.

Denk, H., Eckerstorfer, R.E., Talcott, R.E., Schenkman, J.B. (1977) Biochem. Pharmacol. 26, 1125-1130.

Douglas, R.H., Hultquist, D.E. (1978) Proc. Natl. Acad. Sci. USA 75, 3118-3122.

Early, P.W., Davis, M.M., Kaback, D.P., Davidson, N., Hood, L. (1979) Proc. Natl. Acad. Sci. USA 76, 857-861.

Early, P.W., Huang, H., Davis, M.M., Calame, K., Hood, L. (1980)a Cell 19, 981-992.

Early, P.W., Rogers, J., Davis, M.M., Calame, K., Bond, M., Wall, R., Hood, L. (1980)b Cell 20, 313-319.

Emerson, B.E., Yeakly, J.M., Rosenfeld M.G., Hedjram, F., Guise, J.W. (1989) Nature 341, 76-80.

Enoch, H.G., Strittmatter, P. (1979) J. Biol. Chem. 254, 8976-8987.

Feagin, J.E., Jasmer, D.P., Stuart, K. (1987) Cell 49, 337-345.

Feinberg, A.P., Vogelstein, B. (1984) Anal. Biochem. 137, 266-267.

Fleming, P.J., Dailey, H.A., Cocoran, D., Strittmatter, P. (1978) J. Biol. Chem. 253, 5369-5372.

Franke, W.W., Deumling, B., Ermen, B., Jarasch, E-D., Kleinig, H. (1970) J. Cell. Biol. 46, 379-395.

Fukushima, H., Grinstead, G.F., Gaylor, J.L. (1981) J. Biol. Chem. 256, 4822-4826.

Fukushima, K., Sato, R. (1973) J. Biochem. 74, 161-173.

Galli, G., Guise, J., Tucker, P.W., Nevin, J.R. (1988) Proc. Natl. Acad. Sci. USA 85, 2439-2443.

Gibson, Q.H. (1948) Biochem. J. 42, 13-23.

Goto-Tamura, R., Takesue, Y., Takesue, S. (1976) Biochim. Biophys. Acta 423, 293-302.

Gutmann, A.R., Jandorf, B.J., Bodansky, O. (1947) J. Biol. Chem. 169, 145-152.

Hagenbuchle, O., Tori, M., Schibler, U., Bovey, R., Wellauer, P.K. (1981) Nature 289, 643-646.

Hanahan, D. (1985) in 'DNA cloning, a practical approach' Vol. 1 IRL press 109-135.

Hegesh, E., Avron, M. (1967)a Biochim. Biophys. Acta 146, 397-408.

Hegesh, E., Avron, M. (1967)b Biochim. Biophys. Acta 146, 91-101.

Hegesh, E., Hegesh, J., Kaftory, A. (1986) New Eng. J. Med. 314, 757-761.

Hess, J.F., Fox, M., Scmid, C., Shen, C.J. (1983) Proc. Natl. Acad. Sci. USA 80, 5970-5974.

Hildebrandt, A., Estabrook, R.W. (1971) Arch. Biochem. Biophys. 143, 66-79.

Hino, Y., Asano, A., Sato, R. (1978) J. Biochem. 83, 925-934.

Hollis, G.F., Hieter, P.A., McBride, O.W., Susan, D., Leder, P. (1982) Nature 296, 321-325.

Holloway, P.W. (1971) Biochemistry 10, 1556-1560.

Holloway, P.W., Katz, J.T., (1972) Biochemistry 11, 3689-3696.

Holloway, P.W., Wakil, S.J. (1970) J. Biol. Chem. 245, 1862-1865.

Hultquist, D.E., Passon, P.G. (1971) Nature (New Biol.) 229, 252-254.

Ikawa, Y., Furusawa, M., Sugano, H. (1973) Bibl. Haematol. 39, 955-967.

Ilan, L., Ilan, R., Cuti, D.L. (1981) J. Biol. Chem. 256, 10066-10072.

Ish-Horowicz, D., Burke, J.F. (1981) Nucl. Acids Res. 9, 2989-2998.

Itakura, K., Rossi, J.J., Wallace, R.B. (1984) Ann. Rev. Biochem. 53, 323-356.

Ito, A. (1980)a J. Biochem. 87, 73-80.

Ito, A. (1980)b J. Biochem. 87, 63-71.

Ito, A., Sato, R. (1968) J. Biol. Chem. 243, 4922-4923.

Iyanagi, T., Mason, A.S. (1973) Biochemistry 12, 2297-2308.

Jacq, C., Miller, J.R., Brownlee, G.G. (1977) Cell 12, 109-120

Jaffé, E.R. (1964) The Red Blood Cell (Ed. Bishop, C. and Surgeoner, D.M.) 397-422 Academic Press New York.

Jaffé, E.R., Neumann, G. (1968) Ann. NY Acad. Sci. 151, 795-806.

Kaplan, J.C., Beutler, E. (1967) Biochem. Biophys. Res. Commun. 29, 605-610.

Karin, M., Richards, R.I. (1982) Nature 299, 797-802.

Kensil, C.R., Hediger, M.A., Ozols, J., Strittmatter, P. (1983) J. Biol. Chem. 258, 14656-14663.

Keyes, S.R., Alfano, J.A., Jansson, I., Cinti, D.L. (1979) J. Biol. Chem. 254, 7778-7784.

Kimura, S., Abe, K., Sugita, Y. (1984) FEBS Lett. 169, 443-446.

King, C.R., Piatigorsky, J. (1983) Cell 32, 707-712.

Kitao, T., Sugita, Y., Yoneyama, Y., Hattori, K. (1974) Blood 44, 879-884.

Kondo, K., Tajuma, S., Sato, R., Narita, K. (1979) J. Biochem. 86, 1119-1128.

Kornblihtt, A.R., Vibe-Pedersen, K., Baralle, F.E. (1984)a EMBO J. 3, 221-226.

Kornblihtt, A.R., Vibe-Pedersen, K., Baralle, F.E. (1984)b Nucl. Acids Res. 12, 5853-5867.

Kuma, F., Inomata, H. (1972) J. Biol. Chem. 247, 556-560.

Kuma, F., Prough, R.A., Masters, B.S.S. (1976) Arch. Biochem. Biophys. 172, 600-607.

Lai, C-J., Dhar, R., Khoury, G. (1978) Cell 14, 971-982.

Laski, F.A., Rio, D.C., Rubin, G.M. (1986) Cell 44, 7-19.

Lathe, R. (1985) J. Mol. Biol. 183, 1-12.

Lederer, F., Ghir, R., Guiard, B., Cortial, S., Ito, A. (1983) Eur. J. Biochem. 132, 95-102.

Lee, T-C., Baker, R.C., Stephens, N., Snyder, F. (1977) Biochim. Biophys. Acta 489, 25-31.

Leff, S.E., Evans, R.M., Rosenfeld, M.G. (1987) Cell 48, 517-524.

Lemischka, I., Sharp, P. (1982) Nature 300, 330-335.

Leroux, A., Jumen, C., Kaplan, J.C., Bamberger, J. (1975) Nature 258, 619-620.

Leroux, A., Kaplan, J.C. (1972) Biochem. Biophys. Res. Commun. 49, 945-950.

Lostanlen, D., Lenoir, G., Kaplan, J.C. (1981) J. Clin. Invest. 68, 279-285.

Loverde, A., Strittmatter, P. (1968) J. Biol. Chem. 243, 5779-5787.

Mahler, H.R., Raw, I., Molinari, R., do Amaral, D.F. (1958) J. Biol. Chem. 233, 230-239.

Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Lab. Press: New York.

Mannering, G.J., Kuwahara, S., Omura, T. (1974) Biochem. Biophys. Res. Commun. 57, 476-481.

Marsh, J.B., James, A.T. (1962) Biochim. Biophys. Acta 60, 320-328.

Masters, B.S.S., Baron, J., Taylor, W.E., Isaacson, E.C., Lo Spalluto, J. (1971) J. Biol. Chem. 246, 4143-4150.

Mathews, F.S. (1980) Biochim. Biophys. Acta 622, 375-379.

Mathews, F.S., Levine, M., Argos, P. (1971) Nature (New Biol.) 233, 15-16.

Mathews, F.S., Levine, M., Argos, P. (1972) J. Mol. Biol. 64, 449-464.

Messing, J. (1979) Recombinant DNA Technical Bulletin, NIH Publication No. 77-79, Vol. 2, No. 2, 43-48.

Messing, J., Vieira, J. (1982) Gene 19, 269-276.

Mihari, K., Sato, R., Sakakibara, R., Wada, H. (1978) Biochemistry 17, 2829-2834.

Miyada, C.G., Wallace, R.B. (1987) in Methods in Enzymology 154, 95-107.

Moos, M., Gallwitz, D. (1983) EMBO J. 2, 757-761.

Nabeshima, Y., Kuji-Kuriyama, Y., Marramatsu, M., Ogata, K. (1984) Nature 308, 333-338.

Nagoa, M., Ishibashi, T., Okayasu, T., Imai, Y. (1983) FEBS Lett. 155, 11-14.

Nebert, D.W., Gelboin, H.V. (1968) J. Biol. Chem. 243, 6250-6261.

Nevins, J.R., Darnell, J.E.Jr. (1978) Cell 15, 1477-1493.

Nilsson, A., Johnson, B.C. (1963) Arch. Biochem. Biophys. 101, 494-498.

Nishioka, Y., Leder, A., Leder, P. (1980) Proc. Natl. Acad. Sci. USA 77, 2806-2809.

Noshiro, M., Omura, T. (1978) J. Biochem. 83, 61-77.

Notake, M., Tobimatsu, T., Watanabe, Y., Takahashi, H., Mishina, M., Numa, S. (1983) FEBS Lett. 156, 67-71.

Nugteren, D.H. (1965) Biochim. Biophys. Acta. 106, 280-290.

Okayasu, T., Ono, T., Shinojima, K., Imai, Y. (1976) Lipids 12, 267-271.

Omura, T. Oral presentation at 8$^{th}$ International Congress of Biochemistry, as quoted by Gillette, J.R., Davis, D.C., Sasame, H.A. (1969) Ann. Rev. Pharmacol 12, 57-84.

Orrenius, S., Ernster, L. (1964) Biochem. Biophys. Res. Commun. 16, 60-65.

Ortiz de Montillado, P.R. (ed.) (1986) Cytochrome P-450: Structure, Mechanism and Biochemistry, Plenum, New York.

Oshino, N., Imai, Y., Sato, R. (1966) Biochim. Biophys. Acta. 128, 13-28.

Oshino, N., Imai, Y., Sato, R. (1971) J. Biochem. 69, 155-167.

Oshino, N., Omura, T. (1973) Arch. Biochem. Biophys. 157, 395-404.

Ostertag, W., Meldens, H., Steinheider, G., Kluge, N., Dubi, S. (1972) Nature (New Biol.) 239, 231-234.

Ozols, J. (1970) J. Biol. Chem. 245, 4863-4874.

Ozols, J. (1972) J. Biol. Chem. 247, 2242-2245.

Ozols, J. (1974) Biochemistry 13, 426-434.

Ozols, J. (1975) in First International Conference on Solid Phase Methods in Protein Sequence Analysis (Laursen, R.H., ed.) 221-228 Pierce Chem. Co., Rockford.

Ozols, J., Carr, S.A., Strittmatter, P. (1984) J. Biol. Chem. 259, 13349-13354.

Ozols, J., Gerard, C. (1977)a Proc. Natl. Acad. Sci. USA 74, 3725-3729.

Ozols, J., Gerard, C. (1977)b J. Biol. Chem. 252, 8549-8553.

Ozols, J., Gerard, C., Nobrega, F.G. (1976) J. Biol. Chem. 251, 6767-6774.

Ozols, J., Heinemann, F.S. (1982) Biochim. Biophys. Acta 704, 163-173.

Ozols, J., Korza, G., Heinemann, F.S., Hediger, M.A., Strittmatter, P. (1985) J. Biol. Chem. 260, 11953-11961.

Paltauf, F., Prough, R.A., Masters, B.S.S., Johnston, J.M. (1974) J. Biol. Chem. 249, 2661-2662.

Passon, P.G., Hultquist, D.E. (1972) Biochim. Biophys. Acta 275, 62-73.

Passon, P.G., Reed, D.W., Hultquist, D.E. (1972) Biochim. Biophys. Acta 275, 51-61.

Paulakis, G.N., Lockard, R.E., Vamvakopoulos, N., Rieser, L., Raj Bhandary, U.L., Vournakis, J.N. (1980) Cell 19, 91-102.

Periasamy, M., Strehler, E.E., Garfinkel, L.I., Gubits, R.M., Ruiz Opazo, N., Nadal-Ginard, B. (1984) J. Biol Chem. 259, 13595-13604.

Petragnani, N., Noguiera, O.C., Raw, I. (1959) Nature 184, 1651.

Powell, L.M., Wallis, S.C., Pease, R.J., Edwards, Y.H., Knott, T.J., Scott, J. (1987) Cell 50, 831-840.

Proudfoot, N.J., Brownlee, G.G. (1976) Nature 263, 211-214.

Prough, R.A., Burke, M.D. (1975) Arch. Biochem. Biophys. 170, 160-168.

Raftell, M., Orrenius, S. (1970) Biochim. Biophys. Acta 233, 358-365.

Raw, I., Mahler, H.R. (1959) J. Biol. Chem. 234, 1867-1873.

Raw, I., Molinari, R., do Amaral, D.F., Mahler, H.R. (1958) J. Biol. Chem. 233, 225-239.

Raw, I., Petragnani, N., Camarago-Nogueira, O. (1960) J. Biol. Chem. 235, 1517-1520.

Reddy, W.R., Kupfer, D., Caspi, E. (1976) J. Biol. Chem. 252, 2797-2801.

Remmer, H. (1972) Eur. J. Clin. Pharmacol. 5, 116-136.

Rich, P.R., Tiede, D.M., Bonner, W.D. (1979) Biochim. Biophys. Acta 546, 307-315.

Rollins, T.E., Smith, W.C. (1980) J. Biol. Chem. 255, 4872-4875.

Ross, J., Ikawa, Y., Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 3620-3623.

Sagara, Y., Harano, T., Omura, T. (1978) J. Biochem. 83, 807-812.

Sanger, F., Nicklen, S., Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J., Roe, B.A. (1980) J. Mol. Biol. 143, 161-178.

Sargent, T.D., Jung-Rung, W., Sala-Trepat, J.M., Wallace, R.B., Reyes, A.A., Bonner, J. (1979) Proc. Natl. Acad. Sci. USA 76, 3256-3260.

Sato, T., Friend, C., Deharven, E. (1971) Cancer Res. 31, 1402-1417.

Schafer, D.A., Hultquist, D.E. (1983) Biochem. Biophys. Res. Commun. 115, 807-813.

Schwartz, J.M., Paress, P.S., Ross, J.M., Dipillo, F., Rizek, R. (1972) J. Clin. Investig. 51, 1594-1601.

Schwartz, J.M., Reiss, A.L., Jaffe, E.R. (1983) in (Stanbury, J.B., Wyngaarden, J.B., Fredrickson, D.S., Goldstein, J.L., Brown, M.S., eds.) The metabolic basis of inherited disease, 5th ed. New York: McGraw-Hill, 1654-1665.

Schwarzbauer, J.E., Tamkun, J.W., Lemischka, I.R., Hynes, R.O. (1983) Cell 35, 421-431.

Scott, E.M. (1960) J. Clin. Invest. 39, 1176-1179.

Scott, E.M., Duncan, I.W., Ekstrand, V. (1965) J. Biol. Chem. 240, 481-485.

Scott, E.M., Griffith, I.V. (1959) Biochim. Biophys. Acta 34, 584-586.

Scott, E.M., McGraw, J.C. (1962) J. Biol. Chem. 237, 249-252.

Sharp, P.A. (1981) Cell 23, 643-646.

Shimakata, T., Mihara, K., Sato, R. (1972) J. Biochem. 72, 1163-1174.

Slaughter, S.R., Hultquist, D.E. (1979) J. Cell. Biol. 83, 231-239.

Slaughter, S.R., Williams, C.H., Hultquist, D.E. (1982) Biochim. Biophys. Acta 705, 228-237.

Slightom, J.F., Blech, A.E., Smithies, O. (1980) Cell 21, 627-638.

Sottocasa, G.L., Kuylenstierna, B., Ernster, L., Bergstrand, A. (1967) J. Cell. Biol. 32, 415-438.

Southern, E. (1975) J. Mol. Biol. 98, 503-517.

Spatz, L., Strittmatter, P. (1973) J. Biol. Chem. 245, 792-799.

Stein, J.P., Munjall, R.P., Lagrace, L., Lai, E.C., O'Malley, B.W., Means, A.R. (1983) Proc. Natl. Acad. Sci. USA 80, 6485-6489.

Stoffel, W., Arch, K.L. (1964) Hoppe-Seyler's Z. Physiol. Chem. 337, 123-132.

Strittmatter, P. (1963) "The Enzymes" (ed. Boyer, P.D., Lardy, H.A., Myrback, K.) Academic Press, Inc., New York, Vol VIII, 113-145

Strittmatter, P., Machuga, E.T., Roth, G.J. (1982) J. Biol. Chem. 257, 11883-11886.

Strittmatter, P., Spatz, L., Corcoran, D., Rogers, M.J., Setlow, B., Redline, R. (1974) Proc. Natl. Acad. Sci. USA 71, 4565-4569.

Strittmatter, P., Velick, S.F. (1956) J. Biol. Chem. 221, 253-264.

Strittmatter, P., Velick, S.F. (1957) J. Biol. Chem. 228, 785-799.

Sugiyama, T., Miki, N., Yamano, Y. (1979) Biochem. Biophys. Res. Commun. 90, 715-720.

Tabor, S., Richardson, C.C. (1987) Proc. Natl. Acad. Sci USA 84, 4767-4771.

Takeshita, M., Tamura, M., Kugi, M., Matsuki, T., Yoneyama, Y., Igarashi, T. (1987) Biochem. Biophys. Res. Commun. 148, 384-391.

Takeshita, M., Tamura, M., Yoshida, S., Yubisui, T. (1985) J. Neurochem. 45, 1390-1395.

Tamkun, J.W., Schwarzbauer, J.E., Hynes, R.O. (1984) Proc. Natl. Acad. Sci. USA 81, 5140-5144.

Tanishima, K., Tanimoto, K., Tomoda, A., Mawatari, K., Matsukawa, S., Yoneyama, Y., Ohkuwa, H., Takazakura, E. (1985) Blood 66, 1288-1291.

Townes, P.L., Morrison, M. (1962) Blood 19, 60-74.

Tsurushita, N., Korn, L.J. (1987) Mol. Cell. Biol. 7, 2602-2605.

Ueda, S., Nakai, S., Nishida, Y., Hisajima, H., Honjo, T. (1982) EMBO J. 1, 1539-1544.

Vanin, E.F. (1985) Ann. Rev. Genet. 19, 253-272.

Vanin, E.F., Goldberg, G.I., Tucker, P.W., Smithies, O. (1980) Nature 286, 222-226.

Weissbach, A. (1977) Ann. Rev. Biochem. 45, 25-47.

Yoo, M., Steggles, A.W. (1989) Biochem. Biophys. Res. Commun. 163, 18-24.

Young, R.A., Hagenbuchle, O., Schibler, U. (1981) Cell 23, 451-458.

Yubisui, T., Miyata, T., Iwanaga, S., Tamura, M., Takeshita, M. (1986) J. Biochem. 99, 407-422.

Yubisui, T., Naitoh, Y., Zenno, S., Tamura, M., Takeshita, M., Sakaki, Y. (1987) Proc. Natl. Acad. Sci. USA 84, 3609-3613.

Yubisui, T., Takeshita, M. (1980) J. Biol. Chem. 255, 2454-2456.

Zhang, H., Somerville, C. (1988) Arch. Biochem. Biophys. 264, 343-347.

APPENDIX: SEQUENCE DATA FROM RAT CYTOCHROME $b_5$ GENOMIC CLONES

[I] Genomic clone $Rgb_5$4 - subclone 4.10 (870bp EcoRI fragment): 363bp of sequence

```
5' ATACACAGAGTCTTTACATCAGATGAAGAG(A)AGCATGTCTAGAGTCACGCTGTA
CATCTGAATGCAGCTTCTTT(G)ACTCTGATGGCTGACTGACTAATAAGCTAACCTCTC
N(T)TAGCTCATCTG(C)C(G)TAN(A)GA(N)(T)TAGTANTCATGCAACCTGTTC
                 ∝∝
                 43
(T)TTAAACAG CAT CCT GGT GGG GAA GAA GTC CTA AGA GAG CAA GCT
GGG GGT GAT GCT ACT GAG AAC TTT GAG GAC GTC GGG CAC TCT ACG
                                                  |-SacI-|
GAT GCA CGA GAA CTG TCC AAA ACA TAC ATC ATC GGG GAG CTC CAT
 ∝∝
 85
CCA GTAAGTCATTTCTTGGGGAAATTGCTTCTTCTCTTCCAAGGCTNTTTGACATTGG
GAAGTCTAAATTGATTTAGAATATTTCATTTGAGACAGGCTCTCAGT(AA)CTAGCC
(C)GGTG(G)    3'... EcoRI
                multiple cloning site
```

[II] Genomic clone $Rgb_5$4 - subclone 4.10 (870bp EcoRI fragment): 243bp of sequence

```
EcoRI                    ...5'  CTCCAGGTCTCATCACTGGTTGTACATCAT
multiple cloning site
AGACTCATGGCTTCCCTCTGTATTAACAATGAACTGGAGCATCTTGGAGGACTGAAATG
ACTCTTCGAAGGGAAAGATTCTTCCTGCCGACCTGTGCGCTCGCTCCCTCCCTCCCCTC
CCCAGATATTTCTGACTGTACGGACACAGTTACTTACA(T)TTCTCTAGAAAGAGATTC
GGGCCTAGCCTGGCTTCCACTATCAGTTGTTGTGTCTGCTGACA     3'
```

[III] Genomic clone $Rgb_5 1$ - subclone 1.3 (1.1kb SacI fragment): 386bp of sequence

```
   multiple  |-SacI-|
5' cloning     GAGCTC TCTGTGTAAACTCCGTCCAAGAAGATAGAAGGAGAGCA
   site
CTGTG(A)TTAGTTTTCT(C)ATGGTGACAGCAGGGAAGAGTAGCGAGCGGGAGCACAC
TAGGCGCCTGCATTGCTGCACATGAGAGCATTAGCAAAACTGAGCCAT(C)CTTTTAGC
CAAAGAGGTTCTAGAGGTAAGGAAAATATGCTGTGGTAGTTTTCCATCAGATGGGGTGC
                                                    aa  aa
                                                   107 108
GGTCTCTGCACTGAGGTTCACTCAAACAACTCTTTGCTCATGTGTCTT(T)CTAG T TGG
TGG ACC AAC TGG GTG ATC CCA CGC ATC TCA GCC CTG GTG GTA GCT
                                        aa
                                       133 Ter
CTG ATG TAT CGC CTC TAC ATG GCA GAA GAT TAACCTGTCTGTCCGAAGC
CAAGGAAGGAAAAGACTGCC   3'
```

[IV] Genomic clone $Rgb_5 1$ - subclone 1.2 (1.6kb PstI/BamHI fragment): 497bp of sequence

```
3'   GAAGTATCTGCTGAGTATAGGTGAGAGGTGCTCAAAATGCAGCAAGAAC
(T)CAGCTTNTTTCCTTGCTCTC(T)CAGTAGAGAGGCGCACACAGCGG(C)TG
ANTTCTCTGA(G)GTTGCAATGACTGCA(T)T(G)GACAGTNATGNNCNGTTTC
AGTGCGTAATATGCAGTGATCCTACAGTGTGAGCACAAATGCCCTGATTACTCC
AGCTCTCACCCAGTGCTGGG(A)AAAAGGCCTGTCTGTGCCTCCCCACGCACAG
NAATTGCACTTCTGTTTTCTGCTGTTGAAAACAAAGGGCANACTTAACATAGTG
ATAGTAATGTGTCACAAAGATTTAAAAAAAAAATGGTGGCCTAGAATTTGAATG
CTGTGTGTGACCATTGCTCAGAACGCGACTTTATCTTGTGGAGCACCGATTGTC
CCCTCAAAGTCTGTCAGAGAAAAGGTTTGGGTCCTGTCCTCTTAGGTATTAGTA
                          |-PstI-| multiple
TGTAATGGTTTTGTTGTAAGTCAC CTGCAG  cloning      5'
                                 site
```

[V] Genomic clone $Rgb_5$1 - subclone 1.2 (1.6kb PstI/BamHI fragment): 379bp of sequence

```
   multiple  |-BamHI-|
5' cloning    TGATCC  ACTCTAAACTTTGGGACTGTAGATGGAAATGCTTCTTT
   site
TTTTTTTTTTTTTTTTTTAAGATTTATTTTATTCCATATGAGTACACTGTGGCTGTCTTCA
GACAACACCAGAAGAGGGGATCAGATCTCATTACAGATGGTTGTGAGCCACCATGCGGTT
GCTGGGGATTGAACTGGAAGAGCAGTCAGTGCTCTTAACCGCTGAGCCATCTCTCCAGCC
CGGAAATGCTTCTTTGTTCCTAACTCTACTGTTAATTTTCAAACCACATGGATCAAATTT
NNTTT(T)AACTTTGCTGAATGC(C)TATGCTAATCATTGAATAGTAATTTCAATTCTCA
GTACACGCAGCTTTTGGAAATGAGAC(C)TTCACACA(T)GTCT     3'
```

[VI] Genomic clone $Rgb_5$1 - subclone 1.3 (800bp SmaI/PstI fragment): 164bp of sequence

```
                  ∝∝                                          ∝∝
                  96                                          107
5' TGGTGTTCCAG GAA ACC CTT ATC ACT ACT GTC GAG TCT AAT TCC AG
GTNAGGCAGTCGCTGTCTTTTTTGTTTCGCTTCTT(N)GGTTGCTGTGTTTTGAGTCNA
AGGCTCTTGGAGGGTTCAGGAGCC(N)AGAGCTGCATGGCGCCTGCTG(N)TGTTGA(C)
|-SmaI-| multiple
 CCCGGG  cloning    3'
         site
```

[VII] Genomic clone $Rgb_5$1 - subclone 1.3 (800bp SmaI/PstI fragment): 153bp of sequence

```
5' multiple |-PstI-|
   cloning   GACGTC AAACC(N)A(A)CCAAACCCACATACGTAAAAAGTACCAG
   site
AAGTACTCTTCCCGACACGATC(T)GTACGACCTTCCATCGTGTCTCAACCAATCCGTCT
CGCCAAAGACTCAATTAGCCCAGAAGGACCAGGATGCCACTCAATTCCTCGA(N)   3'
```

[VIII] Genomic clone $Rgb_5$1 - subclone 1.2 (870bp AluI fragment from 1800bp PstI fragment): 450bp of sequence

5′ AATTGCATTACTTGCTGAATGCCTTAGCTAATCATGATAGTAATTCAATTCTCAGTA
GACGTCAGCTTTTTGGAAATGAGACTTTCACACTAGTCCTTTGACTCTTAATGCCAATTC
86 95
CCTTTTTTATCCCTTACAG GAT GAC AGA TCA AAG ATA GCC AAG CCT TCG
GTAAGTTTGTCCAAGTATCACAAAGGAGTGGTGGATAGAGGTAACTATGTATCTACATTG
TATGTANACATGTCAAAATGGGGCAAAGGCACAGCCTACACAGATCTTTCTAGAGAAGGG
GAATATGTAACGTATTGGTTTNAAAANTTAGTTTGGA(C)GTTTTTTGTTTGTTCAGGTT
TTTTTTACTCGTCCACGTGCTTTTGANNNAGTTTGTATGTTACTGTGTAGAACAGAAGTA
|-AluI-| multiple
TTACTCTTTTTAAACATGATAATTATCAATAAATAAATCCAGATAG cloning
site 3′

[IX] Genomic clone $Rgb_5$1 - subclone 1.2 (870bp AluI fragment from 1800bp PstI fragment): 430bp of sequence

multiple |-AluI-|
5′ cloning CTGGAGTAATCNGGGCATTTGTGCTCACACTGTAGGATCACT
site
GCATATTACNCNCTGAAAACAGCNGTGGCTNNCAATNGACTGTCCAGCAGTCATGATCAC
TGCATATTACNCNCTGAAAACAGCNGTGGCTNNCAATNGACTGTCCAGCAGTCATTGCAA
CCTCAGAGAAGCTGGCTGTTCTCATAAATTCTTACTCTAGATTTAATGAGCAAATCTGGG
CGAGAAGGTCAGCTATCTGGATTTTATTTATTGATAATTATGATGTTTAAAAAGAGTCAA
TACTTCTGTTACAAAGTAACATACAAACTCTCCTCAAAAGCACGTGGACGAGTAAAAAAA
AACCTGAACAAAGAAAAAACTCCCAAACTAANCTTTTGCAAACCAATACGTACATAT(T)
CCCTTATCTAGAAGATCTGTGT(A)GGCTGTGACCTT(T)GCCCATTTGACAGTGCTACA
TACANNGATCTGTACTCTATCACACTCTTGGTCTG 3′

[X] Genomic clone $Rgb_5 2$ - 800bp SacI fragment: 380bp of sequence

multiple |-SacI-|
cloning GAGCTC TTAATGTTTTTGCAAAGAACCCAAGTTCAGTTCCCAGCACT
site
GATATTAGGTGGCTCAGAAAAGCCCAGTAACTGGTTCGTGGGTATCTGCTGACCTTCTTCT
GGACCCCGAGAGCAATTGCACTCTTACATATGCAAGCTAC(AC)CGATTGAAAATTATTAA
AAAGAACATATATTTTGAAGTAGAAGAAAA(A)TTCTTTTAGCTTACATATTTTATTCATC
CATGCCTGTGCAGAGAGTGATGATACATTTTGTTGCAGGTGTTGGGGATGATAATGCGAAC
AAAGCAGGAAATT(G)CTCTTCCACATGCAGTGCAACTAGAGAAAGGATCCATAACATAGA
GTAAATNTTCTAGA

[XI] Genomic clone $Rgb_5 2$ - 800bp SacI fragment: 468bp of sequence

multiple |-SacI-|
cloning GAGCTC TTGCAGTTCCATGCTACTTGGCTTTTGACAGTGGT(N)GAAGGA
site
A(T)GCCTGAGAAGAGTT(C)CAGAACCATAATGGTCCGT(T)CCTCATTTTCTTCTGTGC
CCTCTTTGCTACCTCAGCTTTGTCCCCTCTGGCTTGTATGCTGTTCCTTGAGCAATAAA
(A)CTGGATTCTCCCTCACCTGCACTGAAATTCCTCTCGTGAGAAGACTGCTCATCAGCTG
TGCCTGTGGTCGATCGAG(G)C(C)TCGATGCTCTGCGTATGTCATTCTGATTG(C)AGGC
CACATACTCCTGCTGAGAGTGCATGTATGCGCTGCGTATATAGCAGGTGTATCTCGAGTNG
CTGTCTCTATGCGCGCTGTGTCTCTAATAACTACC(C)TATGTT(T)GAA(A)GA(A)T
(C)TGC(G)ACAACCTATGTCTCTGACTGACC(C)TTACTTACAGTCATG(G)CCCCTACC
CCCCTCGGG(G)ACGACCAGCACTCTCTCAGG(T)AAA(A)G