The genetic manipulation of a cDNA encoding the seed storage protein legumin a to alter its amino acid composition

Belfield, Graham Paul

THE GENETIC MANIPULATION OF A cDNA ENCODING THE SEED STORAGE PROTEIN LEGUMIN A TO ALTER ITS AMINO ACID COMPOSITION

GRAHAM PAUL BELFIELD

DEPARTMENT OF BOTANY
UNIVERSITY OF DURHAM
SEPTEMBER 1986

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF
REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE
DEPARTMENT OF BOTANY
UNIVERSITY OF DURHAM

ABSTRACT

The seeds of lugumes are eaten throughout the world, constituting a major input of dietary protein. There can be problems of malnutrition in the more deprived areas where they provide the main source of protein, for though legume seeds can be very rich in protein, they are deficient in some amino acids, mainly methionine, that are essential for a healthy diet. Recombinant DNA technology may offer a solution to this problem by introducing DNA encoding these deficient amino acids into seed storage protein genes, which on reintroduction into the host plant could be grown as a more nutritional crop. The 2s storage proteins of the Brazil nut contain a very high proportion of methionine, therefore DNA encoding 2s protein could be introduced into the sequence of legume storage protein genes. It was proposed that this be attempted, and any constructions that should be produced could be cloned into yeast to detect expression of the mutated genes.

Attempts were made to construct and isolate pUC18 vector clones of Brazil nut DNA, to determine and attempt a rationale for the insertion of DNA from these clones into sequences encoding legumin A by site directed mutagenesis, and to detect the formation of Brazil nut - legumin DNA constructs with radiolabelled probes DNA probes and agarose gel electrophoresis. Initial steps were taken to perform a similar mutation of a vicilin cDNA.

Two pUC18 clones of Brazil nut DNA; - pBnA and pBnB were created and isolated. Clones of legumin - Brazil nut DNA;- pGPB1 were constructed based on an insertional mutation of the legumin cDNA construct pJY8 with Brazil nut DNA from the clone pBnA. They were isolated from the in - situ hybridisation of transformed cells with radiolabelled DNA. Clones of pGPB1 were found by gel electrophoresis and Southern hybridisation to contain the BnA insertion in the correct orientation. Time constraints prevented the cloning of pGPB1 into yeast.

The choice of legumin mutation sites was discussed and the rationale adopted justified on the grounds of restriction site analysis and the restraints imposed by legumin solubility and protein structure. The Results and problems encountered were discussed in some detail. It was suggested that further work should attempt the expression of pGPB1 in yeast.

CONTENTS

CONTENTS (continued)

FIGURES

TABLE

ACKNOWLEDGEMENTS

I would firstly like to thank my Durham based supervisor, Dr. John Gatehouse for introducing me to the techniques of molecular biology and for his helpful advice and constructive critisism. I wish also to thank my I.C.I. based supervisor Dr. Finn Wynne, along with Andy Ryan and JAG for reading the early drafts. I am grateful to Dr. Ron Croy and co-workers for providing me with a large scale prepartion of my subclones and together with JAG for permission to reproduce the imformation in figure 19. Thanks also to Laurence Gatehouse for permission to reproduce his pUC8 data, Mrs Denise Hughes for her drawing of the pUC8 map, Mr Paul Sidney for the photography and finally for the help and advice from the members of labs 117 and D2.

Dedicated to my parents, for constant support and encouragement.

ABBREVIATIONS

The abbreviations used are as recommended by the Biochemical Society (1983). Biochem. J., 209, 1-27. Additions to this list are given below. The one letter notation for amino acids (used in appendicies III and IV) is given in: Biochem. J. (1969), 113, 1-4.

bp	= base pairs
BSA	= bovine serum albumin
kb	= kilobase pairs
LMP agarose	= low melting point agarose
m.p.c.s.	= multi purpose cloning site
SDS	= sodium dodecyl sulphate
SSC	= saline sodium citrate
dCTP	= deoxycytidine 5'-triphosphate
mRNA	= messenger RNA
tRNA	= transfer RNA

CHAPTER ONE

INTRODUCTION

1.1 THE NEED FOR A NUTRITIONALLY IMPROVED LEGUME

The seeds of legumes are commonly eaten throughout the world. They provide us with a high protein source for our diets, without the associated risks to health and energetic inefficiencies (Mantell *et al*. 1985) which arise from consuming animal meat. Furthermore the legumes are a relatively cheap crop to produce, for they need no costly nitrogen fertilisers for healthy growth, due to their symbioses with nitrogen-fixing bacteria of *Rhizobium* species (Sprent 1979). Many legume crops have adapted to the hot, arid environments which are home for much of the world's population, where primarily as a result of climate and poverty, they may be one of only two or three staple foods, constituting the main protein source for the diet. Although most legume seeds contain a very high level of consumable protein (up to 50% of total seed dry weight), their storage proteins do not contain a complete complement of essential amino-acids, and will be primarily deficient in methionine, with secondary amino acid deficiencies in threonine, tryptophan or valine,depending on the host species (Burr 1975). Attempts to breed cultivars with nutritional quality by conventional techniques have yielded little success, yet it is feasible that the deficiency in seed amino acid composition may be rectified using recombinant DNA technology to manipulate the coded product of the storage protein genes.

Genetically attacking the problem of storage protein nutritional imbalance is widely considered to be one of the major immediate applications of genetic engineering (Payne 1983; Shaw 1984; Croy and Gatehouse 1985; Mantell et al. 1985).

Before the potential methods of improvement are considered, it is important to to introduce the form and function of plant storage proteins.

1.2 AN INTRODUCTION TO PLANT STORAGE PROTEINS

The major seed proteins of dicotyledonous plants belong to the globulin solubility fraction (soluble in dilute salt solution but insoluble in water), although significant amounts of water soluble albumin fraction proteins are also present. The globulin fraction has been extensively characterised, particularly in the nutritionally important legumes and oil seeds. Within the storage proteins of the legume family Faboideae exists one of the more complex protein systems, mainly by virtue of extensive posttranslational modification (Croy and Gatehouse 1985), as exemplified by pea (Pisum savitum L.) and broad bean (Vicia faba L.). The properties, synthesis and genetics of pea storage proteins have been well reviewed (Gatehouse et al. 1984; Gatehouse et al. 1985; Ersland et al. 1983; Casey and Domoney 1984; Brown et al. 1981), which make pea a good model legume to work with in nutritional modification studies. The storage proteins of pea consist mainly of two immunologically distinct protein classes of globulin; legumin and vicilin, with a third distinct class convicilin immunologically related to vicilin. Together,

these constitute around two thirds of the total seed protein, and all exist as deposits inside membrane bound organelles termed protein bodies, in the cotyledon. These globulins are generally of high molecular weight and extremely rich in the amino acids arginine, glutamine and asparagine, which may act as nitrogen sources for the developing seedlings (Higgins, 1984). Other dicotyledonous plants may have seeds containing significant amounts of The 2s albumin proteins. These form a class of low molecular weight proteins, principally rich in cysteine and other sulphur containing amino acids including methionine (Ampe *et al*. 1986), the nutritionally important amino acid deficient in pea.

1.3 THE MODIFICATION OF PEA STORAGE PROTEINS

1.3.1 Crop breeding

One way in which nutritional improvements have been made is by crop breeding for an imbalance between the storage proteins legumin[11s] and vicilin[7s] (Derbyshire *et al*. 1976; Payne and Rhodes 1982). The 11s-type proteins contain a higher proportion of methionine than the 7s, therefore increasing ratio of 11s to 7s seed proteins would improve the nutritional balance of the whole seed. increasing the albumin : globulin ratio of seeds has also been suggested and attempted (Croy *et al*. 1984; Croy 1977), however, the improvement would be so small as to be of little significance. Furthermore, crop breeding programs are expensive and may take a number of years before suitable cultivars are developed.

1.3.2 Recombinant DNA technology

Recombinant DNA technology could potentially solve the problem using a number of approaches, all with the singular aim to specifically mutate the storage protein coding sequence with foreign DNA coding for the missing amino acids in substantial amounts. In doing this one must ensure that the mutation does not significantly disrupt either the structural or functional properties of the coded protein, and that the inserted DNA together with any downstream sequence of the gene will on transcription be read in frame. The first consideration presents something of a dilemma: in order to ensure a significant incorporation of DNA coding for the deficient amino-acids, one may need to mutate with DNA of some length, though to avoid disruption of the storage protein on translation it may be only possible to use short DNA sequences. A further problem would be ensuring that the effects of the gene introduced would not be "swamped" by the genes already present, particularly in the cereals which are bred to contain tetra or even octoploid genomes. Unless the returning gene is to be site directed into the genome, its effect would be minimal. Whilst the modification may not be deleterious to the coded protein's form and function, it would be very unlikely that it give it a competitive advantage over the genes already present so one would have to ensure it is finally introduced into a cloning vector of suitably high copy number. An alternative approach is discussed in section 0.4 below.

1.3.3 Oligonucleotide site-specific mutagenesis

One approach that would go a long way to solving this problem would be to synthesise an oligonucleotide that contained the deficient amino-acid coding triplets repetitively. (Sproat and Gait 1984). This could then be incorporated into the main coding sequence using the technique of M13 site-specific mutagenesis (Gillam *et al*. 1980; Gait 1984). The advantages are that one need only mutate with short length DNA sequences that can be constructed to be read in frame and do not require restriction sites for their incorporation. However long stretches of repeating triplets can present a problem in protein expression: for example, consecutive guanines coding for glycine can often slow transcription and therefore expression of the protein, producing what is known as a 'G-C clamp'. Furthermore the process of producing oligonucleotides is generally costly and requires specific equipment that is either very expensive, or cheaper manual apparatus that is problematic and labour intensive.

1.3.4 Site-directed insertion of foreign DNA

A second approach is to incorporate a region of DNA from a protein coding sequence rich in the amino-acids desired from another gene. The foreign DNA can be inserted using the site-directed restriction enzyme method (Mantell *et al*. 1985). This presents the problem of finding suitable restriction sites in regions causing the minimum of disruption while ensuring correct frame reading. Size may not present a problem if the deficient triplets are present in high

concentration in the foreign DNA. This would be a cheaper and more flexible alternative to oligonucleotide synthesis.

1.4 BRAZIL NUT : A SUITABLE LOW-MOLECULAR WEIGHT PROTEIN

It has already been noted that the 2s storage proteins of some plants are generally rich in amino acids not common in the 7s and 11s globulins, so mutation with a region of DNA from a 2s protein coding sequence containing many methionine codons could, subject to the restraints discussed earlier be a practicable solution to the problem of altering 7s or 11s globulin composition. These 2s proteins bear no sequence homology (Ampe *et al*. 1986) to the endogenous 2s proteins of legumes (Gatehouse *et al*. 1985), which may cause problems in control of expression. However, such hybrid genes would have the advantages of improved amino acid composition and by maintaining storage protein controls, potentially good expression efficiencies could be achieved (Gatehouse *et al*. 1985; Croy and Gatehouse 1985). The 2s proteins of the Brazil Nut, (Bn-2s), *Bertholletia excelsa*, are particularly interesting as they contain an abnormally high amount of methionine (Youle and Huang, 1981). This, together with their high cysteine content makes them unique amongst the 2s proteins that have been studied, and potentially of high nutritive value (Ampe *et al*. 1986).

1.5 REINTRODUCTION INTO THE HOST

Inevitably, the mutated genes coding for the improved storage proteins will need to be re-introduced into the host plant. The transformation of plants has been well reviewed (Watson *et al*. 1983; Shaw 1984; Mantell 1985; Old and Primrose 1985) and need only be briefly reviewed here.

1.5.1 *Agrobacterium tumefasciens*

The most practicable approach is to use the parasitic bacterium *Agrobacterium tumefasciens*. This normally forms 'crown gall' tumours on suitable hosts, coded for by the Ti (tumour inducing) plasmid. By removing the region coding for tumour formation and replacing it with the mutated gene, then infecting undifferentiated callus tissue from the host with *Agrobacterium* containing the modified Ti plasmid, one can transform the plant tissue with the altered storage protein coding sequence. The callus could then be induced to regenerate whole plants.

1.5.2 Viral vectors

It has been shown that viral vectors can be used to transform whole plants with a particular gene. This avoids the necessity for plant regeneration, therefore being potentially a useful alternative to using *Agrobacterium*. However, the viruses so far used (Cauliflower Mosaic Virus, Gemini viruses) suffer from two major setbacks;- firstly they have a very small host range, which facilitates the work with only a few plants. Secondly, they have a severe size restraint;the only

genes to be cloned successfully into these gene transfer have been very small (in the order of hundreds of bases), which is useless for nearly all plant transformations.

1.5.3 Direct transformation

It seems now possible to remove cell-walls from plant tissue to release viable naked protoplasts, and under certain conditions, using any of a number of techniques, these can be induced to take up linear DNA without using a vector. One approach is to treat the protoplasts using calcium chloride to disrupt the outer membranes, creating 'leaky' cells which can then take up the DNA. Alternatively one could attempt to micro-inject the DNA into the nucleus of the protoplasts, or place them in an electromagnetic field (electroporation) for a short period which again disrupts the protoplasmic membranes. All these techniques give higher transformation efficiencies than one can expect with Agrobacterium, however they suffer from the problem of recalcitrance. Very few whole plants can as yet be regenerated from single protoplasts, though it is generally thought that this is a technical problem that may be soon overcome.

Presently, no-one has demonstrated a working regeneration scheme for pea in tissue culture (though some workers have achieved transformation) which unfortunately negates prospects of nutritionally improved pea plants until a suitable system is established.

1.6 EXPRESSION IN A SUITABLE HOST

When contemplating manipulation of these storage protein genes, it is important to note that gene transfers leading to efficient, controlled expression would seem more likely to succeed using the same plant species to donate genes for manipulation as to host the altered genes. The modified genes would continue to recognise the endogenous control systems and should be efficiently expressed (Croy and Gatehouse 1985). However, expression may be achieved using a different organism, notably the simple eukaryote, yeast (Struhl 1975). Obtaining expression in yeast would be of greater value than that in the classic prokaryotic cloning organism Escherichia coli. Unlike the latter, the yeast Saccharomyces cerevisiae can perform the post-translational modifications phosphorylation and glycosylation, which may be necessary for true protein function. Also, protein folding and assembly appears to take place in a most "eukaryotic" way in yeast, whereas many eukaryotic proteins form insoluble deposits in E. coli. Furthermore, due to its small genome and short generation times, yeast may be experimentally manipulated as easily as most prokaryotes, giving it a considerable advantage over higher eukaryotes. Yeast therefore is highly suited as an intermediary organism, betwixt E. coli; the cloning organism in which the recombinant gene is constructed, and pea; the mother-host, from which the gene was removed and to where one would eventually hope to return it, once modified. However, it is far from perfect. It does not perform all the posttranslational modifications observed in higher plants, so one might anticipate problems in expression of legume genes,

in light of the complex proteins they encode (Croy and Gatehouse 1985), though some success has been achieved with unmodified vicilin (Watson, personal communication). Yeasts also differ from higher eukaryotes in their post-transcriptional removal of introns; they require a specific sequence to be present within the intron, without which no excision occurs. Plant introns have been shown not to have this recognition sequence, and would require either modification or removal prior to insertion in yeast for expression. These restrictions limit the usefulness of yeast as an primary expression organism, for the mass production of recombinantly produced proteins, as has been attempted with E. coli, that would replace the plant. It is more suited to our needs as a means to test the validity of constructs quickly and cheaply before commitment to a long and costly plant transformation programme. If the mutated protein can be produced in yeast, its physical properties (solubility, etc.) can be checked to make sure that the mutation has not led to a non-viable protein.

1.7 DEFINING THE PROBLEM

The problem set at the start of this dissertation was as follows: To mutate a cDNA coding for a fragment of Legumin A (and Vicilin, time permitting) with the cDNA coding sequence of the 2s protein from Brazil nut. This could then be inserted into a suitable yeast expression vector, to produce the fusion protein. If expression of the modified legumin sequence was detected, this would be proof that such modified proteins could be produced in eukaryotes; the rationale was

valid and more extensive research could be carried out in this area.

1.8 AIMS

1. To subclone two fragments of a cDNA clone encoding the Brazil nut 2s protein into the Pst I endonuclease restriction site in the cloning vector pUC18 from the Pst I endonuclease restriction site in the phage M13.

2. To transform a suitable *E. coli* strain with the pUC18 subclones.

3. To detect the subclones using agarose gel electrophoresis.

4. To determine a region in the Legumin A cDNA JY8 in the cloning vector pUC8 that would be least effected on mutation with Brazil nut DNA in terms of post-expression structure and function.

5. To find suitable restriction sites within this region and also within the Brazil nut DNA fragments for insertion into the legumin cDNA such that the insert and the tail end of the Legumin read in frame.

6. To insert the sequence encoding Brazil nut 2s protein into the Legumin cDNA as above.

7. To transform a suitable *E. coli* strain with the pUC8 constructs.

8. To produce radio-isotope labelled probes for the DNA sequences encoding brazil nut 2s protein and legumin A.

9. To detect the construct using a double colony hybridisation strategy, using the radioactive labelled probes.

10. To confirm the presence, nature and orientation of the mutation with agarose gel electrophoresis, southern blotting and dideoxynucleotide sequencing.

11. To remove the gene, adapting it's terminal ends if required for cloning into a yeast expression vector.

12. Transformation of a suitable strain of Saccharomyces cerevisiae, with the yeast plasmid construct.

13. Detection of expression of the mutated legumin gene, using western blotting and/or labelled antibody techniques.

14. Mutating a clone encoding vicilin from PAD 2.1 using the strategy described (time allowing).

CHAPTER TWO

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Chemicals and biological reagents

All reagents, with the exceptions listed below were produced by BDH Chemicals Ltd., Poole, Dorset, U.K. and were of analytical grade or the best available.

Adenosine 5'-triphosphate (ATP), ampicillin, bovine serum albumin (BSA), dithiothreitol (DTT), ethidium bromide (EtBr), herring sperm DNA, lysozyme, RNase A and tetracycline were from Sigma Chemical Co., Poole, Dorset, U.K.

Sephadex G-50 and Ficoll 400 were from Pharmacia Fine Chemicals, Uppsala, Sweden.

Caesium chloride (CsCl) and sodium chloride (NaCl) were from Koch-Light Ltd., Haverhill, Suffolk, U.K.

Nitrocellulose filters (BA85, 0.45um) were from Schleicher and Schull, Anderman and Co. Ltd., Kingston-upon-Thames, Surrey, U.K.

3MM paper and filter discs from Whatman Ltd., Maidstone, Kent, U.K.

Bacto-Agar was from Difco Laboratories, Detroit, Michigan, U.S.A.

BBL trypticase peptone was from Becton Dickinson and Co., Cockeysville, M.D., U.S.A.

Yeast extract from Bio-Life, Milan, Italy.

Restriction endonucleases were from Bethesda Research Laboratories, (U.K. Ltd., (BRL), Cambridge, U.K., The Boehringer Corporation (London) Ltd., Lewes, East Sussex, U.K., New England Biolabs., CP. Laboratories Ltd., Bishops Stortford, Herts, U.K., and Northumbria Biologicals Ltd.,

Northumbria, England.

T4 DNA ligase and 5-dibromo-4-chloro-3-indoylgalactoside (X gal) were from The Boehringer Corporation (London) Ltd.

Radiochemicals and nick-translation kit (N.5000) were from Amersham International p.l.c., Amersham, Bucks, U.K.

Agarose from Bethesda Research Laboratories (U.K.) Ltd.

Plasmids and NM258 lambda phage DNA were supplied by Dr R.R.D. Croy from communal stocks.

All heat stable solutions were sterilised by autoclaving (15 minutes at 120°C), all others by filtration with the exception of electrophoresis buffers and deionised formamide.

2.1.2 Bacterial strain, plasmid and bacteriophage vectors

The bacterial strain used was a derivation of *E. coli*. The table below lists this strain and the plasmids and bacteriophage used as vectors and/or probes. The sources or references for each are given.

TABLE 1 E. Coli strain, plasmids and bacteriophage

Strain	Genetic Characters	Reference or Source
JM 83	ara, /\(lac-proAB), rpsL, (=strA), 080, lacz /\ M15	Bethesda Research Laboratories (BRL)
Plasmids		
pDUB9	vicilin (in pBR322)	Delauney (1984)
JY8	legumin A (in pUC8)	Yarwood (1985)
pUC18	ApR, lacz	Viera and Messing (1982)
Bacteriophage		
Bn2sJ13	BnA* (in M13mp18)	
Bn2sJ16	BnB* (in M13mp18)	

2.1.3 Glassware and plasticware

All glassware and plasticware used in manipulations of DNA, bacterial cultures and for storage of sterile stock solutions and media were autoclaved prior to use. When ever very good recovery of DNA was required, glassware and plasticware was siliconised using "repelcote" (Hopkin and Williams, Romford, U.K.).

2.1.4 Growth media

All transforming cells were initially grown in liquid Yeast Tryptone (YT) medium:

8 g Tryptic peptone (digest of casein)

5 g Yeast extract

5 g NaCl per litre distilled water.

overnight inoculations were grown in either YT or 2 x YT medium:

16 g Tryptic peptone

10 g Yeast extract

10 g NaCl per litre distilled water.

Appropriate antibiotics were added; For pUC transformed cells, Ampicillin was added to a final concentration of 35 - 50 µg/ml, from a stock solution of 5 mg/ml. stored at -20°C. For pBR322 transformed cells, Tetracycline was also added to a final concentration of 12.5 µg/ml from a stock solution of 12.5 mg/ml. stored in the dark at -20°C.

Transformed cells were plated out on YT agar, with appropriate antibiotics at concentrations given above. Yt agar is YT medium with Bactoagar added at 15 g per litre (1.5%). The insertional inactivation of the lacZ gene in M13 and pUC 8 and 18 on subcloning fragments in or out of these vectors was often used as a screening method for possible recombinants. In these instances, X gal would be added to the agar to a final concentration of 40 µg/ml from a stock solution of 2 mg/ml in dimethyl formamide (dmf). Molten agar was cooled to 55 C before addition of antibiotics / X gal, then poured slowly, flaming the surface if necessary to remove bubbles, allowed to set, inverted, left to dry out at 37°C for 2 days or 55°C for 30 minutes then stored at 4°C prior to use.

2.2 METHODS

2.2.1 Biochemical Techniques

2.2.1.1 Removal of protein by phenol extraction

The method adopted was that recommended by Brawerman et al. (1972). DNA samples were taken up in TE buffer or sterile water, to a final volume of between 200 µl and 2.5 ml, and into Eppendorf tubes or mse tubes depending on the nature and scale of the experiment. An equal volume of phenol was added and the sample was vortexed briefly. An equal volume of Chloroform was added, the sample again briefly vortex mixed, and centrifuged for 3 minutes at room temperature in an Eppendorf centrifuge or for 10 minutes at 1600g in an mse centrifuge. The upper aqueous phase was then transferred to a fresh tube using an automatic pipettor fitted with a disposable tip, discarding the interface and lower organic phase. The procedure was repeated with an equal volume of chloroform. The DNA was then recovered by precipitation with ethanol (q.v. 2.2.1.2).

T.E. Buffer is 10 mM Tris.Cl (pH 8.0), 1 mM EDTA. "Phenol" means phenol equilibrated with TE buffer and containing 0.1% hydroxyquinoline and 0.2% B-mercaptoethanol. "Chloroform" means a 24:1 (v/v) mixture of chloroform and isoamyl alcohol.

2.2.1.2 Precipitation of nucleic acids with ethanol

the volume of DNA solution was estimated and 3 M sodium acetate (pH 4.8) was added to a final concentration of 0.3 M, then mixed well by vortex. Two to three volumes of ice-cold ethanol was added, then mixed well by vortex. The samples

were then stored at -20°C for at least 30 minutes. Samples were then centrifuged in an Eppendorf centrifuge or mse centrifuge for at least 20 minutes at 0°C if possible. The supernatant was discarded and was replaced with an equivalent volume of ice-cold 70% ethanol. The nucleic acid precipitate was disrupted by vortexing, and then the samples were re-centrifuged for 5 minutes. The ethanol, was discarded, the last drops being removed by a 20 µl automatic pipettor. The samples were then dried by desiccation for 5 to 10 minutes. The DNA precipitate was dissolved in the desired volume of TE buffer or sterile water.

2.2.1.3 Preparation of Ribonuclease A (RNase A)

Pancreatic RNase A was dissolved at a concentration of 10 mg/ml in sterile water and heated to 100°C for 15 minutes. It was allowed to cool slowly to room temperature, then was stored at -20°C.

2.2.1.4 Storage of DNA

DNA samples were stored in TE buffer, sterile water or for shorter periods in 70% ethanol at -20°C. When brought out storage for use, they were often warmed to 37°C to speed thawing. Whilst in use, stock samples were kept on ice.

2.2.2 Enzymatic methods used in manipulation of DNA

2.2.2.1 Digestion of DNA with restriction endonucleases

To DNA samples in solution in a 0.5 ml Eppendorf tube were added a 5 x excess in units of the desired restriction enzyme, 5 x reaction buffer and sufficient sterile water to bring the buffer concentration to 1 x (generally to a total volume of 20

µl). If the DNA samples were obtained by small scale DNA preparation, 1 or 2 µl of RNAase would have been added (10 mg/ml) to remove tRNA if the restricted DNA was to be visulised on a gel. The reaction was mixed by shaking, then briefly spun down in an Eppendorf centrifuge. Restrictions were left for 2 to 3.5 hours at 37°C, then the reaction was terminated by heating to 70°C for 5 minutes, by the addition of 0.5 M EDTA (pH 7.5) to a final concentration of 10 mM, or by the addition of a stop dye. Double digestions were usually carried out together at the same time. If lamda DNA was digested for use as a gel size marker, it would be heated to 65 °C for 10 minutes prior to restriction to deanneal it's cohesive ends.

A range of restriction buffers were used that differ principally in salt concentration, depending on the requirements of the restriction endonuclease/s used. Low salt buffer contains no NaCl, 10 mM Tris.Cl (pH 7.5), 10 mM MgCl2 and 1mM Dithiothreitol. Medium salt buffer is identical, save for the inclusion of 50 mM NaCl. High salt buffer is 100 mM NaCl, 50 mM Tris.Cl (pH 7.5), 10 mM MgCl2 and 1mM Dithiothreitol. The enzyme Sma I uses a specialised buffer consisting of 20 mM KCL, 10 mM Tris.Cl (pH 8.0), 10 mM MgCl2 and 1 mM Dithiothrietol. Stop dye was made from of 20% (v/v) glycerol, 10 mM Tris.Cl (pH 8.0), 10 mM EDTA (pH 8), 0.2% (w/v) agarose and 0.1% (w/v) bromophenol blue, xylene cyanol and orange g. This was Autoclaved and forced through a 19 gauge hypodermic needle to form beads.

2.2.2.2 Ligation of DNA

When attempting to ligate two fragments of DNA, they were

first restricted with the appropriate restriction endonucleases. The samples were then phenol extracted and precipitated with ethanol. The samples were taken up with between 20 µl and 100 µl of sterile water, and mixed. A 2 to 5 fold excess of T4 DNA ligase was added along with the appropriate volume of 2 x KLP buffer in 0.5 ml Eppendorf tubes. The volume was adjusted if required with sterile water. The mixture was incubated whenever possible at 15 °C for at least 12 hours, otherwise at 15 °C for 3 hours then room temperature for 1 hour and 37° C for 15 minutes. The samples were then used to transform cells (2.2.3) or were stored at -20 °C.

2 x KLP (Kinase-ligase-polymerase) buffer is made from 100 mM Tris.Cl (pH 7.5), 20 mM MgCl2, 20 mM dithiothrietol and 2 mM ATP.

2.2.2.3 Deleting a plasmid restriction site with S1 nuclease and its subsequent detection by restriction analysis

The method used was adapted from that used by Zeitlin and Efstratiadis and published by Maniatis et al. (1982). 2 µg of plasmid DNA was digested with the restriction endonuclease whose site was to be deleted, to produce "sticky ends", corresponding to the recognition site of that enzyme. The DNA was then phenol extracted and precipitated with ethanol. The Sample was taken up in 400 µl of sterile water, then divided into 4, 0.5 ml Eppendorf tubes, so each contained approx. 500 ng DNA in 100 µl water. To each tube was added 100 µl ice-cold 2 x S1 nuclease buffer containing increasing quantities of S1 nuclease (20, 40, 60 and 80 units for tubes 1 to 4 respectively). The samples were incubated at 37 °C for 30

minutes. The reaction was stopped by the addition of 0.5 M EDTA to a final concentration of 10 mM and 2M tris was added to a final concentration of 50 mM. The samples were again phenol extracted and ethanol precipitated. The "blunted" DNA ends were then "polished"(sic) by the addition of the appropriate number of units of T4 DNA polymerase, with 1mM dNTP's in 2 x KLP buffer. The volume was adjusted to 20 µl with sterile water, and the samples were incubated for 10 minutes at 37°C. The reaction was terminated by heating to 70 °C for 5 minutes. The samples were cooled, and 1 unit of T4 DNA ligase together with ATP to a final concentration of 1mM was added to each. The samples were left to ligate for 2 days at 15 °C, replenishing T4 ligase and ATP after 1 day. The religated plasmid DNA was then used to transform competent cells (2.2.3) and grown on agar plates containing the appropriate antibiotics and X gal. Colonies containing the plasmids were isolated by virtue of a inability to utilise X gal (the site deleted by this method was in the multi-purpose cloning site of the plasmid, inactivating the galactosidase gene, producing a white colony instead of blue). A number of these colonies were used to inoculate overnight cultures from which the plasmid DNA was then isolated (2.2.5), and subsequently restricted with the endonucleases for the deleted site, and other adjacent sites that were considered important for future work. The restrictions were visulised by agarose gel electrophoresis, and those samples that would not digest at the attempted deleted site, but did at adjacent sites were used for further work.

10x S1 nuclease buffer consists of 2M NaCl, 0.5M Na acetate,

10 mM ZnSO4 and 5% glycerol.

2.2.3 Preparation and Transformation of competent cells

The method used was devised by Mandel and Higa (1970). A stab from a glycerol of E coli strain JM83 was used to inoculate 10 ml of YT medium which was grown overnight at 37 °C with shaking. Between 0.1 ml and 1 ml of culture was used to inoculate 10 ml to 100 ml of YT medium, which was incubated at 37 °C on a flask shaker for 1.5 to 4 hours i.e. until the relative absorbance of the culture was between 0.1 and 0.2. The culture was chilled on ice for 10 minutes then centrifuged for 5 minutes at 4000g in a chilled centrifuge. The supernatant was discarded, the cells taken up in half the original volume with ice cold 50 mM CaCl / 10 mM Tris.Cl (pH 8.0) solution and kept on ice for 15 minutes. The centrifugation step was repeated, the cells being taken up in 1/15th. the original volume, and 200 µl aliquots were dispensed into chilled 500 µl Eppendorf tubes. The competent cells were stored for between 12 and 24 hours at 4°C in order to improve their competency.

Samples of DNA in KLP or TE buffer buffer were mixed with 200 µl aliquots of competent cells and stored on ice for 30 minutes, heat shocked for 2 minutes at 42°C, then incubated in 1ml YT medium at 37 °C to enable antibiotic resistance to develop. A portion or all of the transformation mixture was plated out using a glass spreader onto agar containing appropriate antibiotics, then the plates were inverted and left to incubate for 16 to 48 hours at 32°or 37°C.

2.2.4 Agarose gel electrophoresis

Agarose gel electrophoresis was the standard method adopted for the separation, identification and occasionally purification of DNA (Helling *et al.* 1974; Maniatis *et al.* 1982).

2.2.4.1 Gel preparation

A clean plastic gel former was adhered to a 6"x 5"glass plate with vacuum grease. This was placed on a gel table and levelled. A gel comb with the required number of wells was fitted, adjusting the height above the glass plate to around 2 mm. To the appropriate amount of agarose was added 20 ml 10 x Alec's buffer and 190 ml water. This was heated to around 90° C by bunsen, or heated in a microwave oven for five minutes. When the sol had cooled to around 60°C, 20 µl ethidium bromide was added from 10 mg/ml stock solution and mixed by swirling. The sol was poured into the former and comb apparatus, and left to set. A flatbed "submarine" gel tank was filled with 2.2 l of 1 x Alec's buffer solution made up from 10 x Alec's buffer, and containing 200 µl ethidium bromide from 10 mg/ml stock. The gel comb and former apparatus were carefully removed, and the gel, still on the glass plate was then placed on the bed of the tank. The level of buffer was adjusted to around 2mm above the surface of the gel. The horizontal plane of the tank was adjusted to ensure the gel was as level as possible. Samples of restricted DNA in stop dye were loaded into the wells using an automatic pipettor and the power pack, connected to the tank was switched on. The gel was electrophoresed for between 3 and 12 hours at 25 to 125 v.

The ethidium bromide stained DNA was then visulised by placing the gel on a short wave U.V. transilluminator (Sharp et al. 1973), and photographed using a red filter with a 300asa film and type 667 polaroid camera. Occasionally small fragments were isolated on 4% 'newsieve' gels using a minigel apparatus with TBE buffer instead of Alec's buffer, but generally,for observation of fragments under 250bp, 1% gels were used and between 1kbp and 250bp, 0.6% to 0.8% gels were prepared.

10 x Alec's buffer consists of 48.5 g Tris base and 3.7 g EDTA (pH 7.7) per litre. 10 x TBE buffer consists of 108 g Tris base, 55 g boric acid and 20 mM EDTA (pH 8.0).

2.2.4.2 Identification of DNA

Samples were digested with restriction endonucleases that would give information, once visulised, as to its identity. Along with sample DNA, controls and marker DNA were used. Controls usually consisted of restrictions of the DNA molecules that were being manipulated, prior to manipulation, and unrestricted sample DNA. Lambda phage DNA was most commonly used as a DNA standard by digestion with endonucleases that produced known fragment sizes, that could be easily resolved and were in the general size range of the critical sample DNA fragments.

2.2.4.3 Separation of DNA

When attempting to isolate a fragment of DNA from a gel, it was necessary to run the gel tank for as long as possible in order to separate the wanted fragment from the rest of the DNA to avoid contamination of the fragment with other DNA.

2.2.4.4 Purification of DNA

When miniprep DNA (2.2.12) as opposed to maxiprep DNA (appendices) was used in recombinant work, it was often first purified by gel electrophoresis. The samples were first restricted with appropriate restriction enzymes, then run on a gel. The required fragment was identified and isolated (2.2.5), and either used or stored at -20°C in sterile water.

2.2.5 DNA Fragment isolation

The low melting temperature agarose gel method (Weislander 1979), was used for all fragment isolations with one exception, when the freeze elution methodology was used.

2.2.5.1 Low melting temperature agarose gels

A 0.7% agarose gel was prepared using low melting point agarose in a four track gel former, and using Alec's buffer. The gel was cooled to 37°C before pouring. Once set, the gel was was placed in the tank, buffer was added, then the surround was removed. The samples were loaded, and the voltage was increased in steps of 20 v every ten minutes from a starting voltage of 20 v up to 120 v. The gel was carefully removed (as it has not the strength of common agarose), the DNA visualised on a U.V. transilluminator, and the required fragments were removed with a sterile scalpel. Any excess gel was trimmed off, and the fragments were placed in Eppendorf tubes. The agarose was melted by heating to 65°C and 2 to 3 volumes of lmp gel buffer was added, mixed and placed at 65°C for 10 minutes. The solution was then phenol extracted twice, then ethanol precipitated and resuspended in sterile water.

Lmp (low melting point) gel buffer consists of 50 mM Tris.Cl

and 0.5 mM EDTA (pH 8.0).

2.2.5.2 Freeze elution

After suitable restriction, The required fragment was cut from a normal gel with a sterile scalpel. It was then placed in a sterile 50 ml corex tube, to which was added 0.9 ml water and 0.1 ml 3 M Na acetate / 10 mM EDTA solution. The tube was covered in silver foil and left in a dark cupboard for 15 minutes, shaking gently occasionally. A 0.5 ml Eppendorf tube was pierced through the base with a pointed seeker, and was then plugged with siliconised glass wool, then placed in a 1.5 ml Eppendorf tube with the cap removed. The fragment was transferred transferred to the 0.5 ml tube, and was centrifuged in an Eppendorf centrifuge for 15 minutes to melt the fragment. The liquid was transferred to a 1.5 ml tube, to which was added; 5 µl 1M MgCl2, 20 µl 3 M Na acetate (pH 4.8) and 1ml cold ethanol. These were mixed, stored at -20 °C for 30 minutes and then centrifuged for 10 minutes. The precipitate was taken up in sterile water then ethanol precipitated in the usual way.

2.2.6 Radiolabelling DNA by nick translation

2.2.6.1 Procedure

The method used was based on the findings of Maniatis et al. (1975) and Rigby et al. (1977). The required DNA fragment was isolated from a gel, purified and precipitated as described above, then taken up in 35 µl sterile water. From Amersham nick translation kit N.5000 was added 5 µl solution 1

(nucleotides and buffer), 5 µl solution 2 (DNA Polymerase enzymes) and from the -20°C radiochemical store was added 5 ul (approx. 50 µC) dCTP, alpha labelled with Phosphorous isotope 32. The tube was briefly spun in an Eppendorf centrifuge then incubated at 15 °C for between 90 and 120 minutes. A 5ml plastic pipette was clamped vertically in a retort stand, and the bottom end was plugged with glass wool. Swollen sephadex G50 in nick translation buffer was applied to the column with a pastette to a height of around 5 cm. The column was kept wet by regularly topping up a 2 cm head of nick translation buffer. The incubated nick translation reaction was first quenched by addition of 100 µl buffer, then applied to the column. 12, 500 µl aliquots were collected in 1.5 ml Eppendorf tubes (or until the second peak was detected coming off the column. The relative activities in c.p.s. of each tube was measured by placing at a fixed distance from a shielded Geiger - Muller counter. The tubes containing the highest c.p.s. from the first peak were often pooled. 5 µl of the pooled probe was transferred to a liquid scintilation vial and 1 ml of scintilant (POPOP) was added. The activity of the probe in c.p.m. was measured and the total activity of the probe was calculated from the total volume. If this was greater than 1 million counts a minute, the probe was considered 'hot' enough for use. It was labelled and kept in the -20°C radiochemical store.

Nick translation buffer consists of 150 mM NaCl, 10 mM EDTA, 0.1% SDS and 50 mM Tris.Cl (pH 7.5). swollen sephadex G50 can be obtained by storing 2 g of G50 beads in 30 ml nick translation buffer and leaving overnight or autoclaving for 15

minutes at 120°C.

2.2.6.2 Precautions

Whilst incubating, The tube was covered with a lead pot cover, and in transportation in a perspex holder inside a lead pot. All radiochemical manipulations were carried about behind perspex screens. All waste radioactive solutions was disposed down the assigned radioactive waste sink. All contaminated disposables (tips, tubes etc.) were rinsed thoroughly with water in the radioactive waste sink, until no discernible increase in radiation could be detected from them, then disposed in the 'non radioactive' bin. All reusable items (glassware etc.) were similarly rinsed, then placed in the 4% 'decon' solution bath provided. All radioactive samples were clearly labelled with hazard symbols, the exact form of radioactive molecule used and the date. After use of radiochemicals the work area was checked for contamination with a Geiger - Muller counter, and if necessary, spillages would have been cleaned thoroughly with 4% 'decon' detergent.

2.2.7 Transfer of bacteria onto Nitrocellulose filters

2.2.7.1Colony lifts This method was adopted from that described by Hanahan and Meleson (1980). Bacterial colonies to be screened by in situ hybridisation were grown on large rectangular YT amp X gal agar plates (23 cm2). When they had grown to about the size of pinheads, a dampened sterile Schleicher and Schuell BA 85 nitrocellulose filter was carefully lowered onto the surface of the gel until the nitrocellulose became thoroughly moist. Equally carefully,

the filter was peeled off and layed, colony side up on a fresh agar plate. The original master plate was sealed and stored at 37 ° C to allow colonies to reappear. The filter was then treated to bind the bacterial DNA to the nitrocellulose (2.2.8.2).

2.2.7.2 Manual transfer

Bacterial colonies to be screened by colony hybridisation were picked off master agar plates with sterile cocktail sticks or a wire loop, and transferred to a known position on a nitrocellulose disc (Grunstein and Hogness 1975). Occasionally replica plates were made up from these discs to other discs in the same way as is discribed for colony lifts.

2.2.8 Binding DNA to nitrocellulose filters

2.2.8.1 Directly from agarose gels

This method, known as Southern blotting was devised by E. Southern (1975). After electrophoresis was completed, the gel was photographed. It was transferred to a glass baking dish where the excess gel was trimmed off with a scalpel. The DNA was denatured by soaking the gel in 3 volumes of Denaturing solution, for an hour at room temperature with constant shaking. The solution was discarded and replaced with three volumes neutralising solution (I) and was left for an hour at room temperature with shaking. A piece of 3MM paper was wrapped around a glass plate sitting on bungs to form a raised support, placed in a glass baking dish. The dish was filled with 10 x SSC almost to the top of the support and any air bubbles were smoothed out. A square was cut out of a sheet of

saran wrap, so that it would be a few mm's smaller than an agarose gel. The gel was placed on top, inverted, so that the original underside was uppermost, and bubbles were again removed. A piece of nitrocellulose paper (Scleicher and Schuell BA 85) was cut to be around 2mm larger than the gel in both dimensions. The filter was floated on the surface of a solution of 2 x SSC until it was wet completely from beneath. The filter was then immersed for 3 minutes. The nitrocellulose was then placed on top of the gel, all bubbles removed and the position of the wells marked on with a ball point pen. 2 pieces of Whatman 3MM paper, cut to the same size of the gel was placed on top, again removing all bubbles. Disposable nappies were cut (to a height of 5 - 8 cm) to fit just smaller than the 3MM paper and were placed on top. It was ensured that the saran wrap beneath the gel was spread so as to avoid a short circuit of buffer between the 3MM paper beneath the gel and the nappies above. A glass dish was placed on top, and was weighted down by a 2 l flask filled with water. Transfer was left for 24 hours. The towels, 3MM paper and gel were all discarded and the filter was soaked in 6 x SSC at room temperature for 5 minutes. The filter was then layed on a sheet of 3MM paper for an hour to dry at room temperature. The dried filter was then sealed in an envelope of 3MM paper and baked in an 85 ^{0} C oven for an hour. The filter was stored ar room temperature under vacuum until needed for hybridisation.

2.2.8.2 Indirectly from colonies To four large square petri dishes was added (in different dishes) around 100 ml of 10% SDS (Fritch and Boyer, unpublished), denaturing solution,

neutralising solution (II) and 2 x SSC respectively. Sheets of 3MM paper were cut to fit each. Excess solution was removed and colony filters were first layed, colony side up for for 3 minutes on the SDS filter,and then for 5 minutes each on the denaturing, neutralising and SSC filters, avoiding carrying solutions across from dish to dish. The filters were dried and baked as described above in 2.2.6.1.

Denaturing solution is made from 1.5 M NaCl and 0.5 M NaOH. Neutralising solution (I) is made from 1 M Tris.Cl (pH 8.0) and 1.5 M NaCl. Neutralising solution (II) is made from 1 M Tris.CL (pH 8.0) and 0.5 M NaCl. SSC is 0.15 M NaCl, 15 mM Sodium citrate (pH 8.0). SDS is Sodium dodecyl sulphate.

2.2.9 Hybridisation with radiolabelled probes

2.2.9.1 For Southern blots

The filter was placed in a large heat sealable plastic bag, into which 100 ml prehybridising solution (I) was added and heat sealed after all air bubbles were removed. This was incubated at 65 C for an hour in a shaking water bath. The solution was discarded and replaced with 50 ml hybridisation solution (I), containing a suitable amount of probe. (Double stranded DNA probes were boiled for 5 minutes before being added). The bag was resealed and incubated for 4 hours at ^{0}65 C in a shaking water bath. The solution was discarded down the radiochemical waste sink, and the filter was washed with 100 ml 2 x SSC for 30 minutes at 65 C. This was discarded, and the filter was washed twice for 15 minutes at 65^{0}C with 1 x SSC. The filter was then placed on a sheet of 3MM paper and

left to dry at room temperature for an hour, in preparation for autoradiography.

Prehybridisation solution (I) is made from 5 x Denhardt's solution, 5 x SSC and 100 to 200 µg/ml herring sperm DNA. Hybridisation (I) solution is made from 1 to 2 x Denhardt's solution, 5 x SSC, 100µl/ml herring sperm DNA and denatured radioactive probe. Dehardts solution is made from a 20 x stock, kept frozen, which is 0.4% "Ficoll", 0.4% Bovine serum albumin (BSA) and 0.4% Polyvinylpyrrolidone (PVP).

2.2.9.2 For Colony and In situ hybridisations

The filters were placed in large heat sealable bags, with up to three filters per bag. To each bag was added prewashing solution, 50 ml per filter, which was incubated at 42°C for 1 to 2 hours in a shaking water bath. The solution was replaced with 50 ml prehybridising solution (II) per filter, and was incubated at 42°C for 4 to 6 hours. The solution was replaced with 50 ml hybridising solution (II) per filter with probe, and was incubated at 42°C overnight. Filters were washed and dried in the same way as described above for Southern blots, but the solutions included 0.1% SDS, and the washing protocol was first wash;- 2 x 30 minutes at 42°C, second wash;- 3 x 10 minutes at 65°C.

Prewashing solution is 50 mM Tris.Cl, (pH 8.0), 1 M NaCl, 1 mM EDTA and 0.1%SDS. Prehybridising solution (II) is as described for prehybridising solution (I), with the inclusion of 50% formamide for higher stringency. Likewise with hybridising solutions (I) and (II) with the inclusion of 50% formamide and an extra 100 µg/ml herring sperm DNA in solution

(II).

2.2.10 Autoradiography

All filters were autoradiographed as follows. Two large sheets of glass were covered on one face with 3MM paper stuck on with masking tape. The filter was layed on the 3MM covered surface and gently stuck on with tape. Radioactive ink was dotted around the edge of the film in an asymetric pattern, so as to enable easy orientation with the autoradiograph after exposure. The filter side of the plate was covered in saran wrap. Using only the diffuse red light of a dark room safety lamp, a sheet of X-ray film was removed and the box replaced in a dark drawer. The film was sensitized by flashing briefly with an X-ray film flash gun, and placed, flashed side down on to the radioactive filter plate. This was covered with an intensifying screen, the second glass plate, 3MM side down and clamped. The plates were wrapped in a photographic bag and two black plastic bin liners held secure with elastic bands and stored at -80C for between 2 hours and a week. The film was removed under safelight and developed in phenisol developer for 8 minutes, turning the film every 2 minutes. It was then rinsed briefly in water, drained then immersed in fixer for 3 minutes. The film was rinsed again and hung up to dry.

2.2.11 Identification of recombinants

2.2.11.1 Single probe screens

For single probe bacterial hybridisations, the developed film was aligned with the master colony plate and dark spots that

aligned with colonies were judged positive. From these, overnight cultures were set up, from which, plasmid DNA was subsequently extracted, restricted and analysed. For Southern blots, the autoradiograph contained bands corresponding to known positions. The distance of these bands and unknowns from the well line were compared with those from a photograph of the gel showing all the DNA taken before blotting.

2.2.11.2 Double probe screens

The developed films for double probe hybridisations were compared, and any dark spot that coincided was considered positive. Alternatively, colonies were first screened with one probe and positives were replated, then screened with a second probe. Any dark spots on the second film that aligned with colonies were judged positive.

2.2.12 Plasmid DNA preparations

2.2.12.1 Medium scale preparation

Adopted from the method described by Birnholm and Doly (1979). 15 ml Overnight cultures taken from a single colony were pelleted in a Mistral centrifuge at 3600g for 10 minutes. Most of the supernatant was removed, the pellet was resuspended in around 1 ml YT medium and transferred to 1.5 ml Eppendorf.(Invariably 20 ul was retained and stored at 20°C as a backup culture). The bacteria was pelleted by spinning centrifugation for 1 minute in an Eppendorf Centrifuge. All traces of media were removed by draining and pipetting. 200 µl ice-cold solution (I) was added to each pellet which was vortex mixed until fully suspended. Samples were left at room temperature for five minutes, then on ice for two minutes. 400

µl solution (II) was added and mixed gently by inversion. Samples were left on ice for 5 minutes. 300 µl ice-cold solution (III) was added, the tubes were mixed gently by inversion whilst the clot formed and were left on ice for 10 minutes. samples were spun for 10 minutes in an Eppendorf centrifuge, and up to 850 µl of sample supernatant was recovered to fresh tubes. To each sample was added 500 µl isopropanol, which was mixed by vortex and left at room temperature for 5, minutes then centrifuged for 10. Samples were taken up in 200 µl sterile water or TE buffer then phenol extracted, ethanol precipitated, washed in 70% ethanol and desiccated in a vacuum desiccator. Samples were generally taken up in 50µl sterile water or TE buffer and stored at -20 ^{o}C.

Solution (I) is 50 mM glucose, 10 mM EDTA, 25 mM Tris.Cl (pH 8.0 with 2 mg/ml freshly added dessicated lysozyme. Solution (II) is 0.2 M NaOH, 1% SDS and solution (III) is 3 M Na acetate (pH 4.8).

2.2.12.2 Small scale preparation

This method, adopted from that of Birnboim and Doly (1979) by Ish - Horowicz, was essentially the same as the previous with the following alterations. Only 1.5 ml of overnight culture was pelleted: only 100 µl solution (I), 200 µl solution (II) and 150 µl solution (III) was used per sample; there was no isopropanol precipitation step and the ethanol precipitation step requires no addition of 3 M Na acetate.

2.2.12.3 Large scale preparation

Large scale preparations of plasmid DNA were produced by Dr

Ron Croy and workers by ultra centrifugation and separation on a Caesium Chloride gradient, subsequently the method can be found in the appendices.

2.2.13 Microbiological Techniques

2.2.13.1 Aseptic technique

Cultures of bacteria and phage were manipulated using aseptic technique whenever possible. Wire loops were heat sterilised, cocktail sticks autoclaved and glass spreaders sterilised with flaming ethanol. Culture tubes, stock solution bottles, sterile water and TE buffer bottles were flamed at the neck before and after opening. Stock solutions were opened whenever possible in a laminar flow cabinet. All transfers to and from agar plates were carried out so as to minimise possible contamination.

2.2.13.2 Storage

After inoculation, samples were invariably incubated at 37 ^{o}C, but after once grown, cultures grown in liquid and solid media were stored at 4^{o}C, sealed with 'nescofilm' in the latter, for a period of up to 2 months. For longer periods, samples were stored in 50% glycerol v/v YT medium at -80^{o}C.

2.2.13.3 Disposal

Once cultures had served their usefulness, they were disinfected with 25% 'Chloros', and discarded down the drain or in the case of agar plates, sterilised by autoclave, then discarded into an ordinary waste bin.

CHAPTER THREE

RESULTS

3.1 SUBCLONING BRAZIL NUT DNA INTO PUC18 FROM DOUBLE STRANDED M13 CLONES

THE pUC18 Brazil nut subclones were obtained by a "shotgun" cloning approach, where plasmid DNA solutions for J13 and J16, obtained by medium scale preparation from overnight cultures were digested with restriction enzyme Pst I., then ligated with Pst I restricted pUC18 plasmid (which cuts in the multipurpose cloning site;- see appendix IV). The ligation mixtures had been used to transform competent E. coli which were grown on amp X gal media. Large white colonies from each transformation (three J13, three J16) were selected and plasmid DNA was prepared, linearised with restriction enzyme Eco RI and visualised on a 1% gel. figure 1 shows a small increase in plasmid length for samples mpA3, mpB1, mpB2 and mpB3 and a slightly larger increase for samples mpA1 and mpA2. comparison with fragments of known size suggests that the size increases correspond to the Brazil nut DNA / pUC18 recombination event, all other shotgun recombination events would produce a much larger linearised plasmid. Samples mpA1 and mpA2 appear to contain a multiple insertion of the Brazil nut fragment. The confirmation that the size increase was due to an insertion of DNA into the multi purpose cloning site (m.p.c.s.) region of pUC18 was obtained by digestion of samples with the restriction enzyme Pvu II (figure 2). Restriction fragment analysis confirms that samples mpA3, mpB1, mpB2 and B3 display a size increase in the m.p.c.s. fragment of around 160bp, approximately the size of the Brazil nut fragments A and B as shown in figure 4. Samples mpA1 and mpA2 appear to contain 2 BnA insertions. Overnight bacterial

cultures corresponding to samples mpA3 and mpB1 were used to produce the large scale plasmid preparations pBnA and pBnB respectively. Figure 3 shows a Pst I digest of pBnA and pBnB with lambda digested by Eco RI/Hind III. Restriction analysis shows that each subclone contains an insertion at the Pst I site of a fragment around 155 to 166bp in length, thus confirming the successful subcloning of BnA and BnB in pUC18 beyond any reasonable doubt. Figure 4 schematically illustrates the subcloning rationale.

Figure 1:- Gel photograph of Eco RI digested mpA1-3 and mpB1-3 DNA.

i) mpA1 DNA + Eco RI

h) mpA2 DNA + Eco RI

g) mpA3 DNA + Eco RI

f) mpB1 DNA + Eco RI

e) mpB2 DNA + Eco RI

d) mpB3 DNA + Eco RI

c) pUC18 DNA + Eco RI

b) mpA1 DNA (no enzyme)

a) lambda DNA + Hind III

All tracks contained approximately 1µg DNA and digestions were carried out with a 10-fold excess of enzyme. The control in track h was incubated at 37°C with high salt buffer but no enzyme.

Figure 2:- Gel photograph of Pvu II digested mpA1-3 and mpB1-3 DNA.

i) mpA1 DNA + Pvu II

h) mpA2 DNA + Pvu II

g) mpA3 DNA + Pvu II

f) mpB1 DNA + Pvu II

e) mpB2 DNA + Pvu II

d) mpB3 DNA + Pvu II

c) pUC18 DNA + Pvu II

b) mpA1 DNA (no enzyme)

a) lambda DNA + Hind III / Eco RI

All tracks contain approximately 1µg DNA and a 10-fold excess of Enzyme. The control in track h was incubated at 37°C with medium salt buffer but no enzyme.

Figure 1:-

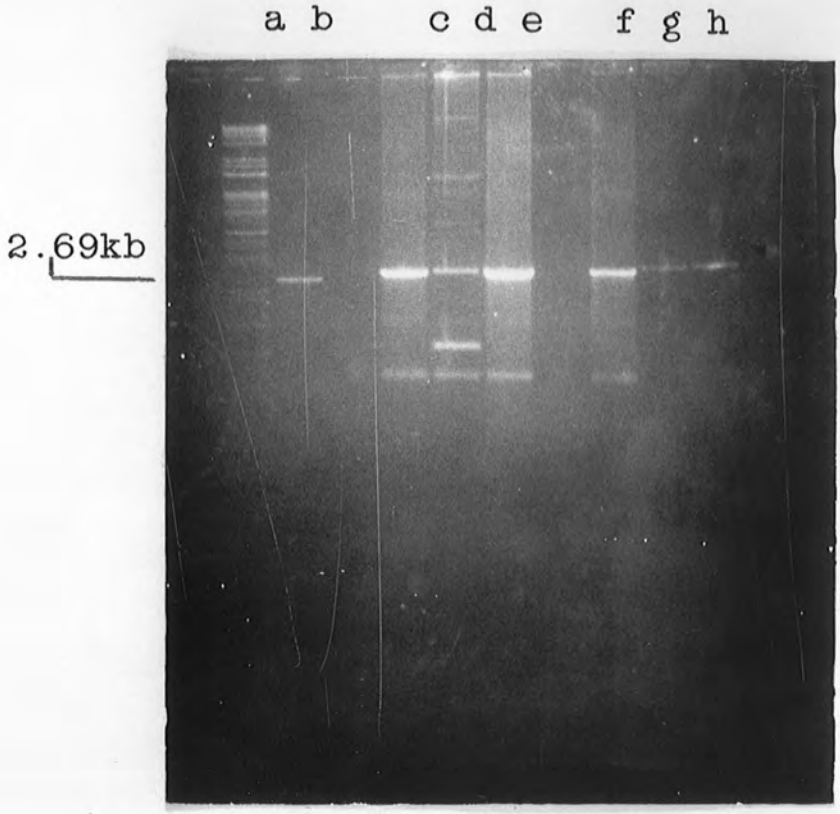
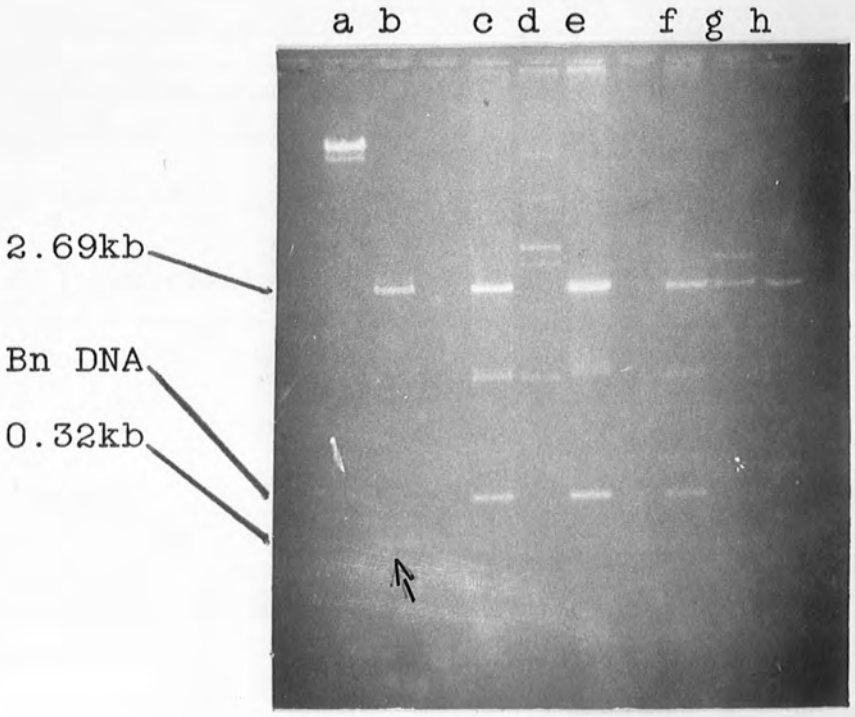


Figure 2:-



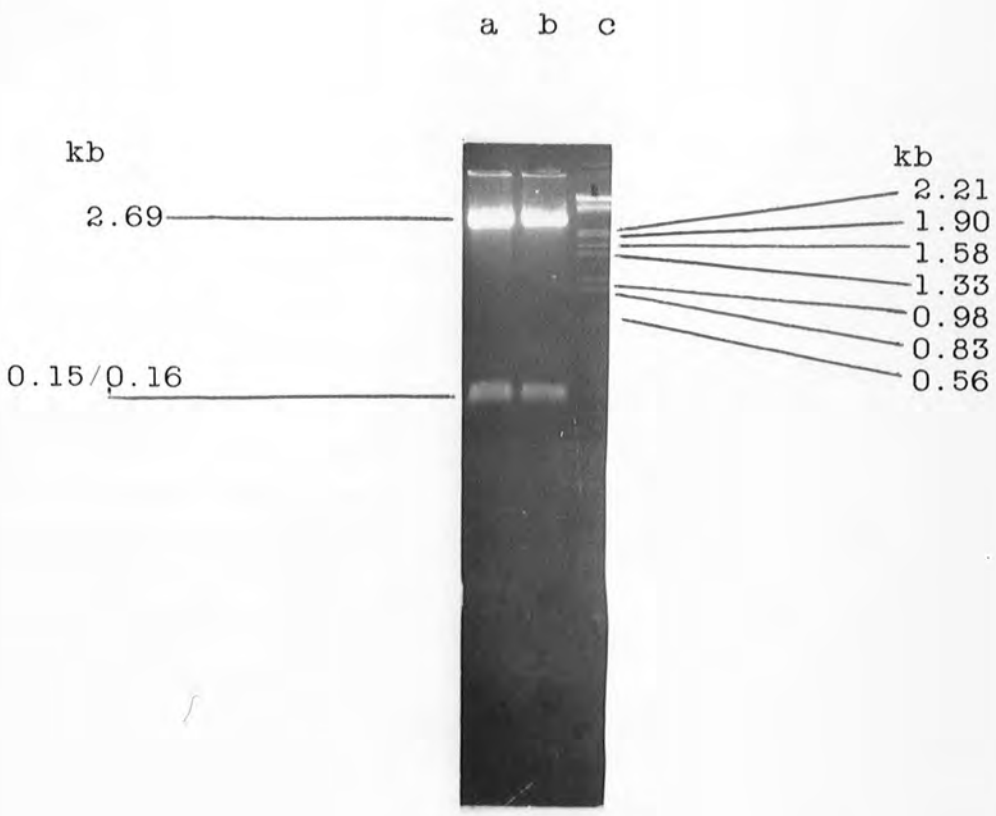


Figure 3:- Gel photograph of Pst I digested pBnA and pBnB DNA.

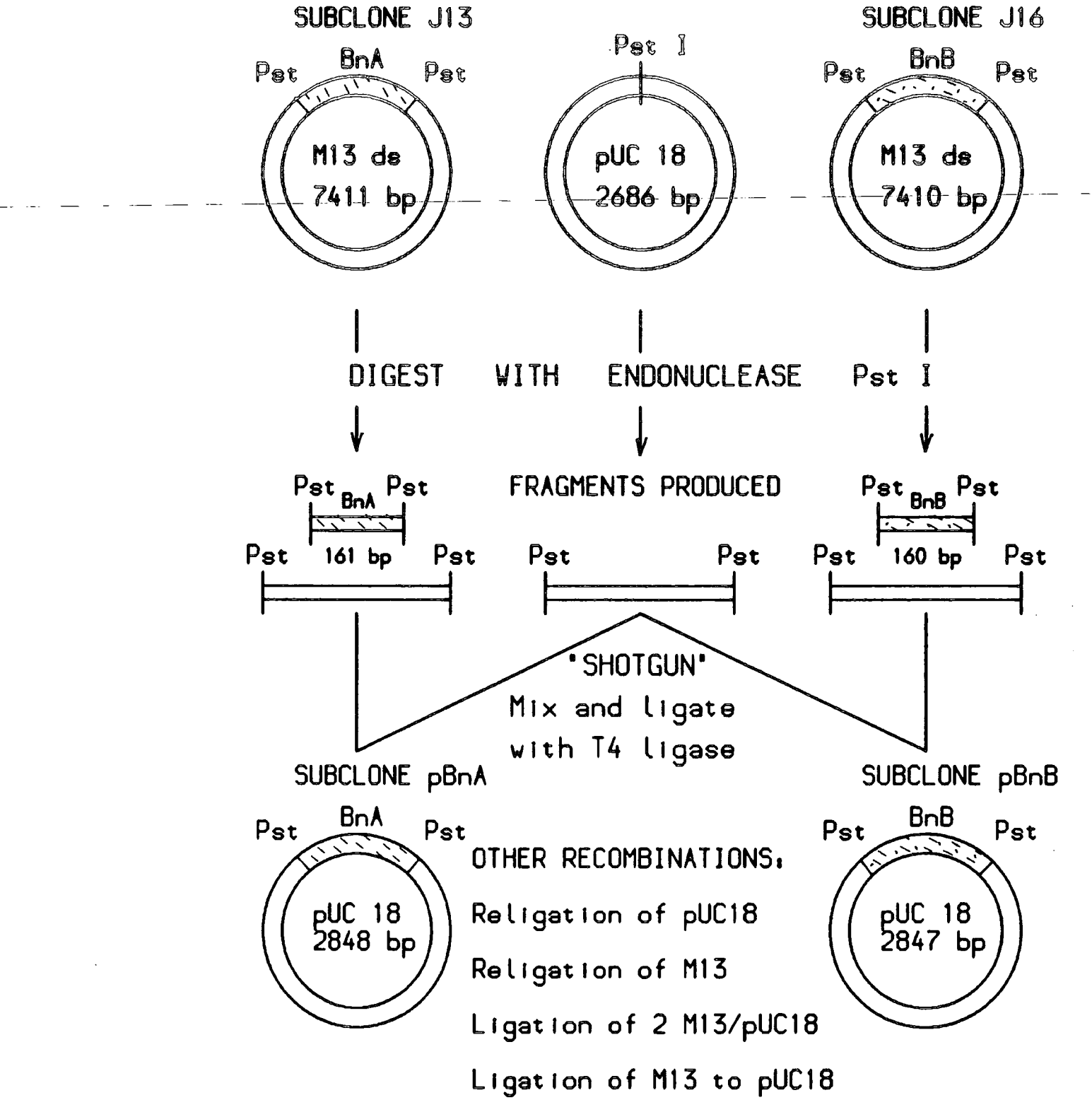
a) pBnA DNA + Pst I

b) pBnB DNA + Pst I

c) lambda DNA +Hind III / Eco RI

pBnA and pBnB tracks contained approximately 10µg DNA and a 2.5-fold excess of enzyme.

Figure 4:- A simple schematic illustration of the subcloning of Brazil nut DNA fragments in pUC18 from double stranded M13 clones.



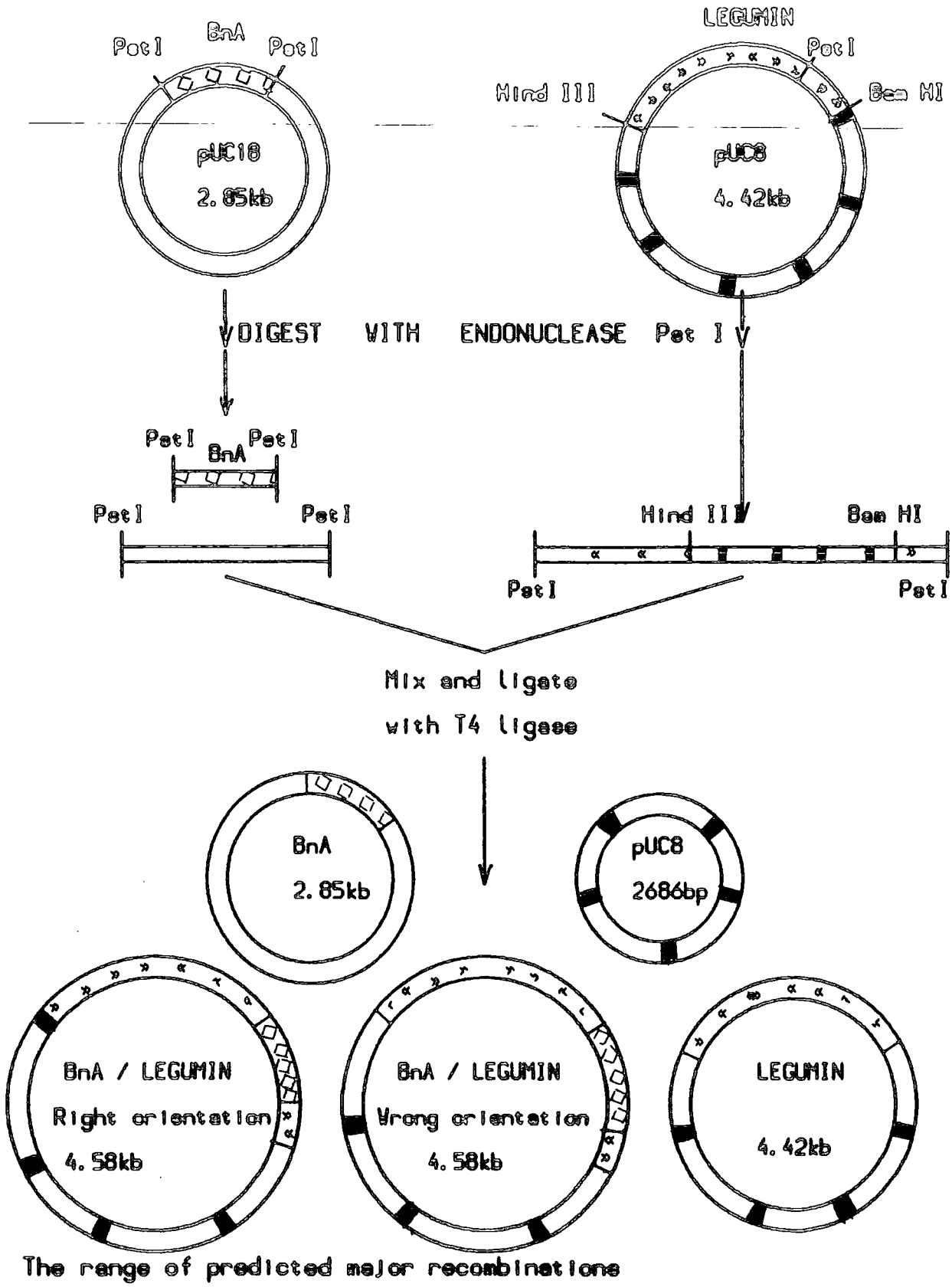
3.2 PRODUCTION OF BRAZIL NUT - LEGUMIN CONSTRUCTS

3.2.1 Construction and screening by "shotgun" cloning - single probe colony hybridisation.

Using the Pst I site rationale illustrated in figure 6, pBnA and JY8 DNA samples were restricted and ligated, and used to transform competent cells of the strain JM 83, producing a wide range of possible recombination events (figure 6). Potential BnA - legumin constructs were selected by growimg the transformation mixtures on agar plates containing ampicillin and X gal, transfering white colonies to nitrocellulose discs. These were screened by colony hybridisation using a radio labelled BnA probe, produced by incorporation of phosphorus 32 labelled cytosine nucleotides by the nick -translation of isolated BnA DNA (figure 5, appendix I). Hybridised colonies were detected by autoradiography (not shown) and from overnight cultures of positives mpLB 1 to 7 , plasmid DNA was prepared, digested with Eco RI and visualised on a 0.7% agarose gel (figure 7), to detect the BnA - legumin constructs by size. Sample mpLB 3 contained DNA of approximately the same size as the legumin control, however further analysis by digestion with Pvu II showed no difference in size of the m.p.c.s. fragment when compared to that of the legumin control (not shown).

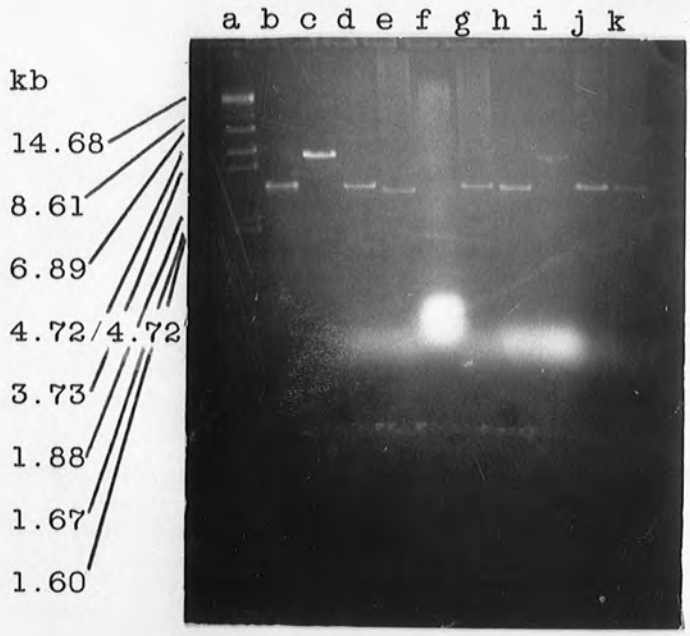
Figure 6:- Shows an illustration of the recombination events expected by the digestion with Pst I and "shotgun cloning" of pJY8 and pBnA.

FIGURE 6.-



The range of predicted major recombinations

Figure 7:- Gel photograph showing digestion of mpLB(1-8) with restriction enzyme Eco RI.



a) lambda DNA + Ava I

b) BnA cont. + Eco RI

c) Leg cont. + Eco RI

d) mpLB 1 DNA (no enzyme)

e) mpLB 7 DNA + Eco RI

f) mpLB 6 DNA + Eco RI

g) mpLB 5 DNA + Eco RI

h) mpLB 4 DNA + Eco RI

i) mpLB 3 DNA + Eco RI

j) mpLB 2 DNA + Eco RI

k) mpLB 1 DNA + Eco RI

All digestions were with approx 1µg DNA with a five-fold excess of enzyme. Undigested control mpLB1 in track d was incubated at 37°C with high salt buffer and no enzyme.

3.2.2 Double probe colony hybridisations

160 white colonies picked off from x - gal plates of BnA - legumin transformation mixtures produced by shotgun cloning were transferred to duplicated positions on two nitrocellulose discs. One disc was hybridised with radiolabelled BnA, the other with labelled legumin DNA. Autoradiographs (figures 8a and 8b) for the filters were aligned to determine colonies that hybridised to both probes. 110 colonies hybridised with the BnA probe, 20 with the Legumin DNA probe and two colonies; G1 and G4 hybridised with both. Figure 9 shows plasmid DNA obtained from overnight cultures of colonies G1 and G4 (definite double positives), B2, B7, B10, G3 and I10 (potential double positives), F2 and H7 (legumin and BnA controls), digested with Eco RI to linearise, and separately with Pvu II to liberate the m.p.c.s. fragment were visualised on a 1% agarose gel. None of the samples appear to contain a legumin DNA - BnA construct though both digestions clearly show on comparison with the controls that G1 and G4 contain both legumin and Brazil nut DNA, suggesting they are either from mixed or co-transformed colonies. In a previous attempt at the double hybridisation screen it was noted that all white colonies selected for the hybridisation screen developed a blue pigmentation on x gal media when stored at 37 ^{o}C for longer than 24hrs. These were streaked out and grown up several times before stable white colonies were produced. It was thought that the legumin construct JY8 could be unstable, and that this would explain the ambiguity of the detection of Brazil nut DNA.

The Eco RI digestion displays an interesting marginal

variation in size that can be explained by examination of the Pvu II digestion. The latter shows that whilst B2, B10, G4 and H7 (BnA control) contain a single Brazil nut insert, B7, G1 and G3 contain two BnA fragments (extrapolated from a calibration curve of known lambda restriction fragments sizes, figure 10, appendix I).

Figure 8a:- An autoradiograph displaying the hybridisation of a Brazil nut DNA labelled probe to colony hybridisation filters.

Figure 8b:- An autoradiograph displaying the hybridisation of a Legumin DNA labelled probe to duplicates of the colony filters shown in figure 8a.

Colonies G1 and G4 hybridise to both probes.

FIG.8A LEG.

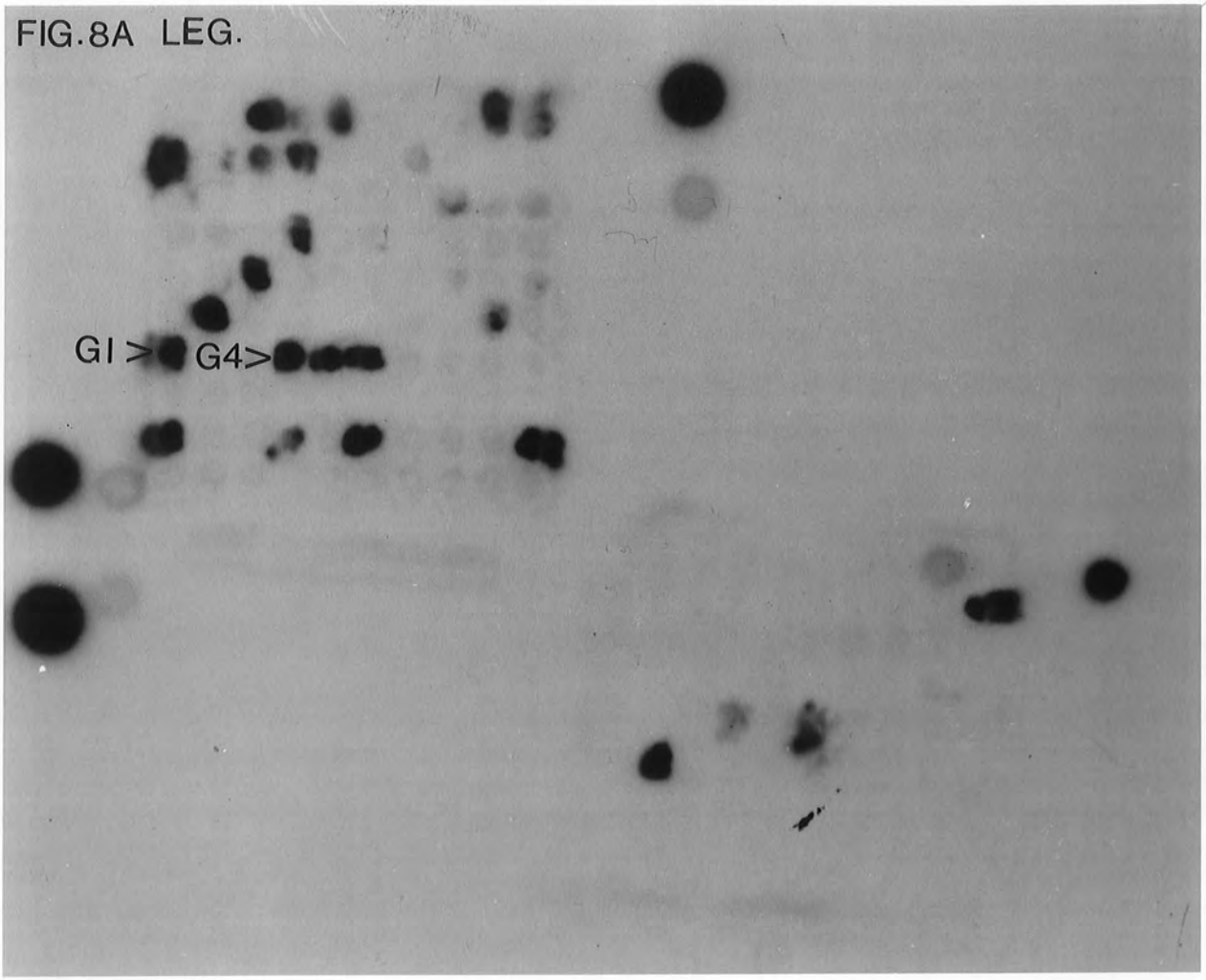


FIG.8B BNA

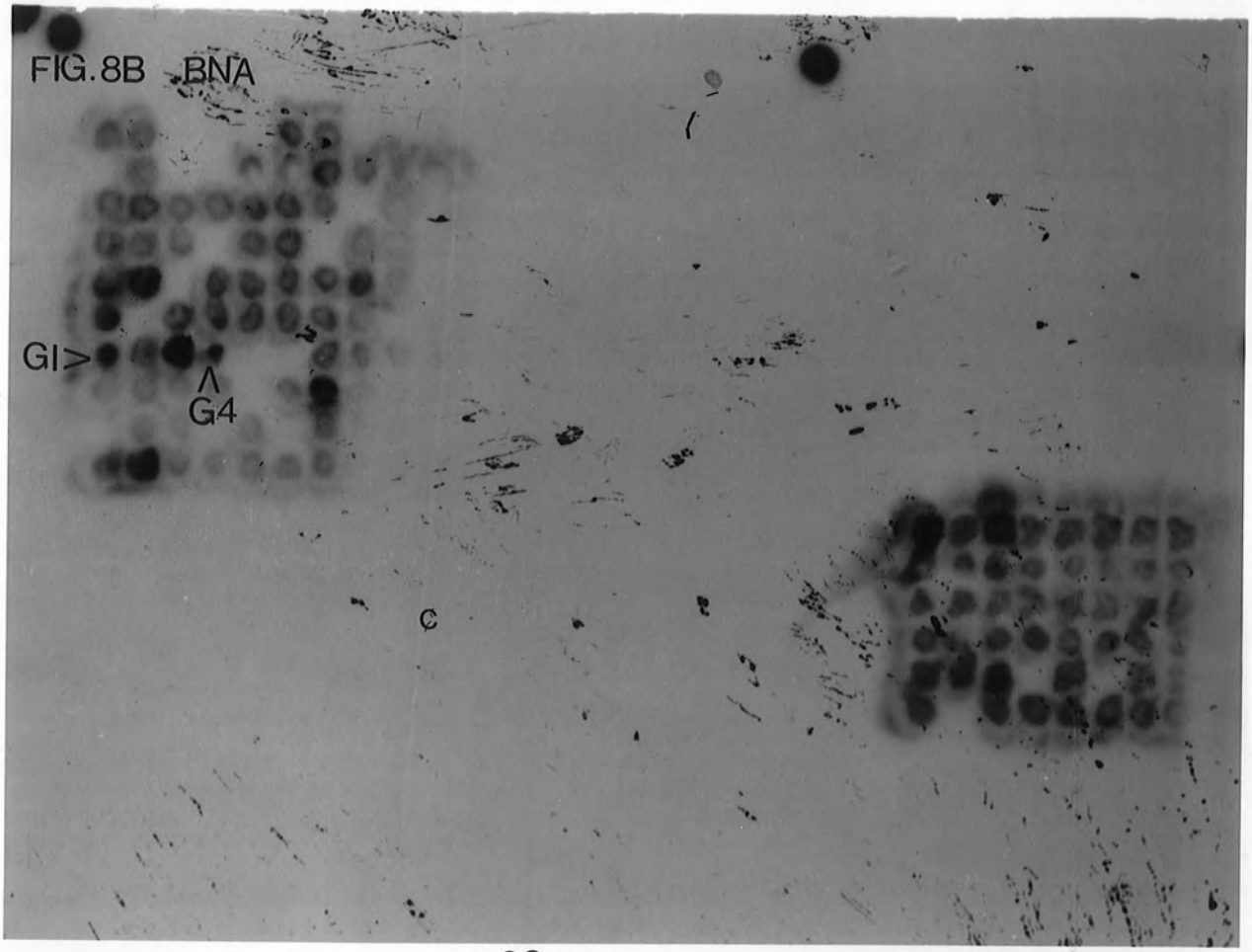
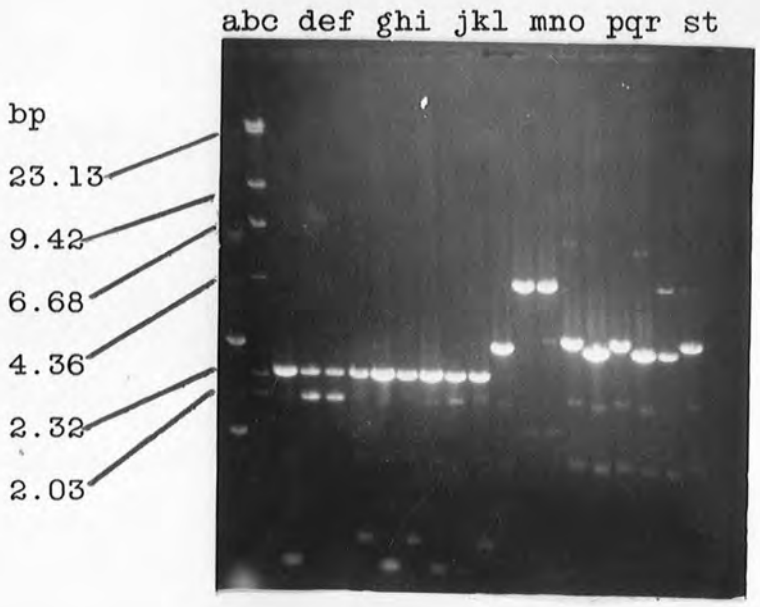


Figure 9:- A digestion of potential BnA - Legumin DNA constructs digested singularly and seperately with restriction enzymes Eco RI and Pvu II.



a) lambda	DNA + Hind III	k) G1	DNA + Pvu II
b) G1	DNA(-enzyme)	l) H7	DNA + Eco RI
c) H7	DNA + Pvu II	m) F2	DNA + Eco RI
d) F2	DNA + Pvu II	n) I10	DNA + Eco RI
e) I10	DNA + Pvu II	o) G3	DNA + Eco RI
f) G3	DNA + Pvu II	p) B10	DNA + Eco RI
g) B10	DNA + Pvu II	q) B7	DNA + Eco RI
h) B7	DNA + Pvu II	r) B2	DNA + Eco RI
i) B2	DNA + Pvu II	s) G4	DNA + Eco RI
j) G4	DNA + Pvu II	t) G1	DNA + Eco RI

All samples contain 1 - 2 µg DNA in a 5 - fold excess of enzyme. Undigested G1 in track b was incubated at 37° C with medium salt buffer and no enzyme.

3.2.3 Fragment isolation and single probe in-situ hybridisation

Figure 11 shows a schematic representation of the fragment isolation - single probe rationale. Subclone pBnA, digested with Pst I was visualised on a 0.7% lmp agarose gel. The BnA restriction fragment was removed, purified and ligated with Pst I digested JY8 legumin DNA then used to transform competent JM83, the whole transformation plated onto two large YT amp X gal plates plates, In - situ copies were made with sheets of nitrocellulose. Figure 12 shows an autoradiograph of the 18 colonies that hybridised with phosphorus 32 radiolabelled Brazil nut DNA probe. Figure 13 shows plasmid DNA purified from overnight cultures of colonies that appeared to correspond to hybridisation events, digested with Pvu II and visualised on a 0.7% gel. 12 samples are shown, A4 and A6 display restriction fragments larger than the standard legumin containing fragments, in addition to these legumin bands. The lambda Hind III digest would suggest the fragment to be around 0.15 kb longer than the equivalent legumin restriction fragment which is the approximate size of the Brazil nut fragment BnA. Therefore, A4 and A6 would appear to have derived from mixed colonies of JM 83 transformed with BnA - legumin constructs and those containing plasmid DNA with legumin alone. The plasmid DNA for 10 more presumptive hybrids was digested with Pvu II and a further three; B5, B7 and B10 displayed evidence for the presence of the BnA - legumin construct in the mixed colony form with legumin plasmid DNA (not shown). Samples of each positive were streaked out on YT amp X - gal plates, single colonies picked

off, their plasmid DNA purified, then digested singularly and separately with restriction enzymes Pvu II and Pst I on a 1% gel. Figure 14 shows these digestions for samples A4(1-5) and A6(1-5). Samples A42, A44, A45 and A63 show pure preparations of the BnA - Legumin construct, A43 a mixed preparation. The Pvu II digest Demonstrates the size increase of the construct containing fragment when compared to the equivalent legumin DNA fragment, and the Pst I digest demonstrating that this size increase is due to a 160bp fragment (appendix II), released on digestion from the Pst I site of the legumin DNA. Similarly, DNA samples of B5(1-5), B7(1-5) and B10(1-2) were purified and restricted, and pure BnA - legumin construct DNA obtained only for B74 (not shown).

Figure 15 shows the digestion of the BnA - legumin constructs A44, A45, A63, and B71 with Dde I. Whilst any positive identification of bands shown by DNA of the right orientation would be masked by the presence of other bands, none of the samples contained restriction bands corresponding to the wrong orientation, thus suggesting that all contain the construct in the correct orientation. This result was confirmed by southern blot (figure 17, appendix I), and the positions of BnA containing fragments highlighted in the acetate overlay of figure 15 (figure 16).

Figure 11:- An schematic representation of the rationale adopted for the production and isolation of BnA - legumin DNA constructs by fragment isolation and single probe in situ colony hybridisation.

FIGURE:- 11

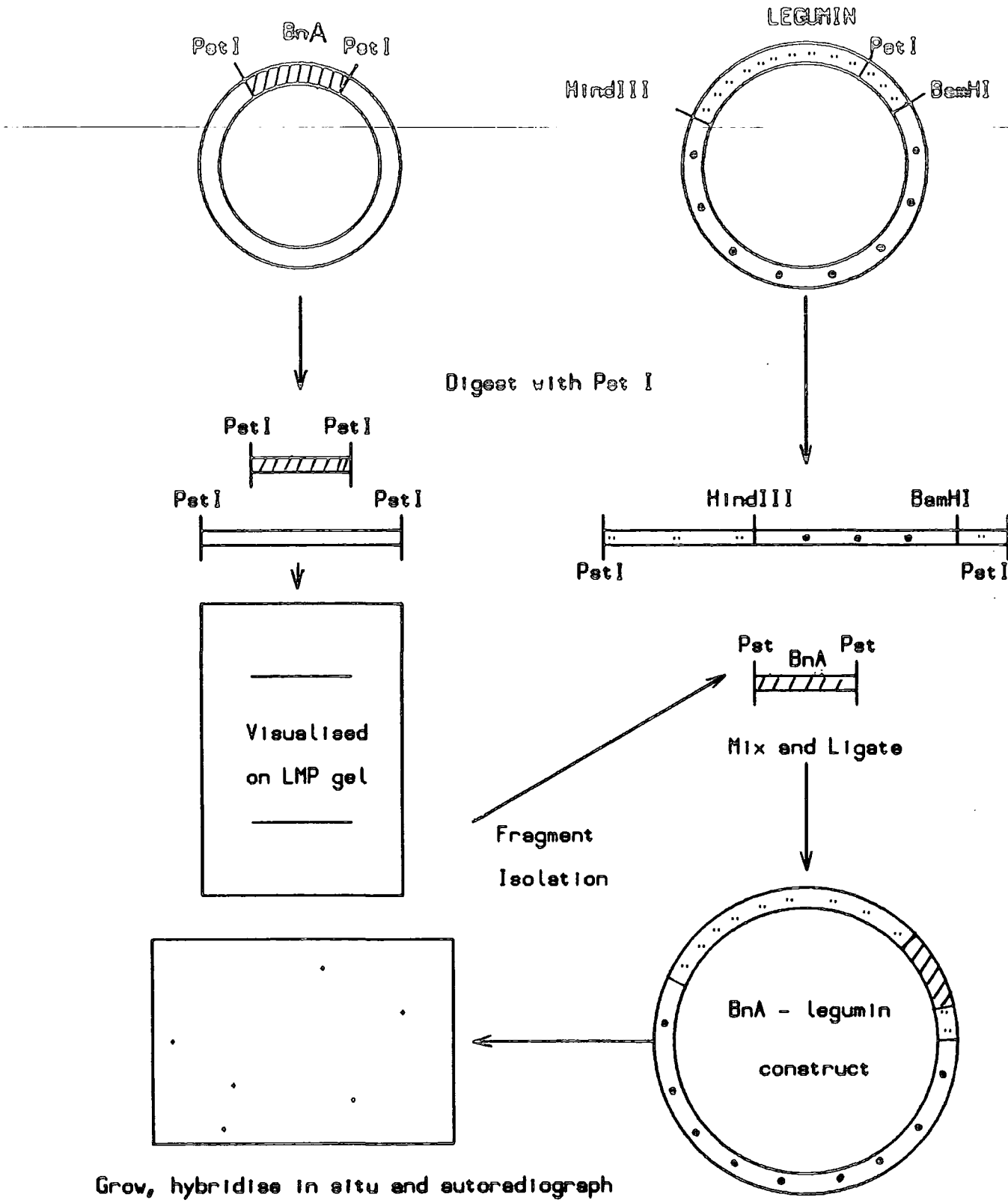

LEGUMIN
PstI
BnA
PstI
PstI
HindIII
BamHI
Digest with Pst I
PstI
PstI
PstI
PstI
HindIII
BamHI
PstI
PstI
Visualised
on LMP gel
Pst
BnA
Pst
Mix and Ligate
Fragment
Isolation
BnA - legumin
construct
Grow, hybridise in situ and autoradiograph

Figure 12:-Autoradiograph showing in situ hybridisation of BnA probe to the whole transformation mixture (Note, half scale and entire autoradiograph not shown).

FIGURE 12

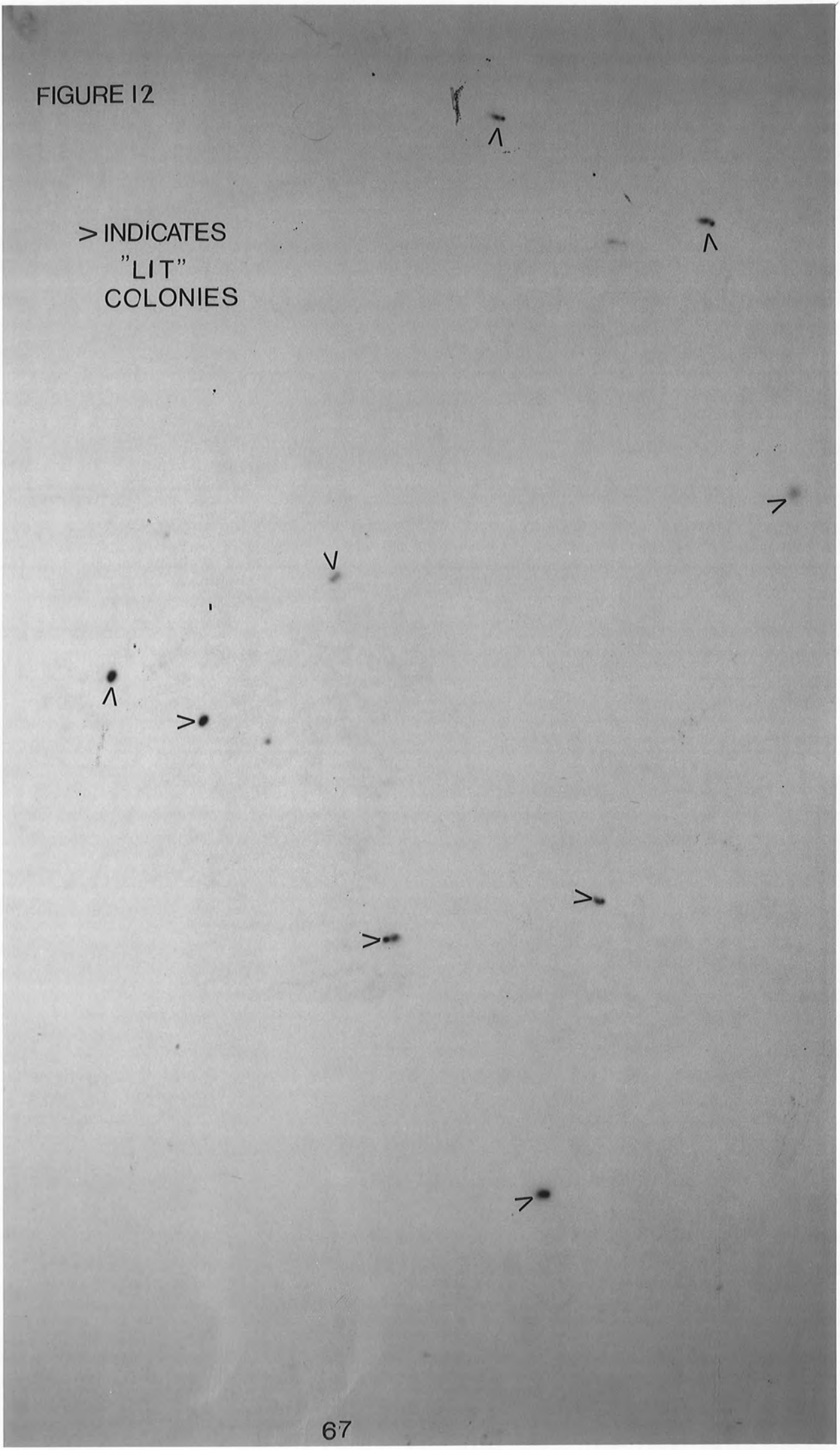


Figure 13:-

A Gel photograph shows a Pvu II digest of DNA from 12 colonies that appeared to hybridise with a radiolabelled BnA probe. (A1-7, A9, B1-4).

a) B4 DNA + Pvu II
b) B3 DNA + Pvu II
c) B2 DNA + Pvu II
d) B1 DNA + Pvu II
e) Lambda DNA+HindIII
f) A1 DNA (-enzyme)
g) pUC18 DNA +Pvu II
h) pBnA DNA + Pvu II
i) JY8 DNA + Pvu II
j) A9 DNA + Pvu II
k) A7 DNA + Pvu II
l) A6 DNA + Pvu II
m) A5 DNA + Pvu II
n) A4 DNA + Pvu II
o) A3 DNA + Pvu II
p) A2 DNA + Pvu II
q) A1 DNA + Pvu II

All tracks contained approximately 1 µg DNA and a five - fold excess of enzyme. The control in track f was incubated at 37 °C with medium salt buffer and no enzyme.

Figure 14:-Gel photograph shows a Pst I digest (tracks a - j) and Pvu II digest (tracks m - v) of DNA from colonies (A41-5, A61-5) purified from from mixed cultures A4 and A6.

a) A65 DNA + Pst I
b) A64 DNA + Pst I
c) A63 DNA + Pst I
d) A62 DNA + Pst I
e) A61 DNA + Pst I
f) A45 DNA + Pst I
g) A44 DNA + Pst I
h) A43 DNA + Pst I
i) A42 DNA + Pst I
j) A41 DNA + Pst I
k) A41 DNA (-enzyme)
l) Lambda DNA + Hind III
m) A65 DNA + Pvu II
n) A64 DNA + Pvu II
o) A63 DNA + Pvu II
p) A62 DNA + Pvu II
q) A61 DNA + Pvu II
r) A45 DNA + Pvu II
s) A44 DNA + Pvu II
t) A43 DNA + Pvu II
u) A42 DNA + Pvu II
v) A41 DNA + Pvu II

All tracks contained approximately 2 µg DNA. Undigested A41 (track k) was incubated at 37 °C with medium salt buffer and no enzyme.

Figure 13

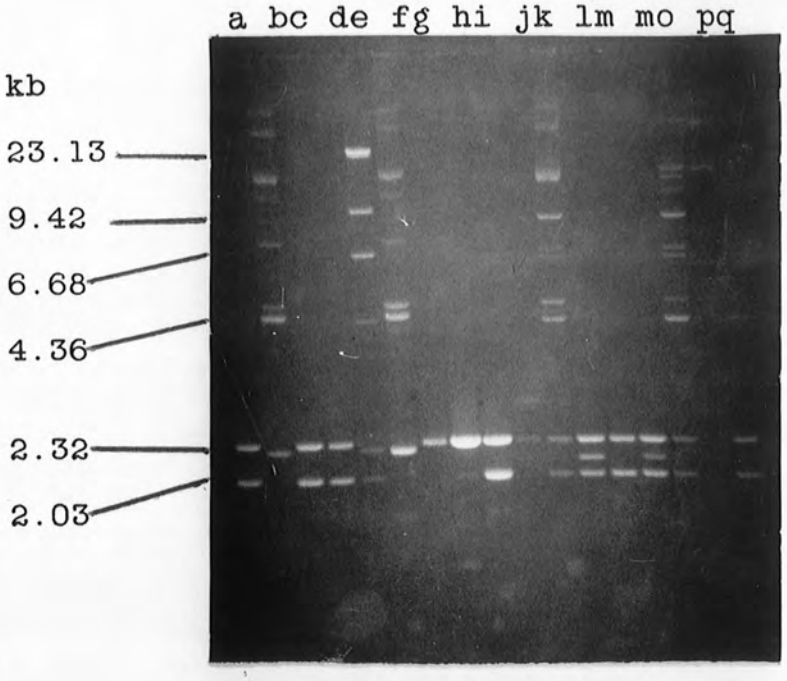


Figure 14

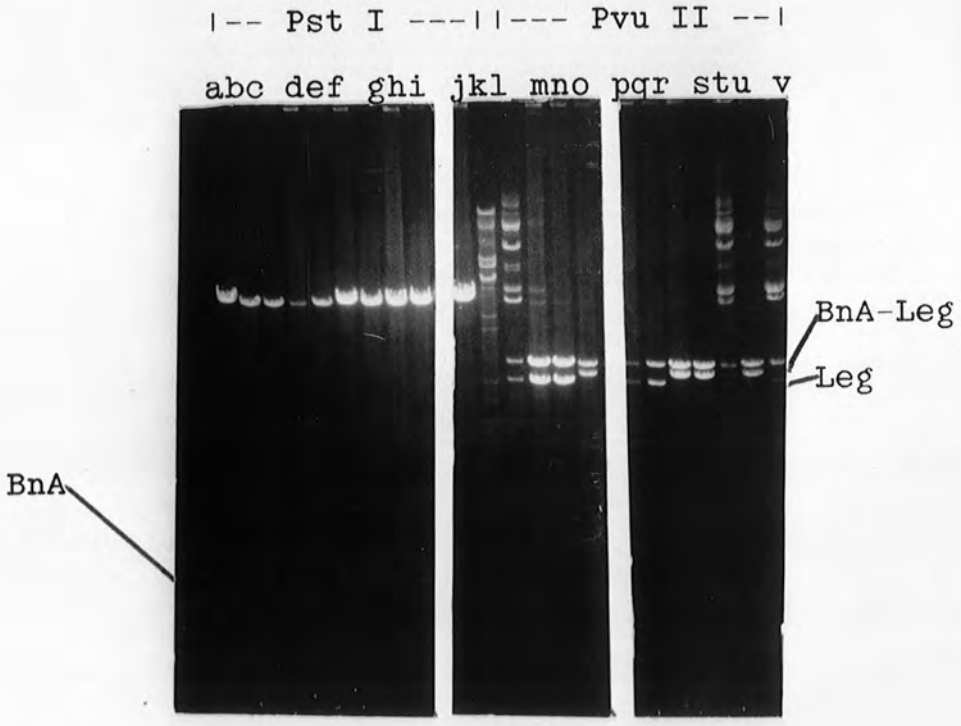


Figure 15:-Photograph shows a Dde I digest of pGPB1 plasmid purified from colonies A41, A44, A45 A63 and B71.

a) Lambda DNA + Eco RI / Hind III
b) pUC 18 DNA + Bst NI

c) pGPB1 (A42) DNA (-enzyme)

d) pBnA DNA + Dde I

e) pJY8 DNA + Dde I

f) pGPB1 (B63) DNA + Dde I

g) pGPB1 (A45) DNA + Dde I

h) pGPB1 (A74) DNA + Dde I

i) pGPB1 (A44) DNA + Dde I

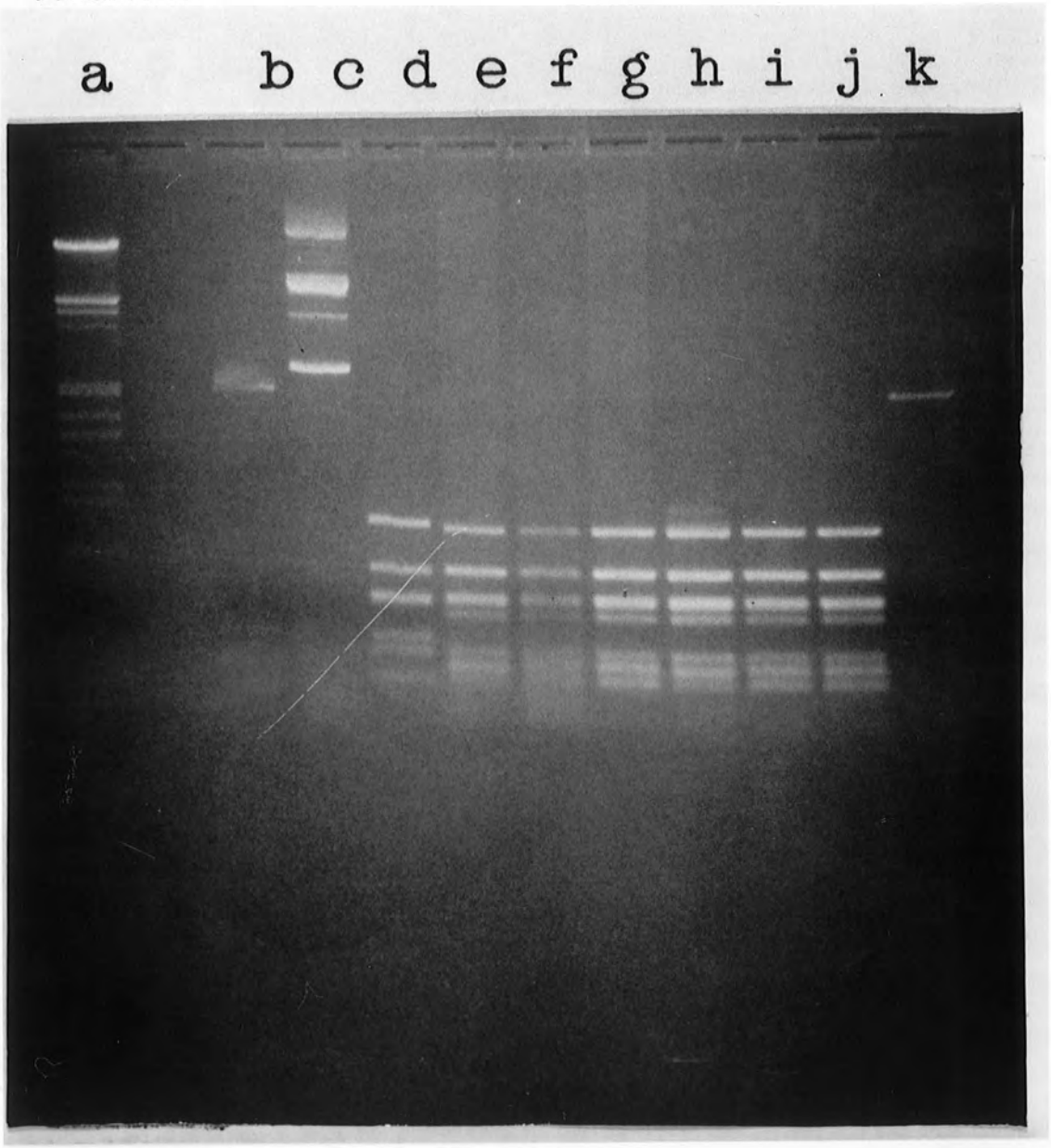
j) pGPB1 (A71) DNA + Dde I

k) pUC 18 DNA + Bst NI

All samples contain approx. 2 µg DNA with a 5 fold excess of enzyme. The pUC18 Bst NI digests (tracks b and k) were incubated at 60 °C. Undigested A42 was incubated at 37 °C with Medium salt buffer and no enzyme. Track h contains a legumin control derived from a mixed colony with Brazil nut construct

Figure 16:-The acetate overlay highlights those fragments that must contain BnA - legumin DNA in the right orientation.

FIGURE 15



3.3 PRODUCTION OF BRAZIL NUT - VICILIN CONSTRUCTS

Some preparation work was carried out for the cloning of Brazil nut into vicilin cDNA; pUC8 plasmid DNA was treated with S1 nuclease to delete the Pst I site (appendix I, figure 18), in order that it may be unique to vicilin, which was to be cloned into the modified vector. Furthermore a vicilin probe, isolated from a digestion of construct pDUB9 and radiolabelled with P32 was made in anticipation of a hybridisation experiment, which was cancelled due to commitment to the legumin - Brazil nut programme.

CHAPTER FOUR

DISCUSSION

4.1 PREVIOUS WORK ON BRAZIL NUT CLONES J13, J16 AND LEGUMIN CONSTRUCT JY8

M13 clones J13 and J16 were taken from a previously prepared cDNA bank. they represent a cDNA subunit encoding Brazil nut 2s protein that had been homopolymerically tailed at the 3 prime ends with 13omer cytosine oligonucleotides, then cloned into plasmids that had been cleaved with Pst I and tailed with 13omer guanine oligonucleotides. The clone was digested with Pst I to release the BNA subunit, which now had created Pst ends and had also cleaved at an internal Pst site to produce two fragments (BnA and BnB). These had been subcloned into the double stranded M13 plasmids J13 and J16, containing BnA and BnB fragments respectively. The sequence of the fragments had been determined (see appendix II). All this work had been carried out prior to this project.

. The legumin construct pJY8 was created by Dr. Jenny Yarwood from two cDNA fragments for the gene legumin A, with a 196bp addition at the 5' end of a short region of genomic DNA The total length is unknown due to an undetermined length of polyA at the 3'end, but it is at least 1730bp long. Further details are described in appendix III.

4.2 ANALYSIS OF LEGUMIN FOR SUITABLE MUTATION SITES

Briefly summarising the objectives and considerations that were taken into account, when determining a rationale for the manipulation of the legumin cDNA:-

The aim was to introduce nutritional improvements into the coding sequence of the seed protein legumin. The legumin

coding sequence had to be mutated by insertion of further genetic information that would lead to a substantial increase in the proportion of deficient amino acids, principally methionine in the coded protein. It had been proposed that this could possibly be achieved by insertion of DNA sequence from a cDNA coding for the methionine rich 2s protein in Brazil nut. Mutation must occur in a directable and controlled manner, at a site and in a region that would not detrimentally alter the character, form or function of the protein coded for by host sequence. The possibilty of this can be minimised by avoiding a shift in the triplet reading frame (which also ensures that the inserted DNA be read correctly), and by ensuring that insertional mutation occurs in a region least likely to be of importance to the structure and function of the coded protein. There could potentialy be a wider range of such regions in the coding sequence of the storage protein legumin than in other functional proteins such as enzymes, as its its function would appear less dictated by its fine structure. Small changes in shape and size resulting from mutatation would be more likely to be accommodated. The precise function of storage proteins is not understood, though what would seem likely to affect function would be a change in solubility or packing ability. This would exclude sites of high or even moderate hydrophobicity, as these are likely to be involved in intra- and intermolecular binding; the first important for establishing major protein structure, the latter for protein packing and maintaining insolubility. Regions of high hydrophilicity must also be avoided, as mutation may increase their solubility and that of the protein as a whole.

The mutation approach used was the site-directed mutation of restriction fragments, which depletes the number of potentialy useful sites further, primarily to those that contain restriction recognition sequences common to both host and insertional DNA. It has already been suggested that mutation with the very short sequences encoding Brazil nut 2s protein would minimise the risk of altering the function of the expressed modified legumin, but predictably these contain few restriction enzyme recognition sites. The Brazil nut DNA fragment in subclone pBnA contained only 23 restriction sites for a total of 14 restriction enzymes. When seen compared to the legumin cDNA which contained over 100 sites for around 40 enzymes, it is not suprising that there are few sites in common with the two sequences that proved useable.

Using the sequence and restriction site data for legumin construct JY8 (see appendix II) provided it was decided that the simplest useful construct would be produced by linearising the legumin construct with either Xho I, Eco RV or Hpa I. All are unique sites that give 'blunt' ended double strand DNA, that cuts in frame i.e. between two amino acid coding triplets. Brazil nut fragment 5 would be cut within the sequence at the Nla IV site, and in the pUC18 m.p.c.s. at the Sma I or Xma I. The double restriction would again cut exactly between coding triplets generating a Brazil nut - pUC18 fragment with 'blunt ends'. This could be ligated to the linearised legumin construct to produce a legumin - BNA construct, that would be perfectly in frame throughout, without the need for linkers, adaptors or enzymatic 'blunting' of any DNA terminal ends. Unfortunately the restriction

enzyme Nla IV which was only produced by New England Biolabs, Beverly, USA had been withdrawn. No equivilent sites existed in either of the two Brazil nut fragments, so an modified rationale was proposed.

Digestion of BnA with Hpa II would generate a 2 base 5' overhang that could be removed with S1 Nuclease to produce a blunt end that would be in frame with the legumin DNA, once ligated in the correct orientation. Analysis of the hydrophilicity index profile for legumin (Croy and Gatehouse, 1985) reproduced in figure 19 (appendix V) suggests that none of these sites would appear particularly suitable. Using the Hpa I site would be especially hazardous as it occurs in the region of greatest hydrophobicity, undoubtedly playing some role in legumin protein - protein interaction.

Linearisation of the Legumin DNA with Pst I and ligatation to the Pst I Brazil nut fragment from subclone pBnA would produce a construct inframe at the legumin 3' - Brazil nut 5' junction, throughout the - Brazil nut DNA, but out of frame at the 3' BnA 5'legumin junction. This would generate a very short stretch of 'missense' with an amber mutation occuring around 40 bases upstream of the native legumin stop signal. Whilst this being undesirable, it may have no significant effect on function of the protein coded, as it would occur very close to the N terminus, and as such is unlikely to remove amino acids that would have played a vital role in protein structure. This is supported by the secondary structure predictions made by Croy and Gatehouse (1985) in figure 19, that suggests the region around it adopts the common tertiary alpha helix. The region also exhibits a

moderate hydrophilicity, which should be acceptable for mutation of its coding sequence with Brazil nut DNA. However, the water soluble nature of 2s protein encoded by the Brazil nut DNA may well effect overall philicity whatever region it occurs in. Unpublished data suggests (J. A. Gatehouse and R. R. D. Croy, pers. comms.) this region has little homology with other related legumins. Such lack of conservation of this region would seem to support the view that it plays no major role in the structure or function of the encoded storage protein.

4.3 ANALYSIS OF RESULTS

4.3.1 Cloning of Brazil nut DNA into pUC18

Two Pst I fragments of DNA encoding a region of Brazil nut 2s storage protein; BnA (161bp) and BnB (160bp) were successfully cloned into pUC18 from d.s. M13. The resulting subclones pBnA and pBnB were demonstrated to each contain a fragment of around 160bp by restriction analysis.

Large white colonies were preferentially chosen from the X gal transformation plate as a visual screening exersize for pUC18 containing small inserts. The other two main white colony producing recombinant events involved much larger fragments, resulting in colonies growing at a slower rate due to their greater replication durations, that appeared to be visually detectable as different. Figure 1 suggests that all samples contain DNA of a greater length than that exhibited by pUC18. Unfortunately the size difference cannot be determined from the imformation presented in this figure due to a failure of

lambda DNA in track a to digest with Hind III. Samples MpA1 and MpA2 in tracks i and h would appear to be of greater length than the other samples. It seems likely on the basis on the small increase in length that the plasmids contain one or several fragments of Brazil nut DNA, though this cannot be confirmed from these results.

The data obtainable from figure two confirms the presence of an insertion of DNA into the region of the pUC18 m.p.c.s.. The known fragment sizes of the pUC18 control in track c was sufficient to confirm the cloning of fragments of Brazil nut DNA as the increased size observed in the vector cloning region could only be due to the insertion of material much less than 2.4 kb in length. This is finally confirmed by a determination of the size of the Pst I fragment released from pUC18 clones in tracks a and b from

interpollation of the lambda digest data shown in in Figure three. The size can only be estimated to around 155 - 165 bp due to the spread of banding on the 0.7% lmp gel (the fragments were to be isolated for other work), though this sufficient. No difference could be observed in the migration distances of the Brazil nut fragment A in track a and fragment B in track B, which is not surprising a their sizes differ by only 1bp.

4.3.2 Mutation of sequence encoding legumin with Brazil nut clone pBnA

The legumin A coding sequence from JY8 was successfully mutated with a 161bp fragment of DNA encoding Brazil nut 2s

storage protein. This was clearly demonstrated by radiolabelled DNA probes and restriction fragment analysis.

Figure 7 gives no positive proof of the presence of a mutated legumin fragment. One sample (track i) contained a band around the size of the legumin control (track c) that was later shown to be a legumin fragment. It was decided that single probe hybridisation approach was too laborious and time consuming, but it was also decided to maintain the 'shotgun' principle in order that high transformation efficiencies might be obtained, which it was thought would suffer if the fragments of interest were isolated from a gel.

Figures 8a and 8b represent the hybridisation of a) BnA and B) Legumin DNA probes to 160 white transformants off an X gal plate. whilst the colonies G1 and G4 hybridised to both probes, restriction analysis (figure 9) revealed that they indeed contained both sequences, but not together. (G1: tracks k and t, G4: tracks j and s). It is likely that theses colonies were mixed i.e. being a mixture of transformed JM83 cells that contain one plasmid with those containing the other. Alternatively these cells could have been co-transformed with a plasmids containing both different inserts. At first examination this seems likely, as such events tend to be rare. However the incorporation of several fragments into a single plasmid is certainly a low frequency event, as illustrated by Eco RI and Pvu II digests of samples G1, G3 (tracks f and o) and B7 (tracks h and q) are slightly larger than that exhibited by the BnA control H7 (tracks c and l). Restriction analysis using values obtained from the

lambda Hind III digest shows that these all contain a double insertion of BnA. This observation is important as it suggests that the legumin - BnA construct must be a very low frequency event for double insertions to be seen more commonly.

Hybridised colonies shown in the autoradiographs in 8a and 8b account for only 128 of the 160 colonies screened. This suggests that as much as 20% of the colonies screened contained neither Legumin or Brazil nut sequences. After restreaked a selection onto agar plates containing ampicillin and X gal a significant proportion showed blue pigmentation, indicative of the presence of intact pUC18. This was also observed in a previous attempt at the double screen approach. It was suggested that this may indicate instability of the legume construct JY8 (J. A. Gatehouse and J. N. Yarwood, pers. comms.) as this phenomenon had been observed by other workers using this plasmid. In retrospect, this seems more than unlikely, as the construct is cloned into a modified pUC8 vector in which 20bp of the multi purpose cloning site (the region between the Hind III and Bam HI sites) was deleted, so that even the complete removal of the whole legumin sequence could not result in complementation of the incomplete genomic JM83 lac Z gene. What seems more likely is that the large scale preparation of JY8 was contaminated with a quantity of uncut pUC vector too small to be directly detected on a gel, but that would on transformation amplify, and by virtue of its smaller size and faster replication times outcompete the larger, slower JY8 plasmids. This process would become more noticable with time, explaining the observation of cells grown on plates containing X gal becoming gradually more pigmented

with time. This does not explain the phenomenon completely, but seems more likely than the instability of the legumin plasmid JY8. It was observed by another worker (L. Briggs, pers. comm.) using the same overnight culture source of JM83 for competent cells (these were systematically replaced after two transformation) that they were contaminated with amp.resistant E. coli cells. This could have lead to a number of non pigmented false positives transferred to the nitrocellulose discs, which would account for the observations drawn above with regard to non hybridising colonies.

After a total of five attempts (twice by single and thrice by double colony hybridisation screen) and the screening of around 600 colonies with no success, the 'shotgun' cloning method was dropped, and an other more conventional cloning stratagy adopted for further work.

In one last attempt to isolate what has been demonstrated must be a very low frequency event, it was decided to screen the whole transformation mixture from a ligation of the Pst I digestion of JY8 and isolated BnA. Isolating the Brazil nut DNA fragment enabled the whole population of cells to be screened in situ with one probe (P32 radiolabelled BnA), simplifing matters considerably, as only colonies that contain BnA in the legumin vector should hybridise with the probe.

Figure 11 demonstrates the results of the in - situ hybridisation (scaled down). Eighteen positives were detected. There were some problems in relating the dots on the autoradiograph to the colonies on the master plate, due mainly to 'wrinkling' of the nitrocellulose in the transfer of

cells from the master plate onto the filter. The difficulty was somewhat lessened by the presence of five blue colonies (contaminants, detected by use inclusion of X gal in the agar that may have possibly arisen by the process described above) which were used as fixed points for aligning those positives identified on the nitrocellulose filter. Perhaps future in - situ hybridisations might be deliberately contaminated with cells transformed with uncut pUC vector as they seem ideally suited as a marker colonies!.

The difficulty of alligning all those colonies that hybrdised the probe can be seen displayed in their plasmid preps (figure 13). Two samples, A4 (track m) and A6 (track l) exhibit an extra band when digested with Pvu II. One band aligns with the legumin control (track b) and the 2.0kb fragment from the lambda Hind III digestion. The extra restriction band occurs approximately halfway between the 2.0kb and 2.3kb fragments of the lambda digest. More detailed restriction analysis revealed the fragment to be around 2.15kb, and would therefore seem likely to represent the legumin - BnA construct present in mixed colony JY8 legumin plasmid. This is confirmed by the results of single cells taken from streaks of A4, A6 and B5, B7 and B10 (also isolated as mixed cultures), as shown in figure 14. The confirmation of the detection was completed by the release of a fragment calculated to be around 160bp on digestion with with Pst I.

Figure 15 shows the digestion of the BnA - legumin constructs A44, A45, A63, and A71 with Dde I in an attempt to determine the orientation of BnA. The pattern of fragments should be different in the case of the two possible orientations of the

BnA fragment within the legumin DNA (figure 16). Restriction fragments of 0.16 and 0.20 kb, that would correspond to a positive identification of BnA in the correct orientation were detected, but unfortunately fragments of these size are produced in the digestion anyway. However, resriction fragments that would be produced by BnA in the wrong orientation of 0.25 and 1.1 kb are clearly not present. This suggests that all constructs, deriving from original colonies A4, A6 and A7 are the same (i.e. all BnA - Legumin plasmid construction pGPB1). containing the Brazil nut fragment BnA in the correct orientatation. This result has been confirmed by southern blot (appendix I, figure 17).

4.3.3 Imperfections in gel photographs

Many of the photographs reproduced in this report show evidence of partially digested DNA in tracks other than the 'no enzyme' control track. Their presence is due to either or a combination of two things; firstly too much DNA may have been added (a common fault with mini-preps as an improvement in recover will result in using more DNA than intended) or factors are carried through from mini preps that effect the ability of the enzyme to digest the DNA. The latter was certainly a problem in early work due to inexperience at the techniques involved, the former tending to be the problem in more recent times. Some photo's also show clouds of tRNA where addition of RNAase had been ommitted. These imperfections are regretable, and I would very much of liked to have time at the end of the project to repeat the early experiments.

4.4 FUTURE WORK

4.4.1 Work that would have been done

This dissertation started with two main objectives: Firstly to mutate a legumin sequence with a region of Brazil nut DNA, and secondly,to demonstrate expression of the modified legumin in Yeast. The first has been largely achieved, though given a little more time I would have wished to check the sequence of the construct by dideoxy-sequencing (Sanger et al. 1977) if time was limited, I would attempt to repeat the work, mutating the sequence encoding the storage protein vicilin. Given another six months, I might attempt the second objective of this project as follows. I would first need to produce enough of the protein to enable it to be characterised. This could possibly be achieved by expression of the constructs in a cell - free *in vitro* translation system, such as the rabbit reticulolysate or wheat germ systems. Once obtained, I would then need to be purify by a combination of biochemical techniques, then characterise it to determine most importantly, its solubility. If the mutations appeared not to have deleteriously altered the function of the protein, the constructs would be linkered, cloned into a suitable yeast expression vector and used to transform *S. cerevisiae*. It is unlikely that I would have the time to develop a system to secrete the modified protein, so that its production might be studied, possibly quantitatively, *in vitro*. In the absence of such a system, I would either look for the production of protein *in vivo* by use of a suitable antibody probe (the gold labelled goat anti rabbit/rabbit anti legumin would aggregate

to legumin depositions, the electron dense gold atoms being clearly visable by transmission electrom microscopy) or by harvesting the cells, purifing and analysis by northern and western blots. Finally, the protein expressed by the yeast could be thoroughly characterised to establish if yeast translation systems are suitable for reasonable expression of mutated eukaryotic protein.

It would be easy to suggest that the second objective of this project might form the basis of a future M.Sc. dissertation.

4.4.2 Advanced research with modified storage proteins

As it stands where I have left it, this project has been a good academic and technical exercise in the fundamentals of practical molecular biology. If the work was to continue, and expression of modified storage proteins was demonstrated in yeast, significant progress could be claimed in the manipulation of yeast expression systems, and a useful route for the study of storage protein engineering would be paved. Success would enable one to justify a commitment to the introduction of modified storage protein sequences into higher plants, and such work may play a significant role in the future of plant molecular biology.

CHAPTER FIVE

BIBLIOGRAPHY

5.1 REFERENCES

AMPE, C., van DAMME, J., de CASTRO, L. A. B., SAMPAIO, M. J. A. M., van MONTAGU, M., and VANDEKERCKHOVE, J. (1986). Primary structures of the 2s storage proteins of Brazil nut. (submitted for publication).

BIRNBOIM, H. C., and DOLY, J. (1979). Nucleic Acids Res. 7, 1513.

BRAWERMAN, G., MENDECKI, J., and LEE, S. (1972). Biochemistry 11, 637.

BROWN, J. W. S., BLISS, F. A., and HALL, T. C. (1981). Theoret. Appl. Genet. 60, 251-9.

BURR, H. K. (1975). in Protein Nutritional Quality of Foods and Feeds, Vol. I ed. M. Freidman, pp. 119-34. Marcel Dekker, New York.

CASEY, R., and DOMONEY, C. (1984). Phil. Trans. R. Soc. Lond. [Biol] 304, 349-58.

CROY, R. R. D. (1977). Ph.D. thesis, University of Aberdeen.

CROY, R. R. D., HOQUE, M. S., GATEHOUSE, J. A., and BOULTER, D. (1984). Biochem. J. 218, 795-803.

CROY, R. R. D., and GATEHOUSE, J. A. (1985). in PLANT GENETIC ENGINEERING, ed. J. H. Dodds, pp. 201-234. Cambridge University Press.

DELAUNEY, A. J. (1984). Ph.D. thesis, University of Durham.

DERBYSHIRE, E., WRIGHT, D. J., and BOULTER, D. (1976). Phytochemistry 15, 3-24.

ERSLAND, D. R., BROWN, J. W. S., CASEY, R., and HALL, T. C., (1983). in Seed Proteins: Biochemistry, Genetics, Nutritive Value, ed. W. Gottschalk and H. P. Muller, pp. 355-376, Martinus Nijhoff, The Hauge.

GATEHOUSE, J. A., CROY, R. R. D., and BOULTER, D. (1984). in Critical Reviews in Plant Science, Vol. I, pp. 287-314. CRC Press, Boca Raton, Fla.

GATEHOUSE, J. A., GILROY, J., HOGUE, M. S., and CROY, R. R. D. (1985). Biochem. J. 225, 239-247.

GAIT, M. J. (1984). in Oligonucleotide Synthersis, ed. M. J. Gait, pp. 6-7. IRL Press, Oxford, Washington DC.

GILLAM, S., ASTELL, C. R., and SMITH, M. (1980). Gene 12, 129-137.

GRUNSTEIN, M., and HOGNESS, D., (1975). Proc. Natl. Acad. Sci. 72, 3961.

HANAHAN, D., and MESELSON, M. (1980). Gene 10, 63.

HELLING, R. B., GOODMAN, H. M., and BOWYER, H. W., (1974). J. Viriol. 14, 1235.

HIGGINS, T. J. V., (1984). Ann. Revue. Plant. Physiol. 35, 191-221.

MANDEL, M., and HIGA, A. (1970). J. Mol. Biol. 53, 154.

MANIATIS, T., JEFFERY, A., and KLEID, D. G. (1975). Proc. Natl. Acad. Sci. 72, 1184.

MANIATIS, T., FRITCH, E. F., and SAMBROOK, J. (1982). Molecular Cloning, ed. T. Maniatis, E. F. Fritch and J. Sambrook. Cold Spring Harbor Laboratory, New York.

MANTELL, S. H., MATTHEWS, J. A., and McKEE, R. A. (1985). in Principles of Plant Biotechnology, ed. S. H. Mantell, J. A. Matthews and R. A. McKee. Blackwell Scientific Publications, Oxford.

OLD, R. W., and PRIMROSE, S. B. (1985). in Principals of Gene manipulation, ed. R. W. Old and S. B. Primrose. Blackwell Scientific Publications, Oxford.

PAYNE, P. I. (1983). in Seed Proteins, ed. J. Daussant, J. Mosse, and J. Vaughan, pp. 223-2253. Academic Press, New York.

PAYNE, P. I., and RHODES, A. P. (1982). in Encyclopedia of Plant Physiology, Vol 14a, ed. D. Boulter and B. Parthier, Springer-Verlag, Birlin.

RIGBY, P. W. J., DICCKERMANN, M., RHODES, C., and BERG, P. (1977). J. Mol. Biol. 113, 237.

SANGER, F., NICKLEN, S., and COULSEN, A. R. (1977). Proc. Natl. Acad. Sci. 74, 5463.

SHARP, P. A., SUGDEN, B., and SAMBROOK, J. (1973). Biochemistry 12, 3055.

SHAW, C. H. (1984). Chemistry and Industry, 3 Dec, 817-824.

SOUTHERN, E. (1975). J. Mol. Biol. 98, 503.

SPROAT, B. S., and GAIT, M. J., (1984). in Oligonucleotide Synthesis, ed. M. J. Gait, pp 83-115. IRL Press, Oxford, Washington DC.

STRUHL, K., CAMERON, J. R., and DAVIS, R. W. (1976). Escherichia coli. Proc. Natl. Acad. Sci. USA 73, 1471-1475.

VIEIRA, J., and MESSING, J. (1982). Gene 19, 259-268.

CHAPTER SIX

APPENDICIES

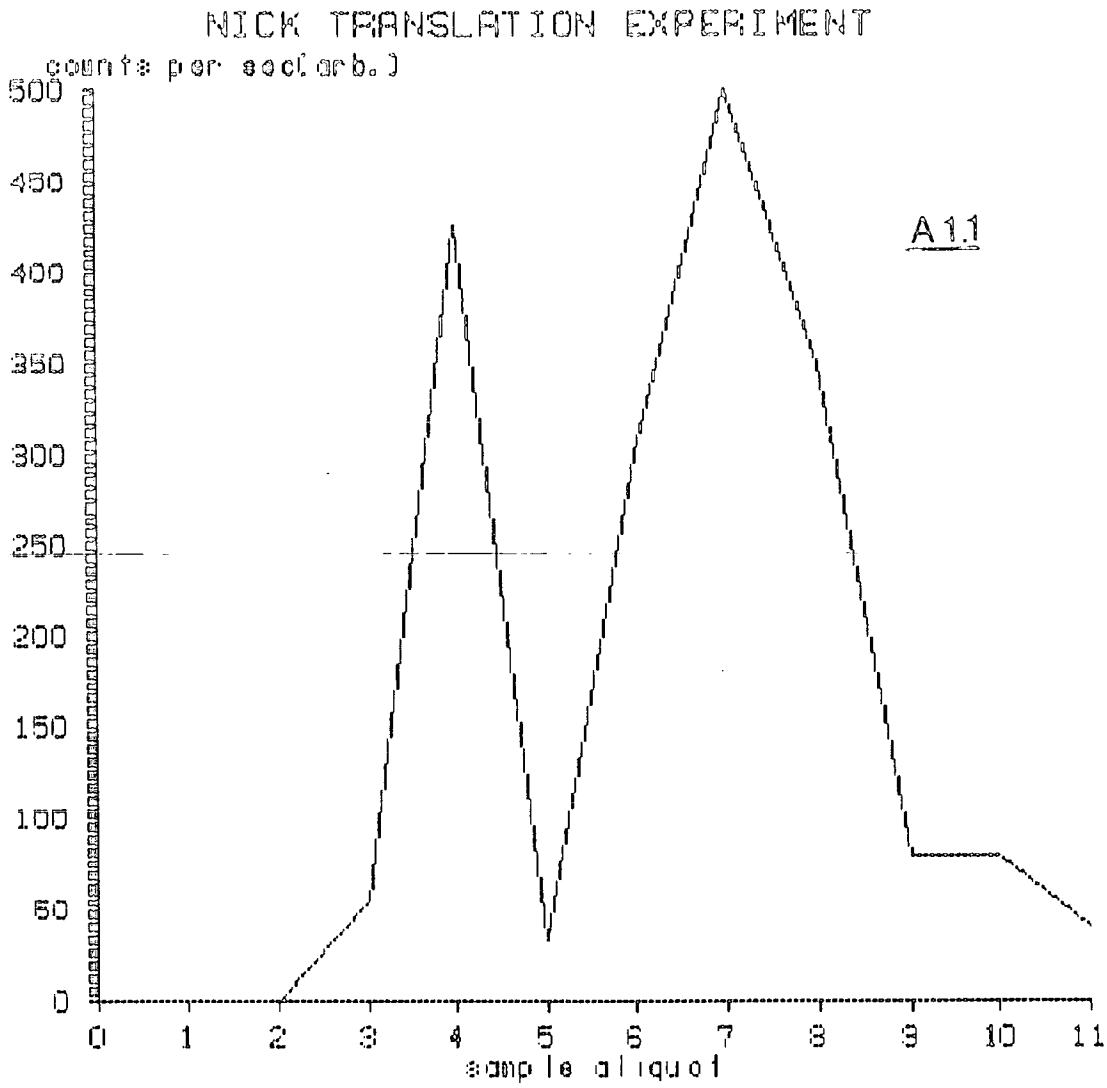
APPENDIX I

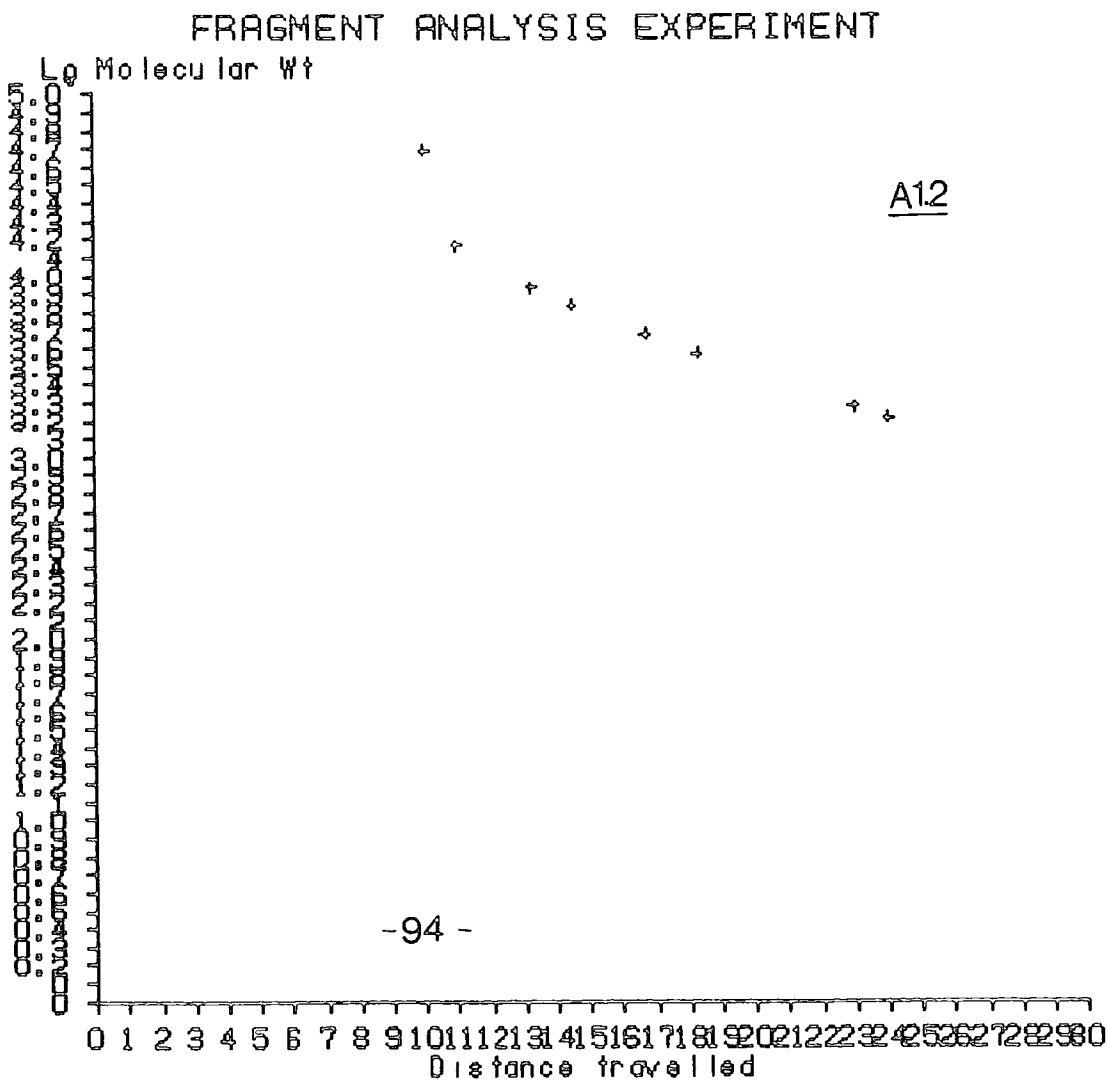
A 1.1 (figure 5) Shows an example of a nick translation plot. The first peak is for the probe, the second unincorporated radio labelled nucleotides.

A 1.2 (figure 10) Shows a lambda - Hind III digest restriction analysis calibration curve.

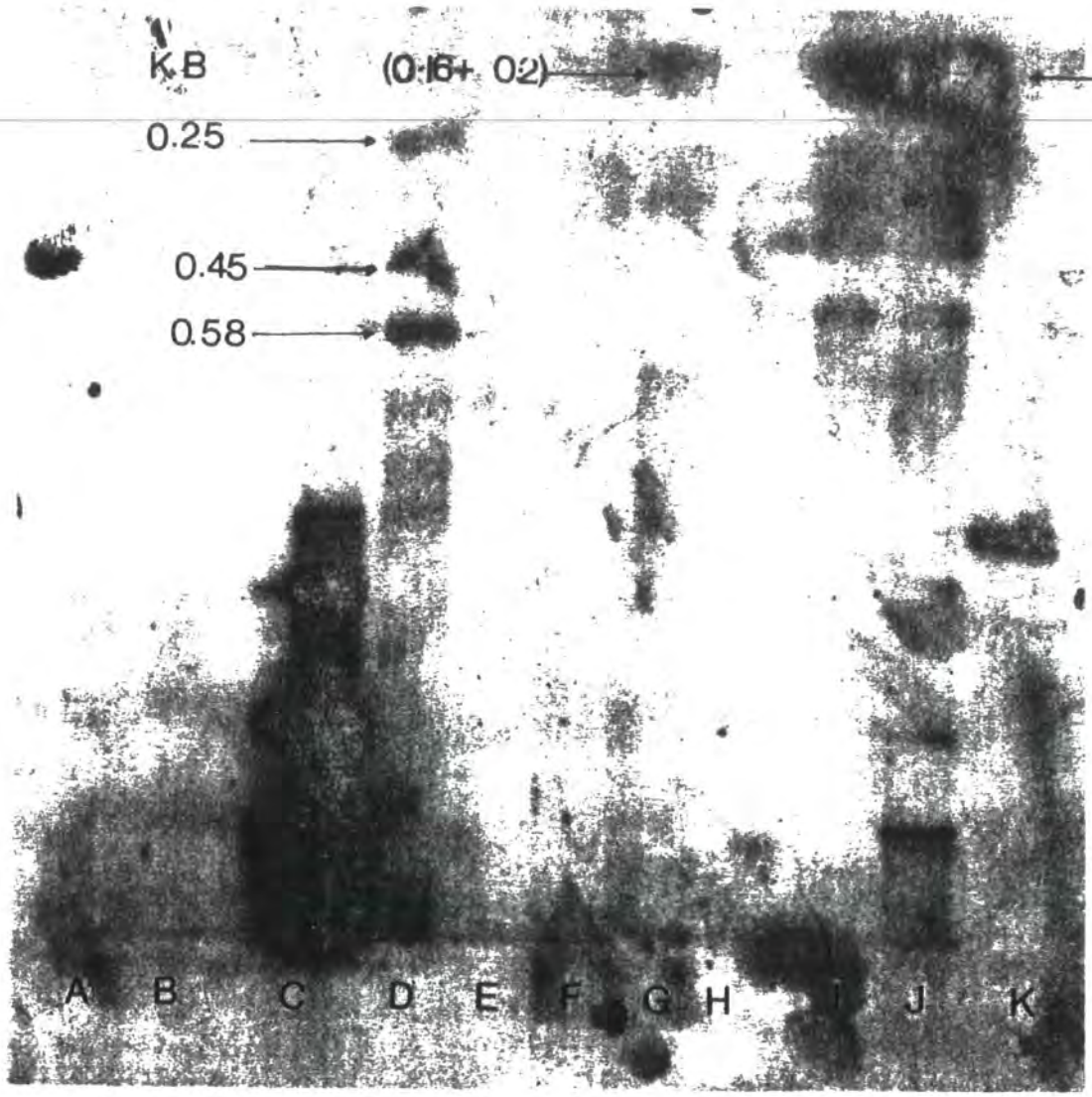
A 1.3 (figure 17) A Southern hybridisation, confirming the oreintation of pGPB1.

A 1.4 (figure 18) Withdrawn.


NICK TRANSLATION EXPERIMENT
counts per sec(arb.)
500
450
400
350
300
250
200
150
100
50
0
A1.1
0 1 2 3 4 5 6 7 8 9 10 11
sample aliquot


FRAGMENT ANALYSIS EXPERIMENT
Lg Molecular Wt
A1.2
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Distance travelled

A 1.3



A 1.4 Figure withdrawn.

A 2. Sequence data DNA clone encoding 2s protein of Brazil nut

A 2.1 Bn2s sequence

A 2.2 Translation of Bn2s sequence

A 2.31 and 2.32 Restriction analysis of Bn2s sequence

A 2.4 resriction map of Bn2s sequence

bn2s.SEQ is 304 base pairs long. A2.1
The formated sequence is 304 base pairs long.
Starting at 1 and ending at 304.

```
        10          20          30          40          50          60
GGGGATGCAG AGACAGCAGA TGCTCAGCCA CTGCCGGATG TACATGAGAC AGATGATGAA

        70          80          90          100         110         120
GGAGAGCCCG TACCAGACCA TGCCCAGGCG GGGAATGGAG CCGCACATGA GCGAGTGCTG

        130         140         150         160         170         180
CGAGCAACTG GAGGGGATGG ACGAGAGCTG CAGATGCGAA GGCTTAAGGA TGATGATGAG

        190         200         210         220         230         240
GATGATGCAA CAGCAGGAGA TGCAACCCCG AGGGGAGCAG ATGCGAATGA TGATGAGGAT

        250         260         270         280         290         300
GGCCGAGAAT CTCCCTTCCC GCTGCAACCT CAGTCCCCAG AGATGCCCCA TGGGCGGCTC

        310         320         330         340         350         360
CCC
```

B:bn2s.SEQ is 304 base pairs long. A2.2
Translation from 1 to 303.

```
     G  D  A  E  T  A  D  A  Q  P  L  P  D  V  H  E  T  D  D  E
      G  M  Q  R  Q  Q  M  L  S  H  C  R  M  Y  M  R  Q  M  M  K
       G  C  R  D  S  R  C  S  A  T  A  G  C  T Stp D  R StpStp R
1    GGGGATGCAGAGACAGCAGATGCTCAGCCACTGCCGGATGTACATGAGACAGATGATGAA

      G  E  P  V  P  D  H  A  Q  A  G  N  G  A  A  H  E  R  V  L
       E  S  P  Y  Q  T  M  P  R  R  G  M  E  P  H  M  S  E  C  C
        R  A  R  T  R  P  C  P  G  G  E  W  S  R  T Stp A  S  A  A
61   GGAGAGCCCGTACCAGACCATGCCCAGGCGGGGAATGGAGCCGCACATGAGCGAGTGCTG

      R  A  T  G  G  D  G  R  E  L  Q  M  R  R  L  K  D  D  D  E
       E  Q  L  E  G  M  D  E  S  C  R  C  E  G  L  R  M  M  M  R
        S  N  W  R  G  W  T  R  A  A  D  A  K  A Stp G StpStpStp G
121  CGAGCAACTGGAGGGGATGGACGAGAGCTGCAGATGCGAAGGCTTAAGGATGATGATGAG

      D  D  A  T  A  G  D  A  T  P  R  G  A  D  A  N  D  D  E  D
       M  M  Q  Q  Q  E  M  Q  P  R  G  E  Q  M  R  M  M  M  R  M
        Stp C  N  S  R  R  C  N  P  E  G  S  R  C  E StpStpStp G  W
181  GATGATGCAACAGCAGGAGATGCAACCCCGAGGGGAGCAGATGCGAATGATGATGAGGAT

      G  R  E  S  P  F  P  L  Q  P  Q  S  P  E  M  P  H  G  R  L
       A  E  N  L  P  S  R  C  N  L  S  P  Q  R  C  P  M  G  G  S
        P  R  I  S  L  P  A  A  T  S  V  P  R  D  A  P  W  A  A  P
241  GGCCGAGAATCTCCCTTCCCGCTGCAACCTCAGTCCCCAGAGATGCCCCATGGGCGGCTC

      P

301  CCC
```

Restriction enzyme is Alu I
AGCT 1 1 match found
145 159

Restriction enzyme is Ava I
CCCGAG 1 CTCGAG 0 CCCGGG 0 CTCGGG 0 1 match found
206 98

Restriction enzyme is Ban II
GGGCCC 0 GGGCTC 0 GAGCCC 1 GAGCTC 0 1 match found
63 241

Restriction enzyme is Bgl I
GCC------GGC 1 1 match found
284 20

Restriction enzyme is Bst NI
CCAGG 1 CCTGG 0 1 match found
83 221

Restriction enzyme is Dde I
CT-AG 2 2 matches found
22 246
268 36

Restriction enzyme is Fnu4 HI
GC-GC 5 5 matches found
99 17
116 30
146 114
260 33
293 11

Restriction enzyme is Foc I
GGATG 6 CATCC 0 6 matches found
2 33
35 99
134 33
167 12
179 57
236 68

Restriction enzyme is Hae III
GGCC 1 1 match found
240 64

Restriction enzyme is Hinf I
GA-TC 1 1 match found
246 58

Restriction enzyme is Hpa II
CCGG 1 1 match found
33 271

Restriction enzyme is Mnl I
CCTC 1 GAGG 4 5 matches found

130	47
177	32
209	25
234	33
267	37

Restriction enzyme is Nco I
CCATGG 1 1 match found

287	17

Restriction enzyme is Pst I
CTGCAG 1 1 match found

147	157

Restriction enzyme is Rsa I
GTAC 2 2 matches found

39	30
69	235

Restriction enzyme is Scrf I
CC-GG 1 1 match found

83	221

Restriction enzyme is Sfan I
GATGC 7 GCATC 0 7 matches found

3	15
18	134
152	31
183	15
198	21
219	62
281	23

Restriction enzyme is Nla III
CATG 4 4 matches found

42	36
78	27
105	183
288	16

Restriction enzyme is Nla IV
GG--CC 2 2 matches found

96	199
295	9

A24

bn25
0
AluI
AvaI
BanI
BglI
BstNI
DdeI
Fnu4HI
FocI
HaeIII
HinfI
HpaII
MnlI
NcoI
PstI
RsaI
ScrFI
SfaNI
NlaIII
NlaIV

A 3 Sequence data for legumin clone pJY8

A 3.1 Sequence translation of pJY8

A 3.2 Restriction analysis of legumin A gene

SEQUENCE TRANSLATION PROGRAMME

Name of file - GBLEG
Comment - J. Yarwood legumin construct
Sequence is 1693 bases in length

A3.1

Sequenced strand

```
AAGCTTCTTGCACTTTCTCTTTCATTCTGTTTTCTACTTTTGGGCGGCTGTTTTGCTTTGAGAGAAGAGCCACAGCAAAATGAGTGCCAGCTAGAACGCCTCGATGCCCTCGAGCCTGAT
K L L A L S L S F C F L L L G G C F A L R E G P G G N E C G L E R L D A L E P D
 S F L H F L F H S V F Y F M A A V L L I E N S H S K H S A S I N A S M P S S L I
  A S C T F S F I L F S T F G R L F C F E R T A T A H I V P A R T P R C P R A I I

AACCGTATAGAATCGGAAGGTGGGCTCATTGAGACTTGCAATCCCAACAACAAGCAATTCCGATGTGCTGGTGTGGCAATCTCTCGTGCTACCCTTCAACGCAACGCCCTTCGCAGACCT
N R I E S E G G L I E T H N P N N K G F R C A G V A I S R A T L G R N A L R R P
 T V I N R K V G S L R L G I P T T S N S D V L V M G S L V L P F N A T P F A D L
  P Y R I G R M A H I D L E S G G G A I P M C N C G N L S C Y P S T G R P S G T L

TACTACTCCAATGCTCCCCAAGAAATTTTCATCCAACAAGGTAATGGATATTTTGGCATGGTATTCCCCGGTTGTCCTGAGACCTTTGAAGAGCCACAAGAATCTGAACAAGGAGAGGGA
Y--Y S--N--A P G E- I F I--G G G N G Y F G H V F P G C P E T F E E P G E S E G G E G
 T T P M L P K K F S S N K V M D I L A M Y S P V V L R P L K S H K N L N K E R D
  L L G C S P R N F M P T R I H I F M M G I P R L S I D L I R A T R I I T R R G T

CGCAGGTACAGAGACAGACATCAAAAGGTTAACCGATTCAGAGAGGGTGATATCATTGCAGTTCCTACTGGTATTGTATTTTGGATGTACAACGACCAAGACACTCCAGTTATTGCCGTC
R R Y R D R H G K V N R F R E G D I I A V P T G I V F W M Y N D G D T P V I A V
 A G T E T D I K R L T D S E R V I S L Q F L L V L Y F G C T T T K T L Q L L P S
  G V G R G T S K G I P I G R G I Y H C S S Y M Y C I L D V G R P R H S S Y C R L

TCTCTTACTGACATTAGAAGCTCCAATAACCAGCTTGATCAGATGCCTAGGAGATTCTATCTTGCTGGGAACCACGAGCAAGAGTTTCTACAATACCAGCATCAACAAGGAGGAAAGCAA
S L T D I R S S N N Q L D Q M P R R F Y L A G N H E G E F L Q Y G H Q Q G G K Q
 L L L T L E A P I T S L I R C L G D S I L L G T T S K S F Y N T S I N K E E S K
  S Y I H I K L Q I P A I S D A I E I L S C M E P R A R V S T I P A S T R R K A R

GAACAAGAAAATGAAGGCAACAACATTTTCAGTGGCTTCAAGAGGGATTACTTGGAAGATGCTTTCAACGTGAACAGGCATATAGTAGACAGACTTCAAGGCAGGAATGAAGACGAAGAG
E Q E N E G N N I F S G F K R D Y L E D A F N V N R H I V D R L Q G R N E D E E
 N K K M K A T T F S V A S R G I T M K M L S T I T G I I I T D F K A G M K T K R
  T R K I R Q Q H F Q M L Q E G L L G R C F Q R E Q A Y S R Q T S R Q E I R R R E

AAGGGAGCCATTGTCAAAGTGAAAGGTGGACTCAGCATCATAAGCCCACCCGAGAAGCAAGCGCGCCACCAGAGAGGCAGCAGACAAGAGGAAGATGAAGATGAAGAAGCAGCCGCGC
K G A I V K V K G G L S I I S P P E K Q A R M Q R G S R Q E E D E D E E K Q P R
 R E P L S K I K V D S A S I A M P R S K R A T R E A A D K R K M K M K R S S R A
  G S M C Q S E R M T Q M M K P T R E A S A P P E R Q Q T R G R I R I R E A A A P

CACCAGAGAGGCAGCAGACAAGAGGAAGAGGAAGATGAAGATGAAGAGAGGCAGCCGCGTCATCAAAGGAGAAGAGGAGAGGAGGAAGAAGAAGACAAGAAAGAGCGCGGCGGCAGCCAA
H Q R G S R Q E E E E D E D E E R Q P R M Q R R R G E E E E E D K K E R G G S Q
 T R E A A D K R K R K M K M K R G S R V I K G E E E R R K K K T R K S A A A A K
  P E R Q Q T R G R G R I R I R E A A A S S K E K R R G G R R R Q E R A R R Q P K

AAAGGCAAAAGCAGAAGGCAAGGAGACAATGGGCTTGAGGAAACAGTTTGCACTGCTAAACTTCGATTGAACATTGGCCCGTCTTCATCACCAGACATCTACAACCCTGAAGCTGGTAGA
K G K S R R Q G D M G L E E T V C T A K L R L N I G P S S S P D I Y M P E A G R
 K A K A E G K E T M G L R K Q F A L L N F D I T L A R L H H Q T S T T L K L V E
  R Q K Q K A R R Q M A I G M S L H C I T S I E H M P V F I T R H L Q P I S M I M

ATCAAAACTGTTACCAGCCTGGACCTCCCAGTTCTCAGGTGGCTCAAACTAAGTGCTGAGCATGGATCTCTCCACAAAAATGCTATGTTTGTGCCTCACTACAACCTGAATGCAAACAGT
I K T V T S L D L P V L R M L K L S A E H G S L H K N A M F V P H Y N L N A N S
 S K L L P A M T S Q F S G G S M I V L S M D L S T K M L C L C L T T T I M Q T V
  Q N C Y Q P G P P S S Q V A Q T K C I A M I S P Q K C Y V C A S L Q P E C K Q Y

ATAATATACGCATTGAAGGGACGTGCAAGGCTACAAGTAGTGAACTGCAATGGCAACACCGTGTTTGATGGAAAGCTAGAAGCCGGACGTGCATTGACAGTGCCACAAAACTATGCTGTG
I I Y A L K G R A R L Q V V N C N G N T V F D G K L E A G R A L T V P Q N Y A V
 I Y T H I R D V Q G Y K I I T A M A T P C L M E S I K P D V H I Q C H K T M L M
  N I R I E G T C K A T S S E L Q M Q H R V I M K A R S R T C I D S A T K L C C G

GCTGCAAAGTCACTAAGCGACAGGTTCTCATATGTAGCATTCAAGACCAATGATAGAGCTGGTATTGCAAGACTTGCAGGGACATCATCAGTTATAAATAATCTGCCGTTGGATGTGGTT
A A K S L S D R F S Y V A F K T N D R A G I A R L A G T S S V I N N L P L D V V
 L Q S H I A T G S H M M I H S R P M I E L V L Q D L Q G H H Q L I I C R M M M L
  C K V T K R Q V L I C S I Q D Q I I S M Y C K T C R D I I S Y K I S A V G C G C

GCAGCTACATTCAACCTGCAGAGGAATGAGGCAAGGCAGCTCAAGTCCAACAATCCCTTCAAATTTCTAGTTCCAGCTCGTCAGTCTGAGAACAGAGCTTCGGCTTAGATTTCGCACCAA
A A T F N L Q R N E A R Q L K S N N P F K F L V P A R Q S E N R A S A I I S M Q
 Q L H S T C R G M R Q G S S S P T I P S N F I F Q L V S L R T E L R L R F R T K
  S Y I Q P A E E I G K A A Q V Q Q S L Q I S S S S S S S V I E Q S F G L D F A P N

ATCAATGAAAGTAATGAATAAGAAAACTAAGGCTTAGATGCCTTTGTTACTTGTGTAAAATAACTCGAGTCATGTACCTTTTTGCGGAAACAGAATAAATAAAAGGTAAAATTTCAGTGC
I N E S N E I E N I G L D A F V T C V K I L E S C T F L R K Q N K I K V K F Q C
 S M K V M N K K T K A I M P L L L V I N N S S H V P F C G N R I N K R I N F S A
  Q I K I I I R K L R L R C L C Y L C K I T R V M Y L F A E T E I I K G K I S V L

TCTAAAAAAAAAA
S K K K
 L K K K
  I K K
```

Restriction enzyme is Acc I
GTAGAC 1 GTCTAC 0 GTATAC 0 GTCCAC 0 1 match found
690

Restriction enzyme is Alu I
AGCT 11 11 matches found
7
94
804
917
1076
1279
1302
1448
1483
1520
1541

Restriction enzyme is Ava I
CCCGAG 1 CTCGAG 1 CCCGGG 0 CTCGGG 0 2 matches found
114
774

Restriction enzyme is Ava II
GGACC 1 GGTCC 0 1 match found
1106

Restriction enzyme is Ban II
GGGCCC 0 GGGCTC 1 GAGCCC 0 GAGCTC 0 1 match found
147

Restriction enzyme is Bsv I
GCAGC 7 7 matches found
802
836
856
896
958
1446
1481

Restriction enzyme is Bcl I
TGATCA 1 1 match found
521

Restriction enzyme is Bgl I
GCC-----GGC 1 1 match found
961

Restriction enzyme is Bst NI
CCAGG 0 CCTGG 1 1 match found
1103

Restriction enzyme is Dde I
CT-AG 9 9 matches found
4
322
756
1119
1134
1141
1338
1531
2549

Restriction enzyme is Eco RV
GATATC 1 1 match found
414

Restriction enzyme is FnuD I
[illegible] matches found
931

Restriction enzyme is Fnu4 HI
GC-GC 13 13 matches found
49
802
836
839
856
896
899
952
955
958
1326
1446
1481

Restriction enzyme is Fsc I
GGATG 2 CATCC 1 3 matches found
275
448
1436

Restriction enzyme is Hae III
GGCC 1 1 match found
1041

Restriction enzyme is Hga I
GACGC 1 GCGTC 1 2 matches found
364
902

Restriction enzyme is Hinc II
GTCAAC 0 GTTAAC 1 GTCGAC 0 GTTGAC 0 1 match found
393

Restriction enzyme is Hind III
AAGCTT 1 1 match found
6

Restriction enzyme is Hinf I
GA-TC 7 7 matches found
135
164
343
400
538
754
1084

Restriction enzyme is HinP I
GCGC 4 4 matches found
786
788
842
950

Restriction enzyme is Hpa I
GTTAAC 1 1 match found

Restriction enzyme is Hpa II
CCGG 2 2 matches found
313

Restriction enzyme is Kpn I
GGTAC 1 TGAGC 1 2 matches found
681
1023

Restriction enzyme is Mbo II
GAAGA 19 TCTTC 1 16 matches found
333
660
714
720
816
822
828
870
876
882
888
916
930
933
936
1047

Restriction enzyme is Mnl I
CCTC 4 GAGG 16 20 matches found
104
113
360
408
595
647
799
810
853
867
873
893
919
924
927
1002
1109
1179
1466
1473

Restriction enzyme is Nci I
CCCGG 1 CCGGG 0 1 match found
312

Restriction enzyme is Nde I
CATATG 1 1 match found
1354

Restriction enzyme is Pae7I
CTCGAG 1 1 match found
114

Restriction enzyme is Pst I
CTGCAG 1 1 match found
1461

Restriction enzyme is Rsa I
GTAC 2 2 matches found
371
452

Restriction enzyme is Sau 3A I
GATC 2 2 matches found
522
1150

Restriction enzyme is Sau96 I
GG-CC 2 2 matches found
1041
1106

Restriction enzyme is Scrf I
CC-GG 2 2 matches found
312
1103

Restriction enzyme is Sfan I
GATGC 3 GCATC 2 5 matches found
108
527
584
663
760

Restriction enzyme is Taq I
TCGA 3 3 matches found
106
115
1028

Restriction enzyme is Tha I
CGCG 4 4 matches found
787
841
901
951

Restriction enzyme is Tth II
CAAACA 1 CAAGCA 1 2 matches found
176
1198

Restriction enzyme is Xho I
CTCGAG 1 1 match found
114

Restriction enzyme is Nla III
CATG 2 2 matches found
302
1146

Restriction enzyme is Nla IV
GG--CC 2 2 matches found
353
729

Following sequences not found.

Aat II	Ahe II	Ahe III	Apa I	Bal I	Bam HI
BstE II	Bst XI	Bvu I	Cla I	Eco RI	Hae II
Hlu I	Mst I	Mst II	Nae I	Nar I	Nco I
Pvu II	Sac I	Sac II	Sal I	Sca I	Sfi I
Sst I	Sst II	Stu I	Tth I	Xba I	Xma I
Nsi I	Bbo I				

A 4 Sequence Data for pUC vectors

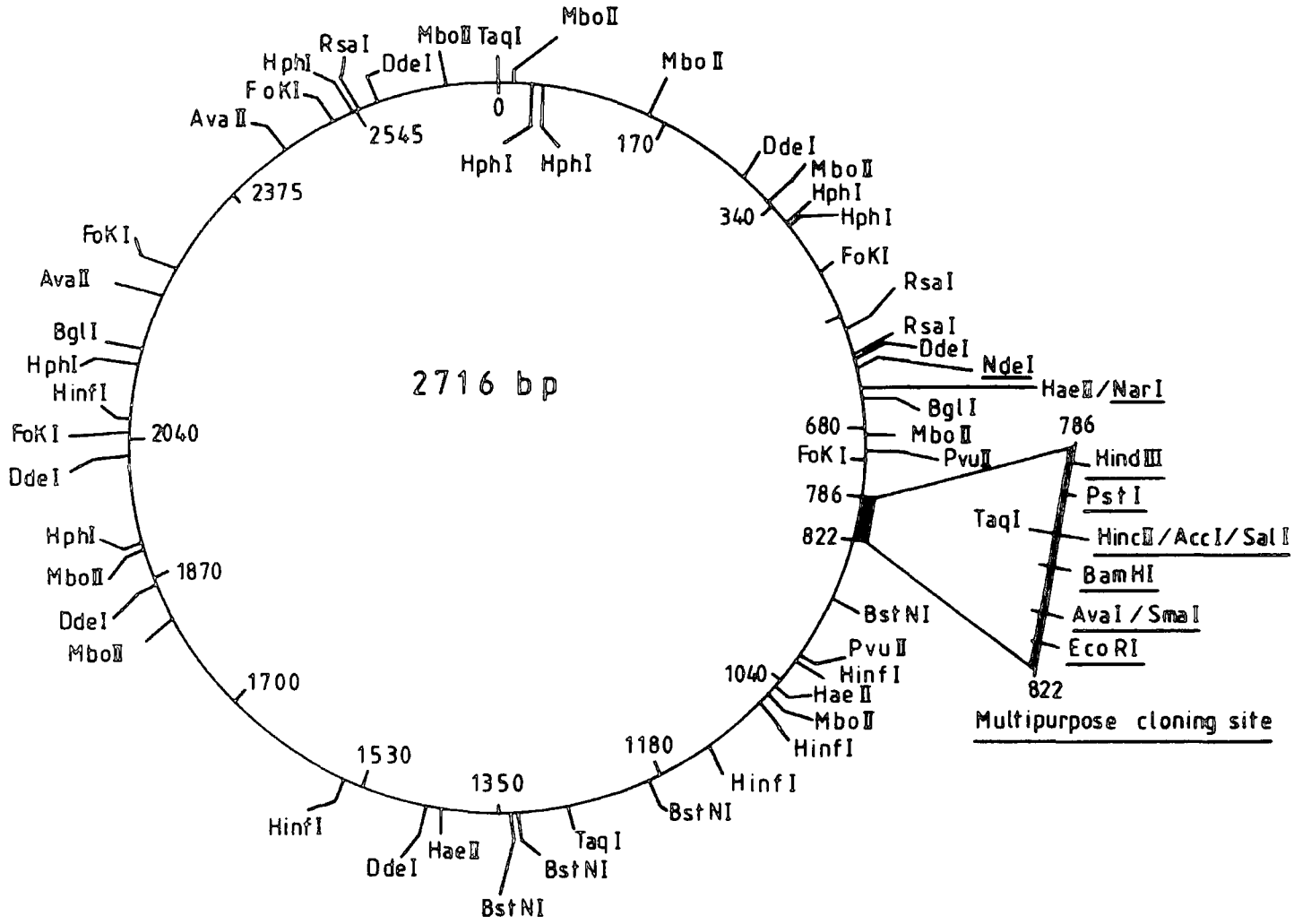
Note. The sequence of pUC8 is included here as data as detailed is not freely available.

A 4.1 Restriction map of pUC8

A 4.21 and 4.22 Sequence data for pUC8

A 4.3 Broad restriction analysis of pUC8

Restriction map of plasmid pUC8



pUC8

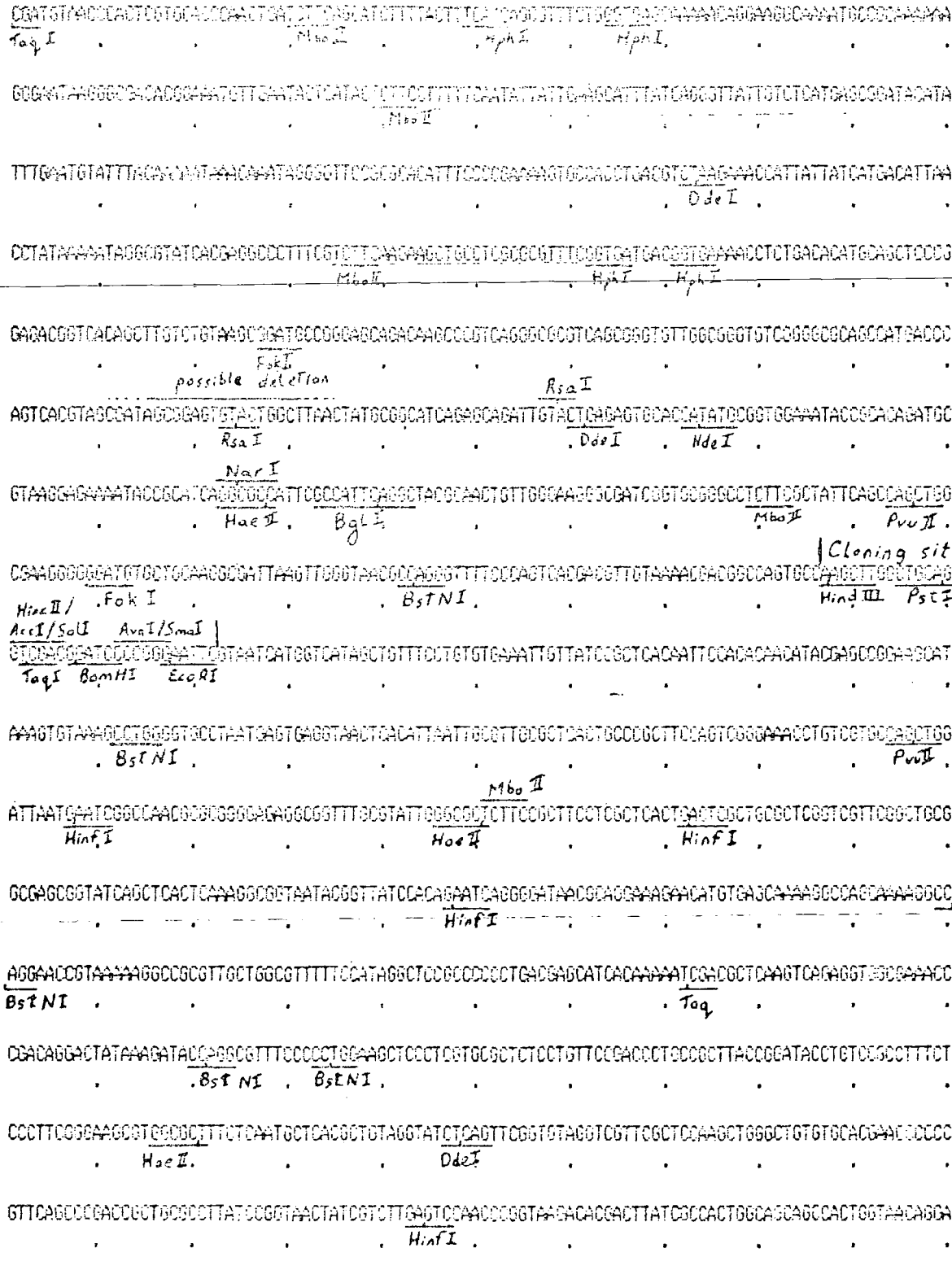


Fig.24. DNA Sequence and Restriction Endonuclease Recognition Sites of Plasmid pUC8

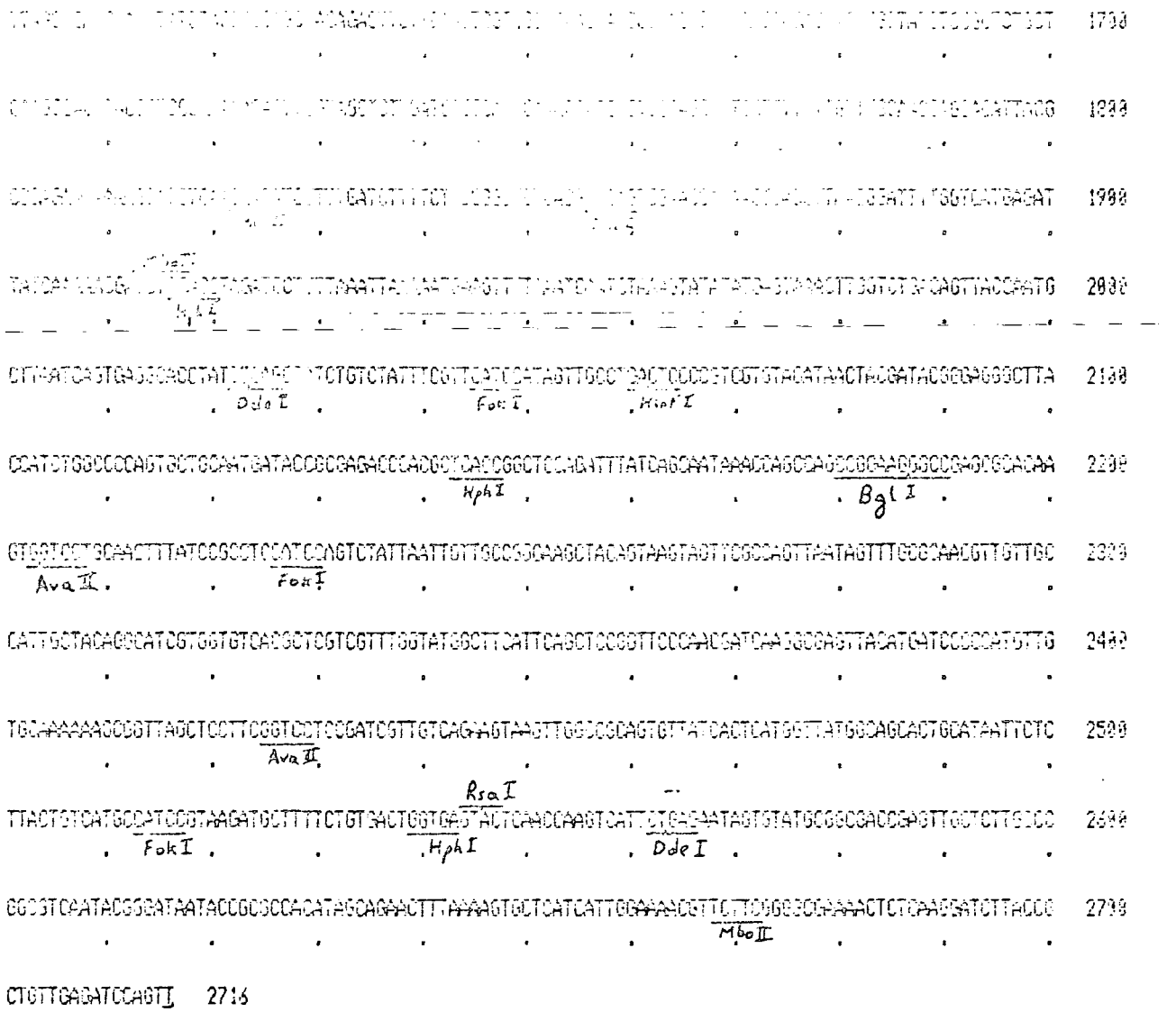


Figure 24 Continued

Table 4.1 A restriction map of plasmid pUC8

The multipurpose cloning site has unique restriction sites for:-

AccI, AvaI, BamHI, EcoRI, HincII, HindIII, PstI, SalI, SmaI

Restriction endonuclease	Number of sites	Position(s)	Fragment sizes* (base pairs)
AccI	1	~803	2716*
AluI	14X		
ApaI	0	-	-
ApyI	up to 4x		
AvaI	1	814	2716*
AvaII	2	2203, 2425	2494*, 222
BalI	0	-	-
BclI	0	-	-
BamHI	1	807	2716*
BcoII	0	-	-
BglI	2	630, 2186	1160, 1556*
BglII	0	-	-
BstEII	0	-	-
BstNI	5	743, 913, 1200, 1321, 1334	2125, 170*, 287, 121 23
ClaI	0	-	-
DdeI	6		
EcoRI	1	817	2716*
EcoRII	5	see BstNI	
EcoRV	0	-	-
FokI	5	430, 710, 2050, 2230, 2515	613, 280, 1340*, 180, 285
HaeII	3	627, 1050, 1420	1923, 423*, 370
HaeIII	17X		
HhaI	17X		
HincII	1	~803	2716*
HindIII	1	787	2716*
HinfI	5		
HpaI	0	-	-
HpaII	13X		
HphI	7		

Table 34 (Continued)

A432

Restriction endonuclease	Number of sites ‡	Position(s)	Fragment sizes* (base pairs)
KpnI	0	-	-
MboI	15X		
MboII	8		
MluI	0	-	-
NarI	1	624	2716*
NciI	6X		
NdeI	1	573	2716*
PstI	1	799	2716*
PvuII	2	696, 996	2416, 300*
RsaI ‡‡	3 or 2	524?, 558, 2545	695, 34?, 1987*
SalI	1	~803	2716*
Sau3A	15X		
Sau96I	6X		
SmaI	1	~814	2716*
SphI	0	-	-
SstI	0	-	-
SstII	0	-	-
StuI	0	-	-
TaqI	3	1, 802, 1272	1445, 801*, 470*
XbaI	0	-	-
XhoI	0	-	-

‡ X indicates these sites are not marked on the wheel map.

* Fragments marked thus contain all or part of the multipurpose cloning site.

‡‡ The RsaI site at position 524 may have been deleted in the construction of pUC8.

A 5Diagram appertaining to discussion.

A 5.1 Secondary structure predictions and hydrophilicity profiles for legumin and vicilin

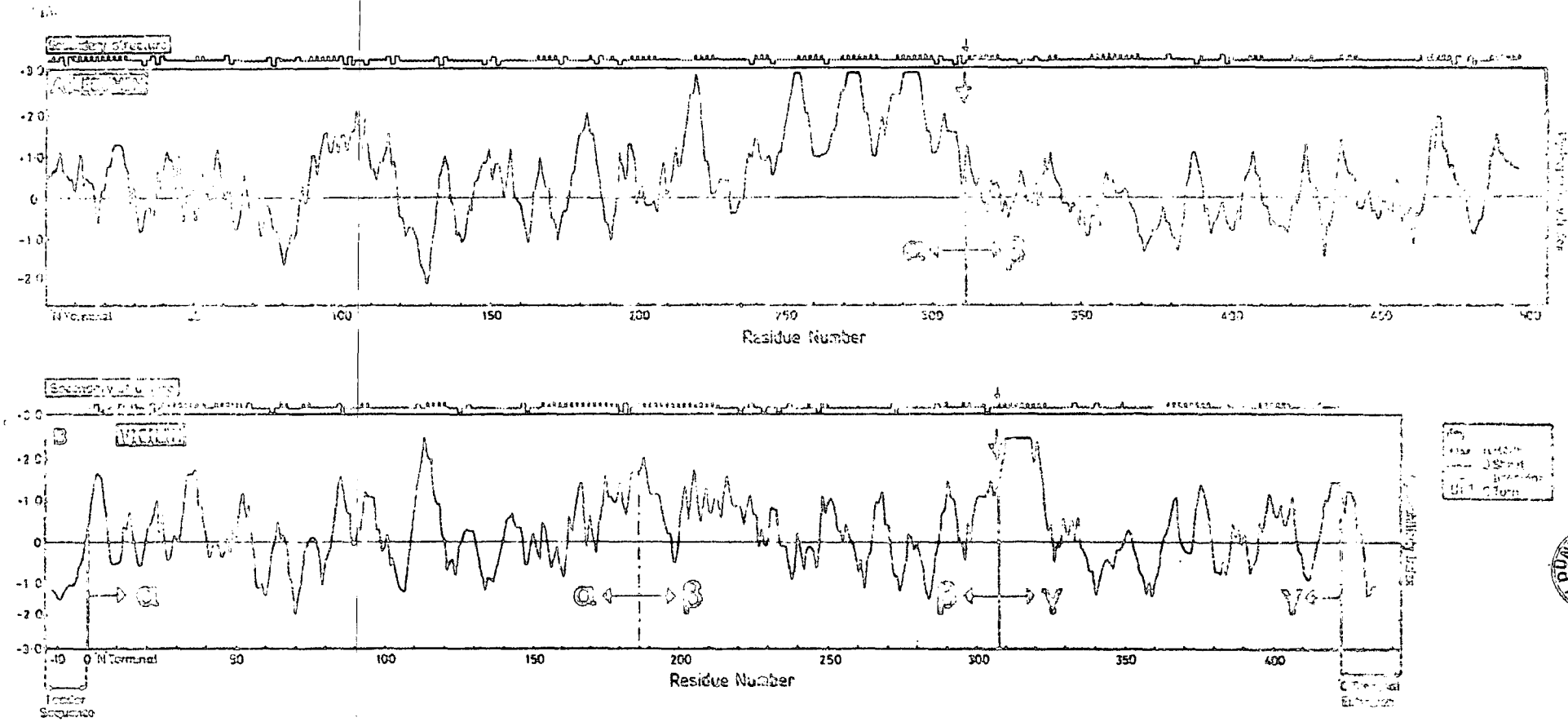


Figure . Secondary structure predictions and hydrophilicity index profiles for the amino acid sequences of (A) legumin and (B) vicilin, precursors predicted by cDNA nucleotide sequences. The computer program used for the secondary structure predictions was based on the analyses of Garnier et al. (1978). Hydrophilicity index profiles were constructed by a program adapted directly from that of Hopp and Wood (1981). The key to the predicted structural features is shown in the box inset.