Durham E-Theses

Cysteine proteinase inhibitors: purification and gene isolation

Fones, Angela Mei Ling

How to cite:

Use policy

CYSTEINE PROTEINASE INHIBITORS:

PURIFICATION AND GENE ISOLATION

Angela Fones Mei Ling

September 1989

A dissertation submitted in partial fulfilment

of requirements for the degree of MSc in

Biotechnology at the Department of Biological

Science, University of Durham, Durham.

Abstract

Genetic engineering has produced a transgenic insect-resistant plant which produces a serine protease inhibitor (Hilder et al., 1987). More insects can be targetted if a plant can be made to produce thiol protease inhibitors. Hence this project attempts to do some preliminary work - purification and gene isolation of thiol protease inhibitors, which is necessary before plants resistant to insects using thiol proteinases as their digestive enzymes can be created.

It is known that cowpea seeds contain thiol protease inhibitors. However, their protein sequences have never been elucidated. In the first part of the project, it was attempted to obtain a pure thiol protease inhibitor. This was so that its sequence could be determined partially to construct an oligo-nucleotide probe to pick the gene coding for it from a cDNA library.

In the second part of the project, it was attempted to isolate the gene coding for the bromelain inhibitor of pineapple, to clone it up, and to sequence it. Though the sequence of this inhibitor has been known for years, its gene has never been isolated or sequenced.

2

CONTENTS

ABBREVIATIONS

The abbreviations used throughout the dissertation were based on those recommended by the Biochemical Society in the Biochemical Journal "Policy of the Journal and Instruction to Authors", vol. 209, pp 1-27, 1983. Other notations used are listed below.

amp	ampicillin
BAPNA	Na-Benzoyl-DL-Arginine P- Nitroanilide
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
IPTG	Isopropyl B-D-thio-galactopyranoside
m.wt., mw	relative molecular weight
oligos	oligonucleotides
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
SDS	sodium dodecyl sulphate
SSC	saline sodium citrate
TEMED	N,N,N',N'-tetramethylethylemediamine
Tris, Tris.Cl	Tris (hydroxymethyl) methylamine
X-gal	5-bromo-4-chloro-3-indolyl-B-galactoside

CHAPTER I
INTRODUCTION

1: INTRODUCTION

1.1: Genetic Engineering to produce insect-resistant crops

Substantial crop losses occur each year due to attacks by insect pests. These are largely controlled by spraying crops with large quantities of agrochemicals, which is not only expensive but can also have detrimental effects on ecosystems. Many cultivars of crop species already exhibit resistance to pests to some extent, but pest evolution forces the plant breeder to continually produce new resistant varieties. Radical changes in the type of resistance conferred are required. Some possibilities have been proposed, involving transfer of genes or groups of genes encoding the following to crop plants; antimetabolite proteins such as lectins or enzyme inhibitors; viral coat proteins; the enzymes chitinase and B-glucanase, the cell wall protein extensin ; and bacterial insecticidal toxins, such as that produced by *Bacillus thuringiensis*. Expression of these products throughout the aerial parts of a plant may enhance resistance but must occur without detrimental effects upon the plant or to the consumer. One disadvantage of expressing an extra gene could be a drain of the plant's resources, hence decreasing its yield, thus one can try to put in a regulatable promoter which switches on the gene only at a certain stage of development when the plant is most susceptible or which expresses the foreign gene only in a certain susceptible organ. Another possibility is to have the gene triggered by a wound or by the infection itself.

A promising set of such genes is the wound-induced protease inhibitors. After certain plants are wounded mechanically or by a

chewing insect, "protease inhibitor inducing factor" (PIIF) is released and this induces the systemic synthesis of several protein protease inhibitors. These inhibitors are directed against insect and microbial proteases and hence protect the plant (Ryan, 1981, 1988).

The bacterium *B. thuringiensis* produces toxins that are toxic to specific insects, depending on the strain used. These bacteria have been used as biological insecticides for over a quarter of a century. The gene coding for its toxins has recently been introduced into tobacco (Vaek et al., 1987) and tomato (Fischoff, 1987). However, the limited range of pests against which it is toxic to is a great disadvantage.

A plant protein that can confer resistance to a wide range of pests is the cowpea trypsin inhibitor. As its metabolic target is the catalytic site of an enzyme, the ability of the insects to evolve a resistance mechanism based on mutation at this site would be minimal. Furthermore, cowpea seeds are not toxic to humans. It has been shown that when the gene encoding this inhibitor is transferred to tobacco, the plant shows enhanced resistance to its own insect pests (Hilder et al., 1987). Such "protease inhibitors" may be able to make plants fight off a much wider range of insects than do the toxins from *B. thuringiensis*. A big advantage of transgenic plants that make their own insecticides is that only insects that attack the plants are affected. Such transgenic plants are now a possibility, encompassing a wide range of plants, due to the rapid development of new techniques in genetic engineering.

1.2: Genetic engineering

Genetic engineering is the artificial recombination of nucleic acid molecules in the test tube, their insertion into a virus, bacterial plasmid, or other vector system, and the subsequent incorporation of the chimaeric molecules into the host organism in which they are capable of continued propagation. The basis of crop breeding is the transfer of genetic information between plants in order to develop desired phenotypes. However, genetic engineering allows directed and highly specific manipulation of the genetic material. Recombinant DNA technology also broadens the possibilities of transferring genes between unrelated organisms, and creating novel genetic information by the specific alteration of cloned genes. It has not yet been perfected to a sufficient level to allow it to be used routinely as a tool in plant breeding but it is likely to become so within the next few decades. The types of breeding situation in which recombinant DNA would be of most value are those involving transfer of a single or small group of genes from one organism to an unrelated organism (even from outside the plant kingdom) where such transfer cannot be achieved by conventional means.

Conventional plant breeding suffers from some drawbacks when breeding for insect resistance in plants. Firstly in relying on sexual fertilisation for improvement of a crop plant, the breeder is limited in the gene pool he has available by the range of plants with which it is sexually compatible. Secondly, conventional plant breeding requires a long "investment" time. On average, 12 generations of backcrossing are necessary following a successful hybridisation of two varieties of small grain cereals.

1.3: Methods of plant genetic engineering

Several different potentially useful systems of introducing foreign genes into plants are currently under active investigation. *Agrobacterium tumefaciens* has provided an effective DNA delivery system for the transformation of several dicot species. Unfortunately, infection in monocots is greatly reduced and apparently lacking in members of the Gramineae. Because of the inability of *A. tumefaciens* to infect gramnineous monocots which are the most important crop plants, exploratory research in the development of other methods for the stable transformation of grasses is being carried out by several groups. A lot of emphasis is placed in transformation by direct delivery of DNA into regenerable intact cells/protoplasts, meristems and embryos. It should be recognized, however, that transformation of one or more cells in multicellular structures will yield chimeral plants that will require a great deal of efffort before being useful for the production of pure lines containing the introduced genes.

1.3.1: Agrobacterium tumefaciens

A. tumefaciens, a gram-negative soil bacterium, is the causative organism of Crown Gall tumour, which it can incite upon the aerial parts of most dicotyledonous plants. This property is conferred upon the bacterium by the presence of a large Ti-(Tumour inducing) plasmid. During the infection process, a defined small part of the Ti-plasmid is transferred to the plant cell, where it becomes inserted into the plant chromosome. This small piece of transferred DNA, the T-DNA, encodes the functions which are expressed as the tumourous phenotype - rapid growth and

11

the production of opines (unusual tumour specific metabolites).
Both these properties are maintained when the tumours are
explanted into tissue culture, where Crown Gall cells proliferate
without added phytohormones. Methods have been devised whereby
the gene responsible for causing tumourous growth has been
removed and foreign ones to be expressed added. Thus infected
plants will express the foreign genes and so will plants
regenerated from this infected tissue.

However, regeneration of normal healthy plants expressing
the foreign gene is not achieved in all infected plant species.
Also, there is a problem of limited host range for the bacterium.
But the recent discovery by scientists at BioTechnica
International (Cambridge, MA) that Agrobacterium can transfer
plasmids other than Ti opens the possibility that its host range
can be extended. Researchers have shown that the mobilization
functions of a small, wide-host-range bacterial plasmid (RSF1010)
can effectively substitute for the Ti plasmid's 25 base pair
direct repeat T-DNA borders (considered essential for transfer to
occur)(Buchanan-Wollaston et al., 1987).

1.3.2: Plant virus vectors

As a means of introducing DNA into the genomes of
dicotyledonous plants, there is currently nothing to rival *A.
tumefaciens*. However, for some applications it may be useful to
use a non-integrating vector which replicates to high copy number
in each cell and disseminates throughout the organism when
introduced into a mature plant. Moreover, a vector for
monocotyledonous crops has not yet been developed. For these

12

reasons the only two known groups of plant DNA viruses, the Caulimoviruses and Geminiviruses are currently under investigation. Researchers have shown that geminivirus-derived vectors should also be useful for heritable gene amplification by the integration of stable master copies of the vector into the plant chromosomal DNA. Production of genetically engineered plants with increased resistance to herbicides, insects or viruses and increased yields of important plant products are applications of potential commercial value.

1.3.3: Other possible vector systems

There are several other potential vectors for gene transfer to plants. They are transposable elements, maize mitochondrial elements, nuclear genomic components, RNA viruses and viroids.

1.3.4: Direct DNA delivery

These methods are advantageous in that they can be used with any crop species.

Researchers have managed to stably transform tobacco protoplasts with Ti-plasmid by incubating protoplasts and DNA in the presence of polyethylene glycol with a post-incubation with high Ca++ concentration (Krens et al., 1982).

Another technique, electroporation, uses electrical pulses of high field strength to permeabilize cell membranes reversibly so as to facilitate the transfer of DNA into cells. This has resulted in stably transformed cells.

Direct DNA transformation into cereal protoplasts using polyethylene glycol treatment or electroporation has been reported for the monocotyledons wheat, Italian ryegrass, maize (Fromm et al, 1986), *Panicum* sp. and rice.

13

The method of microinjection involves the introduction of
DNA solutions under pressure into cellular compartments with
microscopic pipettes. In one study (Crossway et al.,1986), cell
lines cultured from microinjected tobacco protoplasts were shown
to have integrated the foreign DNA sequences into the nuclear
DNA; the average transformation frequency depended on whether the
injections were intranuclear (14%) or cytoplasmic (6%).

All these direct methods share a big drawback: it is
difficult to regenerate mature plants from protoplasts of mono-
cotyledons. Where achieved, it is at low efficiencies. Re-
searchers have managed to regenerate transgenic rice plants
(*Oryza sativa*) (Kyozuka et al., 1987, Toriyama et al., 1988).
Others (Rhodes et al., 1988) have also managed to regenerate
transgenic maize plants by transforming them by electroporation
and growing them over feeder or "nurse" cells. These, however,
were found to be sterile. Some other successful monocot regenera-
tions are for sugar cane (Srinivasan et al., 1986) and *Polypogon*
(Chen et al., 1987). The problem of regeneration of single cells
or protoplasts can, however, be totally avoided by introducing
DNA into whole tissues.

Researchers can introduce DNA into whole cells either by by
using microprojectiles or by microinjection.

In the former method, the projectiles, small particles of
tungsten or gold coated with DNA, are accelerated by a particle
gun or electric discharge so that they penetrate the outer cell
wall and membrane of intact cells (Sanford et al., 1988). This
method has produced stably transformed cells of tobacco and maize
(Klein et al., 1988a,b,c). Recently, Agrocetus, in Wisconsin

(McCabe et al., 1988), produced fertile transgenic soya bean plants by shooting microprojectiles at meristems of immature soya bean seeds. Approximately 2% of shoots derived from these meristems via organogenesis were chimeric for expression of the introduced gene.

Another method which has been suggested to transform cereal plants without involving tissue culture techniques is by De la Pena et al.,1987. They have injected DNA into developing floral tillers of rye plants 14 days before the first meiotic metaphase of the archesporial cells. Seeds were obtained by pair-wise crossing of injected tillers from different plants. Some seeds germinated into plants which expressed the foreign DNA. The team is confident that this transformation procedure can be extended to other cereals, which have a premeiotic development equivalent to that of rye.

Transgenic wheat plants have been obtained by another method in which mature wheat embryos take up DNA by imbibition (Schell, 1987).

Other methods that have potential include gene transfer into pollen (Luo et al., 1988) and microinjection into cells of immature embryos (Neuhaus et al., 1987).

1.4: Natural protease inhibitors

Plants, micro-organisms and animals contain a number of proteins which have the property of forming reversible stoichiometric protein-protein complexes with various proteolytic enzymes, thus bringing about competitive inhibition of their catalytic functions.

The proteolytic enzymes found in nature can be conveniently subdivided into four main groups which are characterized by the nature of their active sites and the reaction mechanism involved. These are the serine proteases (e.g. trypsin, chymotrypsin); the thiol or cysteine proteases (papain, bromelain, ficin); the metalloproteinases (carboxypeptidases A and B, and the aminopeptidases); and the aspartylproteases (pepsin and rennin). Nearly all of these types of enzymes have been shown to be inhibited by proteins or peptides isolated from the cells of animals, plants and micro-organisms. In some cases the proteinase inhibitors exhibit a very narrow range of specificity, being capable of inhibiting only one or two closely related proteinases, whilst others of broad specificity are active against a wide range of different enzymes.

In the plant kingdom most storage organs, such as seeds and tubers, contain from 1 to 10% of their proteins as inhibitors of various types of proteolytic enzymes (Ryan, 1981), and some fruits contain up to 50% of their proteins as inhibitors of serine endoproteinases. The function of the inhibitors in Nature appear to be twofold: (1) to prevent uncontrolled proteolysis within cells, organelles or fluids where limited proteolysis is important to biochemical or physiological processes, and (2) to protect proteins of cells, fluids or tissues from foreign proteolytic enzymes. The specific roles of most known proteinase inhibitors, however, are not well understood.

1.5: Protease Inhibitors as defensive chemicals

Relatively high levels of protease inhibitors are synthesized and stored in plant tissues, where they can interact

with plant pests or pathogens that attempt to consume them. The effects of protease inhibitors on insect digestive enzymes were first researched by Birk and her associates in the early 1960s. From this and other research in many laboratories, it became clear that the defensive role of proteinase inhibitors was only a part of a complex interaction between the many defensive chemicals that are present in plants and the predators and pathogens that attack them. Plants have evolved various chemical weapons of defense against their predators over the years. This has included inhibitors of the digestive proteolytic enzymes of the attacking pests (Broadway et al., 1986).

The digestive processes of higher animals, insects, and microorganisms can differ considerably, both in the classes of the enzymes utilized for protein digestion and in their specificities. Thus in considering inhibitory activity, the mechanistic class and the peptide-bond specificity of the protease must be considered as well as structural aspects of the inhibitor that determine its ability to interact specifically with the enzyme. The association constant of the interaction must be of sufficient magnitude to inhibit the enzyme effectively. Association constants of most protease-inhibitor interactions are from 10^6M to 10^{10}M, and sometimes higher.

The amino acids that comprise the reactive sites of the inhibitors have the potential to be changed by *in vitro* mutagenesis to produce new inhibitory specificities. This provides a basis for future strategies to design various protease inhibitor genes with different active sites to inhibit particular proteo-

lytic enzymes of specific plant pests.

The serine and cysteine endopeptidases are considered to be the most likely targets for inhibition by such engineered protease inhibitors. Both classes of enzymes have been identified in guts of herbivorous insects, and both are secreted by various microorganisms. Digestive roles in insects and microorganisms for the other two classes of proteinases, aspartyl- and metallo- have not been extensively studied and few of their inhibitors are known.

1.6: Purification of protease inhibitors

The isolation and purification of proteinase inhibitors from plants has involved the full range of general techniques for the separation of proteins, but in addition certain novel and special methods have been employed. Many of the plant proteinase inhibitors are extremely resistant to denaturation by heat and this property has frequently been exploited. For example, several isolation procedures include a preliminary step in which the extract or homogenized suspension is heated at 80-100°C for approximately 10 min during which period various contaminating proteins become precipitated and can be removed by filtration or centrifugation. The inhibitors themselves are left essentially unaltered by this process.

Another popular method of purification for these inhibitors involves affinity chromatography on columns containing the proteinase enzymes bound to insoluble resins or polymerized dextrans such as Sepharose. One of the attractions of this method is that crude or only partially purified extracts of plant material may

be applied to the columns at near neutral pH values. As the extract passes through the column the inhibitors form stable complexes with the immobilized proteinases and are retained on the colummn during subsequent washing which removes most of the other impurities. The inhibitors can then be released from their complexes by lowering the pH and increasing the ionic concentration of the eluting buffer. Although this method of purification is attractive because of its relative speed and specificity, it does however suffer from the major drawback that partial proteolysis of the reactive site of the inhibitor may occur. Such cleavage of the reactive site results in the formation of modified forms of the inhibitor which can interfere with the further purification, particularly if a multiplicity of iso-inhibitors are involved.

Ion-exchange chromatography has been widely employed for the separation of different proteinase inhibitors found in plant extracts. The technique has also been adapted for use in the presence of dissociating solvents such as 4M guanidine or 8M urea and this has permitted the separation of the dissociated protomers (sub-units) of the many inhibitors which are normally polymeric in their native form (Melville et al, 1972). The technique of iso-electric focusing has proved to be particularly useful in separating the proteins of families of iso-inhibitors where the individual components sometimes only differ from one another by slight variations in their iso-electric (pI) points.

1.7: Cysteine protease inhibitors

It has been shown that the larvae of Cowpea Weevils (Gate-

house et al., 1985), which thrive on cowpeas, and Colorado Potato Beetles (Wolfson et al., 1987), which consume potato and tomato leaves, employ cysteine proteases as important digestive enzymes. Thus cysteine protease inhibitors have a strong potential to control these insect pests as well as many others that might use cysteine proteases as their digestive enzymes.

Inhibitors of cysteine proteases of animal origin have now been studied in rather great detail, whereas not much is known with respect to these inhibitors from plants. The presence of inhibitors of cysteine proteinases has been demonstrated in barley, wheat, and rye seeds (Fossum, 1970), seeds of legumes (Goldstein et al., 1973, Rele et al., 1980), corn (K. Abe et al., 1980, 1988), rice (M. Abe et al., 1985, 1987a, 1987b, 1988), pumpkin seeds (Zimacheva et al., 1988) and potato tubers (Rodis et al., 1984, Zimacheva et. al, 1984, Brzin et al., 1988). The best studied inhibitor of cysteine proteases was isolated from pineapple stem (Reddy et al., 1975).

1.8: Obtaining the gene

The first step towards producing transgenic insect-resistant crops would be to isolate the gene for the inhibitor from its native organism. When a protein has been known to effectively affect insects and if it has been isolated and its sequence known in full or partially, oligonucleotides to a few consecutive amino acids of the proteins can be constructed. These can be used to probe a cDNA library or genomic library for the gene or if the oligonucleotides are complementary to the ends of the gene, they can be used to isolate the gene from the genome of

the native organism by polymerase chain reaction (see below).

Once the gene has been isolated and further amplified by cloning

into bacteria, it can be sequenced or expressed in yeast and

studied in more detail before inserting into plants.

1.8.1: Polymerase Chain Reaction (PCR)

The polymerase chain reaction is capable of enriching a

specific DNA sequence by a factor of 10^6 within just a few hours,

hence greatly facilitating a variety of subsequent analytical

manipulations.

The polymerase chain reaction amplification involves two

oligonucleotide primers that flank the DNA segments to be ampli-

fied and repeated cycles of heat denaturation of the DNA, anneal-

ing of the primers to their complementary sequences, and

extension of the annealed primers with DNA polymerase. These

primers anneal to opposite strands of the target sequence and are

oriented such that DNA synthesis by the polymerase proceeds

across the region between the primers, hence doubling the amount

of that DNA segment. Since the amplified DNA segment is also

capable of binding primers, each successive cycle doubles the

amount of DNA synthesized in the previous cycle.

Previously, the polymerase used was the Klenow fragment of

Escherichia coli DNA Polymerase I. However, a recently discovered

thermostable DNA polymerase (*Taq*) purified from the thermophilic

bacterium, *Thermophilus aquaticus* , that can survive extended

incubation at 95°C has simplified the procedure tremendously

(Saiki et al., 1988). The process can now be automated as the

polymerase does not have to be replenished after the denaturation

step. The overall performance of the reaction is also improved by

increased specificity, yield, sensitivity, and length of targets that can be amplified.

A significant improvement in specificity is obtained when the temperature of the primer annealing step is raised from 40°C to 55°C. The optimal temperature for annealing, however, depends on the template DNA and the oligonucleotide primers. This can be determined only by trial and error.

The sensitivity of this method is very high. It has been suggested that a single molecule in 10^5 to 10^6 can be amplified.

Whereas the Klenow polymerase could not sustain the exponential accumulation of DNA sequences much greater than 250bp, the *Taq* polymerase, in contrast, can readily effect the synthesis of segments longer than 400bp. With longer primers and longer extension times, this enzyme can even amplify genomic target sequences longer than 3 kb.

1.9: Aims of project

As cysteine protease inhibitors are potentially useful in conferring resistance to plants from insects and microorganisms if expressed, this project attempts to accomplish the following:
(1) purify a cysteine protease inhibitor from cowpea that its sequence might be known and hence its gene picked out from a cowpea cDNA and hence genomic library.
(2) find out the gene sequence of another cysteine protease inhibitor (bromelain inhibitor from pineapple) whose protein sequence has already been elucidated (Reddy et al., 1975).

It is hoped that these genes may some time in the future be inserted into plants to make them resistant to insects and microorganisms utilising cysteine proteases.

22

CHAPTER II
MATERIALS AND METHODS

2: MATERIALS AND METHODS

2.1: Materials

2.1.1: Chemicals and biological reagents

Reagents, unless otherwise specified, were purchased from BDH Chemicals Ltd., Poole, England. The following materials were purchased from the designated sources.

Acrylamide
N,N'-methylene-bis-acrylamide
Ampicillin
Azo-albumin
BAPNA
BSA
Bromelain Inhibitor
EGTA
Ethidium Bromide
L-Cysteine
Lysozyme
Papain
Papain-Agarose
SDS Molecular Weight Markers (Dalton Mark VII-L): Sigma Chemical

Co. Ltd., Poole, Dorset, England

Agarose: Bethesda Research Laboratories, Inc., Cambridge, England.

Bacto-agar: Difco Laboratories, Detroit, Michigan, USA.

Calf intestinal alkaline phosphatase: Boehringer Corporation Ltd., London.

DEAE cellulose: Whatman BioSystems Ltd., Kent, England.

DNA amplification reagents: GeneAmp, Perkin Elmer Cetus.

DNA from pineapple, oligonucleotides: a gift from Dr. Vaughn Hilder

DH5 α competent cells: Bethesda Research Laboratories Life Technologies, Inc., Uxbridge, Middlesex, England.

Dialysis tubing: Medicel International Ltd., London.

IPTG: Northumbria Biologicals Ltd., Northumberland, England.

Nick-translation kit: Amersham Laboratories, Buckinghamshire, England.

Nitrocellulose filters (BA85, 0.45mm): Schleicher and Schull, Anderman and Co. Ltd., Surrey, U.K.

One-Phor-All Buffer Plus: Pharmacia LKB Biotechnology.

Plasmids: Boehringer mannheim GmbH, West Germany.

Restriction enzymes: Northumbria Biologicals Ltd., Northumberland, England.

Sephacryl 200, Sephadex G-50: Pharmacia AB, Laboratory Separation Division, Uppsala, Sweden.

T4 DNA ligases: Boehringer Corporation Ltd., London.

TG2 cells: Amersham Laboratories, Buckinghamshire, England.

Trypticase Peptone: Becton Dickinson, Microbiology systems, Cockeysville, USA.

X-gal: Northumbria Biologicals Ltd.,Northumberland, England.

Yeast extract: Oxoid Ltd., Basingstoke, Hampshire, England

3MM paper: Whatman Ltd., Maidstone, Kent, U.K.

2.1.2: Buffers

2.1.2.A: Phosphate buffer

A stock of 0.2M Phosphate buffer, pH7.6 was made. This was done by adding 43.5ml of 0.2M Na_2HPO_4 to 6.5ml of 0.2M NaH_2PO_4. To make larger volumes, a similar ratio was used, checking with a pH meter. On dilution to lower ionic strengths, it was made sure that the correct pH was obtained by adding more of Na_2HPO_4 to increase alkalinity or by adding more of NaH_2PO_4 to increase acidity.

2.1.2.B: Tris-HCl buffer

To make 50ml of stock 0.1M Tris-HCl buffer, 50ml of 0.1M Tris solution was made and to this was added concentrated HCl, drop by drop with stirring until the correct pH was obtained on the pH meter.

2.1.2.C: TAE buffer

TAE buffer (Alec's buffer) 10X

Tris base -------------- 48.5g

EDTA ------------------- 3.7g

Made up to pH 7.7 with glacial acetic acid and to 1 litre with distilled water.

Running buffer

TAE buffer 1X

2.1.2.D: TE buffer

pH 7.6 TE buffer

Tris.Cl (pH 7.6) ------- 10mM

EDTA (pH 8.0) ----------- 1mM

2.1.2.E: TBE buffer

TBE buffer stock (5X)

Tris-borate ------------ 54g

Boric acid ------------- 27.5g

EDTA (0.5M, pH8.0) ----- 20ml

Made up to 1 litre with distilled water. Diluted to 1X for

working solution.

2.1.2.F: G-50 buffer

NaCl ------------------- 2.63g

EDTA ------------------- 1.12g

Tris.Cl ---------------- 1.82g

SDS -------------------- 0.3g

Made up to 300ml with distilled water and HCl added to pH 7.5.

2.1.3: Solutions

2.1.3.A: Miniprep solutions

Solution 1

EDTA (0.2M) ------------ 5ml

Tris (1.0M, pH 8.0) ---- 2.5ml

D-glucose ------------- 0.9g

Made up to 100ml with distilled water. Made up fresh before use.

Solution 2 (NaOH - SDS)

NaOH (1M) -------------- 2ml

SDS (20%) ------------- 0.5ml

Made up to 10ml with distilled water. The final solution was not

put on ice. This was to prevent the SDS from precipitating out.

The solution was made fresh.

Solution 3 (K(3M)Ac(5M))

KAc (5M) --------------- 6ml

Glacial acetic acid ---- 1.15ml

Made up to 10ml with distilled water.

2.1.3.B: S.O.C.

Trypticase peptone ----- 2.0g

Yeast extract ---------- 0.5g

NaCl (1M) -------------- 1ml

KCl (1M) --------------- 0.25ml

$MgCl_2$, $MgSO_4$(2M) ------- 1ml

Glucose (2M) ----------- 1ml

Made up to 100ml with distilled water. Stored in aliquots at $4^{o}C$.

2.1.3.C: Agarose beads (Agarose gel dye)

The gel dye for the agarose gels was made with the following:

Glycerol --------------- 3.125ml

Tris (50mM, pH 7.7) ---- 2ml

EDTA (0.2M, pH 8) ------ 0.5ml

Distilled water -------- 4.375ml

Agarose ---------------- 20mg

Bromophenol blue ------- 10mg

Xylene cyanol ---------- 10mg

The above mixture was auatoclaved and cooled, then squeezed through a syringe with a fine needle attached.

2.1.3.D: RNase

To prepare a stock RNase solution that was free of DNase, pancreatic RNase was dissolved at a concentration of 10mg/ml in distilled water. This was heated to 100°C for 15min, allowed to cool slowly at room temperature, dispensed into aliquots, and stored at -20°C.

2.1.3.E: Polyacrylamide gel solutions

Main gel stock (mini gel)

Acrylamide ------------ 30g

bis-acrylamide --------- 1g

Made up to 100mls with distilled water.

Stacking gel stock

Acrylamide ------------ 30g

bis-acrylamide --------- 0.344g

Made up to 100ml with distilled water.

SDS Sample Buffer (2X)

Tris ------------------- 4.844g

SDS ------------------- 4g

Sucrose --------------- 20g

Made up to 100ml with distilled water and adjusted to pH 6.8.

Tris/SDS/Glycine Running Buffer Stock (10X)

Glycine --------------- 141g

Tris ------------------- 30g

SDS ------------------- 10g

This was made up to 1 litre with distilled water. Diluted to 1X for working buffer.

Kenacid Blue R Stain

Kenacid R -------------- 1g

Methanol -------------- 1 l

Glacial acetic acid ---- 0.14l

Made up to 2 litres with distilled water.

12.5% Main Gel

The main gels used were 12.5% gels. This was made by adding

the following (for 2 mini gels):

Main gel stock --------- 4.4ml

Tris (1M, pH 8.8) ------ 4.5ml

Distilled water -------- 1.96ml

This mixture was degased in a Buchner flask connected to a vacuum

generated by water pressure. Then the following was added:

Ammonium persulphate --- 0.3ml
(a freshly made
solution of 15mg/ml)

10% SDS ---------------- 0.12ml

TEMED ------------------ 0.004ml

The TEMED was added just before the gel was poured as TEMED

starts polymerization.

Stacking Layer Gel

The following was mixed in a separate buchner flask:

Stacking gel stock ----- 1.0ml

Tris (1M, pH 6.8) ------ 1.25ml

Distilled water -------- 7.32ml

The mixture was degased then the following added:

Ammonium sulphate ------ 0.33g
(Freshly made
at 15mg/ml)
10% SDS ---------------- 0.1ml

TEMED ------------------ 0.01ml

2.1.3.F: 10% SDS

    10g of SDS was dissolved in 90ml of distilled water. This was heated to 68°C to assist dissolution. The pH was adjusted to 7.2 by adding a few drops of concentrated HCl. The volume was adjusted to 100ml. This was then autoclaved.

2.1.3.G: 20X SSC

NaCl ------------------ 17.5g

Sodium citrate -------- 8.8g

Distilled water ------- 80ml

The pH was adjusted to 7.0 with NaOH and the volume adjusted to 100ml. This was then autoclaved.

2.1.3.H: PEG/NaCl

Polyethylene glycol 6000 --- 20g

NaCl ----------------------- 14.6g

Made up to 100ml with distilled water and sterilized by auto-claving.

2.1.3.I: Chloroform

Chloroform ----------------- 96ml

Isoamyl alcohol ----------- 4ml

The above is mixed and stored at room temperature in a closed bottle.

2.1.3.J: Southern filter hybridization solutions

Denhardt's solution (50X)

Ficoll --------------------- 1g

Polyvinylpyrrolidone ------- 1g

Albumin -------------------- 1g

Made up to 100ml with distilled water and stored at -20°C.

Prehybridization solution

50X Denhardt's solution ----------- 10ml

Denatured DNA (Boiled for 10min) -- 10mg

20X SSC -------------------------- 15ml

10% SDS -------------------------- 1ml

Made up to 100ml with distilled water. Freshly prepared before use.

Hybridization solution

50X Denhardt's solution ----------- 10ml

Denatured DNA (Boiled for 10min) -- 10mg

Denatured ^{32}P-labeled probe

DNA (Boiled for 10min) -------------- 0.5-1μg

20X SSC -------------------------- 15ml

10% SDS -------------------------- 1ml

Made up to 100ml with distilled water.

2.1.4: Bacteriological Media

2.1.4.A: YT-amp-Xgal agar

<u>YT agar (1X)</u>

Distilled water -------- 1 000ml

Trypticase peptone ----- 8g

Yeast extract --------- 5g

NaCl ------------------ 5g

Bacto agar ------------ 15g

The above was put in a flask or Schott bottle and autoclaved at 1kg/cm^2 and at 120°C for 15 minutes. After cooling till hand-hot (approximately 70°C), ampicillin and Xgal were added to concentrations of 0.1mg/ml and 0.04mg/ml respectively. Plates were immediately poured.

<u>Ampicillin stock solution</u>

Ampicillin ------------- 0.2g

70% ethanol ----------- 20ml

This was stored at -20°C. 10ml of this was added per litre of YT agar.

<u>X-gal stock solution</u>

X-gal ------------------ 0.2g

Dimethylformamide ------ 10ml

This was stored at -20°C. 2ml of this was added per litre of YT agar.

2.1.4.B: YT-amp media

YT media was made in a similar as above except that no bacto agar was added. 10ml aliquots were put into universal bottles and autoclaved with their caps loose. Once hand-hot, 0.1ml of ampicillin stock was added to each bottle and their caps tightened. These were stored at 4°C till required.

2.1.4.C: 2X YT media

Double the amount of materials as 1X YT media was used. No ampicillin was added.

2.1.4.D: Glucose/minimal medium plates

MgSO$_4$ ------------------------------ 4.92g

Thiamine HCl ------------------------- 6.74g

CaCl$_2$ ------------------------------ 0.29g

Glucose ----------------------------- 4g

The above reagents were each dissolved in 20ml of distilled water and autoclaved and cooled before mixing aseptically with 20ml of an autoclaved mixture of M9 salts and 3g of agar, using the following volumes:

MgSO$_4$ (1M) ----------------------- 1ml

Thiamine HCl (1M) --------------- 1ml

CaCl$_2$ (0.1M) -------------------- 1ml

Glucose (20%) ----------------- 10ml

M9 salts were made as follows:

Na$_2$HPO$_4$ ----------------------- 6g

KH$_2$PO$_4$ ------------------------ 3g

NH$_4$Cl -------------------------- 1g

NaCl ---------------------------- 0.5g

Made up to 1 litre with distilled water and stored at 4°C.

2.1.4.E: H Plates

Trypticase peptone -------- 5g

NaCl --------------------- 4g

Bacto agar -------------- 7.5g

Made up to 500ml with distilled water and sterilized by auto-claving. Plates poured when hand-hot.

2.1.4.F: H Top Agar

Trypticase peptone -------- 2.5g

NaCl ---------------------- 2g

Bacto agar ---------------- 2g

Made up to 250ml with distilled water and aliquoted into 50ml

lots in glass bottles. Sterilized by autoclaving and kept at 4°C.

When ready to use, the agar was melted in a microwave oven.

2.1.4.G: Fresh cells/X-gal/IPTG mix

IPTG (100mM) ---------------------- 40μl

X-gal (2% in dimethylformamide) --- 40μl

Log phase *E. coli* cells ------ 200μl

The X-gal and IPTG solutions was freshly prepared. Log phase *E.*

coli cells were prepared by innoculating 20ml of 2X YT media with

one drop of fresh overnight culture of TG2 cells. This was shaken

at 37°C for approximately 4h.

2.2: METHODS

2.2.1: Protein Purification Methods

2.2.1.A: Extraction

One kilogram of cowpea seeds was ground in a blender and suspended in 2.5l of 100mM Tris-HCl buffer, pH 7.5. This was left overnight at $4^{o}C$ with stirring from a metal propeller.

2.2.1.B: Centrifugation

This was done at approximately 20,000g, the centrifugal field being calculated from the formula:

$$\text{Centrifugal field } (g) = 1\ 118 \times 10^{-8} \times R \times N^2$$

where R is the radius of the head and N is the number of revolutions per minute. These were done in volumes of 250ml or 50ml.

2.2.1.C: Ammonium Sulphate Precipitation

This was done by adding the appropriate amount of ammonium sulphate slowly with stirring from a magnetic stirrer. The amount of ammonium sulphate required to reach the desired saturation was read off the table (Appendix 1, A1.1, Dawson et al., 1986).

2.2.1.D: Dialysis

For dialysis of protein, 1 m lengths of large dialysis tubing was cut and rinsed thoroughly, inside and out with distilled water. These were then boiled in distilled water for 10 mins., cooled and used immediately or stored in distilled water

36

at 4°C. The tubing was always handled with gloves. The dialysis was done by firstly filling the tubes with the material to be dialysed, leaving an inch or two of empty space at one end to ease osmotic pressure. Three knots were tied at each end to ensure no leakage of material. These filled dialysis tubes were then immersed in more than 5X their volume of buffer (41 buckets were ideal for this) and continuous stirring was achieved with a magnetic stirrer. Frequent changes of the buffer was required and around 2 days at 4°C was quite sufficient to get the ionic concentration of the material being dialysed down to the ionic concentration of the surrounding buffer or to get rid of ammonium sulphate salt.

2.2.1.E: Ion Exchange Column

Resin preparation

Pre-swollen DEAE cellulose was used. Approximately 30g of DEAE cellulose was weighed out and stirred with 180ml of 0.2M Phosphate buffer, pH 7.5 (Phosphate buffer at pH 7.5 is used throughout). This was left overnight and the "fines", bits of sloughed off cellulose which floats to the top of the beaker, removed with a 10ml pipette. The wet settled volume of the resin was measured in a measuring cylinder, allowing sufficient time for it to settle using the formula:

$$t = nh$$

where "t" is time in minutes, h is the total height of slurry in the measuring cylinder (cm) and n is a factor between 1.3 and 2.4. The supernatant was removed with any remaining fines to leave a final volume consisting of the "wet settled volume" plus

37

20% buffer.

Column Packing

A 20cm-long column of diameter 1.5cm was packed. This was done in an area free of draughts, direct sunlight and heaters and done as quickly as possible to prevent convection currents in the slurry. The tap of the column was first closed and the stirred slurry quickly poured in. The tap was then opened to allow a drip and the slurry topped up as soon as the buffer at the top of the column was clear. This was repeated until the resin settled up to a height of 20cm. 1cm of clear buffer was left at the top to prevent the column from drying.

Equilibration

The column was connected to a pump at the bottom. The top was fitted with a tight cap connected to a plastic tubing that was immersed at the other end in the buffer to be passed through the column. The level of the buffer had to be higher from the ground than the top of the column to ensure that the column never ran dry. The buffer was pulled into the air-tight column at a rate of 0.5ml per minute.

A buffer of a slightly higher ionic strength was first passed through the column to ensure that the pH of the buffer coming out of the column was the same as the pH of the buffer going in (pH 7.5). After that, around three times the volume of the column of buffer of the similar ionic strength and pH as the buffer the protein was suspended in was poured through.

Loading and washing

After equilibrating, the protein sample was loaded and the

pump connected to a uv-spectrophotometer and a fraction collector respectively. Fractions of 2ml were collected. The uv-spectrophotometer was connected to a chart-recorder so that the peaks showing on the chart could be correlated to the tubes in the fraction collector.

When all of the sample had been loaded, the column was washed with a buffer similar to the equilibrating buffer and similar to the buffer the protein was suspended in.

Eluting

Gradient elution was done with a small glass bridge, with each of its 8cm-long arms immersed in buffers of differing ionic concentration in 250ml beakers. One of the beakers had an ionic strength of 0.5mM and another had 20mM buffer. The plastic tubing leading to the column was put at the bottom of the beaker with lower ionic strength and liquid sucked up through the glass bridge with a syringe to connect the two different buffers. Stirring in the 0.5mM buffer was achieved with a small magnetic stirrer to ensure uniform increase of ionic concentration throughout the buffer in the beaker. Elution was started when the peak from the chart recorder leveled out.

2.2.1.F: Gel Filtration

A 10ml disposable pipette was used for the column (length 20cm, diameter 0.8cm). Sephacryl 200 was used as the resin. A plug of steel wool was used to plug up the opening to prevent resin from coming out. Resin preparation and column packing was done as above. The column was equilibrated with 20ml of 0.1M NH_4HCO_3 manually. The sample was loaded manually when the level

of buffer had just reached the resin. 0.5ml fractions were col-
lected manually in 1.5ml Eppendorf tubes. The pump could not be
used because such small fractions were being collected. 0.1M
NH_4HCO_3 was used to elute the sample once it had been loaded. The
fractions were examined one at a time on a spectrophotometer at
280nm wavelength for the presence of protein.

2.2.1.G: Papain Affinity Column

1ml of papain-agarose resin was put into a 2ml syringe
(length 1.5cm, diameter 0.9cm) with a piece of filter paper at
the bottom to prevent the resin from coming out. The column was
rigged up to the pump for equilibrating using 50mM Phosphate
buffer, pH 7.5 containing 25mM EDTA and 30mM 2-mercaptoethanol to
activate the papain. A flow rate of 0.18ml/min was used. The
column was made air-tight with parafilm wrapped tightly round the
top and the plastic tubing. Sample similar to that loaded onto
the ion exchange column was loaded onto this column. However, it
was brought up to 50mM first by adding the appropriate volume of
0.2M Phosphate buffer and EDTA and 2-mercaptoethanol were added
to 25mM and 30mM respectively to maintain papain in the active
state. After loading, the column was first washed with the equi-
librating buffer followed by another 50mM Phosphate buffer con-
taining only 5mM EDTA to wash away the mercaptoethanol. Elution
was carried out using a 1.7 pH buffer consisting of 0.3M NaCl and
0.01M HCl.

2.2.1.H: BAPNA assays

These were done using the reagents as follows:

Papain (0.1mg/ml) --0.200ml

EDTA/L-Cysteine (37.25mg,30.25mg/ml respectively)---------0.080ml

BAPNA (0.87mg/ml) ---1.000ml

Inhibitor fractions ------------------------------ various amts.

Others (eg. eluting buffer in inhibitor fractions) -- as required.

Phosphate buffer (0.1M) ---------------------- to top up to 4ml.

These were made fresh on the day required. Controls were done

without papain, without inhibitor and without papain or inhi-

bitor. The reaction mixtures were pre-incubated at 28°C for 10

minutes before BAPNA was added. At half-hourly intervals, absorb-

ance at 410nm wavelength was taken. A total of 2 to 2$^1/_2$ hours of

incubation at 28° was done each time.

2.2.1.I: TCA precipitation

This was done to the protein fractions so that the protein

would become concentrated enough to be detected by polyacrylamide

gel electrophoresis. To the sample was added twice its volume of

60% TCA. This was stood on ice for 1h. The samples were then

centrifuged at 4°C and the precipitates washed three times with

80% ethanol. The samples were then dried on a vacuum line and

redissolved in 2X SDS sample buffer to a concentration of 1 -

10μg/μl depending on how pure the protein was expected to be. The

total amount of protein present was estimated from spectro-

photometer readings at 280nm, assuming 1μg of protein per 0.1

unit absorbance. This mixture was boiled for 10 minutes before

loading onto a polyacrylamide gel.

2.2.1.J: Polyacrylamide gel electrophoresis

Mini-gels

Two glass plates were held together with bulldog clips. This

is to allow minimal contact of the gel with the air which might prevent setting. A piece of rubber tubing was inserted between the glass plates and round three sides to hold in the gel. The main gel was pipetted in and a layer of water put on top to level the gel out and to prevent contact of the gel with the air. After the main gel had set, the water was poured away and the stacking layer added as well as the well-former containing slots to form wells. When this had set, the well-former and rubber tubing were removed and the whole apparatus clamped to the reservoir using two bulldog clips for the sides. Running buffer was poured into the upper and lower reservoirs, checking for leaks, and air bubbles directly beneath the gel displaced with a syringe with a bent needle. 10 - 15μl of protein samples were then loaded. Before loading, the protein samples were dissolved in 2X SDS sample buffer to a concentration of 1 to 10μg/μl. This was boiled for 10 minutes, then 0.5μl of bromophenol blue added. After loading, a drop of mercaptoethanol was put into each well. A marker was run beside the samples hence the molecular weights of the proteins could be calculated from a graph of distance migrated vs log mol. wt. After the gel had been run at 90V for around 2h and the dye front was near the edge of the gel, the gel was removed and the stacking layer cut off with a razor. The bottom edge of the gel containing sample 1 was cut at the corner for identification. The gel was then soaked in kenacid blue stain overnight or longer, then transferred to destain and finally to presoak before taking photographs and drying between acetate sheets in a gel drier conected to a vacuum pump.

2.2.2: DNA Manipulation Methods

2.2.2.A: Polymerase Chain reaction

These was done at the different annealing temperatures of 42°C, 46°C, 48°C and 50°C. The reaction mixtures consisted of the following:

C-terminal oligonucleotide (2μg/μl)----- 10μl

N-terminal oligonucleotide (2μg/μl)----- 10μl

Pineapple DNA (0.4μg/μl) ---------------- 2μl

dNTPs mix (1.25mM each dNTP) ----------- 16μl

Reaction buffer ----------------------- 10μl

Water --------------------------------- 51μl

1ul of *Taq* DNA polymerase (final concentration was 5 units/assay) was added after the samples were heated to 95°C. This was to make sure that all proteases that might have been present in the DNA sample were first inactivated. A layer of 100μl paraffin oil was then added to prevent condensation.

The mixtures were put in 1.5 ml Eppendorf tubes which were coated on the outside with petroleum jelly and put in an intel-ligent heating block which was programmed as such:

Step 1: 95°C --------------- 2min

(1X) x°C--------------- 2min

         72°C -------------- 1min

Step 2: 92°C --------------- 1min

(28X) x°C -------------- 2min

         72°C -------------- 1min

Step 3: 92°C -------------- 1min

(1X) x°C -------------- 2min

$72^{\circ}C$ --------------- 5min

     Step 2 was programmed to be repeated 28 times. *x* refers to

the temperature at which annealing was done which was the same

for all steps in a single experiment. The longer final extension

step was to ensure complete polymerization as the polymerase

might have become limiting as the amount of DNA increased. All

the steps could be accomplished within 5h.

2.2.2.B: Agarose Gel Electrophoresis

     2% to 2.5% gels were used. 2% gels were to separate DNA

molecules within the range 3 - 0.1 kb and 2.5% gels were to

separate DNA molecules smaller than 0.1kb.

     The apparatus to hold the gel was prepared by using vacuum

grease to stick the gel mould to a glass slab. The well former

was then balanced on one end such that it was approximately 1mm

above the glass slab. The gel was prepared by dissolving the

appropriate amount of agar (4g for 2% gels, 5g for 2.5% gels) in

a solution of 200ml 1X TAE. This was done by boiling them to-

gether with frequent shaking to prevent the agar from burning.

Once the agar had dissolved, the solution was allowed to cool and

when hand-hot (approximately $70^{\circ}C$), $12\mu l$ of 10mg/ml EtBr was

added. After swirling, this mixture was poured into the apparatus

formed above. After hardening, the mould and well-former were

removed with the help of a razor and the glass slab with the gel

placed in the gel electrophoresis tank. This was then flooded

with 1X TAE buffer to 1mm above the gel surface and the samples

loaded with 1/5th their volume of gel dye. The gel was run at 25V

overnight or at 80V during the day (approximately 5h)

Photographs of the gel were taken with a polaroid camera using uv light (Polaroid 667, ISO 3000 film).

2.2.2.C: Enzymatic restriction of DNA

Restrictions were done by adding 2ul of enzyme (5u/μl) and 2μl of the appropriate restriction buffer (10X) to the DNA sample. This was then made up to 20μl with distilled water and incubated for over 3 h at 37°C. For miniprep DNA samples, 2μl of 1mg/ml RNase was added to remove any tRNA present.

2.2.2.D: Electroelution (Maniatis et al., 1982)

Preparation of bands

An agarose gel was run as normal and the desired bands cut out with a razor under uv light. Exposure to uv light was minimised to prevent denaturation of DNA.

Preparation of Dialysis tubing

Small dialysis tubing of width 1cm when flat was used. This was cut into 10cm lengths and boiled for 10 minutes in a large volume of 2% sodium bicarbonate and 1mM EDTA. The tubings were then rinsed thorougly in distilled water. They were then boiled for a further 10 minutes in 0.001M EDTA. After cooling, they were stored at 4°C submersed in 70% ethanol. Before use, the tubing was rinsed inside and out with distilled water. Gloves were used to handle the tubing.

Electroelution

A dialysis tubing was clipped in the middle and the gel slab inserted into one end with forceps. This section was then filled with sufficient TBE buffer to cover the gel slab. The tubing was

then clipped without leaving any space between the buffer and the clip. The same thing was done for the opposite compartment. This tubing was then immersed in some TBE in a minigel electrophoresis tank. 100V was passed for 30min. The polarity of the current was then reversed for 2min to release the DNA from the wall of the dialysis bag. The dialysis bag was then carefully opened and the buffer containing the eluted DNA recovered. The gel slab was examined under uv light to make sure that all the DNA had eluted. If not, a further few minutes of electroelution was required. The bag was washed out with a small amount of TBE buffer at the end to make sure no DNA was lost. The DNA was then purified.

2.2.2.E: Purification of DNA

To the DNA solution was added an approximately equal volume of phenol. This was vortexed till an emulsion formed, then spun for 2min in a microcentrifuge. The supernatant was transferred to another tube and an equal volume of chloroform added. This was vortexed and spun as above. To the supernatant was added twice its amount of 100% ethanol and 1/10th its volume of 3M NaAc to enable precipitation. This was left precipitating at $-20^{\circ}C$ overnight or at $-80^{\circ}C$ for one hour. The mixture was then centrifuged for 20min and as much of the liquid as possible removed with a pasteur pipette, taking care not to remove the DNA pellet. $200\mu l$ of 70% ethanol was added with vortexing and centrifuging to remove any remaining solutes. This was carefully removed and the pellet vacuum dried in a dessicator. The pure DNA was then resuspended in distilled water or TE buffer. When smaller volumes of DNA had to be purified, TE buffer was added before the addition

of phenol to give a larger volume so that losses through the separation of the aqueous and organic phases were minimized.

2.2.2.F: Ligation of DNA to Vector

Vector Preparation

The vectors used were pUC18 and M13mp19. These were first restricted at certain single restriction sites to form linear lengths of DNA (see section 2.2.2.C). Then, $1\mu l$ of Calf intestinal phosphatase (CIP) was added and incubation continued at 37^oC for 10 min. This was to remove the phosphate groups at the open ends of the restricted DNA to prevent self-ligation. $2\mu l$ of 500mM EGTA was then added to remove all traces of CIP. This mixture was incubated at 65^oC for 1h. The plasmid DNA was then purified.

DNA preparation

The DNA to be inserted was restricted with the appropriate enzyme using the method in section 2.2.2.C. High specific activity (HSA) Hind III restriction enzyme ($40u/\mu l$) was used for the PCR amplified DNA. As the oligonucleotides used in the PCR had attached Hind III cleavage sites at their ends, this restriction served only to restrict the ends of the DNA, hence leaving the amplified gene intact.

Ligation reaction

The following reaction mixture was set up:

Vector ($0.1\mu g/\mu l$)   ----------------- $1\mu l$

DNA sample (> $0.1\mu g/\mu l$) ------------- $1\mu l$

1/10 ATP --------------------------- $1\mu l$

10X Ligation buffer ---------------- $1\mu l$

Ligase (0.25u/μl) ------------------ 1μl

The above was made up to 10μl with distilled water. This was then incubated at 12°C for more than 4 h. Two kinds of controls were also set up, one without the DNA sample to ensure that the phosphate groups had been removed in the vector, and another without the DNA sample or ligase.

2.2.2.G: Transformation with pUC 18

Commercially available DH5α competent cells of library efficiency were used. These were claimed to give more than 1 X 10^8 transformants/μg supercoiled pUC19. 20μl of these were put into Eppendorf tubes. All of the ligation mixture was then added and the mixture shaken gently. This was then put on ice for 30 mins. When the 30 minutes was up, the mixture was heat-shocked for 40 s in a 42°C water bath, care being taken not to shake the tubes. They were then placed on ice for a few minutes to stabilize before 80μl of S.O.C. was added. These were shaken at 225 rpm for 1 hour at 37°C to ensure that the ampicillin resistance genes were expressed before plating out on YT-X-gal-amp plates.

2.2.2.H: Streaking

This was carried out whenever it was required to get individual colonies or colonies from a single parent on a plate. Sterile toothpicks or a metal loop were used. The metal loop was flamed then cooled and the bacteria to be plated out picked up. A zig-zag pattern was drawn on one edge of the plate and the loop flamed again. After cooling, the loop was passed across one or two of the previous lines and a zig-zag pattern drawn on a fresh

edge of the plate. This was repeated one or two more times. In the case of sterile tooth-picks, these were changed after each line was drawn. They were then soaked in ethanol before discarding. The plates were grown overnight at 37°C in an inverted position.

2.2.2.I: Plasmid Isolation (Miniprep)

Bacteria propagation

To 10ml of autoclaved YT medium in Universal bottles was added 100μl of 10mg/ml ampicillin. Bacteria was added either from its colony on a plate or from a previous miniprep (10μl is sufficient). This was grown overnight at 37°C.

Plasmid purification

The bacteria were pelleted down in a high speed centrifuge (30 000g for 10 min). The broth was poured off and the tubes inverted for a few minutes on absorbant paper to dry the pellets. 200μl of 4mg/ml lysozyme in miniprep solution 1 was added, the mixture was then vortexed and stood on ice for 30 minutes. Then 400ul of miniprep solution 2 was added. This was not vortexed but gently shaken. This solution is meant to denature the proteins in solution. After incubation for 5 minutes on ice, 300μl of solution 3 was added. This was incubated for 30 minutes on ice. This solution got rid of any chromosomal DNA. The final solution was spun for 30min at 4°C. The supernatant containing the plasmid DNA was removed and the DNA purified and concentrated as before (section 2.2.2.E), but excluding the addition of 3M NaAc as the KAc present (solution 3) was sufficient to precipitate the DNA.

2.2.2.J: Preparation of competent TG2 *E. coli* cells

The <u>E. coli</u> host strain TG2 is used as it is strongly recommended for transformation with M13. The method used to make competent cells was a quick and easy method which produced competent cells wih only a limited storage life (24h).

A single colony of TG2 was first picked from a glucose/ minimal medium plate. This was grown overnight in 10ml of 2X YT medium at 37°C with shaking. 40ml of 2X YT media in a 250ml flask was innoculated with 2ml of this overnight culture and shaken at 37°C for 2 hours to $OD_{550} \sim 0.3$. The cells were then spun down (3,000g for 5 min). These were then resuspended in 20 ml of sterilized 50mM $CaCl_2$, pre-chilled on ice. After leaving on ice for 20 minutes, the cells were again spun down (3 000g for 2min). They were then resuspended in 4ml of cold 50mM $CaCl_2$. As transformation efficiency increases after standing on ice for a few hours, the cells were left to stand for 3 hours.

2.2.2.K: Transformation with M13mp19

300μl of competent cells were aliquoted into chilled 15ml tubes. Approximately 1ng of DNA was added (4ul of the ligation mixture in section 2.2.2.F gave good results), stirring the cells whilst pipetting. The tubes were rolled gently for a few minutes on ice then left to stand for 40 minutes (on ice). The cells were heat-shocked at 42°C for 45 seconds, care being taken not to shake the tubes. They were then placed on ice for 5 min before plating out. To each tube was added 270ul of fresh cells/ X-gal/ IPTG mix. Then 3ml of molten H top agar was added (kept at 45°C). This mixture was mixed by rolling and immediately poured

onto a prewarmed (37°C) H plate each. When set, the plates were inverted and incubated at 37°C overnight. Clear plaques the next day indicated transformation. These plaques could not be stored as such so phage stocks were made. Single-stranded template DNA was also directly prepared.

2.2.2.L: Preparation of 1ml phage stock

A single TG2 colony was picked from a glucose/minimal medium plate and grown overnight in 10ml of 2X YT medium with shaking. One drop of this was added to 20ml of fresh medium, and shaken at 37°C for 3 hours. 100μl of this and a whole recombinant plaque (using a sterile pasteur pipette for picking up) were added to 1ml of 2X YT media in a sterile 10ml culture tube. This was incubated for 4 hours with shaking at 37°C. This was then transferred to a Eppendorf tube and stored at 4°C, or long term at -20°C. 50ul of this could be used in place of the recombinant plaque in section 2.2.2.M to produce plasmid DNA.

2.2.2.M: Double-stranded and single stranded template DNA preparation from phage

5ml of 2X YT media in a sterile culture tube was inoculated with 50μl of a fresh overnight culture of TG2 cells and a whole recombinant plaque. This was incubated with shaking at 37°C for 5 hours. The cells were spun down (30 000g for 10min) and the supernatant put into five 1.5ml Eppendorf tubes. These could be used to prepare single-stranded DNA immediately or kept at 4°C for a few days. One of the 5 tubes was kept at -20°C as a stock in case more DNA had to be prepared. The precipitate after centrifugation was used for purification of double-stranded DNA by

51

the miniprep method mentioned in section 2.2.2.I., cutting the volumes by half.

To prepare single-stranded DNA from the 4 tubes, 200μl of PEG/NaCl was added to each tube, shaken, and left to stand at room temperature for 15 minutes. This was then centrifuged for 5 minutes and the supernatant discarded. After a further 2 minutes of centrifugation, all traces of PEG was carefully removed with a drawn out Pasteur pipette. 100μl of TE buffer was added to each tube and vortexed for 30s to resuspend the pellet. DNA purification was then carried out as mentioned in section 2.2.2.E. The dry DNA pellets obtained were resuspended in TE buffer in a total volume of 100μl (by transfer from tube to tube). This was centrifuged for five minutes and the supernatant transferred to a clean tube. This was to remove any protein precipitate which would affect the optical density reading. The amount of single-stranded DNA present was then estimated as mentioned below. The yield from a 4ml culture should be around 16-40μg. The estimated sample was ethanol precipitated and resuspended in TE buffer to give 1μg/μl for subsequent sequencing with the automatic DNA sequencer. At least 4μg was needed.

2.2.2.N: Estimation of concentration of DNA

(Maniatis et al., 1982)

1ul of the sample to be estimated was added to 1ml of TE buffer in a quartz cuvette and the OD at 260nm taken. 1 OD of single stranded DNA was equivalent to 40μg/ml. The same reading was equivalent to 50μg/ml of double-stranded DNA and to 20μg/ml of oligonucleotides. The ratio OD_{260}/OD_{280} showed how pure the

DNA was. The ratio should be 1.8 for pure samples and less for impure samples. In such cases, accurate quantitation was not possible.

2.2.2.O: Southern Blot (Maniatis et al., 1982)

This method is useful for locating DNA of a specific sequence on an agarose gel. DNA fragments that have been separated according to size by electrophoresis are transferred to a nitrocellulose filter, and immobilized. The relative positions of the DNA fragments in the gel are preserved during their transfer to the filter. The DNA attached to the filter is then hybridized to ^{32}P-labeled DNA, and autoradiography used to locate the position of any bands complementary to the radioactive probe.

After electrophoresis had been completed, the gel was transferred to a glass baking dish. First, two pieces of Whatman 3MM paper were placed on top of a piece of glass such that they overlapped each other. These were then placed inside a large baking dish and elevated by rubber stoppers. The dish was filled with 10X SSC such that the ends of the paper were in constant contact with it. On top of this, a few more sheets of paper were placed (dampened with 10X SSC), then the agarose gel, with its front side facing downwards. A piece of nitrocellulose filter (presoaked in 2X SSC for a few minutes) was then placed in contact with the gel. A few more sheets of paper (dampened with 2X SSC) were placed on top of this followed by absorbant material such as the filling of babies' napkins. A glass slab and weights were placed on top of this to ensure good contact between the different layers. The objective was to set up a flow of liquid from

the reservoir through the gel and the nitrocellulose paper, so
that the DNA fragments would be eluted from the gel and deposited
onto the nitrocellulose paper. The gel was surrounded with a
water-tight layer of Saran Wrap to prevent short circuiting of
fluid between the absorbant material above the gel and the 3MM
paper beneath it. After 2 days, the nitrocellulose filter was
marked for the wells on the gel and placed in between dry 3MM
paper. It was then dried for 2 hours under vacuum at 80°C.

2.2.2.P: Nick Translation

The DNA labelling kit from Amersham was used to prepare
radioactive probes. To 0.5-1μg of the DNA to be labelled was
added 10 ul of deoxynucleotide triphosphate mix (contains buffer,
dNTPs). Then 5ul of enzyme solution (contains DNA polymerase) was
added, as well as 5μl of 50μCi [^{32}P] dCTP. The reaction mixture
was incubated for 3-4h at 15°C. After phenol extraction, the
sample was poured through a Sephadex G-50 gel filtration column.

2.2.2.Q: Sephadex G-50 gel filtration column

Sephadex G-50 was swollen in G-50 buffer for 1 hour at 65°C.
The resin was then stood to cool at room temperature and the
column packed in a 5ml disposable pipette (length 15cm, diameter
0.7cm). The nick translated reaction mixture was then poured
through followed by more G-50 buffer. Fractions of 5 drops were
collected and tested for radioactivity using a Geiger counter.
The samples were placed at a fixed distance from the counter in
turn and the reading recorded. The first peak of high radio-
activity represented the labeled pUC18 whereas the second peak

represented the free [^{32}P] dCTP. On measuring the first peak on a

liquid scintillation counter, putting 5μl of the pooled fractions

in 1ml of water, a reading of 7.5x10^6 Cerenkov cpm/ug was ob-

tained (Cerenkov radiation is emitted when a charged particle

(beta particle, in this case), travels with a velocity greater

than that of light in a given medium (Rengan, 1986)). This in-

dicated successful labeling. The pooled fractions were then used

for the hybridisation reaction with the nitrocellulose filter.

2.2.2.R: Hybridisation of nitrocellulose filter

Prehybridization

The baked filter was first immersed in 6X SSC for 2 minutes.

The wet filter was then slipped into a heat-sealable plastic bag.

0.2ml of prehybridization fluid warmed to 65°C was then added per

cm^2 of filter. As much air as possible was squeezed from the bag

and the open end sealed with the heat sealer. The bag was incu-

bated for 4 or more hours in a plastic container weighed to the

bottom of a 65°C water bath.

Hybridization

The bag was opened after prehybridization by cutting off one

corner. As much prehybridization solution was squeezed out as

possible. The hybridization fluid was then added to the bag

(50μl/cm^2 of filter). As much air as possible was squeezed out of

the bag and the cut edge sealed. The bag was incubated as before

overnight.

Washing

The bag was cut along the length of three sides and imme-

diately submerged in a tray containing a solution of 0.1X SSC

and 0.1% SDS. After five minutes at room temperature, the solution was changed for a fresh one. This was repeated again after five minutes. The filter was then transferred to a plastic box containing fresh solution and incubated at 65°C for one and a half hours, changing the solution after each half-hour interval. After the final half-hour, 0.1% SSC solution was added instead. This was to remove the SDS. After 5 minutes at room temperature, the filter was placed on a sheet of Whatman 3MM paper and dried at room temperature.

2.2.2.S: Autoradiography

The dried nitrocellulose filter mentioned above was taped to a piece of Whatman 3MM paper, leaving a small border around for marking with radioactive ink. After marking with the ink (a few randomly placed dots), it was put into a plastic bag sealed round three edges. It was not totally sealed as air bubbles trapped could alter the image on the film. This was then placed in an X-ray film holder.

The X-ray film (Fuji film) was first activated by flashing it for a short while. This was then placed on top of the nitrocellulose filter in the X-ray film holder and closed securely. This was done in a dark room with only a red safety light on. The holder was then left at -80°C overnight or longer.

2.2.2.T: Disposal of wastes

Any material that had been in contact with bacteria was disinfected with 25% Chloros or autoclaved before re-using or discarding. Any radioactive material was rinsed thoroughly under running water until no longer radioactive or soaked in Decon.

CHAPTER III
RESULT

3: RESULTS

3.1: PURIFICATION OF A THIOL PROTEASE INHIBITOR FROM COWPEA

The methods used were that of Rele et al.(Rele et al., 1980) with certain modifications as appropriate. The alkaline extract of cowpea seeds was used and the inhibitor attempted to be purified was designated "inhibitor C" in their paper. This was only partially purified by them. However, others (Gatehouse et al., 1985) have managed to get a single polypeptide inhibiting only papain and not trypsin by a similar method. Therefore, ~~it was~~ attempted to follow their work to get a pure thiol protease specific inhibitor in sufficient quantity that its sequence might be determined.

3.1: Steps in protein purification

Two methods were used to purify the protein. The initial steps used were similar but later on, two different columns were used -- a DEAE cellulose ionic exchange column and a papain affinity column.

Steps in common

One kilogram of cowpea seeds were milled and extracted in 2.51 of 100mM Tris-HCl buffer, pH 7.5. This was done overnight with stirring. The extract was then filtered through a piece of muslin. 1.751 of extract was recovered and centrifuged at 20,000g, 4°C, in 250ml bottles for 30 min. The supernatant was precipitated at 0.9 ammonium sulphate saturation overnight at 4°C. The solution was then centrifuged again as before. The precipitate obtained was resuspended in 800ml of 10mM Tris-HCl, pH 7.5. This was put into large dialysis tubing and dialysed against

58

the same buffer to remove the ammonium sulphate. Centrifugation was done as before. To the supernatant was added 2M acetic acid till a pH of 5 was obtained. After centrifugation, the super-natant was put into large dialysis tubing and these were immersed for 3min in a large volume of 10mM sodium acetate, pH 5, heated up to 80°C. These tubes were then immediately cooled in chilled water. The inhibitor was very unstable at pH 5 before heating and it was essential to carry out acid precipitation and heating as rapidly as possible. After cooling, the solution was centrifuged to remove the denatured protein. Ammonium sulphate precipitation was carried out again, this time at 0.95 saturation. This was then centrifuged and the precipitate resuspended in 100ml of 50mM Phosphate buffer, pH 7.5. This was brought down to 5mM by dia-lysing against this buffer. After centrifugation, the fraction was ready for the two columns. It was later discovered that there was a misprint in the paper therefore another dialysis against 0.5mM buffer was required before pouring through the columns.

3.1.2: Results from the ion exchange column

100ml of the fraction, having been dialysed against 5mM buffer, pH 7.5, was poured through the DEAE cellulose column. When 5mM Phosphate buffer was passed through to wash the column, there was a huge absorbance peak on the uv spectrophotometer. Subsequent elution did not yield any peaks as all the protein had passed through with the 5mM buffer. Hence, the washing peak frac-tions were pooled together and dialysed against 0.5mM Phosphate buffer. This was poured through the column and 3 peaks were ob-tained altogether (fig. 1). The first peak was the washing peak,

the second and the third were the peaks supposed to contain the inhibitor. Unfortunately, due to a fault with the fraction collector, the third peak was lost, hence assays were done only on the second peak.

Assay results

The 2ml fractions from the second peak were pooled together and their absorbance at 280nm measured. This was found to be 0.083. BAPNA assays were done with 2ml and 2.72ml of this pooled fraction. This implied a protein concentration of approximately 1.66mg and 2.26mg respectively, judging from its absorbance. Both showed inhibition of 16.8% and 20.5% respectively (fig. 2, table 1a & 1b, fig. 3, table 2a & 2b).

Further purification

This protein (11.9ml of pooled fractions) was further purified by passing it through a Sephacryl 200 gel filtration column Only one protein peak was obtained (fig. 4) The fractions from this peak were pooled together and freeze-dried. 9mg of protein was obtained and further enzyme assays were carried out on this, using up to 2.1mg per assay. No inhibition, however, was observed. This could be due to one of several reasons which are discussed in the next chapter. On SDS polyacrylamide gel electrophoresis, this protein showed a single band at approximately 32,000 relative molecular weight (fig. 5, fig 6, table 3).

3.1.3: Results from the papain column

15ml of 0.5mM extract obtained previously was made up to

Flow rate: 0.5ml/min.

0.5mM – 20mM
Phosphate buffer
gradient.

Ion exchange column – enzyme assays

Absorbance at 410nm

Time/min
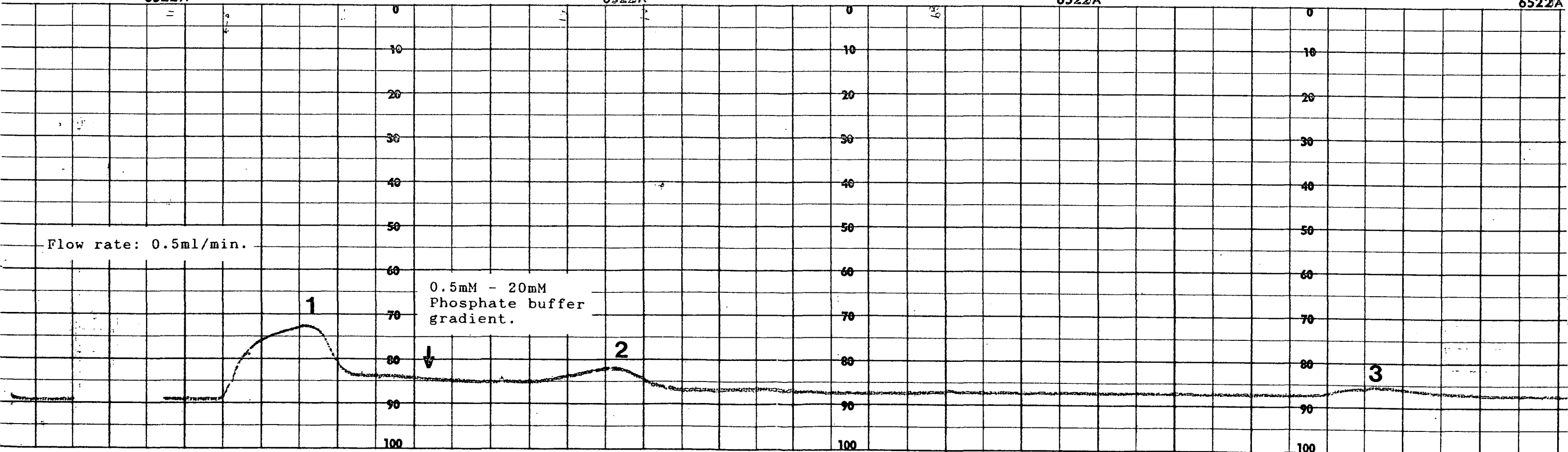
—■— Control 1   —+— 2ml from peak 2   —✳— Control2
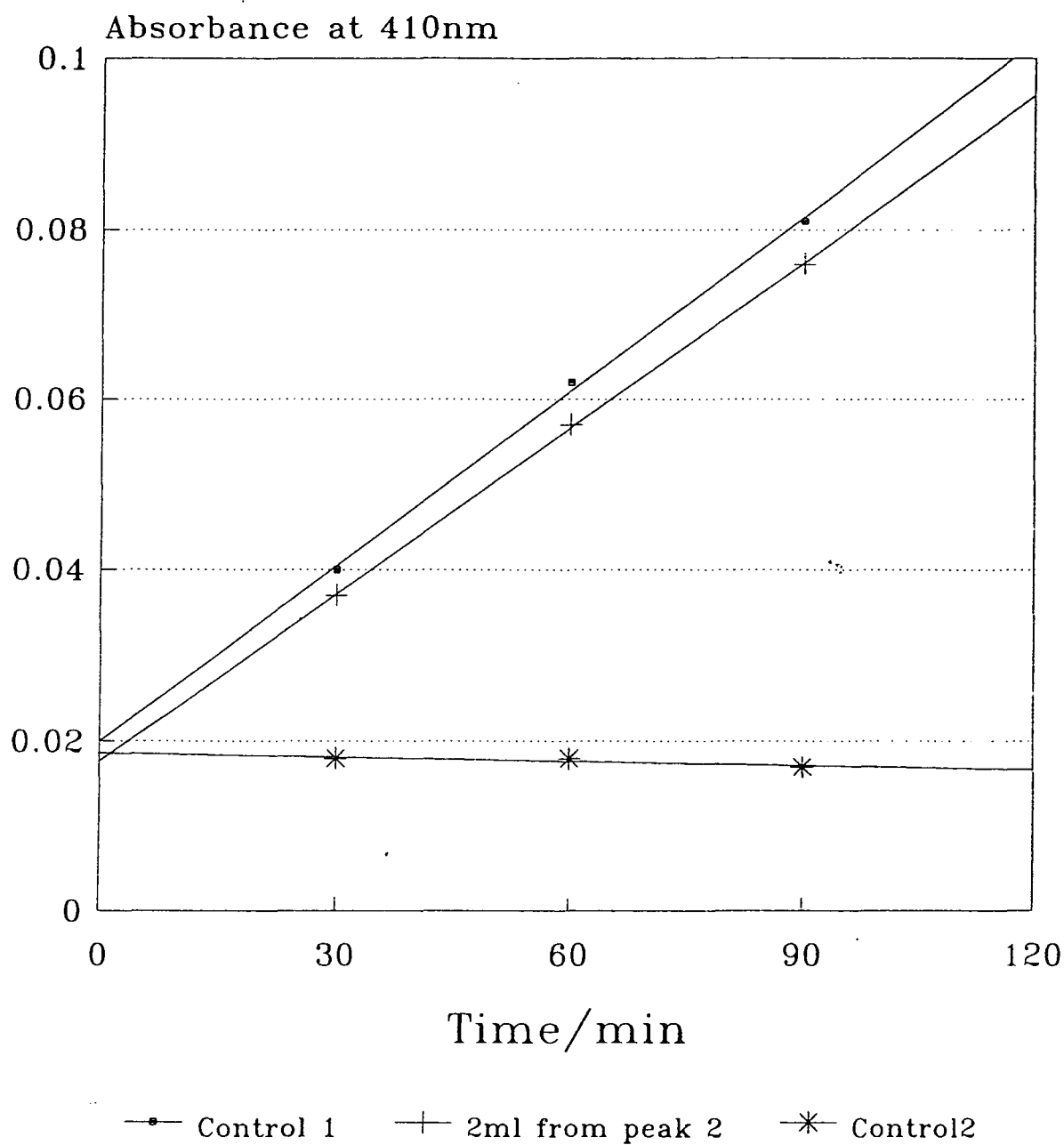
Fig 2: DEAE ion-exchange column - BAPNA assay rate graph of 2ml of fraction from peak 2.

Tube no.	Description
1	Control 1, with papain but without inhibitor.
2	With inhibitor, 2ml of peak 2 from ion-exchange column.
3	Control 2, without papain or inhibitor.

Table 1a: Reaction mixtures in fig. 2.

\ Sample no Time	1	2	3
30	0.040	0.037	0.018
60	0.062	0.057	0.018
90	0.081	0.076	0.017
130	0.116		0.020
150	0.135		0.020
180	0.155		0.018
Rate (10units/min)	7.81	6.50	0.00

Regression Output:

```
Constant              0.014772  0.017666
Std Err of Y Est      0.002715  0.000408
R Squared             0.997017  0.999780
No. of Observation          6         3
Degrees of Freedom          4         1

X Coefficient(s)      0.000781  0.00065
Std Err of Coef.      0.000021  0.000009
```

Table 1b: Absorbance of BAPNA assays at 410nm & regression output of lines obtained. From rates of change in absorbance, it can be seen that tube 2 inhibits papain by 16.8%. ~1.66mg protein was present.

Ion–exchange column – enzyme assays

Absorbance at 410nm

Control   2.72ml from peak 2
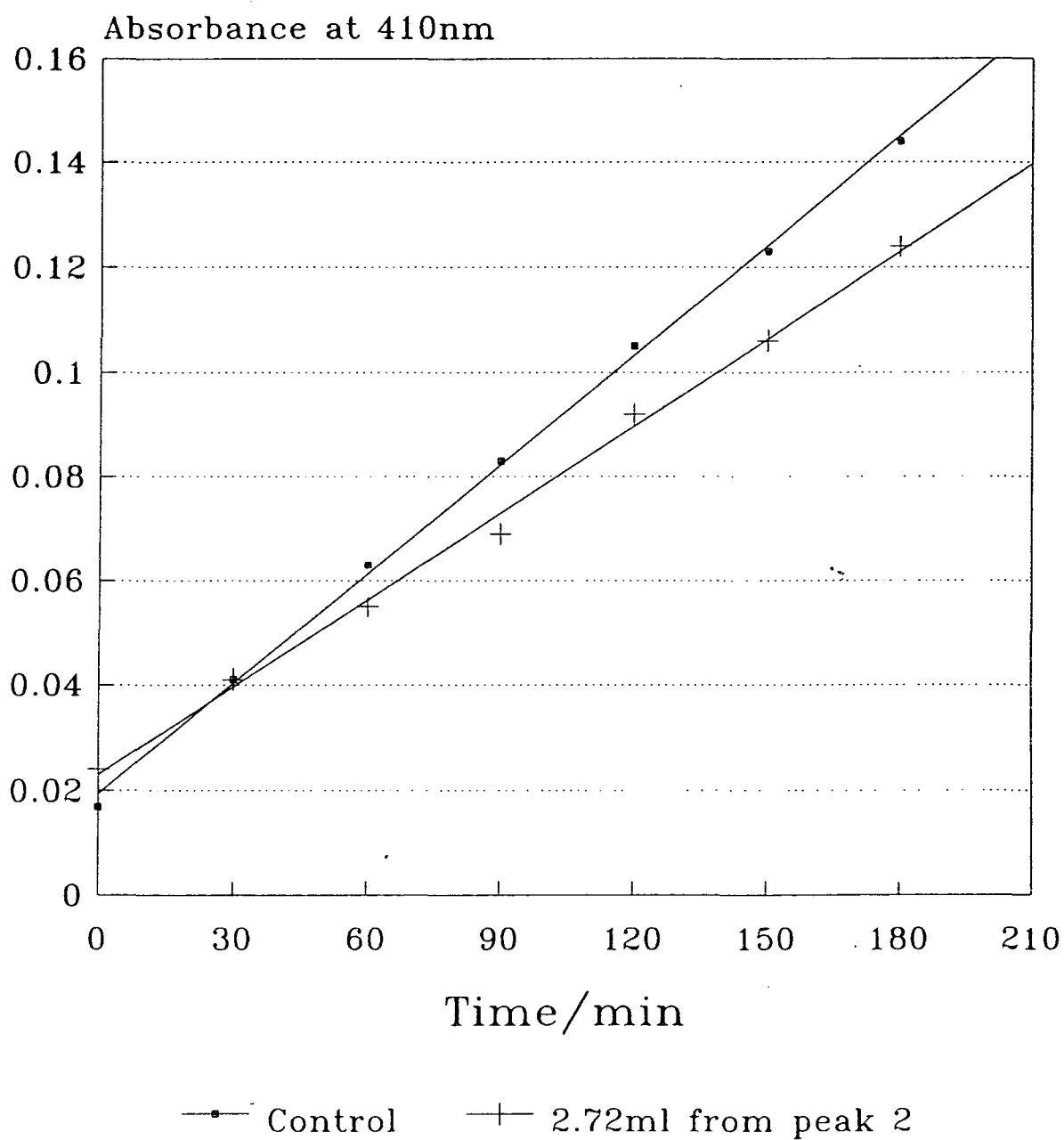
Fig 3: DEAE ion-exchange column - BAPNA assay rate graph of 2.72ml of fraction from peak 2.

Tube no.	Description
1	Control 1, with papain but without inhibitor.
2	With inhibitor, 2.72ml of peak 2 from ion-exchange column.

Table 2a: Reaction mixtures in fig. 3.

\ Sample no Time	1	2
0	0.017	0.024
30	0.041	0.041
60	0.063	0.055
90	0.083	0.069
130	0.105	0.092
150	0.123	0.106
180	0.144	0.124
Rate (10⁻⁵units/min)	6.89	5.48

Regression Output:

	1	2
Constant	0.019285	0.022845
Std Err of Y Est	0.002463	0.002297
R Squared	0.997538	0.996625
No. of Observation	7	7
Degrees of Freedom	5	5
X Coefficient(s)	0.000689	0.000548
Std Err of Coef.	0.000015	0.000014

*Table 2b: Absorbance of BAPNA assays at 410nm
& regression output of lines obtained.
From rates of change in absorbance, it
can be seen that tube 2 inhibits papain
by 20.5%. ~2.26mg of protein was present.*

Gel Filtration Column
Absorbance of 0.5 ml fractions
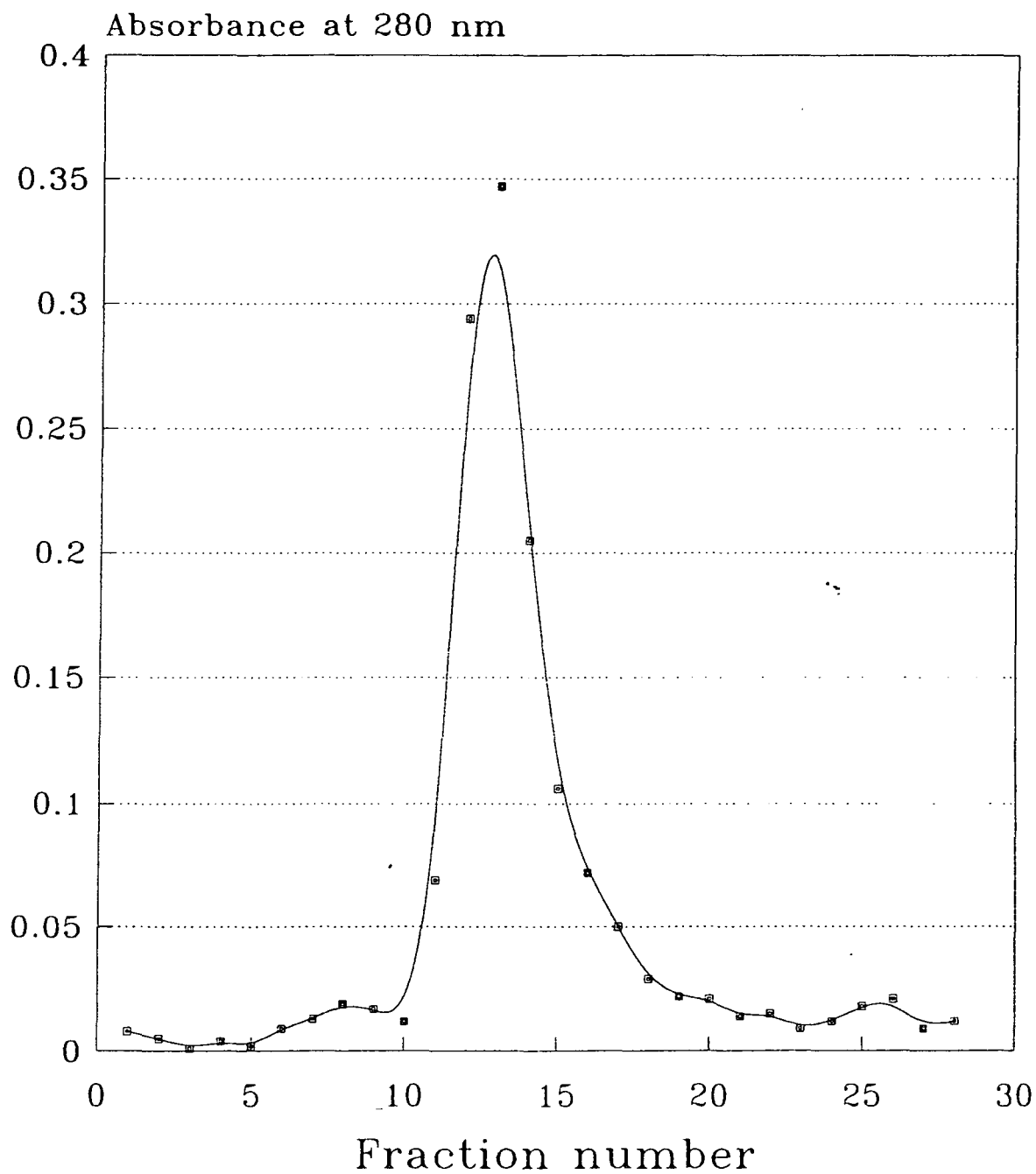
Fig 4: Gel filtration column - absorbance readings of fractions at 280nm.

Marker size/m.wt.

66000
45000
36000
29000
24000
20100
14200
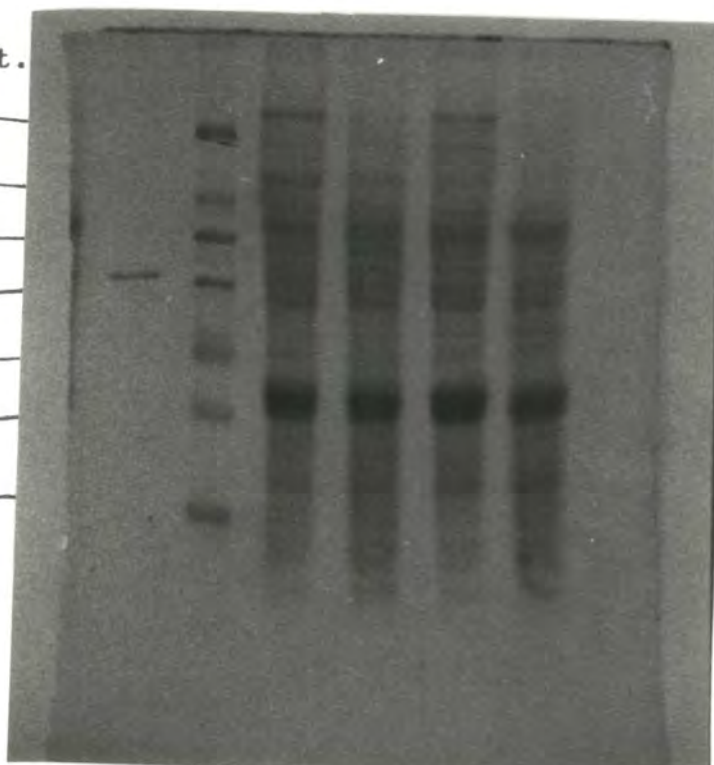
1   2

Well no.	Sample
1	Protein off the Sephacryl 200 gel filtration column after TCA precipitation
2	Sigma SDS molecular markers

Fig 5: Polyacrylamide gel electrophoresis of single protein peak from gel filtration column.
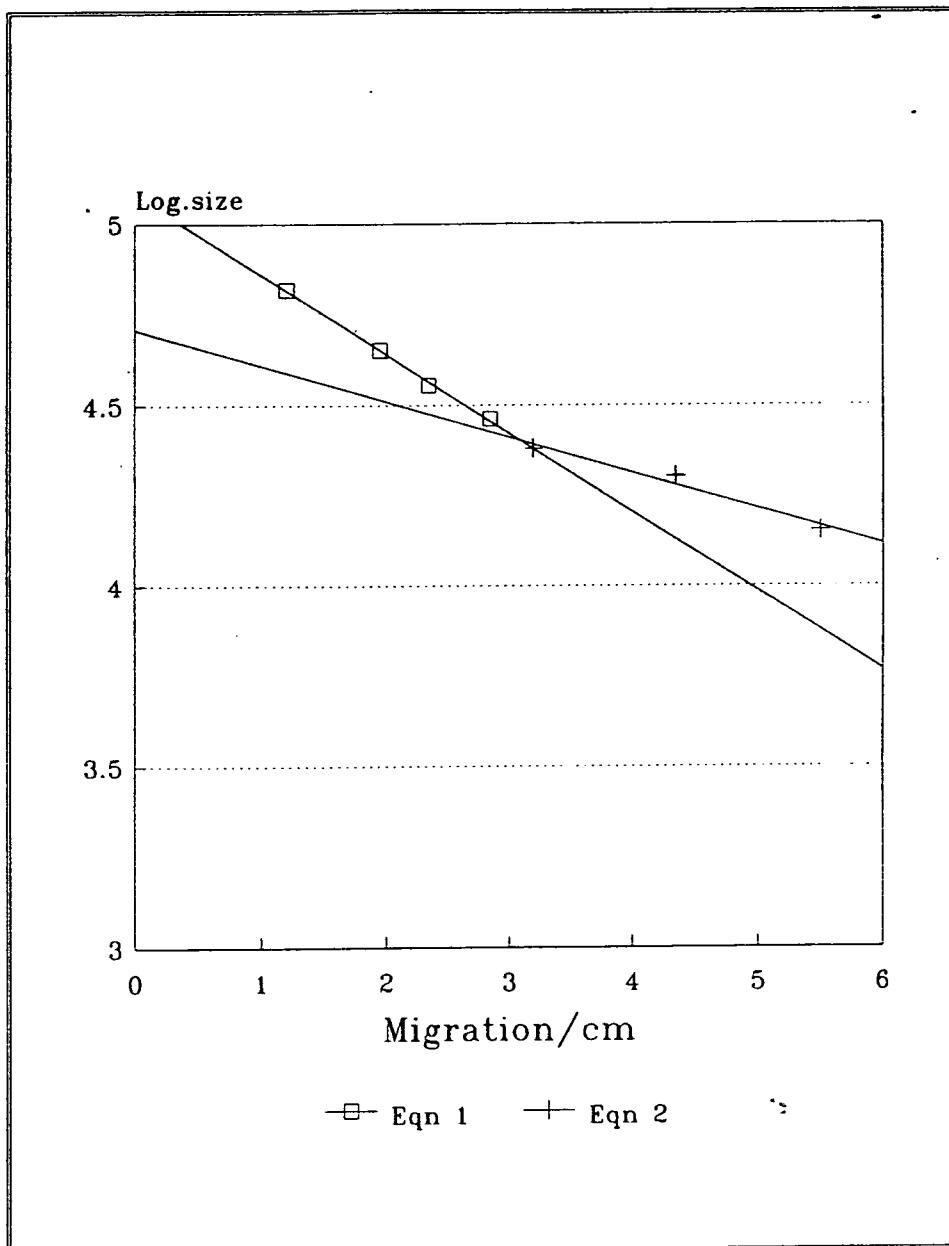
Fig. 6:
Graph of migration/cm vs log/m.wt. for marker in fig. 5.

Marker Migration	Size/나v	Log.size
1.20	66000	4.8195
1.95	45000	4.6532
2.35	36000	4.5563
2.85	29000	4.4624
3.20	24000	4.3802
4.35	20100	4.3032
5.50	14200	4.1523

Sample Migration	Size/나v	Log.size
2.75	30061	4.4780

(Calculated using eqn 1)

Table 3: *Migration distances & sizes of bands in fig. 5. Regression output of fig. 6.*

Using first four points,
Regression Output:

Constant	5.079308
Std Err of Y Est	0.008066
R Squared	0.998135
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s)   -0.21865
Std Err of Coef.   0.006682
Eqn 1: y = -0.21865x + 5.079308

Using last three points,
Regression Output:

Constant	4.709636
Std Err of Y Est	0.030166
R Squared	0.966150
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s)   -0.09909
Std Err of Coef.   0.018548
Eqn 2: y = -0.09909x + 4.709636

68

50mM with 0.2M Phosphate buffer. This was passed through the column as mentioned in section 2.2.1.G. Three peaks were obtained. The first one was obtained when the sample was put in, the second one was obtained when washing buffer 2 was put in, and the third was obtained when the eluting buffer was passed through (fig. 7).

Enzyme assays

Peak 2 fractions were pooled together and called fraction 1 (OD_{280} 0.115). Peak 3 was divided into two and the first part of the peak pooled together and called fraction 2 (OD_{280} 0.156). Fractions from the second part of the peak were pooled together and called fraction 3 (OD_{280} 0.083). BAPNA assays were carried out on 2ml of these fractions. Fraction 1 (~2.30mg protein) showed enhanced enzymic activity (39%) whereas fraction 2 (~3.12mg protein) and fraction 3 (~1.66mg protein) showed inhibition of papain (33.5% & 48.9% respectively) (Fig. 8, table 4a & 4b). On TCA precipitation and polyacrylamide gel electrophoresis, however, it was found that peaks 2 and 3 contained a number of proteins including the one obtained through the ion exchange column (fig.9, fig. 10, table 5). A 52 000 m.wt. protein band was clearly seen. As there was no more material to further purify, the attempt was stopped there.

Just as a further confirmation to make sure that there was inhibitory activity in the extract before pouring through the column, assays was done on some of this solution dissolved in 10mM Tris-HCl donated by Mr A. Remfry. Inhibition was observed for the different volumes used of 0.5ml (32.9%), 1ml (96.0%) and 1.5ml (98.2%) (fig 11, table 6a & 6b).
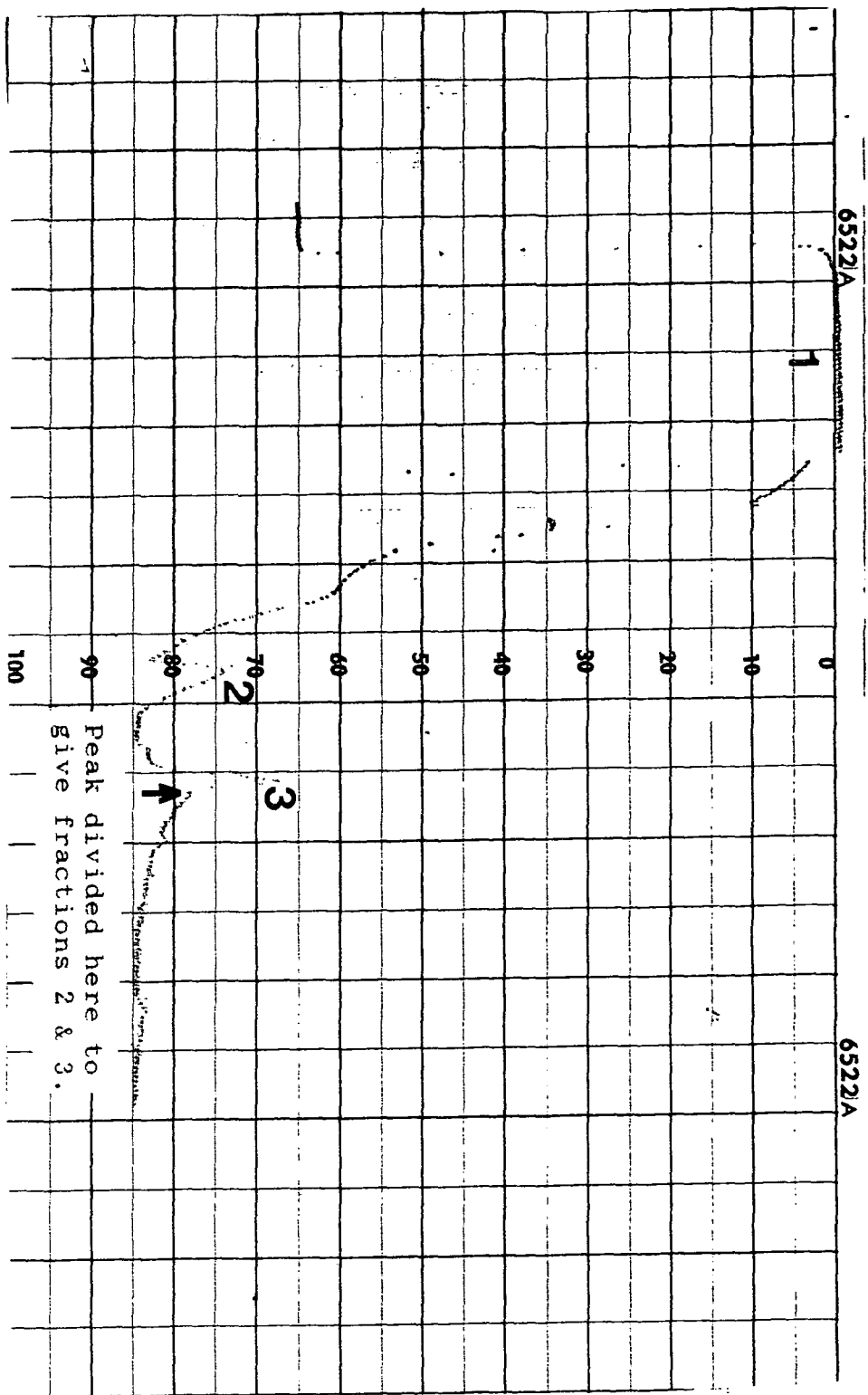
Fig. 7: Papain column - absorbance readings
of fractions at 280nm.

70

Papain affinity column – enzyme assays

Absorbance at 410nm

Time/min
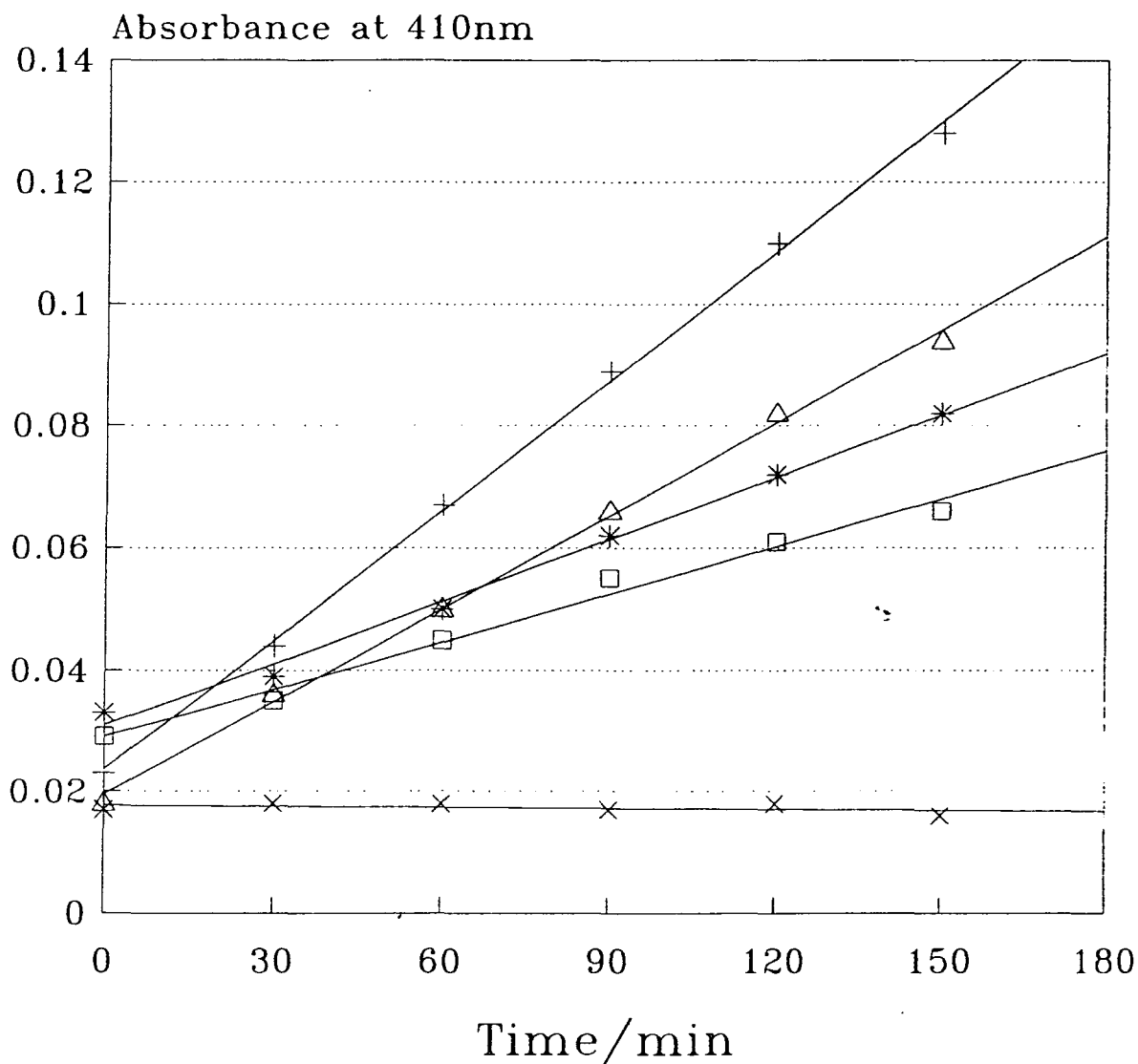
△ Cont1   + Peak1   ＊ Frac2
□ Frac3   ✕ Cont2

Fig 8: Papain affinity column - BAPNA assay rate
graph of 2ml of each fraction.

Tube no.	Description
1	Control 1, with papain but without inhibitor.
2	2ml of fraction 1 from papain column.
3	2ml of fraction 2 from papain column.
4	2ml of fraction 3 from papain column.
5	Control 2, without papain or inhibitor.

Table 4a: Reaction mixtures in fig. 8.

\ Sample no Time abs.	1	2	3	4
0	0.018	0.023	0.033	0.029
30	0.036	0.044	0.039	0.035
60	0.050	0.067	0.050	0.045
90	0.066	0.089	0.062	0.055
130	0.082	0.110	0.072	0.061
150	0.094	0.128	0.082	0.066
Rate (10units/min)	4.93	6.87	3.28	2.52

Regression Output:

Constant	0.019852	0.024113	0.031149	0.029165
Std Err of Y Est	0.001772	0.002563	0.001984	0.001952
R Squared	0.996922	0.996689	0.991349	0.985870
No. of Observation	6	6	6	6
Degrees of Freedom	4	4	4	4
X Coefficient(s)	0.000493	0.000687	0.000328	0.000252
Std Err of Coef.	0.000013	0.000019	0.000015	0.000015

Table 4b: Absorbance of BAPNA assays at 410nm
& regression output of lines obtained.
From rates of change in absorbance, it
can be seen that tube 2 enhances enzymic
activity by 39% whereas tube 3 & 4 inhibit
papain activity by 33.5% and 48.9% respec-
tively (contains ~2.30mg, 3.12mg & 1.66mg
of protein respectively).

Marker size/m.wt

66000
45000
36000
29000
24000

a
b

1 2 3 4 5
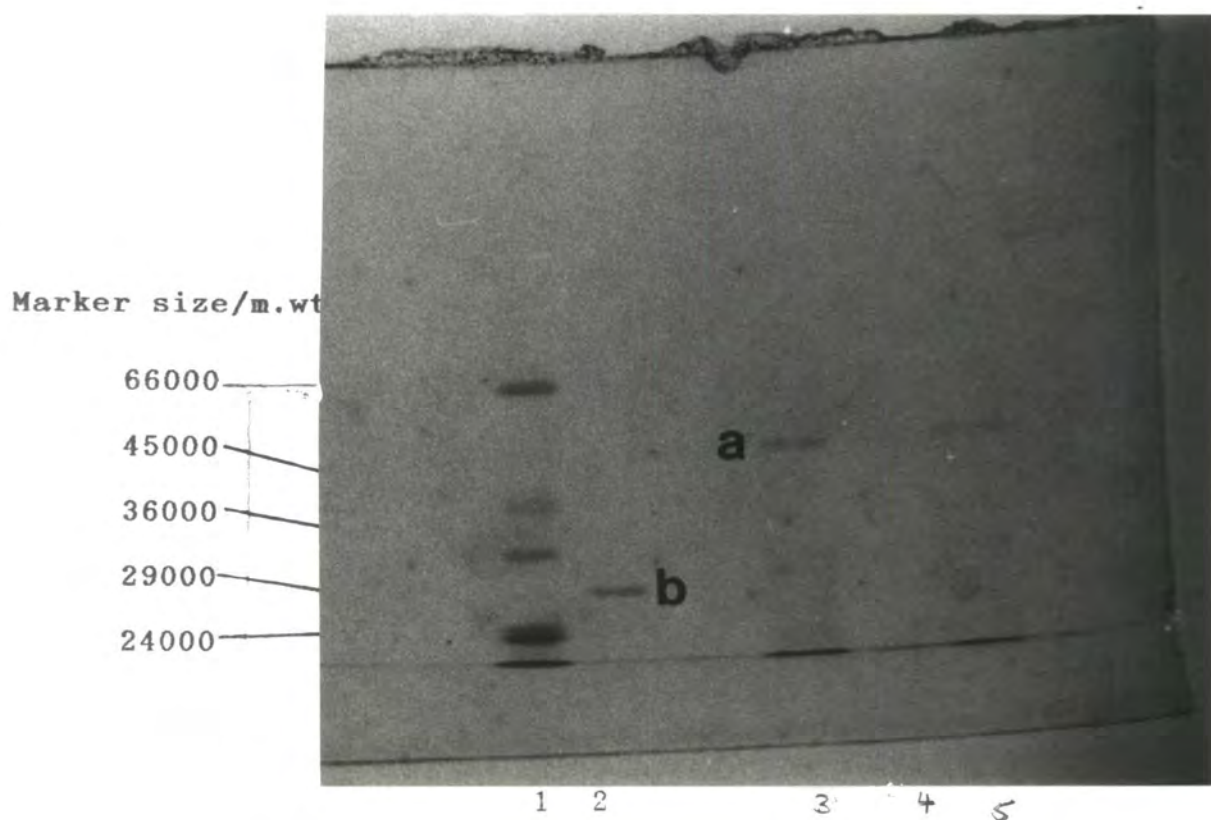
Well no.	Sample
1	Sigma molecular wt. markers.
2	Protein off the Sephacryl 200 gel filtration column after TCA precipitation.
3,5	Protein off the papain column from peak 3, fraction 2.
4	Protein off the papain column from peak 3, fraction 3.

Fig 9: Polyacrylamide gel electrophoresis of protein from gel filtration column & papain affinity column. (bands not very clear in photograph).

73
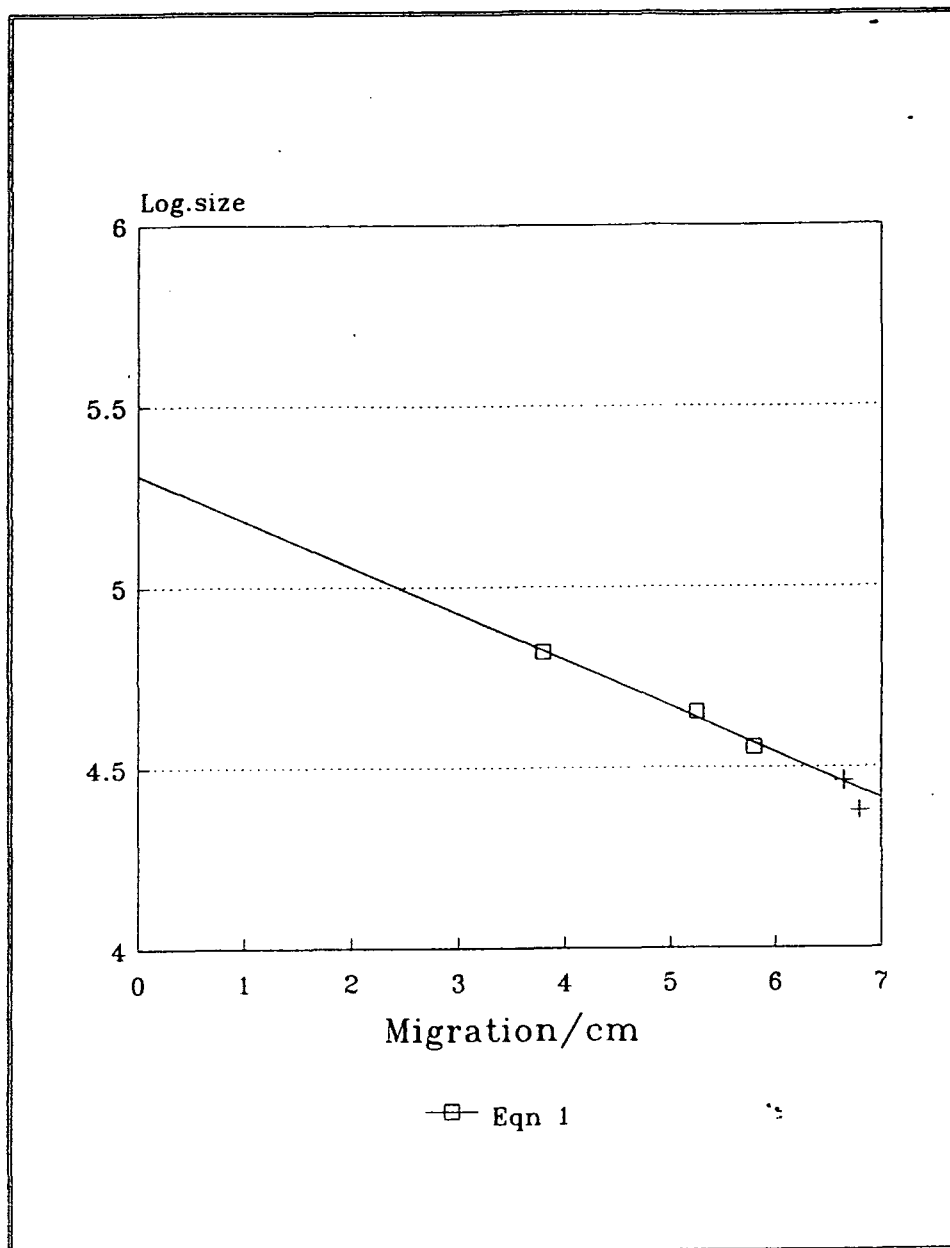
Fig. 10:
Graph of migration/cm vs log/m.wt. for marker in fig. 9.

Marker Migration	Size/mw	Log.size
3.80	66000	4.8195
5.25	45000	4.6532
5.80	36000	4.5563
6.65	29000	4.4624
6.80	24000	4.3802

Band A Migration (major band)	Size/mw	Log.size
4.65	51857	4.7148

Band B Migration	Size/bp	Log.size
6.25	32339	4.5097

(Calculated using eqn 1)

Using first three points,
Regression Output:

Constant	5.310821
Std Err of Y Est	0.019376
R Squared	0.989409
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s)   -0.12817
Std Err of Coef. 0.013260
Eqn 1: y = -0.12817x + 5.310821

Table 5: *Migration distances & sizes of bands in fig. 9. Regression output of fig. 10.*

74

Fig. 11: Partially purified inhibitor before pouring through column - BAPNA assay rate graph.

Tube no.	Description
1	Control 1, with papain but without inhibitor.
2	0.5ml of extract before pouring through any column.
3	1ml of extract before pouring through any column.
4	1.5ml of extract before pouring through any column.
5	Control 2, without papain but with 1ml inhibitor.

Table 6a: Reaction mixtures in fig. 11.

\ Sample no Time abs.	1	2	3	4	5
0	0.017	0.037	0.052	0.067	0.055
30	0.044	0.051	0.060	0.069	0.057
60	0.070	0.071	0.072	0.075	0.060
90	0.095	0.087	0.082	0.080	0.063
120	0.118	0.104	0.092	0.084	0.061
Rate (10^{-4}units/min)	8.43	5.66	0.34	0.15	0.00

Regression Output:

Constant	0.036	0.0512	0.066	0.0182	0.0556
Std Err of Y Est	0.001414	0.000966	0.001080	0.001402	0.001673
R Squared	0.997928	0.997315	0.983009	0.999079	0.794117
No. of Observation	5	5	5	5	5
Degrees of Freedom	3	3	3	3	3
X Coefficient(s)	0.000566	0.00034	0.00015	0.000843	0.00006
Std Err of Coef.	0.000014	0.000010	0.000011	0.000014	0.000017

Table 6b: Absorbance of BAPNA assays at 410nm & regression output of lines obtained. 0.5ml extract inhibited papain activity by 32.9%, 1ml by 96% and 1.5ml by 98.2%.

3.2: FINDING OUT THE GENE SEQUENCE OF A BROMELAIN INHIBITOR FROM PINEAPPLE

Pineapple stem acetone powder is a rich source of compositionally similar but chromatographically distinct polypeptide inhibitors of bromelain. The isoinhibitors have molecular weights of 5 600, and they contain five disulfide bonds and about 50 amino acids each (Perlstein et al., 1973). Primary structural analysis of one of the seven inhibitor fractions (VII) showed extensive microheterogeneity (Reddy et al., 1975). At least five isoinhibitors were present, each composed of two peptide chains joined by disulfide bonds. These chains, designated A and B on the basis of size, comprise 41 and 10 - 11 residues respectively. On the basis of ionization properties and yields of the A and B chains, one of the major inhibitor species in Fraction VII is the covalently linked complex of the two chains shown:

A Chain (A-1): H$_2$N-Asp-Glu-Tyr-Lys-Cys-Tyr-Cys-Ala-Asp-Thr-Tyr-
                              1   5   10

                     Ser-Asp-Cys-Pro-Gly-Phe-Cys-Lys-Lys-Cys-Lys-Ala-
                            15   20

                     Glu-Phe-Gly-Lys-Tyr-Ile-Cys-Leu-Asp-Leu-Ile-Ser-
                            25   30   35

                     Pro-Asn-Asp-Cys-Val-Lys-COOH
                            40

B Chain (B-2): H$_2$N-Thr-Ala-Cys-Ser-Glu-Cys-Val-Cys-Pro-Leu-Gln-
                            1   5   10

                     COOH

The complete amino acid sequence of each of the five components
have been determined and the differences observed between the
structures were accounted for by single substitiutions at four
sites in the sequence.

It was attempted to find out the gene sequence of the inhi-
bitor shown above. First, DNA from young tissue cultured pine-
apple plants was extracted. Then oligonucleotides complementary
to the inhibitor's A-1 and B-1 ends were constructed. It is pos-
tulated that both protein chains are coded for on a single strand
of DNA. DNA coding for the chains are separated by DNA coding
for the proper folding and cleavage of the protein. The con-
structed oligonucleotides were as shown:

```
A-1 N-terminus (179): 5' -CGGAAGCTTGATGAATATAAATGTTATTGTGC
                                   C  G  C  G  C  C  C

A-1 C-terminus (180): 5' -CGGAAGCTTTTTTACACAATCATTGG
                                   C     G  G  G

B-1 N-terminus (181): 5' -CGGAAGCTTACIGCITGTTCIGAATGTGT
                                        C     G  C

B-1 C-terminus (182): 5' -CGGAAGCTTTGIAAIGGACAIACACATTC
                                   G  G     G  C
```

where I = inosine, which will base pair with A ,G ,C or T.
HindIII sites were attached at the beginning of each oligo (first
nine bases) to simplify ligation into a vector plasmid at a later
stage. These oligos were combined together in pairs with the
pineapple DNA and a polymerase chain reaction carried out to
amplify the inhibitor gene. This amplified DNA was then cleaved
with HindIII, as short overhangs (1-2bp) are usually generated by
PCR, and inserted into pUC18 cleaved with a similar enzyme (see
Appendix 2, A2.1, A2.2, for the nucleotide sequence and restric-

tion sites of pUC18, Yanish-Perron et al., 1985). Competent *Escherichia coli* bacterial cells (DH5α) were transformed with this ligation mixture. The plasmids from the transformants containing the correct gene were cleaved with AluI. The correct fragment of DNA was then ligated to M13mp19 vector, using its HincII for insertion (see Appendix 2, A2.3 & A2.4, for the nucleotide sequence and restriction sites of M13mp19, Yanish-Perron, et al., 1985). Restriction with PstI and KpnI showed that certain clear plaques contained the gene. Hence one of these with the gene inserted in the HincII site was chosen and the single stranded DNA sequenced.

3.2.1: Polymerase Chain Reaction

An annealing temperature of 42°C was first tried. There was no amplified band unique to a certain pair of oligos (fig. 12), hence the experiment was repeated at a higher temperature of 50°C. This time round, the non-specific DNA that had been amplified earlier was not amplified. However, there were two distinct bands representing the amplified products of oligos 180 and 181 (fig. 13, fig. 14, table 7). One band (a) represented a size of approximately 160bp which was just the appropriate size for the gene. The other band (b) represented a size of approximately 390bp. This could be a tandem repeat of the gene. A clear band (c) representing approximately 60bp was present in all tracks. This could have been dimers of the oligos. Another amplification at 46°C was carried out. The previous amplification containing the DNA of interest was also reamplified at this temperature. No non-specific DNA was amplified at this temperature, but the pre-

Marker size/bp

1408
517
396
225
75
65

1 2 3 4 5 6
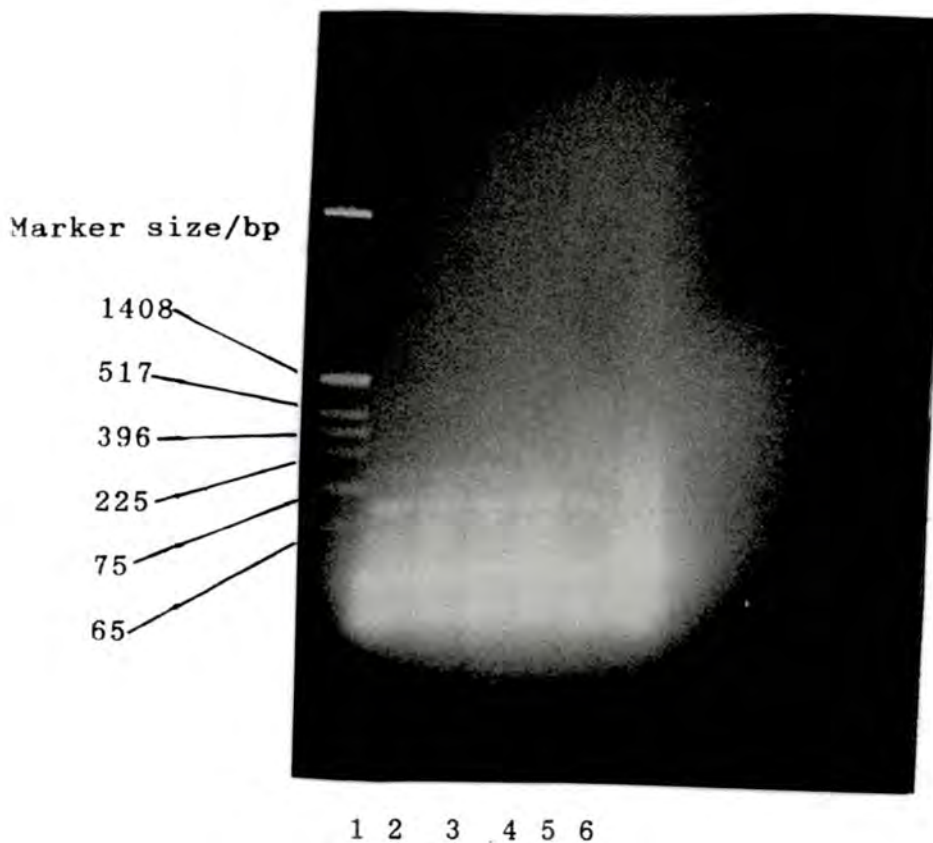
Well no.	Sample
1	Marker consisting of pUC19 cleaved with HinfI.
2,3	Amplification at 42°C using 180 and 181 oligonucleotides.
4,5	Amplification at 42°C using 179 and 182 oligonucleotides.
6	Amplification at 42°C using 179 and 180 oligonucleotides.

Fig.12: PCR at 42°C. No unique amplified band for a certain pair of oligonucleotides.

Marker size/bp

1408	
517	
396	
225	
75	
65	

1 2   3 · 4 5
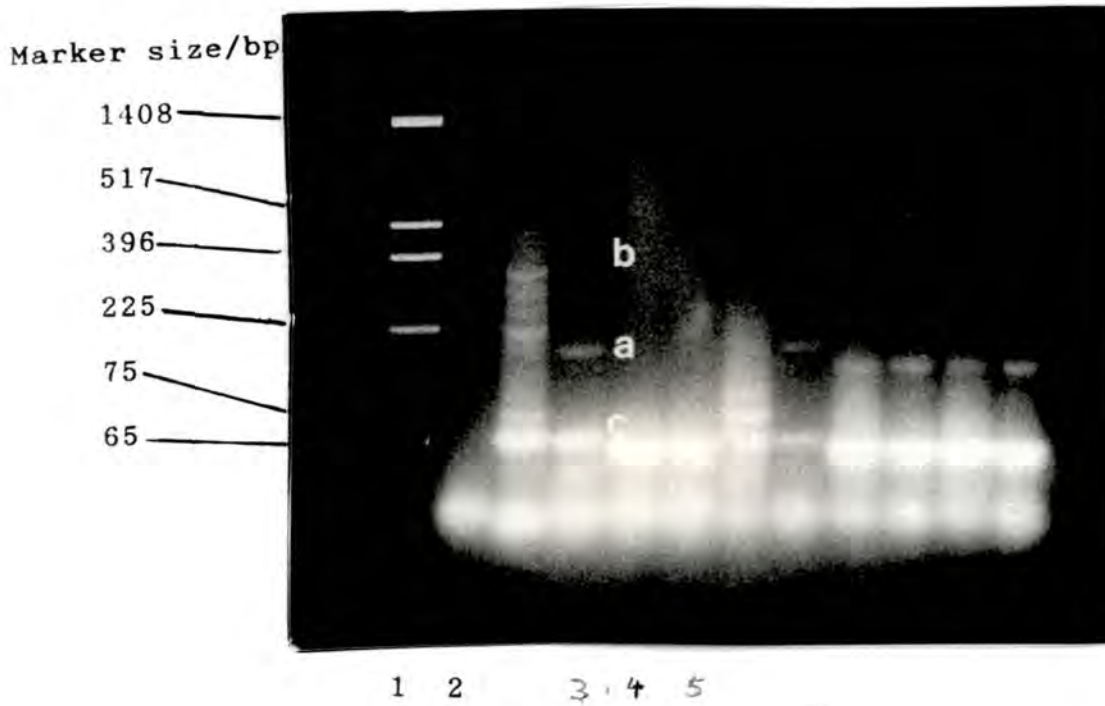
Well no.	Sample
1	Marker consisting of lambda DNA restricted with EcoRI and Hind III.
2	Oligonucleotides 179 and 180.
3	Amplification at 50°C using 180 and 181 oligonucleotides.
4	Amplification at 50°C using 179 and 182 oligonucleotides.
5	Amplification at 50°C using 179 and 180 oligonucleotides.

Fig. 13: PCR at 50°C. Oligonucleotides 180 & 181 give amplified bands of approximately 160bp (band a) & 390bp (band b) (well no.3).

81

Fig.14: Graph of migration/cm vs log/bp for marker in fig.13.
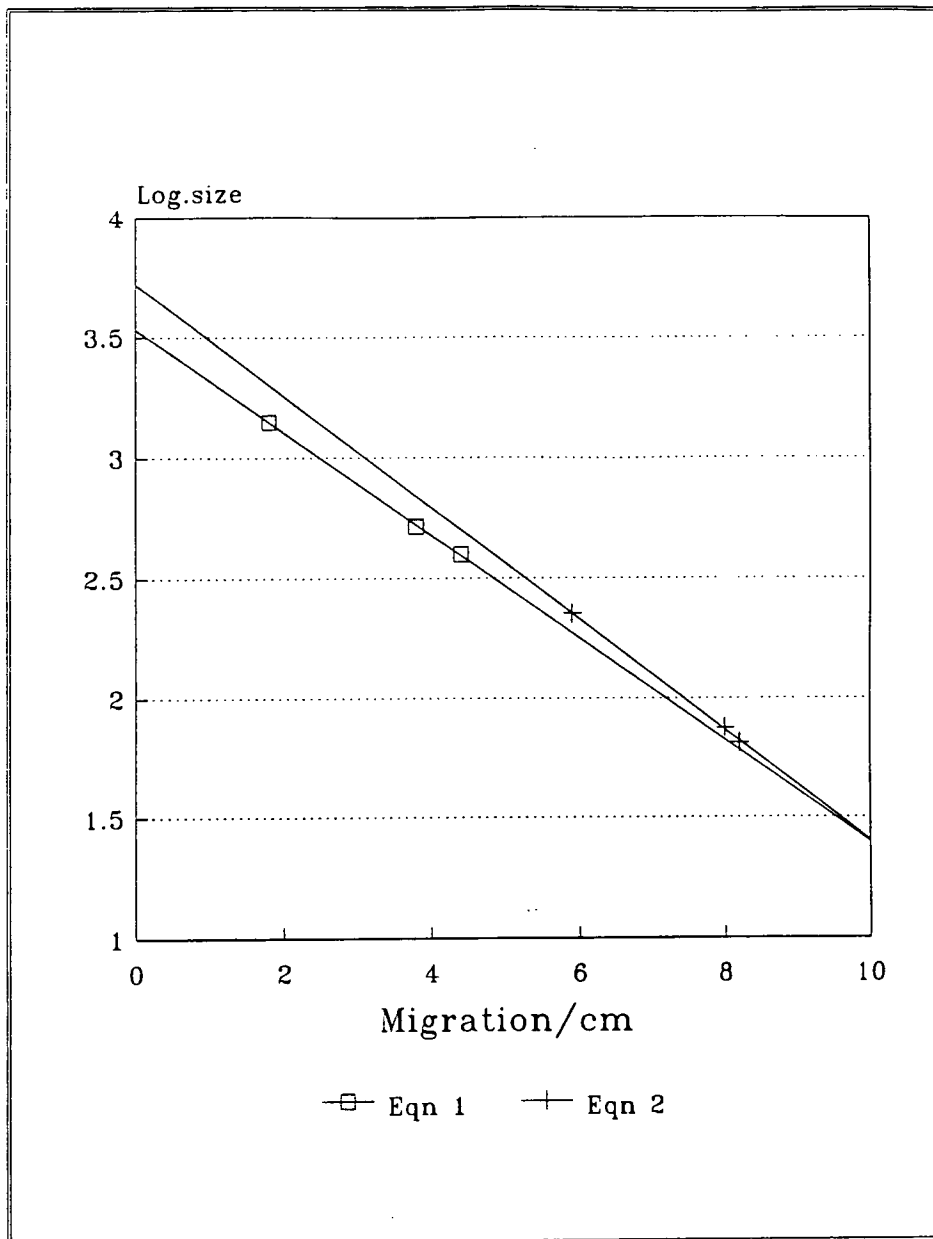
Marker Migration	Size/bp	Log.size
1.80	1408	3.1486
3.80	517	2.7135
4.40	396	2.5977
5.90	225	2.3522
8.00	75	1.8751
8.20	65	1.8129

	Migration	Size/bp	Log size
Band A	6.50	164	2.2140
Band C	8.30	63	1.7973
Oligos	9.90	27	1.4269
(Calculated using eqn 2)			
Band B	4.40	391	2.5924
(Calculated using eqn 1)			

Using first three points,
Regression Output:

Constant	3.530978
Std Err of Y Est	0.008839
R Squared	0.999536
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s)   -0.21331
Std Err of Coef.   0.004591
Eqn 1: y = -0.21331x + 3.530978

Using last three points,
Regression Output:

Constant	3.718688
Std Err of Y Est	0.011242
R Squared	0.999274
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s)   -0.23148
Std Err of Coef.   0.006239
Eqn.2: y = -0.23148x + 3.718688

Table 7: Migration distances & sizes of bands in fig. 13. Regression output of fig. 14.

viously amplified DNA was amplified even more. However, the 390bp band was lost (fig. 15, fig. 16, table 8). The amplified DNA from another PCR was digested with Hind III and run on a gel. The 160bp and the 390bp bands were cut out from the gel and the DNA electroeluted out.

3.2.2: Cloning in pUC18

pUC18 was used as the vector to clone the fragments in *E. coli*. The vector was first cut with HindIII and ligated to the fragments. Commercial competent cells (DH5α) were transformed with this ligation mix and plated on YT-X-gal-amp plates. Plasmid DNA from white recombinant colonies were isolated by minipreps and tested for the presence of the correct fragment using a variety of restriction enzymes, for example, EcoRI (fig. 17, fig. 18, table 9). Recombinant no. 2 contained an insert. HindIII was unable to retrieve the fragments from the transformed bacteria (fig. 19), hence Alu I (fig. 20, fig. 21, table 10) was used. As AluI cuts at the same site as HindIII, and a band corresponding to a fragment of approximately 210bp was obtained, this band was cut out and the DNA electroeluted to be cloned into M13mp19 for easy gene sequencing by the Sanger method (Sanger, 1979).

3.2.3: Cloning in M13mp19

M13mp19 vector DNA was prepared for insertion of the fragment by restricting with Hinc II. This enzyme gives blunt ended fragments, hence the 210bp fragment cleaved by AluI, which was also blunt ended was inserted into this site. A commercial .PN97 strain of *E. coli*, TG 2, was transformed with this ligation mix. Clear plaques indicated transformed bacteria hence minipreps

Marker size/bp

1408

517

396

225

a

1 2 3 4
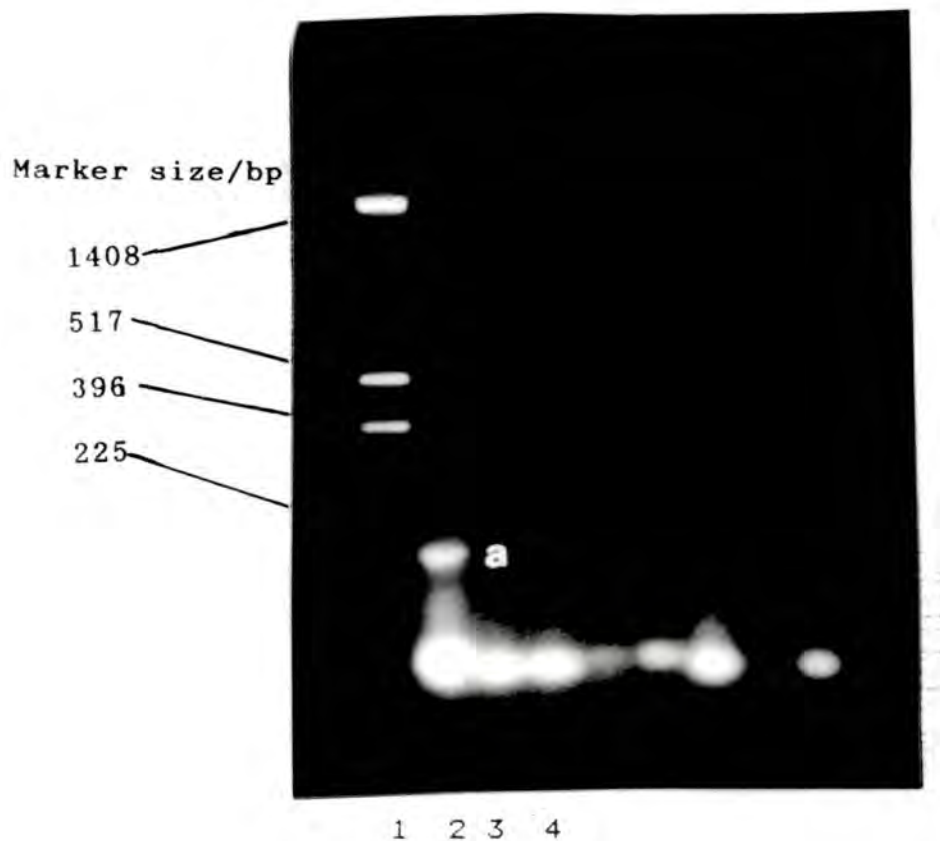
Well no.	Sample
1	Marker consisting of pUC 19 cleaved with Hinf I.
2	50°C amplification reamplified at 46°C.
3	Amplification at 46°C using 179 and 182 oligonucleotides.
4	Amplification at 46°C using 179 and 180 oligonucleotides.

Fig.15: PCR at 46°C. Reamplified sample from fig. 13, well no. 3, gave a band of approximately 180bp (band a).

Fig.16:
Graph of
migration/cm
vs log/bp for
marker in fig.15.
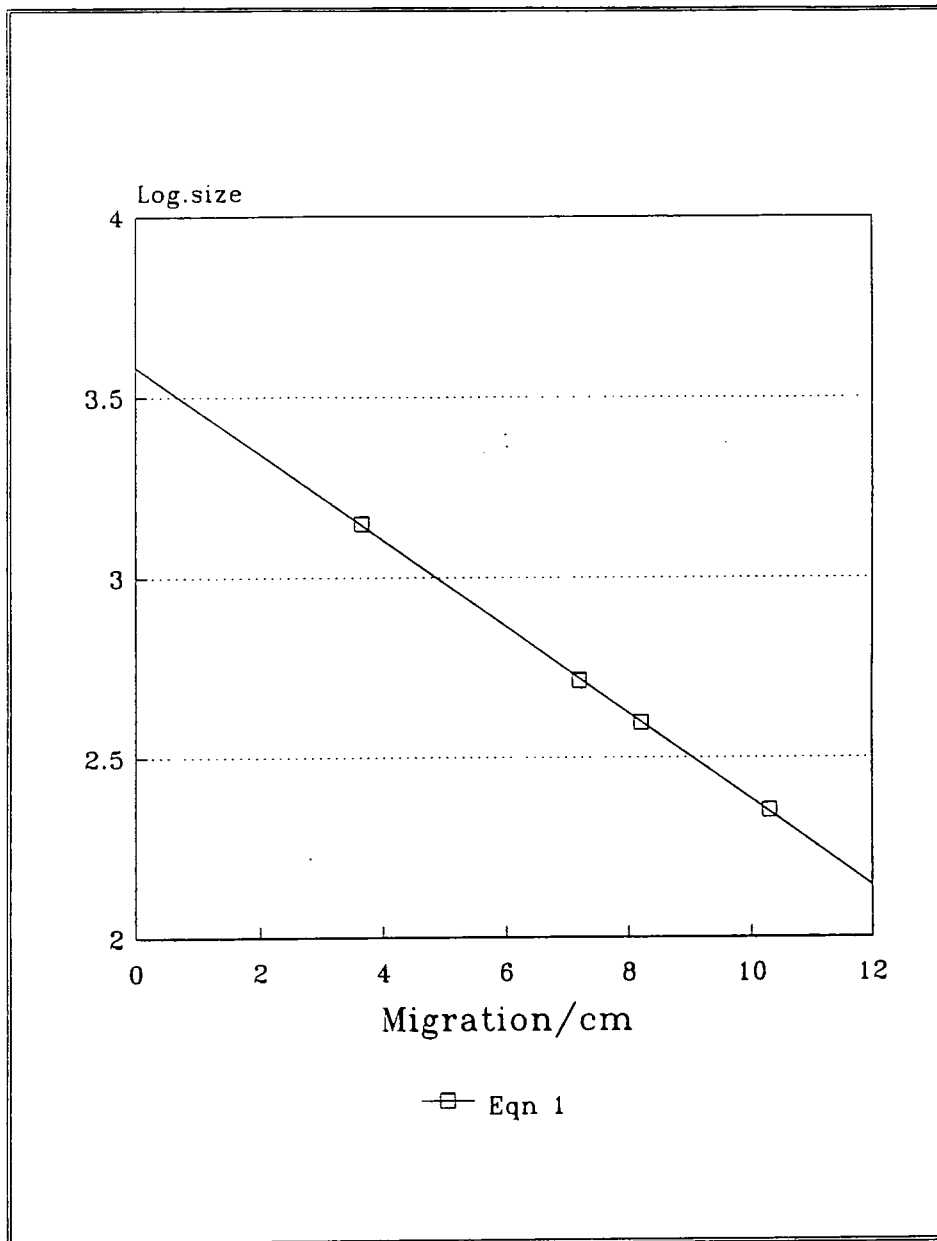
Marker Migration	Size/bp	Log.size
3.65	1408	3.1486
7.20	517	2.7135
8.20	396	2.5977
10.30	225	2.3522

Band A migration	Size/bp	Log size
11.10	179	2.2518

(Calculated using equation 1)

Table 8:
Migration distances
& sizes of bands in
fig. 15. Regression
output of fig. 16.

Regression Output:

Constant	3.582953
Std Err of Y Est	0.005949
R Squared	0.999787
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s)   -0.11992
Std Err of Coef.   0.001236

Eqn 1: $y = -0.11992x + 3.582953$

85

Marker size/bp

2027
1904
1584
1330
983
831
564
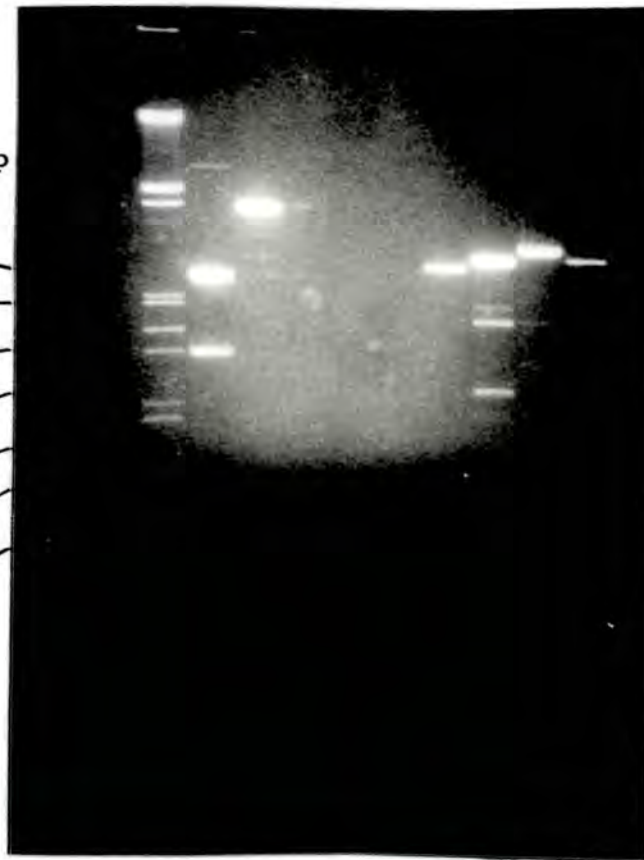
5   4 3 2 1

Well no.	Sample (cleaved with EcoRI)
1	Control, pUC 18 without insert.
2	pUC 18 from recombinant no.2.
3	pUC 18 from recombinant no.3.
4	pUC 18 from recombinant no.4.
5	Marker consisting of lambda DNA restricted with EcoRI and Hind III.

Fig.17: Restriction with EcoRI of pUC18 plasmids
from transformants shows that recombinant
no.2 (in well no. 2) contains an insert.

86

Log.size

Migration/cm

□ Eqn 1   + Eqn 2
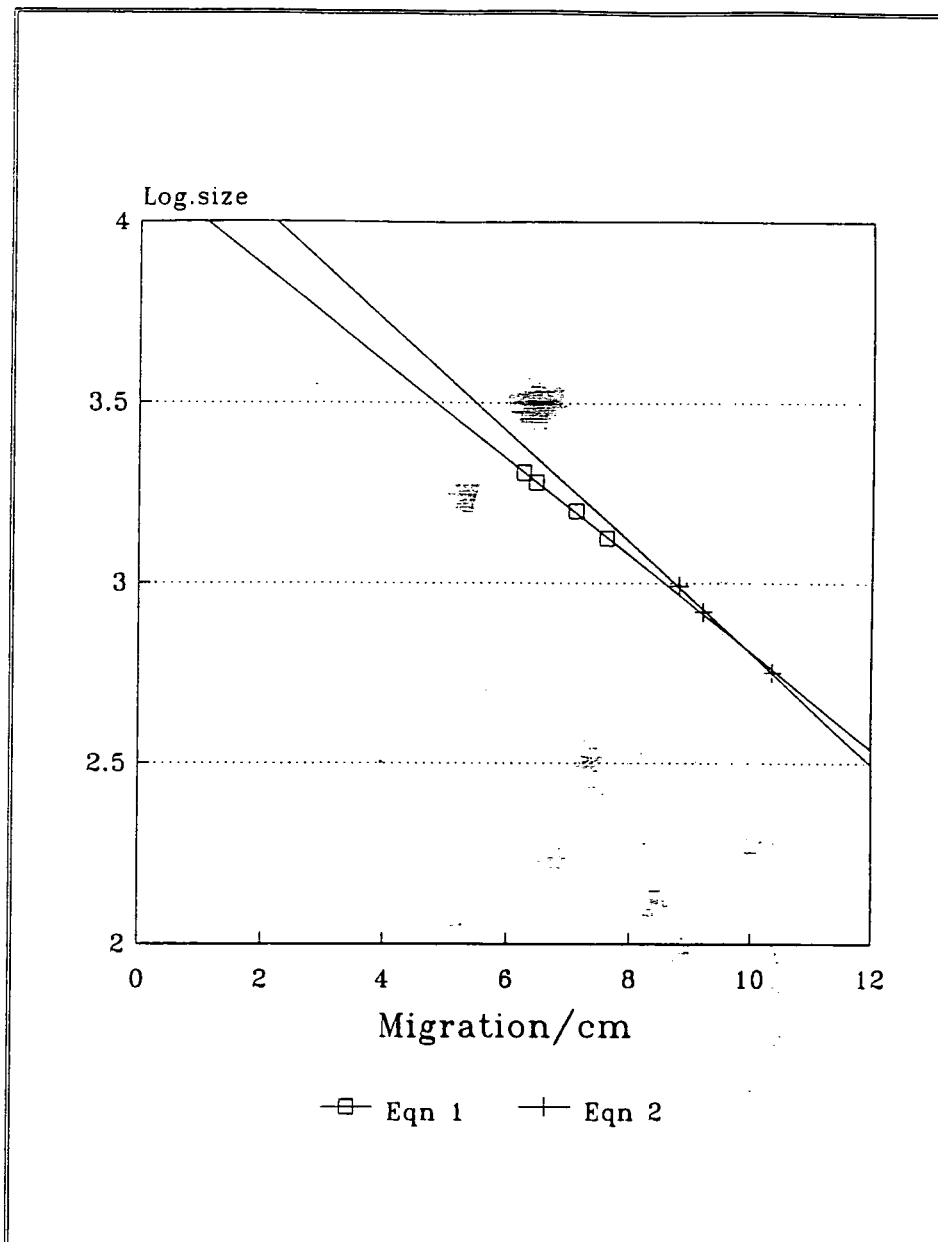
Marker Migration	Size/bp	Log.size
6.25	2027	3.3069
6.45	1904	3.2797
7.10	1584	3.1998
7.60	1330	3.1239
8.80	983	2.9926
9.20	831	2.9196
10.35	564	2.7513

Well no.	Heaviest band migration	Size/bp	Log size
1	5.40	2639	3.4215
2	5.15	2850	3.4549
3	5.40	2639	3.4215
4	5.55	2520	3.4014

(Calculated using eqn 1)

Fig.18:
Graph of migration/cm vs log/bp for marker in fig.17.

Table 9:
Migration distances & sizes of bands in fig. 17. Regression output of fig. 18.

Using first four points,
    Regression Output:

Constant	4.143765
Std Err of Y Est	0.004801
R Squared	0.997754
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s)   -0.13375
Std Err of Coef. 0.004487
Eqn. 1: y = -0.13375x + 4.143765

Using last three points,
        Regression Output:

Constant	4.339308
Std Err of Y Est	0.008405
R Squared	0.997693
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s)   -0.15359
Std Err of Coef. 0.007385
Eqn 2: y= -0.007385x + 4.339308

87
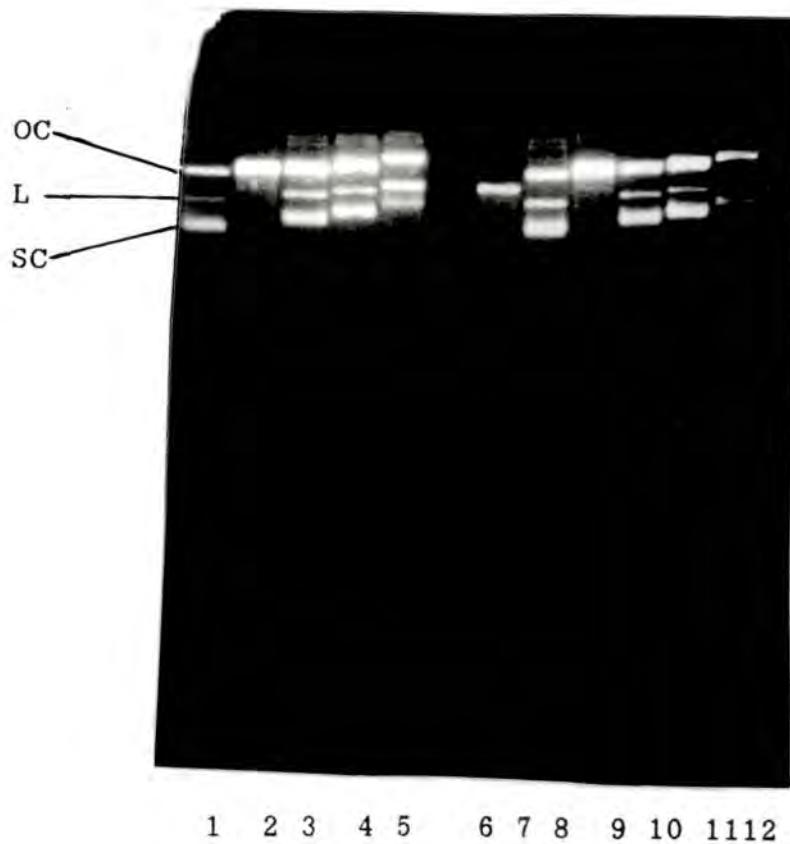
Well no.	Sample
1-5,7-11	pUC18 from transformants restricted with HindIII.
6,12	pUC18 restricted with HindIII.

Fig.19: Restrictions with HindIII of pUC18 plasmids from transformants were only partial. The 3 bands for each track represent the supercoiled form (SC), the linear form (L) and the open circular form (OC).

88

Marker size/bp

2027
1904
1584
1330
983
831
564

1 2 3

←a
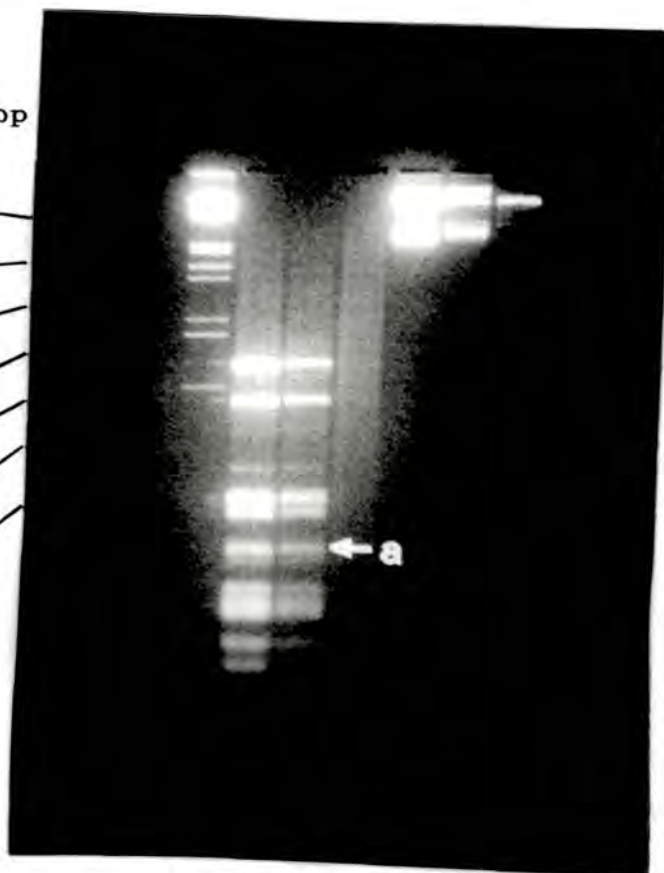
Well no.	Sample
1	Marker consisting of lambda DNA restricted with EcoRI and Hind III.
2-3	pUC 18 from recombinant no.2 restricted with Alu I.

Fig. 20: Restriction with AluI of plasmid from
recombinant no.2. DNA from band a (~ 215bp)
was cloned into pUC18.

Log.size vs Migration/cm

Eqn 1   Eqn 2
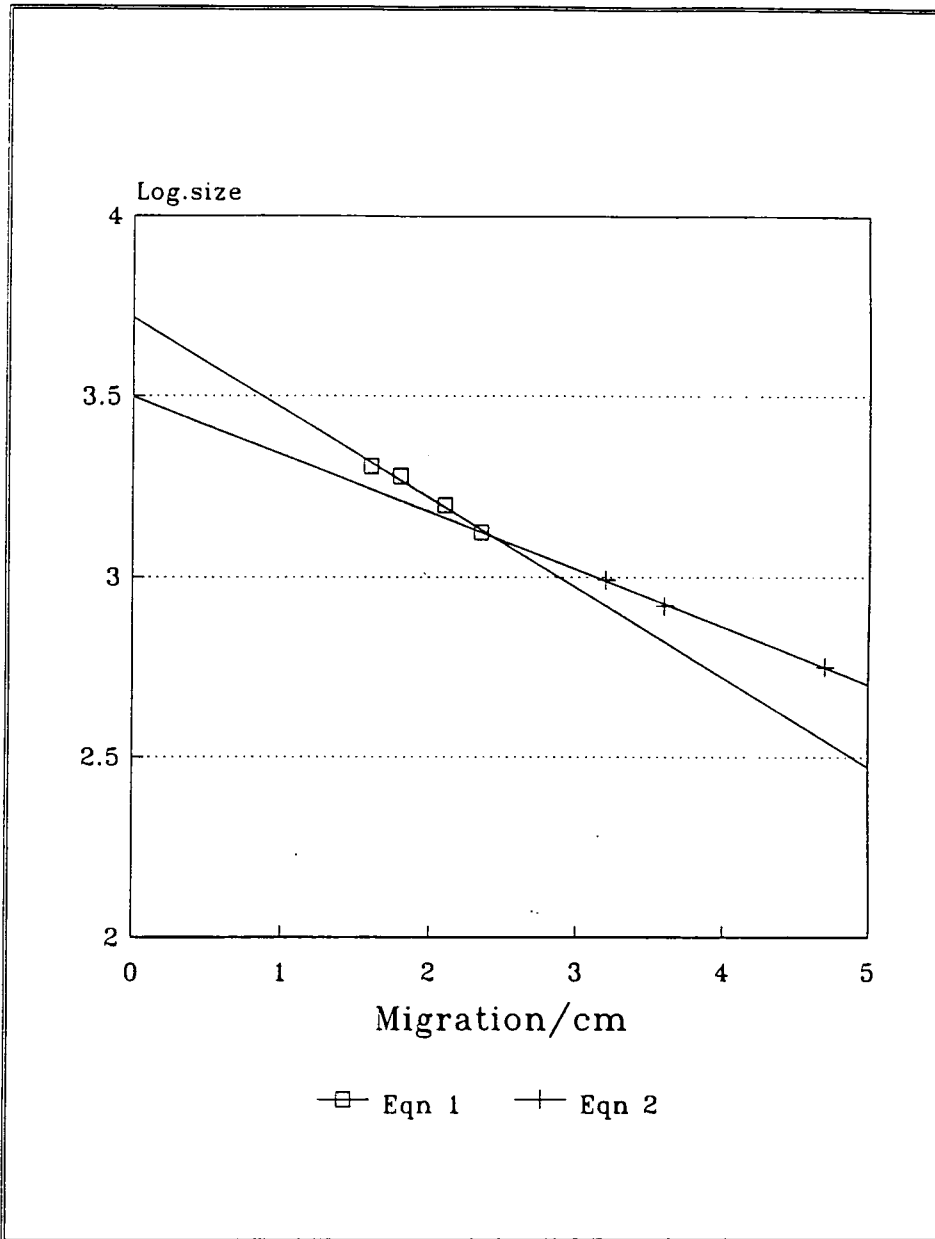
Marker Migration	Size/bp	Log.size
2.35	2027	3.3069
2.55	1904	3.2797
3.00	1584	3.1998
3.35	1330	3.1239
4.45	983	2.9926
4.85	831	2.9196
6.10	564	2.7513

Band A Migration	Size/bp	Log.size
9.00	215	2.3331

(Calculated using eqn 2)

Fig.21:
Graph of migration/cm vs log/bp for marker in fig.20.

Table 10:
Migration distances & sizes of bands in fig. 20. Regression output of fig. 21.

Using first four points,
        Regression Output:

Constant	3.743377
Std Err of Y Est	0.007613
R Squared	0.994353
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s)   -0.18341
Std Err of Coef.   0.009773
Eqn 1: y = -0.18341x + 3.743377

Using last three points,
        Regression Output:

Constant	3.624247
Std Err of Y Est	0.011318
R Squared	0.995817
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s)   -0.14346
Std Err of Coef.   0.009298
Eqn 2: y = -0.14346x + 3.624247

90

from these were tested for the correct fragment by restricting with a mixture of Pst I and Kpn I. Most of the transformants gave a cleaved fragment of a size corresponding to 190bp (fig. 22, fig, 23, table 11). Single stranded DNA was prepared from one of these. The DNA was suspended in 100μl of TE buffer and the OD at 260nm of 1u of this in TE buffer measured. An absorbance of 0.010 was obtained, implying approximately 40μg of purified DNA. As the OD at 280nm was 0.016, this meant that the DNA was not pure nd the estimation of 40μg not reliable. Hence it was reconcentrated to 4μg/μl for sequencing by the automatic DNA sequencer.

3.2.4: The Gene Sequence Obtained

The gene sequence obtained was 169 bases long (fig. 24). It did not correspond to the bromelain inhibitor gene sought after. The last 75 bases of the fragment corresponded to the 987th to 1061st bases of pUC18 (its kanamycin-resistance gene). There was an additional TAG sequence before the M13 sequence started. Before these 78 bases, 91 bases of an unknown gene sequence was present. 13 out of 19 bases in the beginning part of the sequence showed homology to the 181 oligonucleotide mix and 10 out of 17 bases in the ending part of the sequence showed homology to the 180 oligonucleotide mix. Further experiments were done to find out the source of contamination.

3.2.5: Source of contamination

It was strongly suspected that the source of contamination came from the pineapple template DNA. A Southern blot was done of a gel containing a track of PCR-amplified material. Controls were pUC18 with and without the insert. This gel also clearly showed

Marker size/bp

2027
1904
1584
1330
983
831
564
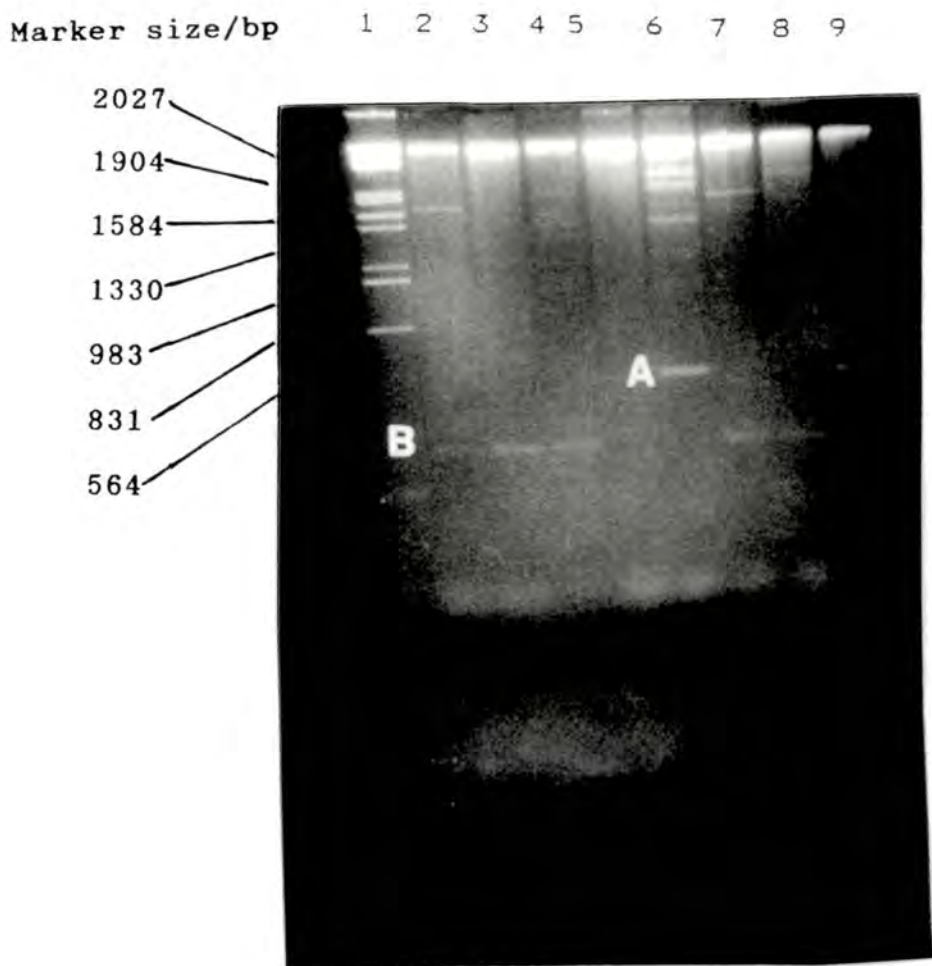
Well no.	Sample
1	Marker consisting of lambda DNA restricted with EcoRI and Hind III.
2-4 6-9	M13 plasmid from recombinants with inserts at Hinc II site and cleaved with Pst I and Kpn I.
5	Control, plasmid from blue colony with no insert cleaved with Pst I and Kpn I.

Fig. 22: Restriction with PstI & KpnI of M13 plasmids from transformants. Band B (~190bp) contained the correct sized insert and was sequenced.

Log.size

Migration/cm

Eqn 1   Eqn 2
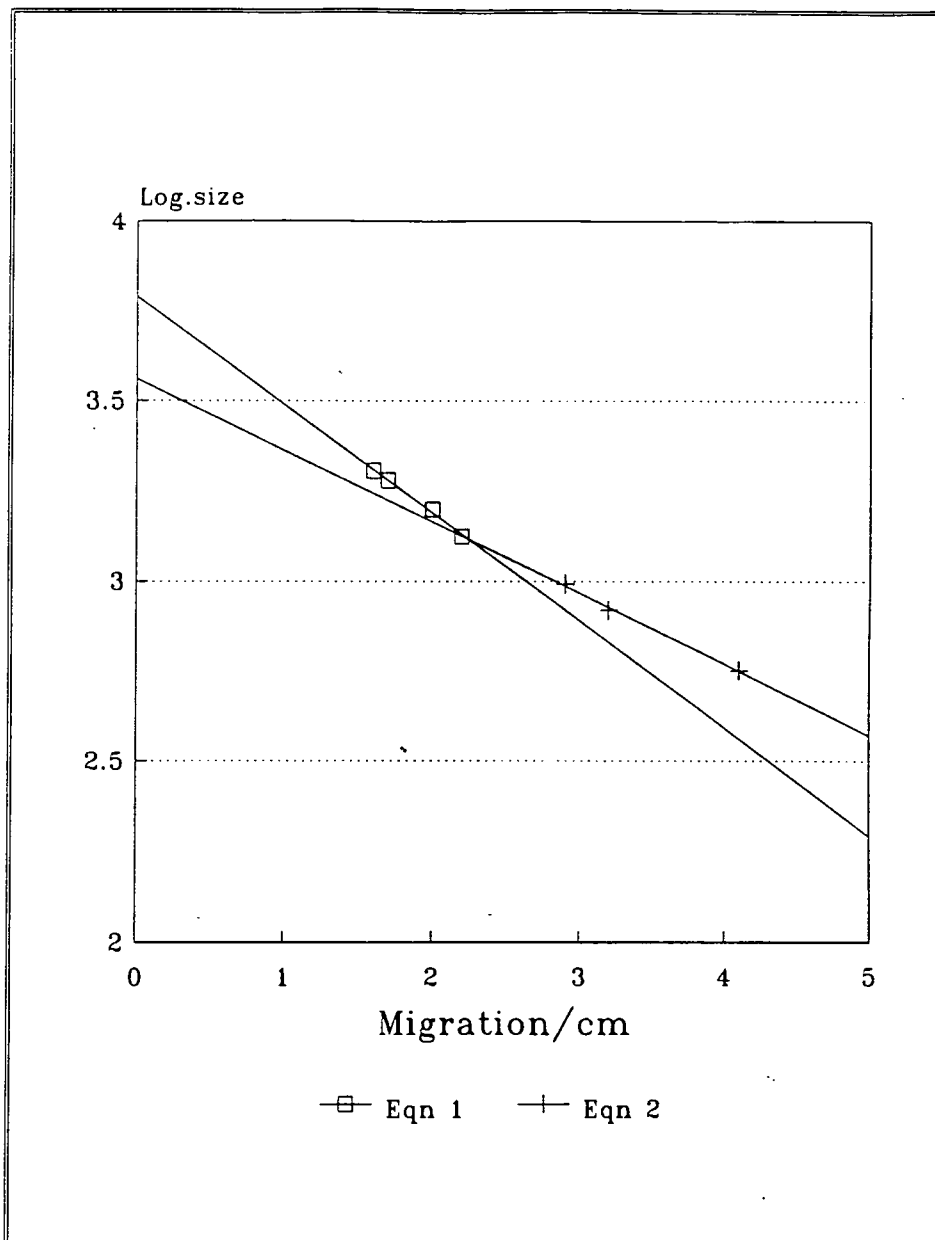
Fig.23:
Graph of migration/cm vs log/bp for marker in fig.22.

Marker Migration	Size/bp	Log.size
1.60	2027	3.3069
1.70	1904	3.2797
2.00	1584	3.1998
2.20	1330	3.1239
2.90	983	2.9926
3.20	831	2.9196
4.10	564	2.7513

Insert A Migration	Size/bp	Log.size
5.10	356	2.5515

Insert B Migration	Size/bp	Log.size
6.50	188	2.2746

(Calculated using eqn 2)

Table 11:
Migration distances & sizes of bands in

Using first four points,
Regression Output:

Constant	3.788839
Std Err of Y Est	0.008452
R Squared	0.993040
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s) -0.29936
Std Err of Coef. 0.017720
Eqn 1: y = -0.29936x + 3.788839

Using last three points,
Regression Output:

Constant	3.560407
Std Err of Y Est	0.009910
R Squared	0.996792
No. of Observations	3
Degrees of Freedom	1

X Coefficient(s) -0.19782
Std Err of Coef. 0.011221
Eqn 2: y = -0.19782x + 3.560407

*Fig. 24: Gene sequence of the DNA fragment
inserted into M13mp19.*

3'- CTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAG
        10   20   30   40   50   60

TCACGACGTTGTAAAACGACGGCCAGTGCCA - 5'
     70   80   90

*Fig. 24a: The first 91 bases of the DNA fragment.
Source not known.*

3'- CTCTCCTGTTCCGACCCTTCCGCTTACCGGATACCTGTGCGCCTTTTTCCCTTCGGGAA
        100   110   120   130   140   150

GCGTGGCGCTTTCTCATGA - 5'
    160

*Fig. 24b: The last 78 bases of the DNA fragment.
Bases no. 92 to 166 correspond to
bases no. 987 to 1061 of pUC18.*

that the insert in pUC18 was the PCR-amplified product (fig. 25).
The Southern blot hybridization and subsequent autoradiograph was
not successful, hence it was attempted to repeat the whole
experiment using a fresh batch of pineapple template DNA. How-
ever, an amplification at 50°C failed to give any amplified prod
ducts. An amplification at 50°C, using 20ng of pUC18 as template
DNA was then attempted. The results were rather conclusive (fig.
26, fig. 27, table 12) as a 160bp band was present, hence proving
that the original pineapple DNA was contaminated with prokary-
otic DNA containing the kanamycin gene.
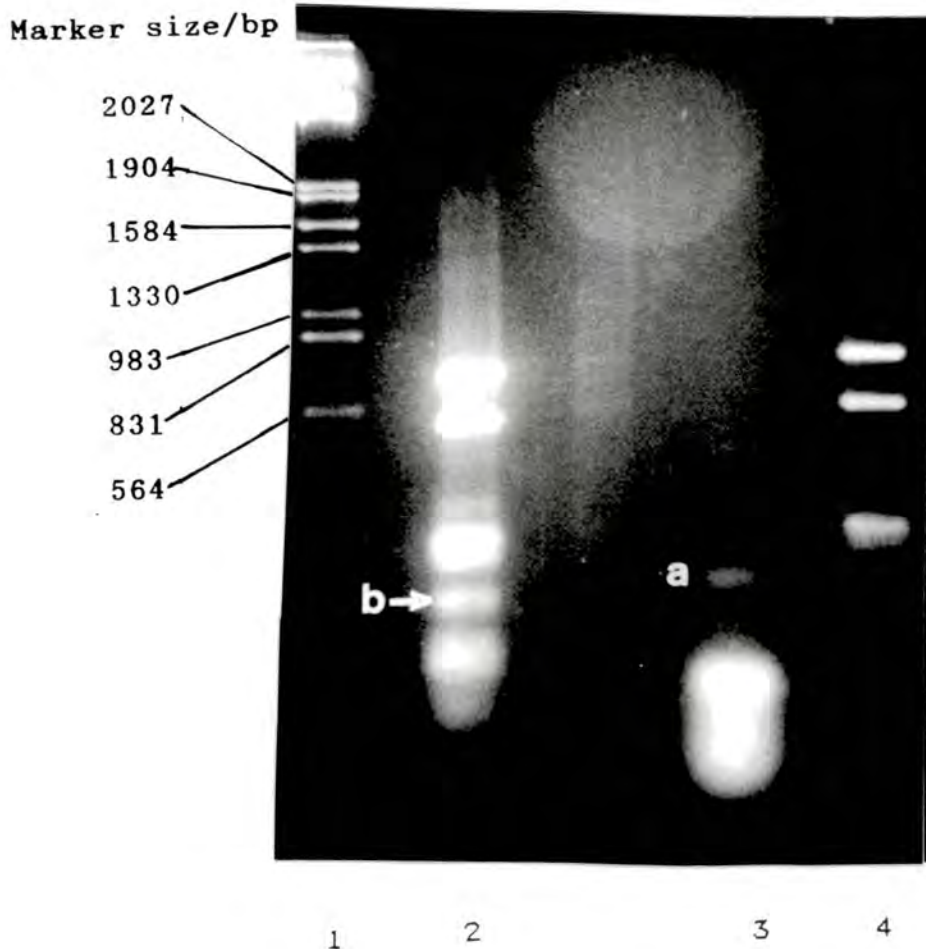
Well no.	Sample
1	Marker consisting of lambda DNA restricted with EcoRI and Hind III.
2	pUC 18 from recombinant no.2. restricted with Alu I.
3	Amplification at 50oC with oligonucleotides 180 and 181.
4	Control, pUC 18 restricted with Alu I.

Fig.25: Gel used for Southern blot. Band a & b are of approximately the same weight implying that the insert, b, and the amplified fragment, a, are the same fragment.

Marker size/bp

2027	
1904	
1584	
1330	
983	
831	
564	
125	

1   2

Well no.	Sample
1	Marker consisting of lambda DNA restricted with EcoRI and Hind III.
2	Amplification of pUC18 at 50°C using 180 and 181 oligonucleotides.

Fig. 26: PCR of pUC18 at 50°C. Band a is 160bp, which corresponds to the band obtained when pineapple DNA was was used as the template.

97

Log.size

Migration/cm

—■— Eqn 1   —+— Eqn 2
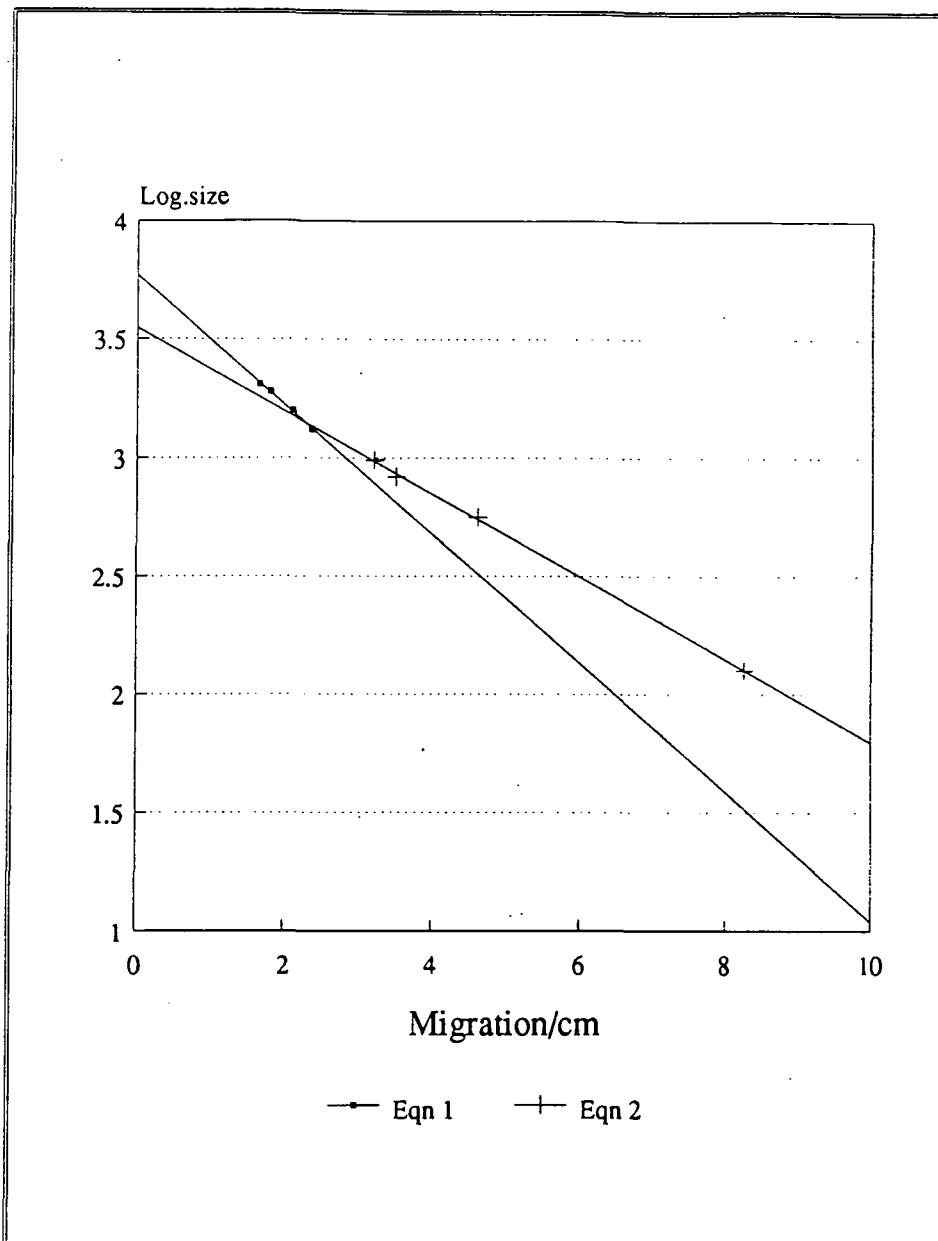
Marker Migration	Size/bp	Log.size
1.65	2027	3.3069
1.80	1904	3.2797
2.10	1584	3.1998
2.35	1330	3.1239
3.20	983	2.9926
3.50	831	2.9196
4.60	564	2.7513
8.25	125	2.0969

Band A Migration	Size/bp	Log.size
7.70	157	2.1957

(Calculated using eqn 2)

Fig.27:
Graph of migration/cm vs log/bp for marker in fig.26.

Using first four points,
Regression Output:

Constant	3.749171
Std Err of Y Est	0.007949
R Squared	0.993844
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s) -0.26412
Std Err of Coef. 0.014697
Eqn 1: y = -0.26412x + 3.749171

Using last three points,
Regression Output:

Constant	3.549235
Std Err of Y Est	0.013396
R Squared	0.999282
No. of Observations	4
Degrees of Freedom	2

X Coefficient(s) -0.17578
Std Err of Coef. 0.003332
Eqn 2: y = -0.17578x + 3.549235

Table 12:
Migration distances & sizes of bands in fig. 26. Regression output of fig. 27.

CHAPTER IV
DISCUSSION

4.1: Protein Purification

4.1.1: Method of Purification

At present the only method to purify thiol protease inhibitors from the alkaline fraction of cowpea is the method described. It is not a very good method, as the yield is very low-2kg of seeds gave only 3.2mg of impure protein for Rele et al, 1980. As alot of protein is needed for assays and feeding trials on insects, a much better method of purification is required, hence the attempts on the papain affinity column. The column, however, bound many proteins as seen in the gels (fig. 9, fig. 10, table 5) and no single band could be narrowed down as the inhibitor.

If enough of the pure inhibitor had been purified, its amino acid sequence could have been determined. Complementary oligonucleotide probes could then have been constructed to fish out the gene. Once this had been done, the gene could be inserted into an expression vector and expressed in yeast, hence much more inhibitor could be obtained in this way and various other tests done on it. If successful in disrupting the growth and physiology of insects and micro-organisms, this gene could then be used to generate resistant transgenic plants.

4.1.2: Assay System

The assay system used involved benzoyl-DL-arginine-*p*-nitroanilide (BAPNA) (Arnon, 1970), modified as in Methods. This system is based upon the cleavage of an amide bond in the small molecular weight synthetic substrate. The free amine is estimated

by the spectrophotometric determination of the *p*-nitroanilide formed during the hydrolysis of BAPNA. This assay system was chosen because it was the most convenient and sensitive one available. There are other systems that could be better like the colorimetric assay for papain using Ac-Phe-Arg-Nan as substrate, which allows the detection of minute quantities of papain (as low as 0.1μg) (Oliva et al., 1988) and the determination by titration using Z-Phe-Arg-NMec as substrate (Brzin et al., 1988). However, such substrates were not readily available.

As a result, alot of the partially purified protein had to be used for assays, leaving little for further purification or other work. The modified assay system used could not be made any more sensitive as maximal periods of 3 h were used and minimal amounts of BAPNA and papain. Any further alterations would have led to greater possibilities of error. Furthermore, as readings were taken from the same samples at intervals, a minimal volume of 4ml was required to prevent dilution effects from the washed cuvettes.

4.1.3: Results from the Ion-exchange Column

Gatehouse et al., 1986, obtained four elution peaks from their ion exchange column. Peaks 2 and 3 showed papain inhibitory activity and no trypsin inhibitory activity. Peak 2 consisted of two proteins when analysed by SDS-PAGE, of which one had approximately 32,000 m.wt. This was of the same weight as the protein obtained from the gel filtration column after pouring through peak 2 from the ion-exchange column (fig. 5, fig. 6, table 3). This, however did not show any inhibition whereas there was inhi-

bitory activity in the fraction before pouring through (fig. 2, table 1a & 1b, fig. 3, table 2a & 2b). As approximately same amounts of protein were used as before, this showed that either there was a trace amount of a second protein that was the inhibitor or that the protein had lost its activity during the interval between the two assays (15 days). As the sample was freeze dried immediately after eluting from the gel filtration column (2 days after the first assay), it could also have been the freeze-drying that inactivated the protein, if it is the inhibitor.

4.1.4: Results from the Papain Column

On Bapna assay, strong inhibitory activity was obtained from the eluting peak of this column (fig. 8, table 4a & 4b). SDS-PAGE showed many faint bands but a strong band at 52,000m.wt. was observed. This could be the main source of the inhibitory activity but could not be proved. The second washing peak showed enhanced enzymatic activity. This could be due to the mercapto-ethanol and EDTA present in the solution or due to some papain leaving the column.

4.1.5: Future Work

As have been shown in the results above, there certainly is an inhibitor of cysteine proteases in cowpea. The problem is in isolating it. If there were just a stretch of amino acids that were known of this inhibitor, an oligonucleotide probe could be constructed to fish the gene out. If the amino acid sequence of the site of plant inhibitors that is responsible for their inhibitory activity on cysteines were known, then this sequence could be used to pick out cysteine inhibitor genes from any plant with

a similar site. Such a common site has been known for years in animals (Barrett, 1987) and have been shown to inhibit cysteines when constructed synthetically (Teno et al., 1987). However, recently it has been shown that this site also exists in a plant cysteine inhibitor (Abe et al., 1987). In their work, a nearly full-length cDNA clone for a cysteine proteinase inhibitor of rice (oryzacystatin) was isolated and the amino acid sequence deduced from it. It was found that it contained the sequence Gln-Val-Val-Ala-Gly which was conserved among most members of the cystatin superfamily (proteinaceous cysteine protease inhibitors of animal origin). If the cysteine protease inhibitor gene in cowpea contains a similar site, this five-amino-acid sequence could be used to isolate it. The bromelain inhibitor, however, is known not to contain this site.

4.2: Gene Isolation

4.2.1: Reason of Contamination

The sensitivity of PCR is its advantage as well as its shortcoming. Trace amounts of contaminants can be detected by this method (Lo et. al, 1988). They suggest the following strategy for detection of contamination. (1) A negative control consisting of all PCR reagents minus the DNA template should be included in every amplification. (2) A likely source of contaminant is plasmid DNA, which is in common use in the laboratory, and such contamination can be proven by amplification with primers that span the vector-insert junction of such plasmid. The precautions in (1) and (2) exclude contamination of the PCR reagents and plasmid contamination of the DNA template. However, contamination of the template other than by the plasmid cannot be excluded.

As it was not expected that a prokaryotic DNA containing a kanamycin gene would have contaminated our template DNA, this could not have been detected any earlier. Future work can be done on the non-contaminated template DNA, but with a lower annealing temperature because $50^{\circ}C$ gave no amplification.

SUMMARY

There were two aims to this project. One was to get a pure thiol protease inhibitor from cowpea seeds and the other to get the gene sequence of a bromelain inhibitor from pineapple.

Two methods were tried to get a thiol protease inhibitor from cowpea. One involved an ion-exchange column and another a papain affinity column. The second peak from the ion-exchange column showed slight inhibition of papain in the BAPNA assays (16.8% for 1.66mg protein and 20.5% for 2.26mg protein, the amount of protein being estimated from absorbance readings). On further purification on a gel-filtration column, no inhibition was observed. The third peak from the papain column was divided into two and called fraction 2 and 3. Fraction 2, containing approximately 3.12mg of protein, inhibited papain activity by 33.5%, whereas fraction 3, containing approximately 1.66mg of protein, inhibited papain activity by 48.9%. Both fractions consisted of a mixture of proteins and there was insufficient material for further purification.

The method used to get the gene sequence of the bromelain inhibitor from pineapple was to first amplify the gene from its genomic DNA using PCR. Two amplified fragments (~160bp and 390bp) were obtained using the oligonucleotides coding for the C-terminus of the long chain and the N-terminus of the short chain of the bromelain inhibitor as primers. The 160bp fragment was inserted into pUC18 and cloned in *E. coli*. It was then inserted into M13mp19 and cloned before sequencing by the Sanger method. The gene obtained was suspected to have come from a prokaryotic

DNA containing a kanamycin-resistance gene. The most probable

source of this contaminant was from the pineapple template DNA

sample.

References

Abe, K. & Arai, S. (1985). Purification of a cysteine proteinase
inhibitor from rice, *Oryza-sativa-*L *japonica. Agric. Biol.
Chem.* **49**, 3349-3350.

Abe, K., Arai, S., Emori, Y., Kondo, H. & Suzuki, K. (1988). The
NH_2-terminal-21 amino cid residues are not essential for the
papain-inhibitory activity of oryzacystatin, a member of the
cystatin superfamily - expression of oryzacystatin cDNA and
its truncated fragments in *Escherichia coli. J. Biol. Chem.*
369, 229-232.

Abe, K., Emori, Y., Kondo, H., Suzuki, K. & Arai, S. (1987b).
Molecular cloning of a cysteine proteinase inhibitor of rice
(oryzacystatin). *J. Biol. Chem.* **262**, 16793-16797.

Abe, K., Kondo, H., Arai, S. (1987a). Purification and
characterization of a rice cysteine proteinase-inhibitor.
Agric. Biol. Chem. **51**, 2763-2768.

Abe, M., Arai, S., Kato, H. & Fujimaki, M. (1980). Thiol-protease
inhibitors occurring in endosperm of corn. *Agric. Biol.
Chem.* **44**, 685-686.

Abe, M. & Whitaker, J.R. (1988). Purification and characteriza-
tion of a cysteine proteinase inhibitor from the endosperm
of corn. *Agric. Biol. Chem.* **52**, 1583-1584.

Arnon, R. (1970). Methods of Enzymology. vol 19. pp 226-244.
Academic Press.

Barrett, A.J. (1987). The cystatins: a new class of peptidase
inhibitors. *TIBS* **12**, 193-196.

Broadway, R.M. & Duffey, S.S. (1986). Plant proteinase inhibitors:
mechanism of action and effect on the growth and digestive

physiology of larval *Heliothis zea* and *Spodoptera exiqua*. *J. Insect Physiol.*32, 827-833.

Brzin, J., Drobnickosorok, M., Kotnik, M., Popovic, T. & Turk, V. (1988). Inhibitors of cysteine proteinases from potato. *Biol. Chem. Hoppe-Seyler.* 369, 233-238.

Buchanan-Wollaston, V., Passiato, J.E., Cannon, F. (1987). The *mob* and *ori* T mobilization functions of a bacterial plasmid promote its transfer to plants. *Nature* 328, 172-175.

Chen, D. & Xia, Z. (1987). Mature plant regeneration from cultured protoplasts of *Polypogon fugax*. Nees ex Streud. *Sci. Sinica* 30B, 698-703. (original not seen, quoted from Vasil, 1988).

Crossway, A. Oakes, J.V., Irvine, J.M., Ward, B. & Knauf, V.C. (1986). Integration of foreign DNA following microinjection of tobacco mesophyll protoplasts. *Mol. Gen.Genet.* 202, 179-185.

Dawson, R.M.C., Elliott, D.C., Elliott, W.H. & Jones, K.M. (1969). *Data for Biochemical Research.* Clarendon Press, Oxford.

De la Pena, A., Lorz, H. & Schell, J. (1987). Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nature* 325, 274-276.

Fischoff, D.A. et al. (1987). Insect tolerant transgenic tomato plants. *Bio/Technology.* 8, 807-813.

Fossum, K. (1970). *Acta Pathol. Microbiol. Scand.* Section B, 28, 741-754. (original not seen, quoted from Zimacheva et al., 1988.)

Fromm, M.E., Taylor, L.P., & Walbot, V. (1986). Stable transformation of maize after gene transfer by electroporation. *Nature* **319**, 791-793.

Gasser, C.S. & Fraley, R.T. (1989). Genetically engineering plants for crop improvement. *Science* **244**, 1293-1299.

Gatehouse, A.M.R., Butler, K.J., Fenton, K.A. & Fatehouse, J.A. (1986). Partial characterization of a major proteolytic enzyme in the larval gut of *Callosobruchus maculatus*. *Entomol. Exp. Appl.* **39**, 279-286.

Goldstein, Z., Trop, M. & Birk, Y. (1973). *Nature New Biol.* **246**, 29-30.

Hilder, V.A., Gatehouse, A.M.R., Sheerman, S.E., Barker, R.F. & Boulter, D. (1987). A novel mechanism of insect resistance engineered into tobacco. *Nature* **330**, 160-163.

Klein, T.M., Fromm, M.E., Gradziel, T. & Sanford, J.C. (1988c). Factors influencing gene delivery into *Zea-mays* cells by high-velocity microprojectiles. *Bio/Technology* **6**, 559-563.

Klein, T.M., Fromm, M.E., Harper, E.C., Maliga, P., Sanford, J.C. & Svab, Z. (1988a). Stable genetic transformation of intact *Nicotiana* cells by the particle bombardment process. *Proc. Natl. Acad. Sci. USA.* **85**, 8502-8505.

Klein, T.M., Fromm, M., Sanford, J.C., Schaaf, S., Sletten, M., Tomes, D. & Weissinger, A. (1988b). Transfer of foreign genes into intact maize cells with high-velocity microprojectiles. *Proc. Natl. Acad. Sci. USA.* **85**, 4305-4309.

Krens, F.A., Molendi, J.L., Wullems, G.A. & Schilper, R.A. (1982). *in-vitro* transformation of plant protoplasts with

Ti-plasmid DNA. *Nature* 296, 72-74.

Kyozuka, J., Hayashi, Y. & Shimamoto, K. (1987). High frequency plant regeneration from rice protoplasts by novel nurse culture methods. *Molec. Gen. Genet.* 206, 408-413.

Lo, Y-M.D., Mehal, W.Z. & Flemming, K.A.(1988). False-positive results and the polymerase chain reaction. *Lancet* #8612, 679.

Luo, Z.-x. & Wu, R. (1988). *Plant Mol. Biol. Reporter* 6, 165 (original article not seen, quoted from Gasser et al., 1989.)

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

McCabe, D.E., Swain, W.F., Martinell, B.J. & Christou, P. (1988). Stable transformation of soybean (<u>Glycine max</u>) by particle acceleration. *Bio/Technology* 6, 923-926.

Meeusen, R.L. (1989). Insect control with genetically engineered crops. *Ann. Rev. Entomol.* 34, 373-381.

Melville, J.C. & Ryan, C.A. (1972). Chymotrypsin Inhibitor I from Potatoes: large scale preparation and characterization of its subunit components. *J. Biol. Chem.* 247, 3445.

Oliva, M.L.V., Sampaio, C.A.M. & Sampaio, M.U. (1988). Purification and partial characterization of a thiol proteinase inhibitor from *Enterolobium contortisiliquum* beans. *Biol. Chem. Hoppe-Seyler* 369, 229-232.

Perlstein, S.H. & Reddy, F.J. (1973). *J. Supramol. Struct.* 1, 249-254. (original not seen, quoted from Reddy et al.,

1974.)

Reddy, M.N., Keim, P.S., Heinrikson, R.L. & Kezdy, F.J. (1975). Primary structural analysis of sulfhydryl protease inhibitors from pineapple stem. *J. Biol. Chem.* 250, 1741-1750.

Rele, M.V., Vartak, H.G. & Jagannathan, V. (1980). Proteinase inhibitors from *Vigna unguiculata* subsp. *cylindrica*. I. Occurrence of thiol proteinase inhibitors in plants and purification from *Vigna unguiculata* subsp. *cylindrica*. *Arch. Biochem. Biophys.* 204, 117-128.

Rhodes, C.A., Lowe, K.S., Ruby, K.L. (1988). Plant regeneration from protoplasts isolated from embryonic maize cell cultures. *Bio/Technology* 6, 56-60.

Rodis, P. & Hoff, J.E. (1984). Naturally occurring protein crystals in the potato, inhibitor of papain, chymopapain and ficin. *Plant Physiol.* 74, 907-911.

Ryan, C.A. (1981). Proteinase inhibitors. In *The Biochemistry of Plants, A Comprehensive Treatise* vol. 6. (ed. P.K. Stumpf & E.E. Conn), pp. 351-371. Academic Press, New York.

Ryan, C.A. (1989). Proteinase Inhibitor Gene Families: Strategies for Transformation to Improve Plant Defences Against Herbivores. *BioEssays* 10, 20-24.

Ryan, C.A. (1988). Molecular biology of wound-inducible proteinase inhibitors in plants. *Plant Cell & Environment* 11, 345-349.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A. (1988). Primer-

directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239, 487-491.

Sanford, J.C. (1988) The Biolistic Process. *Trends in Biotech.* 6, 299-302.

Sanger, F., Nicklen, S. & Coulson, A.R. (1979). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74, 5463-5467.

Schell, J.S. (1987). Transgenic plants as tools to study the molecular organization of plant genes. *Science* 237, 1176-1183.

Srinivasan, C. & Casil, I.K. (1986). Plant regeneration from protoplasts isolated from protoplasts of sugarcane. *J. Pl. Phsiol.* 126, 41-48.

Teno, N., Tsuboi, S., Okada, Y., Itoh, N. & Okamoto, H. (1987). Amino acids and peptides: XV. Synthesis of Gln-Val-Val-Ala-Gly and derivatives, a common sequence of thiol proteinase inhibitors and their effects on thiol proteinase. 30, 93-98.

Topfer, R., Gronenborn, B., Schell, J. & Steinbiss, H.-H. (1989). *Plant Cell* 1, 133. (original article not seen, quoted from Gasser et al., 1989.)

Toriyama, K., Arimoto, Y., Uchimiya, H. & Hinata, K. (1988). Transgenic rice plants after direct gene transfer into protoplasts. *Bio/Technology.* 6, 1072-1074.

Vaek, M., Reynaerts, A., Hofte, H., Jansens, S., De Beuckeleer, M., Dean, C., Zabeau, M., Van Montague, M. & Leemans, J. (1987). Transgenic plants protected from insect attack. *Nature* 328, 33-37.

Vasil, I.K. (1988). Progress in the regeneration and genetic manipulation of cereal crops. *Bio/Technology* **6**, 397-402.

Wolfson, J.L. & Murdock, L.L. (1987). Suppression of Colorado potato beetle growth and development by digestive proteinase inhibitors. *Entomol. Exp. Appl.* **44**, 235-240.

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

Zimacheva, A.V., Ievleva, E.V. & Mosolov, V.V. (1984). Isolation and characterization of inhibitors of cysteine proteinases from potato tubers. *Biokhimiya* **49**, 1153-1158. (original not seen, quoted from Zimacheva et al., 1988.)

Zimacheva, A.V., Ievleva, E.V. & Mosolov, V.V. (1988) Cysteine proteinase inhibitor from pumpkin seeds. *Biochemistry-USSR* **53**, 640-645. (Translated from *Biokhimiya* **53**, 740-746.)

Final concentration of ammonium sulphate, % saturation at 0°C

Initial	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
	g solid ammonium sulphate to add to 100 ml of solution																
0	10.7	13.6	16.6	19.7	22.9	26.2	29.5	33.1	36.6	40.4	44.2	48.3	52.3	56.7	61.1	65.9	70.7
5	8.0	10.9	13.9	16.8	20.0	23.2	26.6	30.0	33.6	37.3	41.1	45.0	49.1	53.3	57.8	62.4	67.1
10	5.4	8.2	11.1	14.1	17.1	20.3	23.6	27.0	30.5	34.2	37.9	41.8	45.8	50.0	54.4	58.9	63.6
15	2.6	5.5	8.3	11.3	14.3	17.4	20.7	24.0	27.5	31.0	34.8	38.6	42.6	46.6	51.0	55.5	60.0
20	0	2.7	5.6	8.4	11.5	14.5	17.7	21.0	24.4	28.0	31.6	35.4	39.2	43.3	47.6	51.9	56.5
25		0	2.7	5.7	8.5	11.7	14.8	18.2	21.4	24.8	28.4	32.1	36.0	40.1	44.2	48.5	52.9
30			0	2.8	5.7	8.7	11.9	15.0	18.4	21.7	25.3	28.9	32.8	36.7	40.8	45.1	49.5
35				0	2.8	5.8	8.8	12.0	15.3	18.7	22.1	25.8	29.5	33.4	37.4	41.6	45.9
40					0	2.9	5.9	9.0	12.2	15.5	19.0	22.5	26.2	30.0	34.0	38.1	42.4
45						0	2.9	6.0	9.1	12.5	15.8	19.3	22.9	26.7	30.6	34.7	38.8
50							0	3.0	6.1	9.3	12.7	16.1	19.7	23.3	27.2	31.2	35.3
55								0	3.0	6.2	9.4	12.9	16.3	20.0	23.8	27.7	31.7
60									0	3.1	6.3	9.6	13.1	16.6	20.4	24.2	28.3
65										0	3.1	6.4	9.8	13.4	17.0	20.8	24.7
70											0	3.2	6.6	10.0	13.6	17.3	21.2
75												0	3.2	6.7	10.2	13.9	17.6
80													0	3.3	6.8	10.4	14.1
85														0	3.4	6.9	10.6
90															0	3.4	7.1
95																0	3.5
100																	0

Initial concentration of ammonium sulphate, % saturation at 0°C

A1.1: Ammonium sulphate saturation chart (reproduced from Dawson et al., 1969).

A2.1: Nucleotide sequence of pUC19. The only difference between pUC18 and pUC19 is in the reversed sequence between 393 and 458 bases.

```
          10        20        30        40        50        60        70        80
TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT

          90        100       110       120       130       140       150       160
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGA

          170       180       190       200       210       220       230       240
GCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC

          250       260       270       280       290       300       310       320
ATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGG

          330       340       350       360       370       380       390       400
GGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATT

          410       420       430       440       450       460       470       480
CGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCT

          490       500       510       520       530       540       550       560
GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATG

          570       580       590       600       610       620       630       640
AGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

          650       660       670       680       690       700       710       720
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCG

          730       740       750       760       770       780       790       800
GTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGG

          810       820       830       840       850       860       870       880
AAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCC

          890       900       910       920       930       940       950       960
GCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG

          970       980       990       1000      1010      1020      1030      1040
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC

          1050      1060      1070      1080      1090      1100      1110      1120
GGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG

          1130      1140      1150      1160      1170      1180      1190      1200
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGAC

          1210      1220      1230      1240      1250      1260      1270      1280
TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG

          1290      1300      1310      1320      1330      1340      1350      1360
GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG

          1370      1380      1390      1400      1410      1420      1430      1440
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGA

          1450      1460      1470      1480      1490      1500      1510      1520
AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT

          1530      1540      1550      1560      1570      1580      1590      1600
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTA

          1610      1620      1630      1640      1650      1660      1670      1680
TATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA

          1690      1700      1710      1720      1730      1740      1750      1760
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT

          1770      1780      1790      1800      1810      1820      1830      1840
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTC

          1850      1860      1870      1880      1890      1900      1910      1920
CTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG

          1930      1940      1950      1960      1970      1980      1990      2000
CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA

          2010      2020      2030      2040      2050      2060      2070      2080
ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA

          2090      2100      2110      2120      2130      2140      2150      2160
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGC

          2170      2180      2190      2200      2210      2220      2230      2240
TTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC

          2250      2260      2270      2280      2290      2300      2310      2320
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT

          2330      2340      2350      2360      2370      2380      2390      2400
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTC

          2410      2420      2430      2440      2450      2460      2470      2480
ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAT

          2490      2500      2510      2520      2530      2540      2550      2560
ACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTA

          2570      2580      2590      2600      2610      2620      2630      2640
TTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATC

          2650      2660      2670      2680      2690      2700      2710      2720
ATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC
```
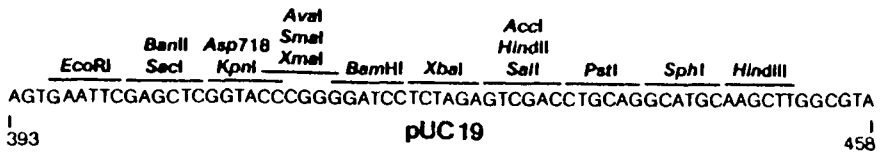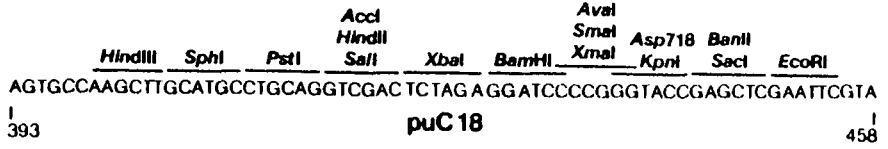
HindIII	SphI	PstI	Accl HindII SalI	XbaI	BamHI	Aval SmaI Xmal	Asp718 KpnI	BanII SacI	EcoRI

AGTGCCAAGCTTGCATGCCTGCAGGTCGAC TCTAG AGG ATCCCCGGG TACCGAGCTCGAA TTCGTA

393   puC18   458

EcoRI	BanII SacI	Asp718 KpnI	Aval SmaI Xmal	BamHI	XbaI	Accl HindII SalI	PstI	SphI	HindIII

AGTG AATTCGAGCTCGGTACCCGGGGATCC TCTAGAGTCGACC TGCAGGCATGCAAGCTTGGCGTA

393   pUC19   458

ENZYME	SITE		POS.		POS.		POS.
ACCIA	(GTCGAC)	(429)	429				
ACYIA	(GGCGCC)	(235)	235				
ACYIB	(GGCGTC)	(2235)	2235				
ACYID	(GACGTC)	(2617)	2617				
ALUI	(AGCT)	(43)	43	(19)	62	(245)	307
		(96)	403	(45)	448	(22)	470
		(95)	565	(64)	629	(118)	747
		(226)	973	(136)	1109	(257)	1366
		(521)	1887	(100)	1987	(63)	2050
ASUI	(GGGCC)	(284)	284	(1534)	1820		
AVAIA	(CCCGGG)	(412)	412				
AVAIIB	(GGTCC)	(1837)	1837	(222)	2059		
BAMHI	(GGATCC)	(417)	417				
BBVIA	(GCAGC)	(41)	41	(1172)	1213	(3)	1216
		(206)	1422	(694)	2116		
BBVIB	(GCTGC)	(254)	254	(73)	327	(303)	630
		(81)	711	(18)	729	(419)	1148
		(602)	1750				
BGLI	(GCCNNNNNGGC)	(245)	245	(1568)	1813		
DDEIA	(CTAAG)	(2622)	2622				
DDEIB	(CTCAG)	(1081)	1081	(409)	1490	(166)	1656
DDEIC	(CTGAG)	(171)	171	(2025)	2196		
ECORI	(GAATTC)	(396)	396				
ECORIIA	(CCAGG)	(354)	354	(479)	833	(121)	954
ECORIIB	(CCTGG)	(545)	545	(422)	967		
FNUDII	(CGCG)	(2)	2	(2)	4	(103)	107
		(545)	652	(2)	654	(198)	852
		(581)	1433	(330)	1763	(493)	2256
		(332)	2588				
FNU4HIA	(GCCGC)	(850)	850	(155)	1005	(1084)	2089
		(351)	2440				
FNU4HIB	(GCGGC)	(150)	150	(582)	732	(1479)	2211
HAEIB	(AGGCCA)	(819)	819	(11)	830		
HAEIA	(TGGCCT)	(1282)	1282				
HAEIIA	(GGCGCC)	(235)	235				
HAEIIC	(GGCGCT)	(680)	680	(370)	1050		
HAEIII	(GGCC)	(287)	287	(102)	389	(257)	646
		(174)	820	(11)	831	(18)	849
		(434)	1283	(458)	1741	(80)	1821
		(267)	2088				
HGAIA	(GACGC)	(908)	908	(578)	1486		
HGAIB	(GCGTC)	(108)	108	(2128)	2236		
HGIAIA	(GTGCAC)	(177)	177	(943)	1120	(1246)	2366
HGIAIB	(GAGCTC)	(402)	402				
HGIAIC	(GTGCTC)	(2281)	2281				
HHAI	(GCGC)	(3)	3	(103)	106	(130)	236
		(21)	237	(331)	568	(65)	653
		(28)	681	(33)	714	(270)	984
		(67)	1051	(100)	1151	(174)	1325
		(109)	1434	(393)	1827	(93)	1920
		(337)	2257	(332)	2589		
HINCIIA	(GTCGAC)	(429)	429				
HINDIII	(AAGCTT)	(447)	447				
HINFIA	(GAATC)	(641)	641	(140)	781		
HINFIB	(GACTC)	(706)	706	(988)	1694		
HINFIC	(GAGTC)	(427)	427	(750)	1177		
HPAII	(CCGG)	(48)	48	(34)	82	(331)	413

ENZYME	SITE		POS.		POS.		POS.
HPAII	(CCGG)	(111)	524	(489)	1013	(147)	1160
		(26)	1186	(190)	1376	(404)	1780
		(34)	1814	(67)	1881	(110)	1991
		(242)	2233				
HPHIIA	(GGTGA)	(12)	12	(9)	21	(2152)	2173
		(241)	2414				
HPHIB	(TCACC)	(1550)	1550	(227)	1777	(622)	2399
KPNI	(GGTACC)	(408)	408				
MBOI	(GATC)	(277)	277	(141)	418	(955)	1373
		(75)	1448	(11)	1459	(8)	1467
		(78)	1545	(12)	1557	(105)	1662
		(341)	2003	(18)	2021	(46)	2067
		(258)	2325	(17)	2342	(36)	2378
MBOIIA	(GAAGA)	(1456)	1456				
MBOIIB	(TCTTC)	(291)	291	(394)	685	(862)	1547
		(755)	2302	(78)	2380	(109)	2489
MNLIA	(CCTC)	(29)	29	(260)	289	(132)	421
		(274)	695	(283)	978	(878)	1856
		(206)	2062				
MNLIB	(GAGG)	(662)	662	(259)	921	(324)	1245
		(400)	1645	(81)	1726	(947)	2673
MBTI	(TGCGCA)	(256)	256	(1663)	1919		
PSTI	(CTGCAG)	(435)	435				
PVUI	(CGATCG)	(276)	276	(1790)	2066		
PVUII	(CAGCTG)	(306)	306	(322)	628		
RSAI	(GTAC)	(168)	168	(241)	409	(1769)	2178
SACI	(GAGCTC)	(402)	402				
SALI	(GTCGAC)	(429)	429				
SFANIA	(GATGC)	(78)	78	(130)	208	(1948)	2156
SFANIB	(GCATC)	(153)	153	(76)	229	(665)	894
		(1052)	1946	(440)	2386		
SMAI	(CCCGGG)	(412)	412				
SPHI	(GCATGC)	(441)	441				
TAQI	(TCGA)	(400)	400	(30)	430	(476)	906
		(1444)	2350				
XBAI	(TCTAGA)	(423)	423				
XHOIIA	(GGATCC)	(417)	417				
XHOIIB	(AGATCC)	(1458)	1458	(98)	1556	(785)	2341
XHOIIC	(GGATCT)	(1447)	1447	(97)	1544	(780)	2324
NDEI	(CATATG)	(183)	183				
NARI	(GGCGCC)	(235)	235				
MBTI	(TGCGCA)	(256)	256	(1663)	1919		

THESE ENZYMES DID NOT APPEAR.....

ACCIB	ACCIC	ACCID	ACYIC	ASUII
AVAIB	AVAIC	AVAID	AVAIIA	AVAIII
AVRII	BALI	BCLI	BGLII	BBTEIIA
BBTEIIB	BBTEIIC	BBTEIID	CLAI	DDEID
HAEIA	HAEIC	HAEIIB	HAEIID	HGIAID
HINCIIB	HINCIIC	HINCIID	HINFID	HPAI
SACII	XHOI	XHOIID	XMAIII	NCOI

A2.3:
Nucleotide sequence
of M13mp18, using
the original HincII
site as reference
point.

```
          10        20        30        40        50        60        70        80
AATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAATGAAAATATAGCTAAACAGGTTATTGA

          90       100       110       120       130       140       150       160
CCATTTGCGAAATGTATCTAATGGTCAAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTACATGGAATGAAA

         170       180       190       200       210       220       230       240
CTTCCAGACACCGTACTTTAGTTGCATATTTAAAACATGTTGAGCTACAGCACCAGATTCAGCAATTAAGCTCTAAGCCA

         250       260       270       280       290       300       310       320
TCCGCAAAAATGACCTCTTATCAAAAGGAGCAATTAAAGGTACTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCT

         330       340       350       360       370       380       390       400
GGTTCGCTTTGAAGCTCGAATTAAAACGCGATATTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGCAATCCGCT

         410       420       430       440       450       460       470       480
TTGCTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTCTGAACTGTTTAAAGCA

         490       500       510       520       530       540       550       560
TTTGAGGGGGATTCAATGAATATTTATGACGATTCCGCAGTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCCC

         570       580       590       600       610       620       630       640
CTCTGGCAAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAACGAGGGTTATGATAGTG

         650       660       670       680       690       700       710       720
TTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAATGTGGTATTCCTAAATCTCAACTG

         730       740       750       760       770       780       790       800
ATGAATCTTTCTACCTGTAATAATGTTGTTCCGTTAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACGTCCTGACTG

         810       820       830       840       850       860       870       880
GTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAATTAAACCATCTCAAGCCCAATT

         890       900       910       920       930       940       950       960
TACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCTTATTCACTGAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGA

         970       980       990      1000      1010      1020      1030      1040
ATATCCGGTTCTTGTCAAGATTACTCTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCT

        1050      1060      1070      1080      1090      1100      1110      1120
CTTTCAAAGTTGGTCAGTTCGGTTCCCTTATGATTGACCGTCTGCGCCTCGTTCCGGCTAAGTAACATGGAGCAGGTCGC

        1130      1140      1150      1160      1170      1180      1190      1200
GGATTTCGACACAATTTATCAGGCGATGATACAAATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTC

        1210      1220      1230      1240      1250      1260      1270      1280
AAAGATGAGTGTTTTAGTGTATTCTTTCGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTTTACCC

        1290      1300      1310      1320      1330      1340      1350      1360
GTTTAATGGAAACTTCCTCATGAAAAAGTCTTTAGTCCTCAAAGCCTCTGTAGCCGTTGCTACCCTCGTTCCGATGCTGT

        1370      1380      1390      1400      1410      1420      1430      1440
CTTTCGCTGCTGAGGGTGACGATCCCGCAAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTAT

        1450      1460      1470      1480      1490      1500      1510      1520
GCGTGGGCGATGGTTGTTGTCATTGTCGGCGCAACTATCGGTATCAAGCTGTTTAAGAAATTCACCTCGAAAGCAAGCTG

        1530      1540      1550      1560      1570      1580      1590      1600
ATAAACCGATACAATTAAAGGCTCCTTTTGGAGCCTTTTTTTTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCAAT

        1610      1620      1630      1640      1650      1660      1670      1680
TCCTTTAGTTGTTCCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAACCCCATACAGAAAATTCAT

        1690      1700      1710      1720      1730      1740      1750      1760
TTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGGTTACGCTAACTATGAGGGTTGTCTGTGGAATGCTACAGGCGTT

        1770      1780      1790      1800      1810      1820      1830      1840
GTAGTTTGTACTGGTGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAAATGAGGGTGG

        1850      1860      1870      1880      1890      1900      1910      1920
TGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCTCCTGAGTACGGTGATACACCTA

        1930      1940      1950      1960      1970      1980      1990      2000
TTCCGGGCTATACTTATATCAACCCCTCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCT

        2010      2020      2030      2040      2050      2060      2070      2080
TCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAATAGGTTCCGAAATAGGCAGGGGGCATTAACTGT

        2090      2100      2110      2120      2130      2140      2150      2160
TTATACGGGCACTGTTACTCAAGGCACTGACCCCGTTAAAACTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGT
```
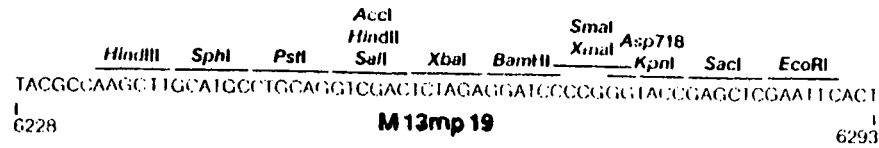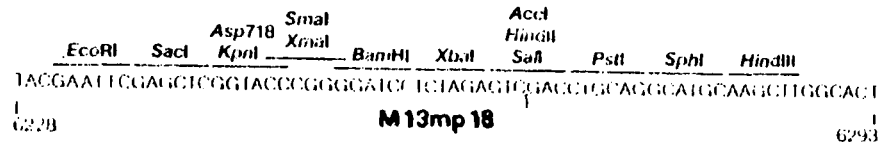
```
        2170      2180      2190      2200      2210      2220      2230      2240
ATGACGCTTACTGGAACGGTAAATTCAGAGACTGCGCTTTCCATTCTGGCTTTAATGAAGATCCATTCGTTTGTGAATAT

        2250      2260      2270      2280      2290      2300      2310      2320
CAAGGCCAATCGTCTGACCTGCCTCAACCTCCTGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGA

        2330      2340      2350      2360      2370      2380      2390      2400
GGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTG

        2410      2420      2430      2440      2450      2460      2470      2480
ATTTTGATTATGAAAAGATGGCAAACGCTAATAAGGGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGAC

        2490      2500      2510      2520      2530      2540      2550      2560
GCTAAAGGCAAACTTGATTCTGTCGCTACTGATTACGGTGCTGGTATCGATGGTTTCATTGGTGACGTTTCCGGCCTTGC

        2570      2580      2590      2600      2610      2620      2630      2640
TAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAATGGCTCAAGTCGGTGACGGTGATAATTCACCTT

        2650      2660      2670      2680      2690      2700      2710      2720
TAATGAATAATTTCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCTTTAGCGCTGGTAAA

        2730      2740      2750      2760      2770      2780      2790      2800
CCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTTTTATATGTTGCCACCTT

        2810      2820      2830      2840      2850      2860      2870      2880
TATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCAGTTCTTTTGGGTATTCCGTT

        2890      2900      2910      2920      2930      2940      2950      2960
ATTATTGCGTTTCCTCGGTTTCCTTCTGGTAACTTTGTTCGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTAAGA

        2970      2980      2990      3000      3010      3020      3030      3040
TAGCTATTGCTATTTCATTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTTGTGGGTTATCTCTCTGATATTAGC

        3050      3060      3070      3080      3090      3100      3110      3120
GCTCAATTACCCTCTGACTTTGTTCAGGGTGTTCAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTTATGTTATTCT

        3130      3140      3150      3160      3170      3180      3190      3200
CTCTGTAAAGGCTGCTATTTTCATTTTTTGACGTTAAACAAAAAATCGTTTCTTATTTGGATTGGGATAAATAATATGGCT

        3210      3220      3230      3240      3250      3260      3270      3280
GTTTATTTTGTAACTGGCAAATTAGGCTCTGGAAAGACGCTCGTTAGCGTTGGTAAGATTCAGGATAAAATTGTAGCTGG

        3290      3300      3310      3320      3330      3340      3350      3360
GTGCAAAATAGCAACTAATCTTGATTTAAGGCTTCAAAACCTCCCGCAAGTCGGGAGGTTCGCTAAAACGCCTCGCGTTC

        3370      3380      3390      3400      3410      3420      3430      3440
TTAGAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTGGGCGCGGTAATGATTCCTACGATGAAAATAAAAAC

        3450      3460      3470      3480      3490      3500      3510      3520
GGCTTGCTTGTTCTCGATGAGTGCGGTACTTGGTTTAATACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGA

        3530      3540      3550      3560      3570      3580      3590      3600
TTGGTTTCTACATGCTCGTAAATTAGGATGGGATATTATTTTTCTTGTTCAGGACTTATCTATTGTTGATAAACAGGCGC

        3610      3620      3630      3640      3650      3660      3670      3680
GTTCTGCATTAGCTGAACATGTTGTTTATTGTCGTCGTCTGGACAGAATTACTTTACCTTTTGTCGGTACTTTATATTCT

        3690      3700      3710      3720      3730      3740      3750      3760
CTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAGCCCTAC

        3770      3780      3790      3800      3810      3820      3830      3840
TGTTGAGCGTTGGCTTTATACTGGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATTATGATT

        3850      3860      3870      3880      3890      3900      3910      3920
CCGGTGTTTATTCTTATTTAACGCCTTATTTATCACACGGTCGGTATTTCAAACCATTAAATTTAGGTCAGAAGATGAAA

        3930      3940      3950      3960      3970      3980      3990      4000
TTAACTAAAATATATTGGAAAAGGTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAGCATTTTACATATAGTTA

        4010      4020      4030      4040      4050      4060      4070      4080
TATAACCCAACCTAAGCCGGAGGTTAAAAGGTAGTCTCTCAGACCTATGATTTTGATAAATTCACTATTGACTCTTCTC

        4090      4100      4110      4120      4130      4140      4150      4160
AGCGTCTTAATCTAAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAAAATTAATTAATAGCGACGATTTACAGAAGCAA

        4170      4180      4190      4200      4210      4220      4230      4240
GGTTATTCACTCACATATATTGATTTATGTACTGTTTCCATTAAAAAAGGTAATTCAAATGAAATTGTTAAATGTAATTA

        4250      4260      4270      4280      4290      4300      4310      4320
ATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCTCAGGTAATTGAAATGAATAATTGGCCTCTGCGCGATTTT
```

A2.3: *The only difference between M13mp18 and M13mp19 is in the reversed sequence between 6228 and 6293 bases.*

```
         4330      4340      4350      4360      4370      4380      4390      4400
       GTAACTTGGTATTCAAAGCAATCAGGCGAATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACTGTATATTCATC

         4410      4420      4430      4440      4450      4460      4470      4480
       TGACGTTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCTAATAATTTTGATATGGTGGTTCAATTC

         4490      4500      4510      4520      4530      4540      4550      4560
       CTTCCATAATTCAGAAGTATAATCCAAACAATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGAT

         4570      4580      4590      4600      4610      4620      4630      4640
       GATAATTCCGCTCCTTCTGGTGGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATTAATAACGTTCG

         4650      4660      4670      4680      4690      4700      4710      4720
       GGCAAAGGATTTAATACGAGTTGTCGAATTGTTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTATTGACG

         4730      4740      4750      4760      4770      4780      4790      4800
       GCTCTAATCTATTAGTTGTTAGTGCACCTAAAGATATTTTAGATAACCTTCCTCAATTCCTTTCTACTGTTGATTTGCCA

         4810      4820      4830      4840      4850      4860      4870      4880
       ACTGACCAGATATTGATTGAGGGTTTGATATTTGAGGTTCAGCAAGGTGATGCTTTAGATTTTTCATTTGCTGCTGGCTC

         4890      4900      4910      4920      4930      4940      4950      4960
       TCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCGTTCGGTA

         4970      4980      4990      5000      5010      5020      5030      5040
       TTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTGTCTGTGCCACGT

         5050      5060      5070      5080      5090      5100      5110      5120
       ATTCTTACGCTTTCAGGTCAGAAGGGTTCTATCTCTGTTGGCCAGAATGTCCCTTTTATTACTGGTCGTGTGACTGGTGA

         5130      5140      5150      5160      5170      5180      5190      5200
       ATCTGCCAATGTAAATAATCCATTTCAGACGATTGAGCGTCAAAATGTAGGTATTTCCATGAGCGTTTTTCCTGTTGCAA

         5210      5220      5230      5240      5250      5260      5270      5280
       TGGCTGGCGGTAATATTGTTCTGGATATTACCAGCAAGGCCGATAGTTTGAGTTCTTCTACTCAGGCAAGTGATGTTATT

         5290      5300      5310      5320      5330      5340      5350      5360
       ACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTTTACTCGGTGGCCTCACTGATTATAA

         5370      5380      5390      5400      5410      5420      5430      5440
       AAACACTTCTCAAGATTCTGGCGTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATT

         5450      5460      5470      5480      5490      5500      5510      5520
       CCAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGG

         5530      5540      5550      5560      5570      5580      5590      5600
       TGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTC

         5610      5620      5630      5640      5650      5660      5670      5680
       TCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCAC

         5690      5700      5710      5720      5730      5740      5750      5760
       CTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGAC

         5770      5780      5790      5800      5810      5820      5830      5840
       GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTG

         5850      5860      5870      5880      5890      5900      5910      5920
       ATTTATAAGGGATTTTGCCGATTTCGGAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCT

         5930      5940      5950      5960      5970      5980      5990      6000
       TGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCGCTGGTGAAAAGAAAAACCACCCTG

         6010      6020      6030      6040      6050      6060      6070      6080
       GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA

         6090      6100      6110      6120      6130      6140      6150      6160
       AAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCG

         6170      6180      6190      6200      6210      6220      6230      6240
       GCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTCGAGCT

         6250      6260      6270      6280      6290      6300      6310      6320
       CGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGG

         6330      6340      6350      6360      6370      6380      6390      6400
       GAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCG
```

```
         6410      6420      6430      6440      6450      6460      6470      6480
       CACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGC

         6490      6500      6510      6520      6530      6540      6550      6560
       CGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACGGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGAT

         6570      6580      6590      6600      6610      6620      6630      6640
       GCGCCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTA

         6650      6660      6670      6680      6690      6700      6710      6720
       CTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTCCTATTGGTT

         6730      6740      6750      6760      6770      6780      6790      6800
       AAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATA

         6810      6820      6830      6840      6850      6860      6870      6880
       CAATCTTCCTGTTTTTGGGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTACCGTTC

         6890      6900      6910      6920      6930      6940      6950      6960
       ATCGATTCTCTTGTTTGCTCCAGACTCTCAGGCAATGACCTGATAGCCTTTGTAGATCTCTCAAAAATAGCTACCCTCTC

         6970      6980      6990      7000      7010      7020      7030      7040
       CGGCATTAATTTATCAGCTAGAACGGTTGAATATCATATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCTTTTG

         7050      7060      7070      7080      7090      7100      7110      7120
       AATCTTTACCTACACATTACTCAGGCATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATCCTTGCGTTGAAATA

         7130      7140      7150      7160      7170      7180      7190      7200
       AAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTTTTGGTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATT

         7210      7220      7230      7240      7250      7260      7270      7280
       GCTTAATTTTGCTAATTCTTTGCCTTGCCTGTATGATTTATTGGATGTT
```


EcoRI	SacI	Asp718 Kpn I	Smal Xmal	BamHI	XbaI	AccI HincII SalI	PstI	SphI	HindIII

TACGGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACT
6228   **M13mp18**   6293

HindIII	SphI	PstI	AccI HincII SalI	XbaI	BamHI	Smal Xmal Asp718 Kpnl	SacI	EcoRI

TACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCACT
6228   **M13mp19**   6293

Enzyme	Site		Pos.		Pos.		Pos.
AccIA	(GTCGAC)	(6263)	6263				
AhaIIC	(GGCGCC)	(6000)	6000				
AluI	(AGCT)	(39)	39	(24)	63	(140)	203
		(26)	229	(104)	333	(600)	933
		(854)	1497	(29)	1516	(1446)	2962
		(313)	3275	(336)	3611	(484)	4095
		(1330)	5425	(204)	5629	(331)	5960
		(93)	6053	(64)	6117	(95)	6212
		(25)	6237	(45)	6282	(93)	6375
		(111)	6486	(180)	6666	(63)	6729
		(220)	6949	(27)	6976	(201)	7177
AvaIA	(CCCGGG)	(6246)	6246				
AvaIC	(CTCGGG)	(5824)	5824				
AvaIIA	(GGACC)	(5913)	5913				
BalI	(TGGCCA)	(5079)	5079				
BamHI	(GGATCC)	(6251)	6251				
BanIA	(GGCACC)	(5676)	5676	(454)	6130	(334)	6464
BanIB	(GGTACC)	(6242)	6242				
BanIC	(GGCGCC)	(6000)	6000				
BanID	(GGTGCC)	(1248)	1248	(5228)	6476		
BanIIB	(GAGCTC)	(6236)	6236				
BanIID	(GGGCTC)	(5642)	5642				
BbVIA	(GCAGC)	(931)	931	(4604)	5535	(516)	6051
		(303)	6354	(73)	6427		
BbVIB	(GCTGC)	(1366)	1366	(1154)	2520	(611)	3131
		(1739)	4870	(1052)	5922		
BglI	(GCCNNNNNGGC)	(6430)	6430				
BglII	(AGATCT)	(6934)	6934				
ClaI	(ATCGAT)	(2526)	2526	(4355)	6881		
DdeI	(CTNAG)	(233)	233	(865)	1098	(272)	1370
		(46)	1416	(367)	1783	(63)	1846
		(15)	1861	(15)	1876	(24)	1900
		(72)	1972	(42)	2014	(303)	2317
		(15)	2332	(15)	2347	(15)	2362
		(998)	3360	(652)	4012	(27)	4039
		(39)	4078	(14)	4092	(28)	4120
		(160)	4280	(600)	4880	(381)	5261
		(672)	5933	(575)	6508	(399)	6907
		(153)	7060	(128)	7188		
DraI	(TTTAAA)	(189)	189	(283)	472	(4149)	4621
		(2162)	6783	(290)	7073		
EcoRI	(GAATTC)	(6230)	6230				
EcoRIIA	(CCAGG)	(5940)	5940	(196)	6136		
EcoRIIB	(CCTGG)	(1013)	1013	(952)	1965	(4032)	5997
		(330)	6327	(127)	6454		
FnuDII	(CGCG)	(43)	43	(304)	347	(771)	1118
		(57)	1175	(1290)	2465	(889)	3354
		(54)	3408	(190)	3598	(353)	3951
		(361)	4312	(681)	4993	(495)	5488
		(24)	5512	(20)	5532	(496)	6028
		(2)	6030	(658)	6688	(63)	6751
Fnu4HI	(GCNGC)	(931)	931	(435)	1366	(27)	1393

Enzyme	Site		Pos.		Pos.		Pos.
Fnu4HI	(GCNGC)	(891)	2284	(3)	2287	(24)	2311
		(45)	2356	(164)	2520	(611)	3131
		(1739)	4870	(629)	5499	(14)	5513
		(22)	5535	(387)	5922	(129)	6051
		(303)	6354	(73)	6427		
FokIA	(GGATG)	(3546)	3546	(3697)	7243		
FokIB	(CATCC)	(239)	239	(6121)	6360		
HaeIIA	(GGCGCC)	(6000)	6000				
HaeIIB	(AGCGCC)	(5558)	5558	(8)	5566		
HaeIIC	(GGCGCT)	(6445)	6445				
HaeIID	(AGCGCT)	(2709)	2709	(329)	3038		
HaeIII	(GGCC)	(1395)	1395	(849)	2244	(309)	2553
		(2527)	5080	(158)	5238	(106)	5344
		(69)	5413	(311)	5724	(214)	5938
		(98)	6036	(257)	6293	(102)	6395
		(117)	6512	(169)	6681	(341)	7022
HgaIA	(GACGC)	(526)	526	(1637)	2163	(315)	2478
		(758)	3236	(3450)	6686		
HgaIB	(GCGTC)	(4082)	4082	(1075)	5157		
HgiAIA	(GTGCAC)	(4742)	4742				
HgiAIB	(GAGCTC)	(6236)	6236				
HgiAIC	(GTGCTC)	(5464)	5464				
HhaI	(GCGC)	(44)	44	(966)	1010	(74)	1084
		(92)	1176	(293)	1469	(725)	2194
		(272)	2466	(244)	2710	(329)	3039
		(56)	3095	(312)	3407	(190)	3597
		(714)	4311	(683)	4994	(495)	5489
		(13)	5502	(9)	5511	(22)	5533
		(26)	5559	(8)	5567	(434)	6001
		(28)	6029	(65)	6094	(331)	6425
		(21)	6446	(115)	6561		
HincIIA	(GTCGAC)	(6263)	6263				
HindIII	(AAGCTT)	(6281)	6281				
HinfI	(GANTC)	(136)	136	(80)	216	(274)	490
		(21)	511	(212)	723	(1287)	2010
		(486)	2496	(348)	2844	(413)	3257
		(160)	3417	(324)	3741	(96)	3837
		(234)	4071	(45)	4116	(232)	4348
		(771)	5119	(209)	5328	(46)	5374
		(63)	5437	(328)	5765	(22)	5787
		(253)	6040	(221)	6261	(342)	6423
		(261)	6684	(19)	6903	(137)	7040
HpaII	(CCGG)	(314)	314	(651)	965	(129)	1094
		(829)	1923	(454)	2377	(18)	2395
		(156)	2551	(818)	3369	(472)	3841
		(176)	4017	(1596)	5613	(545)	6158
		(89)	6247	(215)	6462	(18)	6480
		(357)	6837	(123)	6960	(60)	7020
HphIA	(GGTGA)	(1375)	1375	(398)	1773	(135)	1908
		(489)	2397	(144)	2541	(39)	2580
		(39)	2619	(6)	2625	(2221)	4846

Enzyme	Site	Pos.		Pos.		Pos.
HphIA	(GGTGA)	(270) 5116	(589)	5705	(241)	5946
		(33) 5979	(1025)	7004		
HphIB	(TCACC)	(1502) 1502	(1132)	2634	(2288)	4922
		(2108) 7030				
KpnI	(GGTACC)	(6242) 6242				
MboI	(GATC)	(1381) 1381	(332)	1713	(507)	2220
		(4032) 6252	(153)	6405	(96)	6501
		(434) 6935				
MboIIA	(GAAGA)	(2217) 2217	(1694)	3911	(2479)	6390
MboIIB	(TCTTC)	(781) 781	(3293)	4074	(196)	4270
		(666) 4936	(318)	5254	(332)	5586
		(917) 6503	(301)	6804		
MnlIA	(CCTC)	(254) 254	(119)	373	(187)	560
		(27) 587	(68)	655	(383)	1038
		(49) 1087	(143)	1230	(66)	1296
		(21) 1317	(8)	1325	(19)	1344
		(71) 1415	(90)	1505	(391)	1896
		(48) 1944	(75)	2019	(243)	2262
		(6) 2268	(404)	2672	(4)	2676
		(217) 2893	(158)	3051	(269)	3320
		(31) 3351	(351)	3702	(604)	4306
		(391) 4697	(74)	4771	(149)	4920
		(5) 4925	(421)	5346	(69)	5415
		(265) 5680	(340)	6020	(235)	6255
		(277) 6532	(423)	6955		
MnlIB	(GAGG)	(484) 484	(141)	625	(747)	1372
		(359) 1731	(102)	1833	(15)	1848
		(15) 1863	(15)	1878	(129)	2007
		(312) 2319	(15)	2334	(15)	2349
		(15) 2364	(4)	2368	(967)	3335
		(685) 4020	(799)	4819	(15)	4834
		(612) 5446	(947)	6393	(117)	6510
		(575) 7085	(105)	7190		
MstI	(TGCGCA)	(6424) 6424				
MstII	(CCTNAGG)	(6507) 6507				
NaeI	(GCCGGC)	(5612) 5612				
NarI	(GGCGCC)	(6000) 6000				
NciIA	(CCCGG)	(6246) 6246				
NciIB	(CCGGG)	(1923) 1923	(4324)	6247	(590)	6837
NdeI	(CATATG)	(2722) 2722	(1080)	3802	(3043)	6845
NlaIII	(CATG)	(149) 149	(47)	196	(910)	1106
		(193) 1299	(499)	1798	(235)	2033
		(123) 2156	(698)	2854	(677)	3531
		(87) 3618	(99)	3717	(1461)	5178
		(1043) 6221	(55)	6276	(580)	6856
NlaIV	(GGNNCC)	(1061) 1061	(187)	1248	(292)	1540
		(10) 1550	(252)	1802	(248)	2050
		(323) 2373	(18)	2391	(3252)	5643
		(12) 5655	(21)	5676	(190)	5866
		(134) 6000	(130)	6130	(112)	6242
		(9) 6251	(213)	6464	(12)	6476
PstI	(CTGCAG)	(6269) 6269				
PvuI	(CGATCG)	(6404) 6404				
PvuII	(CAGCTG)	(5959) 5959	(93)	6052	(322)	6374
RsaI	(GTAC)	(173) 173	(107)	280	(741)	1021

Enzyme	Site	Pos.		Pos.		Pos.
RsaI	(GTAC)	(143) 1164	(604)	1768	(27)	1795
		(93) 1888	(16)	1904	(65)	1969
		(163) 2132	(1334)	3466	(201)	3667
		(522) 4189	(190)	4379	(1004)	5383
		(102) 5485	(758)	6243	(599)	6842
		(322) 7164				
SacI	(GAGCTC)	(6236) 6236				
SalI	(GTCGAC)	(6263) 6263				
Sau96I	(GGNCC)	(5723) 5723	(190)	5913	(24)	5937
		(458) 6395				
ScrFI	(CCNGG)	(1013) 1013	(910)	1923	(42)	1965
		(3975) 5940	(57)	5997	(139)	6136
		(110) 6246	(1)	6247	(80)	6327
		(127) 6454	(383)	6837		
BfanIA	(GATGC)	(25) 25	(363)	388	(965)	1353
		(3496) 4849	(1696)	6545	(13)	6558
BfanIB	(GCATC)	(3978) 3978				
SmaI	(CCCGGG)	(6246) 6246				
SnaBI	(TACGTA)	(1267) 1267				
SphI	(GCATGC)	(6275) 6275				
TaqI	(TCGA)	(336) 336	(790)	1126	(381)	1507
		(441) 1948	(579)	2527	(927)	3454
		(239) 3693	(971)	4664	(1018)	5682
		(552) 6234	(30)	6264	(618)	6882
XbaI	(TCTAGA)	(6257) 6257				
XhoIIA	(GGATCC)	(6251) 6251				
XhoIIB	(AGATCC)	(2219) 2219				
XhoIID	(AGATCT)	(6934) 6934				
XmnI	(GAANNNNTTC)	(357) 357	(2288)	2645		

These enzymes did not appear.....

AatII	AccIB	AccIC	AccID	AhaIIA
AhaIIB	AhaIID	ApaI	AvaIB	AvaID
AvaIIB	BanIIA	BanIIC	BclI	BamHI
BstII	BstEII	EcoRV	HgiAID	HincIIB
HincIIC	HincIID	HpaI	MluI	NcoI
NotI	NruI	NsiI	SacII	ScaI
StuI	Tth111I	XhoI	XhoIIC	XmaIII