STUDIES ON THE TARGETING AND PROCESSING OF PRORICIN.

Michael Westby BSc. (Hons) Dunelm

A thesis submitted for the degree of
Doctor of Philosophy

Department of Biological Sciences
University of Warwick
Coventry.
U.K.

September, 1991

CONTENTS

FIGURE LEGENDS.

TABLES.

ACKNOWLEDGMENTS.

I would like to thank my colleagues at Warwick, past and present, who have inspired, supported and teased me. In not mentioning any by name I do not want to minimise their contribution to this work in any way (for it is considerable) but rather I do this for fear of omission. However, I would like to thank my supervisor Lynne Roberts personally, both for giving me the chance to taste real Italian pasta and for critical reading of this manuscript whilst in the throws of motherhood.

I would also like to say "Grazie!" to all my friends at the Istituto Biosintesi Vegetali, Milan, in particular to Aldo Ceriotti who refused to recognise a bad result when he saw one. Why he ever wanted me back a second time will remain a mystery to me.

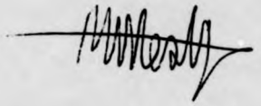
I am grateful to Matt Busby for growing the castor bean plants and to Jonathon Cook for preparing the Xenopus egg cell-free extract.

Thanks to Penny, Sean, Alicia and Richard for living with me so harmoneously over these past years and to Mum, Dave and Marian for their unfailing love and support.

I acknowledge the Science and Engineering Research Council for funding this project and the European Molecular Biology Organisation for providing a short term fellowship which funded my stay in Milan.

DECLARATION.

All the results presented in thesis were obtained by the authour unless specifically indicated in the text. This work has not been used in any previous application for a degree.



Mike Westby

ABBREVIATIONS

aa	amino acid
$A^{(x)}$	absorbance of a solution at (x)nm
APS	ammonium persulphate
ATP	Adenosine triphosphate
bp	base pairs
cDNA	Complementary DNA
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
Ci	Curie
ConA	Concanavalin A
cpm	Counts per minute
dATP	Deoxyadenosine triphosphate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	dithiothreitol
EDTA	Ethylenediaminetetra-acetic acid
endoH	endo-N acetylglucosaminidase H
FDA	Fluorescein diacetate
GTP	Guanosine triphosphate
GUS	beta-glucuronidase
hr(s)	hour(s)
Ig	immunoglobulin
IPTG	Isopropylthiogalactoside
kb	Kilobases

kD	kilo Daltons
mg	milligrammes
min(s)	minute(s)
ml	millilitres
mM	millimolar
mRNA	messenger RNA
ng	nanogrammes
nos	nopaline synthase transcription terminator
npt2	neomycin phosphotransferase II
OD	optical density
OH	oocyte homogenate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PHA	Phytohaemagglutinin
PMSF	Phenylmethylsulphonyl fluoride
PEG	Polyethylene glycol
RF	replicative form (of M13)
RIP(I/II)	ribosome inactivating protein (types I & II)
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	Ribonuclease inhibitor
rpm	revolutions per minute
SDS	Sodium dodecyl sulphate
ss	single-stranded
TCA	Trichloroacetic acid
TEMED	N,N,N,N' tetramethylenediamine

Th	mutated linker with thrombin recognition sequence
Tris	Tris-hydroxymethylaminomethane
tRNA	transfer RNA
Tween 20	Polyoxyethylene sorbitan monolaurate
µCi	microcurie
µg	microgrammes
µl	microliters
UTP	Uridine triphosphate
(v/v)	volume/volume
(w/v)	weight/volume
Xa	mutated linker with Factor Xa recognition sequence
X-gal	5-bromo-4-chloro-3-indolyl-B-D-galactoside

To Mum and Dad.

Out, out brief candle,
Life's but a walking shadow, a poor player
That struts and frets his work upon the stage
And then is heard no more, it is a tale
Told by an idiot, full of sound and fury,
Signifying nothing.

(Shakespeare, from Macbeth)

"Courage!" he said, and pointed toward the land,
"This mounting wave will see us shoreward soon."
In the afternoon they came unto a land,
In which it seemed always afternoon.

(Tennyson, The Lotos Eaters)

SUMMARY.

The targeting and processing of proricin, a precursor form of the highly cytotoxic castor bean lectin, ricin, was studied in this thesis. Targeting signals are present on plant seed storage protein precursors which first direct them into the ER-lumen (signal peptides) and then sort them to the protein bodies. Computer analysis of the 35aa residue N-terminal presequence of preproricin identified two residues as being likely sites of signal peptidase cleavage. Using site-directed mutagenesis, chimeric genes were constructed between the N-terminal region of preproricin and npt2. Expression of one such chimera in rabbit reticulocyte lysate (containing microsomal membranes) showed that the first 22aa N-terminal residues of preproricin function as a signal peptide. The ricin/npt2 chimeras were then transiently expressed in tobacco protoplasts in order to determine whether the 35aa preproricin presequence contained any vacuolar targeting information. Evidence is presented which suggests that this region is not responsible for targeting proricin to the protein bodies.

Recombinant proricin was synthesised by injection of proricin transcripts into Xenopus oocytes. Proricin synthesised in this way possessed lectin activity but did not depurinate mammalian ribosomes, indicating that the A chain domain has no RNA N-glycosidase activity when in the context of the precursor. In order to demonstrate the conversion of an inactive enzyme precursor to an active mature enzyme (the conversion of proricin to ricin), an assay was designed to purify the proricin processing enzyme from the protein bodies of dry castor bean seeds. Radiolabelled proricin was made in vitro by expression of proricin transcripts in wheat germ lysate. This was then used as a substrate to identify processing activity present in protein body extracts from dry castor bean seeds. Processing of the 63kD labelled proricin polypeptide to products of 30kD was observed by SDS-PAGE analysis and fluorography of assay samples. However, the levels of enzyme activity in the dried seed were too low to enable purification of the proricin processing enzyme. Therefore, an alternative strategy was designed to enable in vitro proteolytic cleavage of proricin. Site-directed mutagenesis was used to design proricin linkers encoding a recognition sequence for either Factor Xa or thrombin. Proricin clones containing these mutant linkers were expressed in E.coli, wheat germ lysate and a Xenopus egg-cell free extract. When expressed in Xenopus egg cell-free extract, the proricin mutants were shown to have lectin activity, indicating that the B chain domain was biologically active. Evidence for Factor Xa-specific cleavage of proricin with a mutated linker is presented.

CHAPTER 1

INTRODUCTION.

1.1 SUMMARY

Since its identification by Stillmark in 1888 as a protein extract from castor beans that caused erythroagglutination, ricin has had a highly profiled and diverse history. Its name has cropped up in connection with therapeutic agents designed to fight cancer and in the more sinister field of international espionage: the murder of the Bulgarian dissident, Markov, in London in 1979 was thought to have been caused by a ricin-tipped umbrella.

Ricin can perhaps best be defined as the highly cytotoxic 7S lectin that accumulates in the seeds of the castor bean, Ricinus communis. This description covers the three features of ricin which account for its widespread scientific appeal; namely, it is a potent toxin, a sugar-binding lectin and, as it accumulates in the seed, it can easily be obtained in large amounts. As will be discussed in this introduction, ricin is initially synthesised as a preproprotein and the object of this study was to examine the post-translational targeting and processing of this precursor.

The introduction has therefore been divided as follows. Plant seed storage proteins will first be considered with particular attention being paid to the subclass of the 7S lectins, into which ricin falls. This first section will concentrate on their synthesis as preproproteins, post-translational modifications and subcellular sites of storage, as these are of particular relevance to this work. In section 1.3 a general scheme for the intracellular vesicular routing of proteins will be proposed, based largely on pioneering work in animal and yeast systems. The areas in which plants now appear to differ from this scheme will be discussed. Finally, ricin will be described in detail before this introduction section is completed with a summary of the aims of this work.

1.2 PLANT SEED STORAGE PROTEINS

1.2.1 INTRODUCTION

In addition to supplying the genetic information necessary for the successful development of a new plant, parent spermatophytes (seed plants) must also lay down reserves within the seeds. These reserves must include a source of energy, organically-fixed nitrogen, pre-formed carbon chains and minerals: all of which will enable the germinating progeny to survive until their root and photosynthetic systems are sufficiently developed to sustain further growth.

The reserves commonly found in seeds are carbohydrates, proteins, lipids and phytins (a source of metal ions and phosphorus). The storage proteins are a major component of these reserves and may account for up to 80% of the total protein in mature cells. They are a source of organically-fixed nitrogen, having a high content of arginine, asparagine/aspartate and glutamine/glutamate residues - for example, these residues constitute 36% of the amino acid residues in the legumin storage protein from soybean, Glycine max (see Wilson, 1986).

1.2.2 CLASSIFICATION OF PLANT STORAGE PROTEINS

In 1924, Osborne defined a tetrapartite classification scheme for plant storage proteins that was based on their relative solubilities in different solvents. The four groups were as follows;

ALBUMINS - soluble in H_2O at neutral/slightly acidic pH
GLOBULINS - soluble in salt solutions
GLUTELINS - soluble in strongly acidic or basic solutions
PROLAMINS - soluble in ethanol (aqueous alcohol)

Generally speaking, glutelins and prolamins are the major storage proteins in monocotyledonous plants (monocots.) whereas dicotyledonous plants (dicots.) have globulins as their primary storage proteins. Exceptions to this rule include oat, whose major storage proteins are globulins (see Higgins, 1984), and the castor bean, that store 2S albumins as a significant source of organically-fixed nitrogen and sulphur.

Due to their commercial importance, the storage proteins of cereals and legumes have been the most widely studied monocot. and dicot. storage proteins respectively.

1.2.3 STORAGE PROTEINS OF CEREALS

In cereals the endosperm is the major seed storage tissue. Prolamins are the major group of proteins in cereals such as barley, wheat, maize and sorghum, where they can account for 80 to 100% of the protein in the endosperm protein bodies (Pernollet, 1978). Rice, however, stores predominantly glutelins. Oats have both glutelins and globulins as storage protein.

1.2.4 STORAGE PROTEINS OF DICOTS.

The main storage tissue in legumes is the cotyledonary tissue, though in plants with a developed endosperm, this may be the site of storage protein deposition. The globulins found

in legume cotyledons fall into two types; the 11S legumins and the 7S N-glycosylated vicilins. The relative amounts of these two types of globulins varies between different species.

(i) Legumins

As illustrated by their high average sedimentation coefficient legumins are large molecules, consisting of 6 heterodimers held together by non-covalent forces. Each heterodimer is made up of a large 40kD acidic subunit and small 20kD basic subunit, covalently attached by a single inter-chain disulphide bridge. Many different isoforms of both the acidic and basic subunits exist as the legumins are encoded by a multi-gene family. However, Staswick et al (1981) showed that the association of the a- and b- isoforms of the soybean legumin, glycinin, was non-random and it is now known that this is because they are initially synthesised as a single 60kD precursor (see Nielsen, 1984) which is later processed to its mature two-chain form, after the formation of the disulphide bond.

Storage proteins of the legumin type have been identified in non-leguminous dicots. and include the 11S globulin from pumpkin (see Akazawa and Hara-Nishimura, 1985). The monocot. seed globulin from oat also resembles the 11S legumins and contains acidic and basic subunits that are synthesised as a single polypeptide precursor (Brinegar and Peterson, 1982).

(ii) Vicilins

As mentioned above, vicilins are N-glycosylated. This co-translational modification occurs as the protein is translocated into the lumen of the ER. Differential oligosaccharide processing of these glycan residues partly explains the complex pattern of vicilin forms within a single plant species. Sun et al (1974) and Thanh and Shibasaki (1978) have shown that vicilin from Phaseolus vulgaris (phaseolin) and Glycine max respectively contain three major

subunits (a, a' and B) ranging from 40-60kD in size. Badenoch-Jones et al (1981) have shown that at least 10 subunits of pea vicilin can be distinguished. However, Chrispeels et al (1982) have demonstrated that 4 subunits (with molecular weights of 75, 70, 50 and 49kD) are initially synthesised and the six smallest subunits arise by specific proteolysis of the 49/50kD subunits.

All vicilins described to date are from the Leguminosae family, with the exception of acalin A from cotton (see Wilson, 1986).

(iii) Convicilin

Croy et al (1980) have isolated a third type of storage globulin from pea. Convicilin is a 290kD tetrameric protein - each subunit having a molecular weight of 71kD.

(iv) Albumins

In addition to the globulins, 7S lectins and 2S albumins are present in significant quantity in some plant species. They may be considered to be storage proteins as they contain organically-fixed nitrogen, accumulate during seed development, and are rapidly degraded during germination.

(v) 7S Lectins (see review by Etzler, 1985: Sharon & Lis, 1989)

The 7S lectins can generally be classified as plant storage proteins in that they accumulate during seed development, they have an amino acid composition containing relatively high proportions of fixed nitrogen and are rapidly degraded during germination. They are also, as we shall see, initially synthesised as preproprotein precursors, transported along the same intracellular routes and are deposited in the same organelles as other plant storage proteins. However, these proteins do not simply act as stores of organically-fixed nitrogen. They belong to a diverse class of proteins called lectins, found in all classes and families of organisms and which are classified functionally by their ability to bind to

carbohydrate moieties (Sharon and Lis, 1989). Lectins are also multi-valent (each molecule can bind more than one sugar molecule) and this property results in agglutination when intact cells and lectins are mixed together, since all cells have a 'sugar coating' of glycoproteins, glycolipids or polysaccharide. Lectins can be sub-divided into six specificity groups according to the monosaccharide that is the best competitive inhibitor of erythroagglutination.

A number of roles have been proposed to explain their widespread presence in plants: to date lectins have been found in more than 1000 plant species (reviewed by Eztler, 1986). However, Chrispeels and Raikhel (1991) have most recently used the examples of the Phaseolus lectins and the chitin-binding lectins to propose a common ancestory and role for this family of seed proteins in dicots and monocots respectively. They propose that their primary function is in plant defence, by for example, discouraging predation by animals (feeding studies have attributed the toxicity of many seeds to their lectin contents) or limiting hyphal growth upon fungal penetration of the plant cells and release of lectins from ruptured vacuoles.

The most abundant source of lectins are the 7S lectins in mature seeds and particularly in the seeds of the Leguminosae where they can constitute up to 10% of the total protein (Pusztai et al, 1979). Typically legume lectins exist as tetramers with subunits of 25-35kD. The subunits may or may not be identical. For example, Concanavalin A (ConA) from the jackbean, Canavalia ensiformis, is assembled from a single 25.5kD subunit, whereas pea lectin (and other lectins from the Vicieae tribe) is composed of two non-identical subunits of 17kD & 6kD (Trowbridge, 1974) and is termed a two-chain lectin. There is a high degree of protein sequence homology between all the legume lectins, but this homology is not immediately apparent. For example, for correct alignment the B chain of the

two chain lectins must be placed alongside the amino-terminus of the single chain lectins, followed by the A chain. On the other hand, Hemperely and Cunningham (1983) used the term 'circular homology' to describe how the primary amino acid sequences of the lectins from the Dioclea tribe align with those of the one and two chain lectins. These apparently cryptic homology patterns are a consequence of different post-translational modifications of the lectin precursors and are discussed below in 1.2.6.

Other 7S lectins include the lectins from Abrus pretorius and Ricinus communis (see review by Olsnes and Pihl, 1982). In both species there are two closely related lectins: a heterodimeric cytotoxin (abrin and ricin) and a relatively non-toxic tetrameric agglutinin which resembles a dimerised form of the toxin. In these lectins, only one of the two non-identical subunits has sugar-binding properties (the B chain): the other subunit is the toxic moeity (A chain) and will be discussed in more detail in section 1.4.

Lectins frequently exist in a number of different isoforms, termed isolectins, which can arise from expression of more than one gene within a lectin gene family. For example, phytohaemagglutinin (PHA) from Phaseolus vulgaris contains both erythro-binding (PHA-E) and leuco-binding (PHA-L) subunits encoded by two closely-linked genes (Hoffman and Donaldson, 1985) which can assemble randomly to form 5 isolectins. Most 7S lectins are glycoproteins, an exception being the mature form of ConA, although even this is initially synthesised as a glycoprotein (Herman, Shannon and Chrispeels, 1985).

(vi) 2S Albumins

Youle and Huang (1982) identified three low-molecular-weight albumins, all with apparent molecular weights of 12kD, in the endosperm of castor bean seeds. The amino acid composition of these 2S albumins reveals a 30% content of

glutamate residues and 8.5% cysteine (see Wilson, 1986) suggesting that these are important stores of nitrogen and sulphur in the castor bean.

Other 2S-albumins include the napins isolated from rapeseed, Brassica napus (Lonnerdal and Janson, 1972). The cDNA sequence of these proteins reveals that they too are synthesised as a single polypeptide precursor (Crouch et al, 1983).

The 2S albumins, protease inhibitors, prolamins, alpha-amylase- and trypsin-inhibitors from cereals have 3 homologous domains (see Kreis et al, 1985) and it is believed that they all arose from a common ancestral gene.

1.2.5 SUBCELLULAR SITE OF STORAGE PROTEIN DEPOSITION - THE PROTEIN BODIES

Numerous immunocytochemical and cell fractionation studies have identified the site of storage protein deposition within a cell to be the single-membrane-bound organelles, termed aleurone grains or protein bodies (reviewed by Pernollet, 1978). In dicots. they range between 1 and 22 um in diameter. Pernollet (1978) identified three types of protein bodies defined by their internal morphology: those with a homogeneous, granular internal morphology; those with phytin globoids; and, as in the case of the well-characterised protein bodies of the castor bean, those that contain both globoids and crystalloid inclusions (Youle and Huang, 1976: Tully and Beevers, 1976). Gifford and Bewley (1983) showed that these crystalloid inclusions contain deposits of 11S globulin, the major storage protein of the castor bean.

As many dicot. storage proteins are glycosylated it has been possible to propose a common route of intracellular transport for these proteins from the ER lumen, via the the Golgi apparatus to the protein bodies (defined here as the

secretory pathway). This is because oligosaccharide side chains are modified post-translationally in an orderly and unidirectional way: each of the various modification enzymes involved has been localised to a particular vesicular compartment along the secretory pathway (see review by Kornfield and Kornfield, 1985). The oligosaccharide side chains of glycoproteins therefore provide a biochemical 'footprint' of their route through the cell. In this way, Chrispeels (1983ab) showed that PHA passed through the Golgi apparatus en route to the protein bodies in developing Phaseolus seeds. Immunocytochemical data of Greenwood and Chrispeels (1985) in Phaseolus supports this observation, as does that of Herman and Sharon (1984 & 1985) who showed that the Golgi apparatus was involved in the transport to the protein bodies of the lectins alpha-galactosidase haemagglutinin in soybean and ConA in jackbean.

However, the transport of monocot storage proteins to the protein bodies is far from clear. There is evidence from studies of prolamin deposition in maize and wheat to suggest that protein bodies arise from budding of the endoplasmic reticulum (ER) (Larkins and Hurkman, 1978: Parker, 1982). This would imply that some monocot storage proteins may not pass along the secretory pathway en route to pre-existing protein body organelles but instead, accumulate in regions of the ER which delimit and bud to form '*de novo*' protein bodies. However, this work is open to criticism as cross-contamination of fractions or misinterpretation of immunocytochemical data when considered in isolation is possible. One factor which has hindered work in this area is that most monocot storage proteins are not glycosylated and therefore their intracellular route cannot be traced biochemically as it has been for the dicot proteins. Indirect evidence is now available in support of this secretory pathway-independent route for monocot storage

proteins from the expression of zeins and alpha-gliandins in Xenopus oocytes (Wallace et al, 1988: Simon et al, 1990). In both studies the prolamins became aggregated and remained in the ER. However, Simon and co-workers found that the closely-related gamma-gliandin was efficiently secreted from the oocytes. This data may be interpreted to imply that there is not a single (common) route of protein storage in monocots, or it may be argued that the proteins simply behave differently when expressed in a heterologous system.

Chrispeels and Tague (1990) choose to distinguish between the two types of protein bodies described above by terming them 'ER-derived vesicles' and 'protein storage vacuoles'. However, for the purpose of this thesis, the general term 'protein bodies' will be used as it is now generally accepted that all protein bodies in dicots are both derived from, and the progenitor of, the large central vacuole (Bollini and Chrispeels, 1979: Nishimura and Beevers, 1978).

1.2.6 SYNTHESIS AS PREPROPROTEINS

A common feature of many of the storage proteins is that they are synthesised as precursors, or 'preproproteins'.

Lord and Robinson (1986) distinguish between two different proteolytic processing steps that lead to the mature protein. The first is the co-translational cleavage of the N-terminal signal peptide that directs the storage protein into the lumen of the endoplasmic reticulum (ER) (see next section for more detail). The second step involves endoproteolytic cleavage(s) of the segregated protein to yield its mature form.

Evidence that this second modification step occurs in the protein bodies, the site postulated for the processing of pea storage proteins by Chrispeels et al (1982), has been presented for the the 11S globulins from pumpkin (Hara-Nishimura and Nishimura, 1987). They isolated an enzyme activity from

vacuoles of developing pumpkin cotyledons which was able to process their storage protein precursors in vitro. They also found that 11S globulin precursors were cleaved to their mature sizes after incubation with an extract from castor bean seeds, suggesting that this activity may be present in different species.

A common feature of these protein body-localised processing events is the observation that cleavage of the precursor polypeptide backbone usually occurs on the carboxyl side of an asparagine residue (Higgins, 1984).

(i) Legumin processing

Study of cDNA sequences available for many storage proteins illustrates that proteolytic processing does not always involve a single cleavage of the polypeptide precursor. In the case of the legumin-like proteins there is removal of a linking peptide that lies between the mature polypeptides in the proprotein. In the legumin precursor from pea, processing involves the removal of a 6aa residue linker (Boulter, 1984) whereas glycinin precursor from soybean loses a 4aa residue linker (Nielsen, 1984). In each case the carboxy-terminal residue of the linker is an asparagine residue. It should be added that a 12aa residue carboxy-terminal peptide is also lost from the basic subunit of soybean (Nielsen, 1984). In this case an asparagine-specific endoprotease is not thought to be involved.

(ii) Vicilin processing

The vicilin-like precursors undergo different degrees of proteolytic processing. For example, a 4aa residue carboxy-terminal propeptide is thought to be cleaved from beta-phaseolin, a subunit of vicilin from Phaseolus (Sandro Vitale, personal communication). On the other hand, the 50kD subunit of pea vicilin can be processed at two internal sites to release subunits of 19kD, 13.5kD and 12-14kD. Processing at

only one of these sites will result in subunits of intermediate size (see Boulter, 1984: reviewed by Lord and Robinson, 1986). The processing sites in the 50kD subunit have been shown to lie on the carboxyl side of an asparagine. In vicilin isoforms that do not have an asparagine residue in either of these positions cleavage does not occur at that position. This explains the apparent partial processing of the 50kD subunit and the fact that both the full-length and intermediately-sized subunits constitute a significant proportion of the mature vicilin.

(iii) Lectin processing

Asparagine -specific cleavages are also observed in the processing of some legume 7S lectins. Amino acid sequencing of the two-chain lectins from the Vicieae tribe of legumes reveals that proteolytic cleavage of the prolectin molecule in each case occurs on the carboxyl side of an asparagine residue (summarised by Yarwood et al, 1988). However, there are examples of single chain lectins in this tribe. The protein sequences of a single chain lectin from Lathyrus sphaericus (Richardson, Yarwood and Rouge, 1987) and two single chain isolectins from Lathyrus nissolia (Yarwood et al, 1988) identifies an asparagine residue in a position equivalent to the Asparagine residue lying on the carboxyl side of the cleavage site in the two chain lectins. One explanation for the absence of cleavage at this site in the single chain Viciae lectins is that the L.sphaericus and L.nissolia species lack an asparagine -specific processing enzyme. An alternative suggestion made by Yarwood and co-workers was based on the observation that in the precursors of two chain lectins, the asparagine at the cleavage site formed part of a putative asparagine-linked glycosylation signal (i.e. Asn-X-Thr/Ser:see section 1.3.4). No such signal was present in the sequences of the single-chain lectins at this position. As this region of the protein precursor must be accessible for a processing

enzyme to cleave the polypeptide backbone of the two-chain Vicieae lectins, it is reasonable to suggest that it is near the surface and is glycosylated. In their mature form, however, the two chain lectins are not glycosylated at this asparagine residue and so Yarwood et al propose that a deglycosylation event at this site might be a prerequisite for endoproteolytic processing of the Vicieae lectins. They therefore proposed that a glycosylation/deglycosylation may be a pre-requisite for endoproteolytic processing of the two chain lectins.

In support of this the complex processing of the single chain lectin ConA from jackbean is known to involve such a glycosylation/deglycosylation step which is followed (sometime later) by four asparagine -specific cleavages and a unique polypeptide ligation event (Carrington et al, 1985: Bowles et al, 1986: reviewed by Bowles and Pappin, 1988). Also the derived protein sequence from pea seed lectin cDNA (Higgins et al, 1983) reveals the asparagine residue at the end of its 6aa linker as being a potential site for asparagine-linked glycosylation. However, to date there is no direct evidence that shows a link between glycosylation and endoproteolytic processing.

Lectins are not proteolytically processed in every case. For example, there is no evidence for the proteolytic processing of PHA in the protein bodies of Phaseolus.

(iv) 2S albumins

From the predicted amino acid sequence of napin precursor, Crouch et al (1983) suggested that there is a 21aa residue N-terminal signal peptide followed by a region of 38aa residues that has a net negative charge, before the start of the mature small subunit. They also proposed that there is a 20aa residue linker between the small and large subunits which again has a net negative charge. Both of these regions are thought to be cleaved from the proprotein in the protein bodies. None of

these proteolytic events however, appear to involve a peptide bond cleavage on the carboxyl side of asparagine.

1.2.7 SUMMARY

It is clear that many storage proteins are formed as preproprotein precursors which become N-glycosylated and are finally sequestered and processed to their mature form in the protein bodies. In many cases there is a role for a protein body-localised class of asparagine-specific endoprotease(s) in the removal of propeptides. The ability of a protein body soluble matrix fraction isolated from one species to process a storage protein precursor from another species suggests that this processing machinery may be highly conserved. However, very little else is known about this processing step and no asparagine-specific enzymes have, to the author's knowledge, been characterised.

1.3 INTRACELLULAR TRAFFICKING OF NEWLY-SYNTHESISED PROTEINS

1.3.1 INTRODUCTION

As has been shown, seed storage proteins are made as preproproteins that are sequestered in the ER before being transported to the protein bodies. What route they follow and what distinguishes them from other synthesised proteins will now be discussed.

Many of the steps involved in the intracellular routing of proteins were initially elucidated by studying animal and yeast systems. Only recently have similar studies in plants begun to reveal many common features and show that much of the transport machinery is conserved throughout the eukaryotes.

1.3.2 MEMBRANE TRANSLOCATION AS A FIRST STEP IN COMPARTMENTALISATION

(i) The Paradox

Order is imposed within a cell by the segregation of proteins necessary for a specific reaction within membrane-bound compartments: this may be the periplasmic space of E.coli or, in the case of the eukaryotic cell, by segregation within organelles. Boundaries are maintained by the impermeability of water soluble proteins to the membrane lipid bilayer. However, an apparent paradox arises when considering that, with the exception of a relatively few proteins encoded and expressed within mitochondria and chloroplasts, all proteins share a common site of synthesis, that is on cytosolic ribosomes. Therefore, the proteins destined for organellar locations must at some stage cross an apparently impermeable membrane barrier.

But what distinguishes membrane-segregated from cytosolic proteins?

(ii) Topogenic signals

The signal hypothesis was first proposed by Blobel and

Dobberstein (1975), who suggested that secretory proteins had specific signals which interacted with specific receptors on the membrane of the endoplasmic reticulum (ER). Blobel later generalised the signal hypothesis (Blobel, 1980) by using the term 'intracellular protein topogenesis' to include the targeting of proteins across other intracellular membranes. He proposed that topogenic sequences within the protein are recognised by specific receptors on the target membrane, during (co-translational) or shortly after (post-translational) their synthesis. The last statement takes into account the observations *in vitro* that ER translocation of secretory proteins is closely linked to translation whereas proteins targeted to other organelles (e.g. mitochondria and chloroplasts) can be imported post-translationally. For a recent review of membrane targeting signals see Von Heijne (1990).

(ii) Translocation-competent state.

Zimmerman and Meyer (1986) suggested that an essential property of any protein that is translocated is that is must be in an 'unfolded state' - that is, a polypeptide with secondary but no tertiary structure. They concluded that factors that prevent or promote protein 'unfolding' form components of the translocation systems.

In light of the observations that preproalphafactor from yeast and human glucose transporter can cross the ER membrane post-translationally, Schatz (1986) proposed that the more commonly observed tight coupling between synthesis and translocation of secretory proteins was merely an operational one to ensure that proteins were not allowed to fold before translocation. Walter and Lingappa (1986) favour this view of ribosome-dependent versus ribosome-independent translocation as reflecting the requirements of the maintenance of a 'translocation competent state'.

Wiech et al (1990) have recently collated these theories and define a translocation competent state as meaning that the protein must be water-soluble, possess secondary but no tertiary structure, and interact with cytosolic 'chaperones'. This last term has been used to describe proteins that affect the folding of others without playing any part in their final structure (see Ellis and Hemmingsen, 1989). The signal recognition particle (see below) and members of the heat shock protein family (hsp70) (Chirico et al, 1988) act as such chaperones in the translocation of proteins across the ER lumen.

1.3.3 TRANSLOCATION ACROSS THE ER MEMBRANE

The translocation across the ER membrane has been reviewed by Walter and Lingappa (1986) and, more recently, by Pugsley (1990). Walter and Lingappa suggest that the two questions to be addressed when considering ER translocation are:

(i) how is the target membrane recognised, and
(ii) once targeted to the membrane surface, how is the protein translocated across the membrane?

(i) Recognition of Target membrane.

Proteins destined for the secretory pathway are distinguished by the presence of signal peptides. These signals are discrete N-terminal extensions that are cleaved from the mature protein during translocation (for review of signal peptides see von Heijne, 1990). They have very little or no amino acid (primary) sequence homology but comparison of many known signal peptides has elucidated three functional domains: a positively-charged amino-terminus, or n-region; a hydrophobic h-region; and a carboxy-terminal c-region which contains polar residues and contains the signal peptide cleavage site (von

Wiech et al (1990) have recently collated these theories and define a translocation competent state as meaning that the protein must be water-soluble, possess secondary but no tertiary structure, and interact with cytosolic 'chaperones'. This last term has been used to describe proteins that affect the folding of others without playing any part in their final structure (see Ellis and Hemmingsen, 1989). The signal recognition particle (see below) and members of the heat shock protein family (hsp70) (Chirico et al, 1988) act as such chaperones in the translocation of proteins across the ER lumen.

1.3.3 TRANSLOCATION ACROSS THE ER MEMBRANE

The translocation across the ER membrane has been reviewed by Walter and Lingappa (1986) and, more recently, by Pugsley (1990). Walter and Lingappa suggest that the two questions to be addressed when considering ER translocation are:

(i) how is the target membrane recognised, and
(ii) once targeted to the membrane surface, how is the protein translocated across the membrane?

(i) Recognition of Target membrane.

Proteins destined for the secretory pathway are distinguished by the presence of signal peptides. These signals are discrete N-terminal extensions that are cleaved from the mature protein during translocation (for review of signal peptides see von Heijne, 1990). They have very little or no amino acid (primary) sequence homology but comparison of many known signal peptides has elucidated three functional domains: a positively-charged amino-terminus, or n-region; a hydrophobic h-region; and a carboxy-terminal c-region which contains polar residues and contains the signal peptide cleavage site (von

Heijne, 1985). The n- and h-regions are known to be involved in membrane targeting - the hydrophobic core forming an alpha helix. The c-region often contains alpha-helix breaking residues such as glycine and proline, and residues at positions -3 and -1 relative to the cleavage site tend to be small and uncharged. All plant storage proteins synthesised as preproproteins carry N-terminal extensions that fit this description.

The signal peptide interacts with a cytosolic factor, the signal recognition particle (SRP). SRP, an 11S small cytoplasmic ribonucleoprotein, was purified by Walter and Blobel (1980) and consists of a 7S RNA component and 6 polypeptide chains. During protein synthesis, SRP recognises the signal peptide as it emerges from the ribosome, binds to the nascent chain and arrests further polypeptide elongation. Krieg et al (1986) showed that the 54kD monomeric polypeptide of SRP contained the signal peptide binding site. Sequencing of cDNA encoding the 54kD protein (Bernstein et al, 1989: Romish et al, 1989) has revealed a methionine-rich binding pocket which may accommodate the hydrophobic signal sequence. The deduced amino acid sequence also reveals a potential GTP binding site. Pugsley (1990) suggests that GTP binding/hydrolysis at this site may affect the affinity of SRP for the signal peptide.

The SRP/nascent polypeptide/ribosome complex then interacts with another characterised protein factor, the SRP receptor (Walter and Blobel, 1981) or docking protein (Meyer et al, 1982). This is a heterodimeric integral membrane protein consisting of a 69kD trans-membrane alpha chain and a 30kD beta chain. A hydrophilic cytosolic domain of the alpha subunit strongly resembles a nucleotide binding site and Walter and Lingappa (1986) suggest that the RNA component of SRP may interact with this domain.

At this stage the ribosome binds to the membrane in a functional way. Meyer's group have recently purified a 180 kD integral ER membrane protein which binds to ribosomes and that they propose is the ribosome binding factor (unpublished).

The next step in the translocation process is the GTP-mediated release of the SRP from the signal peptide (Connolly and Gilmore, 1989). This allows the recycling of SRP, its receptor and the resumption of peptide elongation. A second membrane protein, identified by Weidmann *et al* (1987) and termed the Signal Sequence Receptor (SSR), (see also Walter, 1987) binds the signal peptide and may be involved in the release of the SRP from the signal peptide (reviewed by Rapoport, 1990).

(ii) Translocation of the nascent peptide.

Although signal peptides are hydrophobic and can insert into lipid monolayers (Cornell *et al*, 1989), the mature portion of the protein is unlikely to pass through the lipid bilayer. Instead, integral membrane proteins are thought to assemble to form a channel, or translocon, through which the protein is translocated (see Pugsley, 1990). The SSR mentioned above may form part of this translocon (Rapoport, 1990).

Plant storage preproprotein precursors can be translocated across mammalian membranes and the involvement of the SRP in this translocation has been demonstrated using globulin precursors from *Vicia faba* (Bassuner *et al*, 1984). This suggests that similar mechanisms are involved in the targeting to and translocation across the plant cell ER membrane although there is no direct data on the plant components themselves.

1.3.4 CO-TRANSLATIONAL MODIFICATIONS.

As the nascent polypeptide emerges into the lumen of the ER, co-translational cleavage of signal peptide and core asparagine-linked glycosylation takes place catalysed by the membrane-associated hexameric signal peptidase and a glycosyl-transferase complex respectively. The glycosylated asparagine residue forms part of a tripartite recognition sequence, Asn-X-Ser/Thr (where X can be any amino acid), in animals and higher plants. The glycan, composing 3 glucose, 9 mannose and 2 N-acetylglucosamine residues, is transferred from a lipid donor to the side chain of the asparagine. Sequential modifications of the oligosaccharides occurs both in the ER lumen and in the Golgi (reviewed by Kornfield and Kornfield, 1985).

1.3.5 PROTEIN FOLDING IN ER.

(i) Disulphide bridge formation.

The oxidising environment of the ER lumen promotes the spontaneous formation of disulphide bridges between intra- and inter-chain cysteine residues. A soluble ER resident protein, Protein Disulphide Isomerase (PDI), has been identified (Freedman, 1984). It catalyses the formation of productive disulphide bridges by lowering the activation energy of reactions involving free thiol groups and/or disulphide chain exchanges. Its ability to catalyse the formation of the correct disulphide bridges within a newly-synthesised protein was illustrated by Bulleid and Freedman (1988). They observed that the wheat storage protein, gamma gliadin, could be imported into PDI-depleted microsomes but was not correctly disulphide bonded. Addition of purified PDI to the microsomes during their preparation resulted in the correctly folded form of the protein.

(ii) The involvement of Binding Protein.

Another resident ER protein that has an effect on the folding and assembly of proteins is the binding protein (BiP). This protein binds tightly to ATP, shows great homology with the hsp70 chaperone family, and is likely to act in a similar way (Munro and Pelham, 1986). A simple functional model for BiP is that it acts as a reversible detergent which binds hydrophobic residues exposed on the surface of misfolded monomers or unassembled oligomeric subunits. The binding of BiP shields the hydrophobic regions, and thus avoids protein precipitation. This interaction could be reversed by the hydrolysis of BiP-bound ATP, which may also provide the energy necessary for correct folding or assembly. Munro and Pelham (1986) showed that BiP (from S.cerevisiae) bound to immobilised immunoglobulin heavy chain could be released upon the addition of ATP but not ADP, indicating that hydrolysis of the high energy phosphate bond is the driving force for dissociation. In plants, a BiP homologue has been reported in tomato (see Pelham, 1989)

1.3.6 PROTEIN EXPORT FROM ER.

Proteins destined for compartments other than the ER (lysosome/plant vacuole, secretory vesicles, or those destined for extracellular secretion) exit the ER and pass along the route of 'Biosynthetic Protein Transport' (Pfeffer and Rothman, 1987) . Pfeffer and Rothman propose that the pathway for biosynthetic protein transport is constitutive, non-selective and unidirectional. Proteins deposited in the ER lumen are transported to the Golgi and then to the cell surface in a series of transport steps involving the budding and fusion of transport vesicles from a donor to an acceptor organellar compartment. They use the term 'bulk flow' to describe the constitutive, non-selective nature of the protein trafficking.

They also review evidence which suggests that the cytosolic factors involved in vesicle trafficking are conserved in all eukaryotes. A feature of the vesicles is that they are coated on their cytosolic face but the coat does not contain clathrin. Clathrin is involved in receptor-mediated endocytosis and, therefore, Pfeffer and Rothman suggest that the distinction between clathrin and non-clathrin coating may represent the difference between selective and non-selective transport.

A consequence of the bulk flow secretory pathway is that proteins whose destinations are not the extracellular medium, such as resident ER/Golgi proteins and those destined for the lysosomes/plant vacuole, require signals to retard them or direct them away from this pathway. Pfeffer and Rothman suggest that these need not be contiguous blocks (such as the signal peptide) and may instead be represented by 'signal patches' formed as the protein assembles into its tertiary and quaternary structures.

Evidence that the same route of biosynthetic protein transport exists in plants was first shown by the observation that many seed storage proteins are glycosylated and the protein-borne oligosaccharides contain Golgi-specific modifications (Chrispeels, 1983ab). Also Greenwood and Chrispeels (1985) observed immunocytochemical localisation of phaseolin and phytohaemagglutinin in the ER and Golgi of developing bean cotyledons.

Recent evidence that transport along this pathway is by default has been provided by the work of Dorel *et al*, (1989), Iturriaga *et al* (1989), and Denecke *et al* (1989). Dorel *et al* (1989) expressed a chimeric gene, containing the promoter and signal peptide coding sequence of a vacuolar seed protein fused to the coding sequence of a cytosolic seed albumin, in transgenic tobacco. They found that the chimera was correctly expressed and glycosylated but was not associated with any

organellar fraction. They concluded that the chimera had been secreted though this was not directly demonstrated. More convincing evidence was provided by Iturriaga *et al* (1989) and Denecke *et al* (1989) who both fused genes for known signal peptides onto marker enzymes and expressed them in tobacco. Denecke and his colleagues used electroporation of protoplasts and transient expression assays to show the secretion of the marker enzymes from the transformed protoplasts into the medium.

1.3.7 CONTROL OF ER EXPORT (see review by Pelham, 1989)

Given that export from the ER can occur at a similar rate to bulk flow (Weiland *et al*, 1987: Karrenbauer *et al*, 1990), what is the explanation for the observed variation in half-times of secretion between exported proteins? Pfeffer and Rothman (1987) suggest two factors that may be involved. The first is that the rates of export reflect the variability in the rates of protein folding/assembly, as only correctly folded and assembled proteins are exported (reviewed by Rose and Doms, 1988). For example, Vitale *et al* (1991) have expressed beta-phaseolin in *Xenopus* oocytes and found that only correctly assembled trimers were secreted. Furthermore, they observed that unassembled subunits were rapidly degraded in the ER. Secondly, the high surface area to volume ratio of the ER may result in the membrane acting as a absorption matrix. The rate of movement of exported proteins through the lumen would then be determined by the number and strength of their hydrophobic/electrostatic interactions with this matrix.

The final question to be addressed when considering export from the ER is the observed retention of resident luminal ER proteins. Munro and Pelham (1986) identified a conserved 4aa C-terminal sequence, lysine-aspartate-glutamate-leucine (KDEL), present on three major ER resident proteins;

BiP, PDI and endoplasmin (grp94). They have subsequently shown that removal of this sequence from BiP leads to its slow secretion. In converse experiments the tetrapeptide sequence, fused to the C-terminus of a secretory protein, results in its retention in the ER (Munro and Pelham, 1986). This data suggests that KDEL functions as an ER retention signal. Interestingly, the sequencing of BiP from the yeast S.cerevisiae (Rose and Misra, 1989) revealed histidine-DEL (HDEL) which Dean and Pelham (1990) have recently shown to be the ER retention signal in this yeast species.

Many more ER resident proteins have now been examined and show a KDEL (or KDEL related) C-terminus (for a list see Pelham, 1989). In the higher plant proteins so far examined, BiP cloned from tomato appears to have a HDEL signal (see Pelham, 1989) whereas the auxin binding protein from Zea mays has a KDEL C-terminus (Inohara et al, 1990).

This ER retention signal is thought to interact with a specific receptor in a post-ER compartment which then returns the resident lumenal protein to the ER (Munro and Pelham, 1987: Warren, 1987). Putative KDEL receptors have been identified in animals and yeast. Using an anti-idiotypic antibody approach Vaux et al (1990) identified a 72kD integral membrane protein which bound anti-anti-KDEL immunoglobulin and was concentrated in a post ER compartment. Pelham's group, on the other hand, used a genetic approach in S.cerevisiae and suggest that the ERD2 gene product, a 25kD protein with 7 predicted membrane-spanning domains, is the HDEL receptor (Lewis et al, 1990: Semenza et al, 1990). They have also isolated a human cDNA ERD2 homologue indicating that this gene product is highly conserved (Lewis and Pelham, 1990).

The ability of plant systems to recognise and

specifically retain KDEL-tagged proteins was recently shown by Herman et al, (1990). They used site-directed mutagenesis to modify the C-terminus of the legume vacuolar protein phytohaemagglutinin (PHA) such that the four terminal aa residues were K-D-E-L. When expressed in transgenic tobacco they found a significant proportion of the mutant localised in the ER. However, some of the PHA-KDEL was not retained. Their results, together with those of Zagouras and Rose (1989) who observed retardation (rather than retention) of two KDEL-tagged animal secretory proteins expressed in mammalian tissue culture cells, suggest that there are other factors responsible for efficient retention. Herman et al (1990) suggest that the reason for incomplete retention was that the KDEL signal was not being presented optimally on their PHA-KDEL chimera.

1.3.8 VACUOLAR TARGETING

It is now thought that the trans-Golgi network (TGN), a series of tubulo-vesicular extensions of the trans-Golgi, is the site where such sorting takes place (reviewed by Griffiths and Simons, 1986).

(i) Targeting to the Lysosome

The mannose-6-phosphate pathway in animals enables lysosomal enzymes to reach their correct sub-cellular destination (Kornfield and Mellmam, 1989). Lysosomal enzymes have an as yet uncharacterised signal which is recognised by a phosphotransferase localised in the early Golgi. Upon recognition, the enzyme phosphorylates the lysosomal protein on a mannose residue. This is then recognised by the mannose-6-phosphate receptor in the TGN (Gueze et al, 1985) and targeted to the lysosome. The search for the phosphorylation signal (see Baranski et al, 1990) is underway.

(ii) Yeast Vacuolar Targeting

Targeting to the yeast vacuole has also been

investigated. Two vacuolar enzymes, carboxypeptidase Y (CPY) and proteinase A (PrA), are both synthesised as preproproteins. The N-terminal propeptide (a stretch of amino acid residues located immediately downstream of the signal peptide) was shown to contain vacuolar sorting information and fusion of this region from preproCPY (Johnson *et al*, 1987: Valls *et al*, 1987) or preproPrA (Klionsky *et al*, 1988) to the yeast secretory protein invertase resulted in targeting of the fusion protein to the vacuole. Valls *et al* (1990) have since restricted the vacuolar targeting signal to four contiguous amino acids within the propeptide of CPY - that is, glutamine-arginine-proline-leucine (QRPL). The large numbers of different cytosolic and membrane proteins involved in yeast vacuolar sorting is illustrated by the isolation and characterisaton of a host of sorting mutants which fall into 40 vacuolar protein sorting (VPS) complementation groups (Rothman *et al*, 1989).

(iii) Plant Vacuolar/Protein Body Targeting

The story in higher plants is much less clear. As Bowles (1990) points out, most studies on vacuolar trafficking to date have involved seed storage protein deposition in dicotyledons. These appear to follow the biosynthetic pathway as described above. Hoffman *et al* (1987) and Wilkins *et al* (1990) have both expressed monocot storage proteins in tobacco and showed that that they are correctly targeted to the vacuole. However, as discussed in Section 1.2 there is evidence to suggest that some monocot seed storage proteins may not pass along the biosynthetic pathway. Targeting of proteins to the vacuole in monocots therefore, remains unclear (see review by Bowles, 1990).

(iv) Vacuolar/Protein Body targeting in dicotyledons.

As discussed earlier, sugar modifications specific for the late (trans-) Golgi appear on many glycosylated seed storage proteins suggesting that, as with animal (lysosomal)

and yeast systems, targeting occurs in a late compartment along the biosynthetic transport pathway. However, treatment of developing cotyledons or seeds with the glycosylation inhibitor tunicamycin has shown that glycosylation is not required for the correct targeting of some seed storage proteins (Bollini et al, 1985: Lord, 1985b). More recently, site-directed mutagenesis has been used to remove the glycosylation signals present in PHA (Voelker et al, 1989). When this mutant was expressed in transgenic tobacco the non-glycosylated PHA mutant was correctly targeted to the protein bodies.

One protein whose intracellular transport is inhibited by tunicamycin treatment is the precursor form of the jackbean lectin, Concanavalin A (Faye and Chrispeels, 1987). However, the post-translational modifications of this lectin are complicated and Bowles (1990) suggests that the inhibition of glycosylation results in an active lectin species which binds to glycoproteins at the luminal face of the RER.

Correct tissue-specific expression, glycosylation and targeting to the protein bodies is observed when the plant storage proteins B-phaseolin (Sangupta-Gopalan et al, 1985), PHA (Sturm et al, 1988) and pea vicilin (Higgins et al, 1988) are expressed in the seeds of transgenic tobacco. Furthermore, the potato tuber vacuolar protein patatin was correctly glycosylated and transported to the vacuoles of transformed tobacco leaf protoplasts (Sonnewald et al, 1990). This is an interesting observation as vacuoles in parenchymal tissue are essentially lytic/lysosome-like compartments whereas the vacuolar-derived protein storage vesicles (protein bodies) are primarily storage compartments. This data therefore suggests a common mechanism for vacuolar sorting in different tissues and between different plant species.

Tague and Chrispeels (1987) have shown that when PHA is expressed in yeast it is correctly targeted to the vacuole.

This suggests that there may be common mechanisms for vacuolar targeting in yeast and higher plants.

Expression of B-phaseolin in Xenopus laevis oocytes (Vitale et al, 1986), PHA in chinese hamster ovary (COS) cells (Voelker et al, 1986) and B-phaseolin in a baculovirus-infected insect cells (Bustos et al, 1988) all lead to their secretion, indicating that different sorting mechanisms are involved between plants and animals, and contrasting with the apparently common mechanism of ER retention (see section 1.2.7 above).

(v) The search for a plant vacuolar targeting signal in yeast.

Following their successful vacuolar targeting of PHA in yeast (Tague and Chrispeels, 1987), Chrispeels and his colleagues fused portions of cDNA encoding the N-terminal regions of PHA to the secreted yeast protein invertase (Tague et al, 1990). They localised the vacuolar targeting information to a region representing residues 14 to 23 of mature PHA. They further defined the signal as being centred around a Glutamine residue at position 19. By comparison to other legume lectins, they showed that this glutamine residue was conserved and the residues in positions 19 to 22 showed similarity to the QRPL vacuolar targeting sequence described above (Valls et al, 1990).

However, although mutation of this Glutamine 19 caused secretion of the PHA/invertase hybrids, it had no affect on the correct targeting of the complete PHA polypeptide when expressed in yeast. Thus it seems other regions in the mature PHA molecule may act as a targeting signal, possibly in synergy with the N-terminal propeptide.

(vi) Clathrin coated vesicle transport to vesicles.

Further evidence for receptor-mediated transport of plant proteins to the vacuole is provided by Harley and Beevers (1989). They showed that clathrin-coated vesicles were involved in the transport of pea lectin to the storage vacuoles in pea.

Clathrin is known to be involved in receptor-mediated vesicular trafficking (see review by Robinson and Depta, 1988).

(vii) Plant defence-related proteins

Both the wound induced proteinase inhibitors from tomato leaves (Graham *et al*, 1985) and the basic isoforms of the pathogen-induced B-1,3-glucanases from tobacco (Van den Bulcke *et al*, 1989) are targeted to the vacuole. Graham *et al* (1985) found that, in addition to a signal peptide, the primary translation product of the tomato proteinase inhibitor contained a 19aa residue N-terminal extension that was cleaved in the vacuole. On the other hand, the vacuolar forms of B-1,3-glucanase are synthesised in tobacco with a cleavable C-terminal propeptide (Shinshi *et al*, 1988). Do these extensions hold vacuolar targeting information?

In support of this, Reikhel and co-workers (Bednarek *et al*, 1990: Wilkins *et al*, 1990) have recently shown in monocots that barley lectin precursor contains a C-terminal extension which is necessary for vacuolar targeting.

1.3.9 SUMMARY

It is clear that intracellular sorting in plant systems shares some common features and mechanisms with the well-characterised pathways in animals and yeasts. In particular, evidence now exists for common mechanisms of membrane insertion, ER retention and vesicular traffic. However, there appear to be some features unique to the Plant Kingdom - such as the oligosaccharide-independent targeting of proteins to the vacuole which is not necessarily mediated by an N-terminal sorting sequence. The nature and location of plant protein body and vacuolar targeting signals are currently poorly understood.

1.4 RICIN AND THE CASTOR BEAN

1.4.1 INTRODUCTION

Ricin is isolated from the seeds of the castor bean, Ricinus communis L. (family, Euphorbiaceae). The castor bean plant is grown in many countries as it is a hardy plant and a good source of oil; a dry seed typically contains 100mg of lipid (Roberts and Lord, 1981a). The fruiting body is 3-4cm in diameter, spined and contains three seeds. Seeds are brown in colour with black flecks. The endosperm is the main storage organ and is the most prominent internal tissue - the cotyledons being poorly developed.

The protein bodies of the endosperm are the main sites of protein storage and their protein components have been well characterised (Youle and Huang, 1976: Tully and Beevers, 1976). They are divided into crystalloid and soluble matrix fractions. The crystalloid component consists of 11S globulins whereas the 7S lectins and 2S albumins form the main part of the matrix. All of these storage proteins are disulphide-bonded heterodimers, initially synthesised as preproproteins and are transported along the secretory pathway en route to the protein bodies (Lord, 1985). Storage protein synthesis occurs at a late stage in seed development; maximal synthesis occurring at the time of testa development (Roberts and Lord, 1981a).

1.4.2 11S GLOBULINS.

The 11S globulins consist of a heterogenic group of closely-related non-glycosylated polypeptides which can account for up to 80% of the protein storage reserve in the mature seed (Tully and Beevers, 1976). They are legumin-like, consisting of six non-covalently bound subunits, each subunit being a disulphide-linked heterodimer with a large, acidic A chain and a smaller, basic B chain (Gifford and Bewley, 1983). The A and

B chains are synthesised as a single polypeptide. Cleavage of this precursor has been demonstrated *in vitro* using a matrix fraction from castor seed protein bodies as a source of processing activity (Fukasawa *et al*, 1988).

1.4.3 2S ALBUMINS.

The small 2S albumins are allergenic (Youle and Huang, 1978) and are a major store of nitrogen and sulphur, containing a high proportion of glutamate and cysteine residues (see Wilson, 1986). A protein sequence has been determined for a 2S albumin and both cDNA and genomic clones have been isolated (Sharief and Li, 1982: Irwin, 1989). Analysis of the cDNA sequence reveals that the 2S albumins are synthesised as a single precursor, the two subunits being separated by a 3aa residue linker which terminates in asparagine. Irwin demonstrated the processing of the albumin precursor *in vitro* using a matrix fraction of protein bodies isolated from dried seeds. Irwin (1989) also presented evidence for a unique polycistronic mRNA species which encodes two closely-related albumin precursors.

1.4.4 7S LECTINS - RICIN

Ricin and *Ricinus communis* agglutinin (RCA) are the two closely-related lectins found in the matrix fraction of the castor seed protein bodies. Ricin (RCA II) is a 62kD monovalent lectin which agglutinates cells only weakly, but is one of the most cytotoxic proteins known. For instance, Balint (1974) comments that 1kg of purified ricin would be sufficient to kill 3.6 million people. It is a heterodimeric molecule with a 265aa residue A chain linked by a single disulphide bond to a 260aa residue B chain. Both chains contain two N-glycosylation recognition sequences (Asn-X-Ser/Thr). The A chain exists in both singly and doubly glycosylated forms: the composition of

each oligosaccharide moiety being (N-acetylglucosamine)$_2$ (xylose)$_1$(fucose)$_1$(mannose)$_3$ (see Lord *et al*, 1987). The B-chain is glycosylated with two oligosaccharide groups composed solely of (N-acetylglucosamine) and (mannose) and it also contains four intra-chain disulphide bridges.

RCA (RCA I) is a divalent lectin with strong agglutinating properties. Structurally it is composed of two ricin-like heterodimers linked by non-covalent forces. Both A and B chains of RCA closely resemble their respective subunits in ricin and immuno cross-reactivity has been shown to exist. However, the RCA B chain contains a fucosylated oligosaccharide chain not present in ricin B chain (Lord and Harley, 1985).

The A chains of both lectins have RNA specific N-glycosidase activity which allows them to modify and thereby to inactivate eukaryotic 60s ribosomal subunits (see Section 1.4.5). The B chains contain one high affinity and one low affinity sugar binding site specific for galactose and N-acetylgalactosamine.

Whole ricin has been crystallised (Villafranca and Robertus, 1977) and the 3D structure determined (Montfort *et al*, 1987). A schematic representation is shown in Figure 1.4.4. An interesting feature of the B chain is that it is a bilobal structure - each domain containing a galactose binding site and two disulphide bridges. It is suggested that this is a result of successive gene duplications of an ancestral galactose-binding peptide (Villafranca and Robertus, 1981).

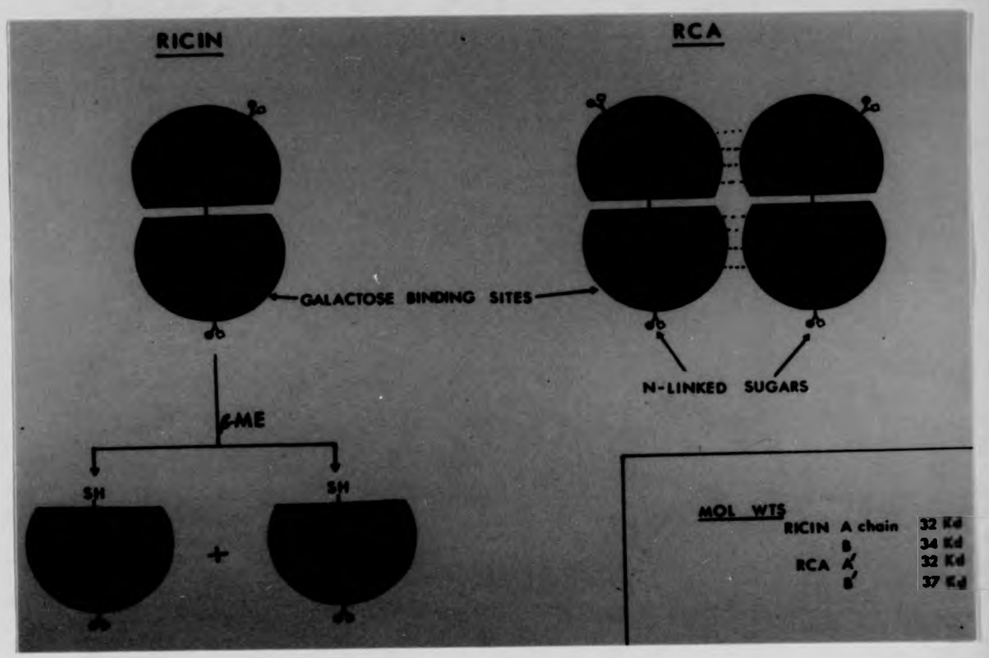
Ricin and RCA can be purified from a crude seed homogenate by affinity chromatography on Sepharose 6B; the lectins being selectively retained by interaction with the beta-galactose residues exposed on the matrix. A number of different isolectin forms exist (Lin and Li, 1980) as a result of expression from small multi-gene families (Roberts *et al*, 1990).

FIGURE 1.4.4 **SCHEMATIC REPRESENTATION OF THE 3D STRUCTURE OF RICIN AND RCA.**

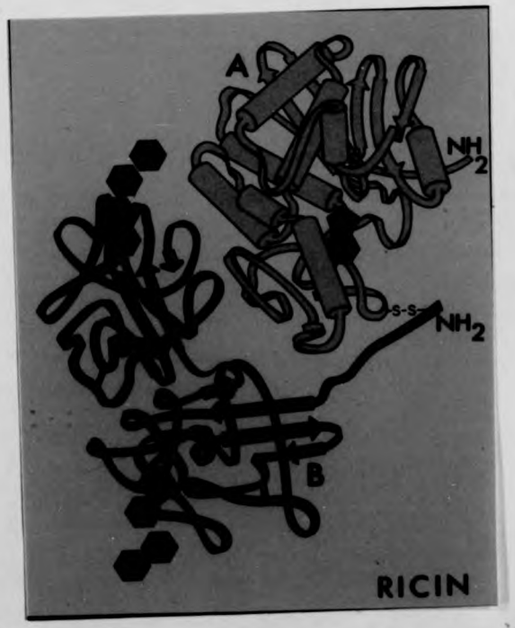
A/ Schematic representation of the castor bean 7S-lectins, ricin and RCA. Ricin is a heterodimer with a 32kD A chain and 34kD B chain, linked by a single disulphide bond. Both chains are N-glycosylated. The A chain (effectomer) enzymatically modifies eukaryotic 60S ribosomes and the B chain (haptomer), with its two sugar binding sites, binds to most eukaryotic cells, making ricin a potent cytotoxin. RCA resembles a dimeric form of ricin, with two A' chains and two B' chains, the two A'B' pairs being held together by non-covalent forces. The tetrameric configuration of RCA makes it a strong cell agglutinator, as the B chains can bind to sugar residues exposed on the surfaces of different cells. RCA however, is relatively non-cytotoxic. (Figure reproduced by permission from J.M.Lord.)

B/ Diagramatic representation of the 3D structure of ricin elucidated by X-ray crystallography (Montfort *et al*, 1987). The amino-terminus of A chain (grey) and B chain (blue) is shown (NH_2): the N-linked glycans are represented by hexagons: alpha-helices are represented as tubes and beta-sheets are shown by arrows. The A chain has both alpha-helical domains and beta-sheets. The B chain consists of two beta-sheeted domains, each containing a sugar-binding site, which have arisen as a result of gene duplication (Villafranca and Robertus, 1981).

A



B



Complete cDNA clones for both ricin and RCA have been isolated and sequenced (Lamb *et al*, 1985: Roberts *et al*, 1985). Amino acid predictions reveal that the A chains of the two lectins are 93% homologous and the B chains 84% homologous. Using clones derived from the ricin cDNA, soluble recombinant A and B chains have been synthesised separately in *E.coli* (O'Hare *et al*, 1988: Hussain *et al*, 1989). However, the recombinant B chain is extremely unstable when synthesised in this unglycosylated form. In an endeavour to find an expression system capable of producing a stable recombinant B chain, B chain clones have also been expressed in yeast and *Xenopus* oocytes using the 35aa residues from the N-terminus of preproricin to direct the B chain into the ER lumen (Richardson *et al*, 1988ab). Unfortunately, the recombinant B chain made in yeast was not stable because it contained high mannose glycans added to the protein by yeast sugar modification enzymes which are dissimilar to those in higher eukaryotic organisms. However, the recombinant B chain synthesised in *Xenopus* oocytes appears to be both stable and biologically active.

Both lectins are synthesised as single-chain preproproteins that have a 35aa residue N-terminal extension and a 12aa residue linking peptide which separates the sequences of the mature A and B chains. The N-terminal extension contains a ER signal peptide which mediates the co-translational translocation of nascent preproproteins across the ER membrane (Roberts and Lord, 1981b). It is at this stage that signal peptide cleavage and core-glycosylation takes place. The proteins are disulphide bonded and transiently sequestered in the ER lumen (Butterworth and Lord, 1983). Direct evidence for the lectin precursors being transported along the protein biosynthetic (secretory) pathway involving Golgi-localised oligosaccharide modifications is available (Lord, 1985ab: Foxwell *et al*, 1985). Lord (1985b) also showed

that the protein body targeting signal was not oligosaccharide-specific by showing the correct deposition of unglycosylated proricin in the protein bodies of tunicamycin-treated developing seeds.

Once sequestered in the protein bodies, proricin is processed by endoproteolytic removal of the 12aa residue linker (Harley and Lord, 1985). This linker terminates with an asparagine residue. Furthermore, the 35aa residue N-terminal extension of preproricin has a carboxy-terminal asparagine and, as will be shown, computer predictions strongly suggest that the whole extension is not removed by the signal peptidase in the ER. It is therefore likely that proteolytic processing also involves the removal of an N-terminal peptide.

1.4.5 RICIN - A TYPE II RIP

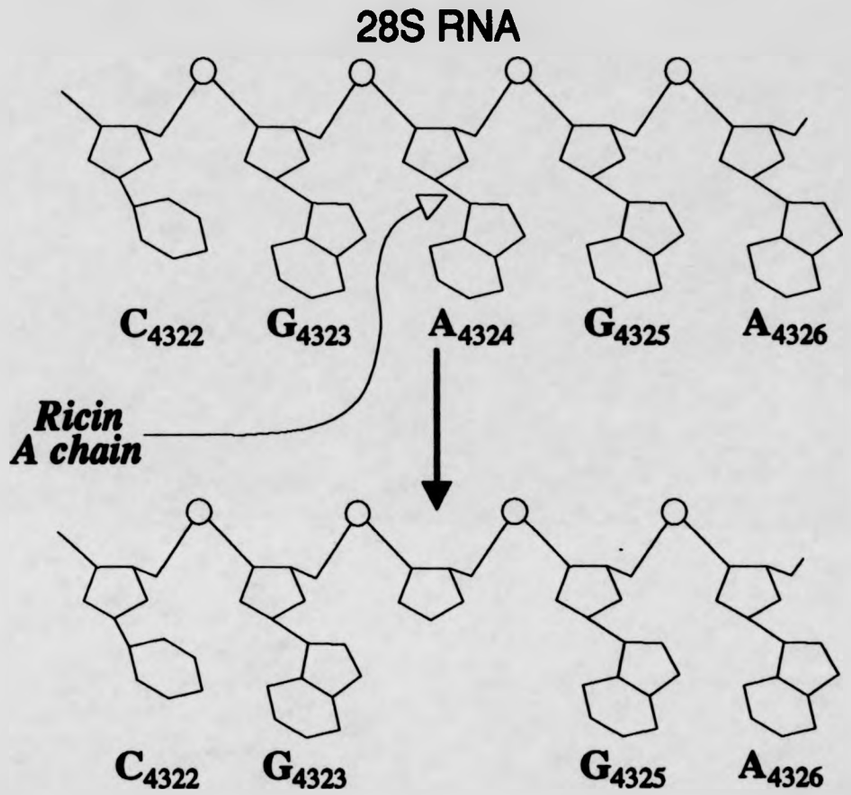
Ricin is part of a large family of toxic lectins termed Ribosome Inactivating Proteins (RIPs). Ricin A chain is the toxic moiety (effectomer) and it enzymatically inactivates 60s ribosomal subunits thus inhibiting eukaryotic protein synthesis (Olsnes *et al*, 1975: reviewed by Olsnes and Pihl, 1982). Protein initiation and elongation are inhibited in ribosomes pre-incubated with ricin (Osborn and Hartley, 1990). Its ability to arrest protein synthesis in rabbit reticulocyte lysates has been used in double message translation assays to allow characterisation of ricin A chain mutants *in vitro* (May *et al*, 1989).

The A chain is an RNA-specific N-glycosidase. It specifically depurinates a single adenine (A^{4324}) in rat 28S rRNA (Endo and Tsurugi, 1987: Endo *et al*, 1987) (see Figure 1.4.5). The depurination makes the RNA sugar-phosphate backbone susceptible to mild base hydrolysis, a feature which has been exploited experimentally to diagnose A-chain catalysed modification of ribosomes. The reagent commonly used is acetic

FIGURE 1.4.5 SITE OF ACTION OF RICIN ON RAT 28S rRNA.

The site of action of reduced ricin A chain on rat 28s rRNA. The A chain specifically depurinates a single adenine nucleotide in the 28S rRNA component of 60S eukaryotic ribosomal subunits (position 4324 in rat 28srRNA). This depurination makes the sugar-phosphate backbone of the 28S rRNA susceptible to mild base hydrolysis. Treatment with acetic aniline for instance releases a 650 nucleotide fragment, which can be resolved by denaturing gel electrophoresis and is diagnostic for ricin-catalysed ribosomal modification.

(Figure reproduced by permission of J.Cook.)



aniline (see May *et al*, 1989). The residues Glu^{177} and Arg^{180} in ricin A chain are involved in its enzymatic activity (Frankel *et al*, 1990: Chaddock,J. personal communication). It is unclear from the published literature whether the precursor forms of ricin have RNA N-glycosidase activity. However, the A chain alone is not toxic to intact cells. The B chain has a mediating role in the cytotoxicity of ricin as it binds to glycoproteins bearing exposed galactose residues on the cell surface. The endocytosis of ricin has been extensively reviewed (see Lord *et al*, 1987) and further discussion is not relevant here.

RIPs have been identified in other plants and bacteria and fall into two categories: type one RIPs which are non-cytotoxic and resemble ricin A chain; and type two RIPs which are cytotoxic heterodimers resembling whole ricin.

1.4.6 RICIN AND IMMUNOTOXINS

Their potent toxicity has made RIPs popular candidates for therapeutic purposes in selective cell destruction (reviewed by Lord *et al*, 1987). The term 'immunotoxins' has been coined to describe conjugates of immunoglobulins (Igs) raised against epitopes on the surfaces of malignant cells coupled to RIP derivatives. Initial tissue culture studies suggested that Ig:whole ricin conjugates were toxic but non-specific since the B chain binds opportunistically to most cell types. Conversely, Ig:Ricin A chain conjugates had a high specificity but variable toxicity. The lack of cell specificity and higher toxicity with the Ig:whole ricin conjugates have been attributed to the B chain. The binding of the B chain to animal cell surfaces has been discussed. The higher toxicity may be the result of a potentiating activity of the B chain on the intracellular membrane translocation of the A chain into the cytosol.

Current research therefore, is aimed at genetically engineering the B chain to remove the galactose binding sites but maintain its putative ability to potentiate the internalisation of the A chain. However, problems have arisen when trying to express the B chain or derived mutants. Glycosylation and the correct formation of the 4 intrachain disulphide bonds seem essential for solubility of the product (Richardson et al, 1989: Richardson et al, 1991). This is thought to be the reason that expression in the ER lumen of Xenopus laevis oocytes has proved to be the only successful heterologous system for this polypeptide to date. Once synthesised, the B chain requires stabilisation in solution by the presence of the A chain and galactose. It has therefore been considered for sometime that expression of the precursor form of ricin might be a better means of providing a soluble B chain, providing that a successful means of releasing the B chain from the A component is found.

1.4.7 SUMMARY

The castor bean seed storage proteins have been extensively studied and their subcellular sites of storage have been determined. In particular, the cytotoxic lectin, ricin, has been the focus of intense interest. Ricin cDNA, genomic and protein sequences have been reported. The three dimensional structure of mature ricin has also been determined. Biologically active recombinant ricin A chain has been successfully expressed and purified. The mode and site of action of this A chain on eukaryotic ribosomes has been investigated and its use in therapeutic agents has been documented. However, immunotoxins containing the A chain alone appear to have variable toxicity. Research in this field has now turned to look at the role of B chain in the internalisation of ricin into the cell. The B chain of ricin

has been expressed in bacterial, yeast and animal heterologous systems with variable success. It appears glycosylation and correct intra-chain disulphide bond formation of the expressed B chain are required for its solubility. To date the only useful system found for the expression of B chain has been in Xenopus oocytes. Ricin B chain is also stabilised in solution in the presence of the A chain and therefore co-expression of the single chain ricin precursor (containing both A and B chains) may prove to be a way of successfully expressing soluble B chain.

As with many plant storage proteins, ricin is initially synthesised as a preproprotein. It shares many common features with other seed storage proteins and so provides a good model for studying both intracellular targeting (to the protein bodies) and post-translational processing of these proteins.

Ricin has an RNA-specific N-glycosidase activity. Many enzymes are initially synthesised as inactive zymogen precursors and require endoproteolytic processing (which often requires the removal of one or more propeptides). By studying the enzyme activities of preproricin before and after cleavage it may be possible to determine whether ricin activation occurs in a similar way.

1.6 AIMS OF THE PROJECT

The aims of the work described here are summarised below:

1. To determine the function of the N-terminal propeptide of proricin.

2. To determine the biological activities associated with proricin.

3. To purify and characterise an (asparagine-specific) endoprotease(s) from castor bean endosperm protein bodies which might post-translationally process *in vitro* synthesised proricin.

4. To compare the RNA-specific N-glycosidase activity of proricin before and after linker cleavage.

CHAPTER 2

METHODS.

2.1 METHODS OF DNA AND RNA MANIPULATION

2.1.1 GROWTH AND MAINTENANCE OF BACTERIAL CULTURES.

Details of the Escherichia coli (E.coli) strains used in these studies are shown in Table 2.1.1.

(i) Maintenance.

Colonies were streaked onto glucose/minimal media plates (1xM9 salts, 1mM $MgSO_4$, 1mM thiamine-HCl, 0.1mM $CaCl_2$, 0.2% w/v glucose, 1.5% Bacto-agar: 10xM9 salts is 60g Na_2HPO_4, 30g KH_2PO_4, 10g NH_4Cl, 5g NaCl per litre). The plates were incubated inverted overnight, at 37C, before being sealed with parafilm. They were stored at 4C for periods of up to one month before re-plating.

(ii) Growth on solid media.

For the selection of cells transformed with plasmid DNA, bacteria were suspended in L-Broth (1% w/v Bacto-tryptone, 0.5% w/v Bacto-yeast extract, 1% w/v NaCl) and spread evenly onto L-Amp plates (L-Broth made 1.5% w/v Bacto-agar and 100µg/ml Ampicillin). The plates were incubated inverted overnight at 37C. Single colonies were picked from these plates and streaked onto fresh L-Amp media. These colonies could be stored for up to two weeks at 4C.

(iii) Growth in liquid media.

To create an overnight culture, 10mls of L-Broth was inoculated with a single colony picked from a plate, and incubated overnight at 37C, 200rpm, in an orbital shaker. To select for bacterial colonies transformed with plasmids bearing antibiotic resistance to ampicillin, the antibiotic was added to the L-Broth before incubation to a final concentration of 100µg/ml.

TABLE 2.1.1 **STRAINS OF E.COLI.**

Strain	Genotype
BMH71:18#	K12(lac-pro)SupE thi/F'proA^{+}B^{+} lacI^{q} lacZ deltaM15
BMH71:18mutL#	as above but also mutL::Tn10
JA221*	hsd^{+} trpE5 leuB6 lacY recA1/F' lacI^{q} Lac^{+} Pro^{+}
TG1 (Amersham)	K12, (lac-pro)SupE thi hsdD5/F'proA^{+}B^{+} traD36 lacI^{q} lacZ deltaM15

\# supplied by Michael Neuberger of the MRC Laboratories, Cambridge, UK. (ref. Kramer, Kramer and Fritz, 1984)

* supplied by Inouye,M., Department of Biochemistry, Robert Wood Johnson Medical School at Rutgers, New Jersey 08854 (ref. Takagi *et al*, 1988)

2.1.2 PREPARATION OF PLASMID DNA.

(i) Small-scale preparation of plasmid DNA - "mini-preps".

The alkaline-lysis method of Maniatis et al (1989) was used for the rapid preparation of plasmid DNA.

(ii) Large-scale preparation of plasmid DNA.

The method used was based on the method described by Holmes and Quigley (1981). An overnight culture was grown from a colony transformed with the plasmid of interest. The following day, 400mls of SOB medium (2% w/v Bacto-tryptone, 0.5% yeast extract, 10mM NaCl, 2.5% KCl, 10mM $MgCl_2$: made 100ug/ml Ampicillin) was inoculated with 4mls of this overnight culture. The 400ml culture was incubated overnight as described above. The cells were then pelleted at 5,000rpm for 10min at 4C, and resuspended in 7.2mls SET (25mM Tris-HCl pH8.0, 10mM EDTA, 15% w/v sucrose). To this, 7.2mls lysozyme solution (4mg/ml lysozyme in SET: made fresh) was added and the suspension incubated at room temperature for 5min. Then 12mls of 10% w/v Triton X-100 was added and the solution boiled over a bunsen flame until it became gelatinous and stuck to the walls of the container. The container was briefly immersed in a boiling water bath, before plunging it into icy water.

The gelatinous cell mass formed after these treatments was centrifuged at 20,000rpm for 30min at 4C. The supernatant from this centrifugation was transfered to a fresh tube and 0.5 volumes 7.5M ammonium acetate added. The solution was incubated on ice for 20min to precipitate protein, and was then centrifuged at 15,000rpm for 10min at 4C. The protein pellet was discarded, 0.7 volumes isopropanol added, and the nucleic acid left to precipitate at -20C for 10 minutes. The precipitated nucleic acid was pelleted at 15,000rpm for 10min at 4C. The pellet was drained and the nucleic acid resuspended in 4mls Low TE (10mM Tris-HCl pH8.0, 0.1mM EDTA).

The supercoiled plasmid DNA was then purified away from

RNA, chromosomal and nicked-plasmid DNA by resolution on a Caesium chloride gradient. 4.3g caesium chloride was dissolved in and 0.5ml ethidium bromide solution (5mg/ml in water) added to the 4ml isolated nucleic acid. The solution was sealed in a 4.5ml Beckman 'quick-seal' ultracentrifuge tube and centrifuged at 55,000rpm, 20C for 6 hours in a Beckman Vti65 rotor. The lower, supercoiled plasmid band was removed using a syringe and needle, and the ethidium bromide removed by three extractions with isopropanol which had first been equilibrated with caesium chloride-saturated TE (10mM Tris-HCl pH8.0, 1mM EDTA). Four volumes of Low TE was then added to the aqueous phase, the solution made 0.3M with respect to sodium acetate, and the DNA precipitated by adding 2 volumes of absolute ethanol and incubating at -20C for 2 hours to overnight.

Finally, the DNA was pelleted at 15,000rpm for 10min at 4C, the pellet washed in 70% ethanol, dried *in vacuo*, and resuspended in 400µl water.

(iii) Single-stranded (ss) M13 preparations.

The single-stranded (+ strand) form of vectors and constructs derived from the bacteriophage M13 (Messing, 1986) was prepared according the method of Schreier and Cortese (1979).

(iv) Replicative form (RF) M13 preparations.

2ml of 2xTY/magnesium (1.6% w/v Bacto-tryptone, 1% w/v yeast extract, 0.5% w/v NaCl, made 10mM $MgCl_2$), was inoculated with an M13 plaque taken from a fresh plate - using a glass pipette to remove the plaque in the form of an agar plug. 0.01 volumes of an overnight culture of *E.coli* TG1 cells was added, and the culture incubated in orbital shaker for 5-6 hours at 37C and 300rpm. The cells were then pelleted in a microfuge for 5min and the supernatant, containing infectious ssM13 particles, was stored overnight at 4C.

400ml of 2xTY/magnesium was inoculated with 0.01 volumes of an overnight culture of *E.coli* TG1 cells and incubated at

37C, 200rpm, until the cell culture was in mid-log phase (taken to be when the optical density of the culture at 600nm reached approximately 0.5). The suspension of infectious ssM13 particles was then added, and the culture grown for a further 4 hours at 37C, 300rpm. The RF form was then prepared as described for large-scale plasmid preparations above.

2.1.3 STORAGE OF DNA.

All preparations of purified plasmid and M13 (ss and RF) DNA were stored frozen in H_2O or TE at -20C.

2.1.4 QUANTITATIVE ESTIMATION OF DNA.

(i) Spectrophotometric estimation.

The concentration of DNA in an aqueous solution was estimated by measuring the absorbance at 260nm (A_{260}) and converting this value to mg/ml, viz; 1 OD unit is equivalent to 40ug/ml for double-stranded (ds) DNA or 30ug/ml for ssDNA. As an indication of purity, the level of protein contamination in a DNA solution was monitored by determination of the A_{260}:A_{280} ratio. A solution of pure nucleic acid gives an A_{260}:A_{280} ratio of 1.8 (Maniatis et al, 1989).

(ii) Gel analysis.

To quantitate fragments of plasmid DNA, a known aliquot was run on an agarose gel (as described below) alongside 500ng of bacteriophage lambda DNA that had been previously digested with HindIII and EcoRI and heated at 65C for 5min to dissociate the cos sites. The bands were visualised by placing the gel on a UV transilluminator, and the intensity of the fragment compared to those of the lambda bands. The concentration of DNA was then determined by reference to Table 2.1.4.

TABLE 2.1.4

SIZES AND CONCENTRATIONS OF LAMBDA FRAGMENTS AFTER EcoR1/HindIII DIGESTION.

Size of Fragment(bp)		% of total bp	ng DNA/500ng
21,226		43.7	218
5,148	run as 1 band	20.9	105
4,973			
4,268		8.8	44
3,530		7.3	37
2,027		4.2	21
1,904		3.9	20
1,584		3.3	17
1,375		2.8	14
947		2.0	10
831		1.7	9
564		1.2	6
125		0.3	2

2.1.5 RESOLUTION OF DNA BY GEL ELECTROPHORESIS.

(i) Agarose/TBE gels (Maniatis et al, 1989).

DNA samples were mixed with 0.1 volumes 10xLoading buffer (15% w/v Ficoll 400, 5mM EDTA, 0.1% Bromophenol blue) and loaded into the wells of a 11x14cm^2 horizontal slab gel comprising of 0.7%-1.5% w/v agarose in 1xTBE (10xTBE is 108g Tris-HCl, 55g boric acid, 9.3g EDTA per litre) and containing 0.5µg/ml ethidium bromide. The gels were run submerged in 1xTBE, in a BRL Series H5 electrophoresis tank at 60mA for 2-4 hours. Resolved DNA bands were visualised by placing the gel on a UV transilluminator.

(ii) Polyacrylamide gel electrophoresis.

Both buffer-gradient (Biggin et al, 1983) and linear (Maniatis et al, 1989) 6% w/v polyacrylamide gels were used for the resolution of DNA bands synthesised in sequencing reactions. The solutions necessary for preparation of a buffered-gradient gel are as follows:

40% acrylamide (1000ml)
380 g acrylamide
20 g bis-acrylamide
H_2O to 1000ml
Filtered through Whatman No1. paper
Stored in the dark at 4C.

0.5xTBE, 6% acrylamide gel mix (500ml)
75 ml 40% acrylamide
25 mls 10xTBE
230 g urea
Stored for 1-2 months in the dark at 4C.

(con. overleaf)

5xTBE, 6% acrylamide gel mix (200ml)

30 ml 40% acrylamide
100 ml 10xTBE
92 g urea
10 mg bromophenol blue
Stored for 1-2 months in the dark at 4C.

To prepare a 40cm x 20cm x 0.35mm gel, 30ml of 0.5xTBE gel mix and 7ml of 5xTBE gel mix were aliquoted into separate beakers. 60µl and 14µl of a freshly prepared solution of 25% w/v APS in water were added to the 0.5xTBE and 5xTBE gel mixes respectively. Polymerisation of the acrylamide was initialised by adding 60µl and 14µl of TEMED to the respective solutions and mixing. Immediately, 22ml of 0.5xTBE gel mix was taken up into a 50ml plastic syringe. The remainder was taken up into a 20ml syringe before slowly drawing up the 7ml of 5xTBE gel mix into the same syringe, being careful not to mix the two layers that formed. When this operation had been completed, three or four bubbles of air were drawn through the 20ml syringe in order to form a gradient between the two buffers. This solution was then carefully dispensed in between two taped glass plates separated by spacers, which were held at an angle of 45° to the vertical. After the dispensation of the gradient the plates were lowered onto the bench and the 50ml syringe picked up. The plates were then again raised to 45° and the contents of this second syringe were used to complete the gel. During the pouring of the gradient it was important to maintain an even flow of fluid in order to guard against the formation of bubbles. Once the gel had been poured, a suitable gel comb was fitted and the gel left in a horizontal position to polymerise.

For linearly-buffered sequencing gels, the following solution, suitable for 40cm or 50cm long gels, was prepared:

Linear gel mix (50ml)

7.5 ml 40% acrlyamide
5 ml 10xTBE
23 g urea
H_2O to 50 ml

To initiate polymerisation, 100µl of 25% w/v APS and TEMED were added, and the solution was immediately poured between prepared gel plates held at an angle of 45° to maintain an even flow. A comb was fitted and the gel left to polymerise.

All sequencing gels were run in 1xTBE, at a constant power of 40W for 2.5 - 5 hours. After running, the gel was fixed in 10% methanol, 10% acetic acid, for 15min, and dried onto 3MM paper on a heated vacuum gel drier. The dried gel was exposed to Fuji RX X-ray film in a lightproof cassette before developing the film as described in section 2.3.6.

For the sizing of synthetic oligonucleotides, phosphorylated in the presence of [^{32}P]ATP, 16% polyacrylamide gels were used. The 16% gel mix was made as shown below:

16% gel mix

40% acrylamide stock	80 ml
10xTBE	20 ml
urea	92 g
H_2O to	200 ml

Stored at 4C in the dark for 1-2 months.

To initialise polymerisation, 90µl 25% w/v APS and 90ul TEMED were added to 45ml of gel mix, and the gel poured between two taped 40x20cm^2 sequencing plates.

Before loading, the sample was denatured at 80C for 5min in formamide dye (95% formamide, 20mM EDTA, 0.1% w/v bromophenol blue, 0.1% w/v xylene cyanol). The gel was run at 40W in 1xTBE until the bromophenol dye had progressed two-

thirds through the gel. The gel was then fixed in 10% methanol, 10% acetic acid, covered with clingfilm and exposed to Fuji RX X-ray film.

2.1.6 GEL ELECTROPHORESIS OF RNA.

This was used for the resolution of radiolabelled mRNA species synthesised *in vitro*. A 15% formaldehyde, 1.5% agarose gel was made by first boiling 1.5g of agarose in 150ml H_2O and 20ml 10xMOPS buffer (0.2M MOPS, 50mM sodium acetate, 10mM EDTA - NaOH pH7.0). Once dissolved, the agarose solution was cooled to 60C and 30ml of filtered formaldehyde was added (40% w/v commercially available formaldehyde solution, filtered through Whatman No.1. paper before use). The gel was finally poured into a standard 15cm flat-bed gel plate fitted with a comb, and the gel allowed to set in a fume cupboard.

The samples were prepared by mixing 5µl of the RNA sample with 15µl of Denaturation solution (made immediately before use by mixing 500µl deionised formamide with 150µl filtered formaldehyde and 100µl 10xMOPS buffer). This was heated at 60C for 5min, cooled to room temperature and 2ul RNA loading buffer (50% glycerol, 0.2% w/v bromophenol blue) added. Before loading the samples onto the gel the wells were flushed out with Denaturation solution.

The gel was run submerged in 1xMOPS buffer, in a fume hood at 40mA for 2-3 hours. The gel was then dried down onto Whatman 3MM paper and the dried gel exposed to Fuji RX X-ray film.

2.1.7 GEL ISOLATION OF DNA FRAGMENTS.

Two methods of gel isolation were regularly used. For efficient recovery of DNA fragments, the low-melting point agarose gel method of Higuchi *et al* (1981) was found to be the most reliable and was followed exactly. However, for a more

rapid recovery of DNA fragments where high recoveries were not essential, the DNA was eluted onto filter paper exactly as described by Dretzen et al (1981).

2.1.8 PARTITIONING AND PRECIPITATION OF DNA.

All phenol and phenol/chloroform extractions were performed exactly as described by Maniatis et al (1989). Ethanol precipitation of nucleic acid, using 0.1 volumes 3M sodium acetate pH6.0 or 0.5 volumes of 7.5M ammonium acetate, was also as described by Maniatis et al (1989). Isopropanol precipitation was used to selectively precipitate large nucleic acid fragments away from short-chain RNA species and free nucleotides. For this, the DNA solution was made 0.6M with respect to sodium acetate pH6.0, 0.6 volumes isopropanol was then added and the DNA precipitated at room temperature for 5min. The precipitated DNA was pelleted, washed, dried and resuspended as described by Maniatis et al (1989).

2.1.9 RESTRICTION AND MODIFICATION OF DNA.

(i) Restriction digests.

Restriction enzymes were used according to the suppliers instructions. Incubations with DNA were carried out in either the buffers supplied with the enzyme or in the appropriate 1x salt buffer described by Maniatis et al (1989).

(ii) Phosphatasing reactions.

The removal of 5' phosphate groups from DNA fragments was done according to the method of Maniatis et al (1989). The reaction was carried out in 1xCIP buffer (made as a 10x stock, viz.; 0.5M Tris-HCl pH9, 10mM $MgCl_2$, 1mM $ZnCl_2$, 10mM spermidine) with 1 unit calf intestinal phosphatase (cip) per ug DNA. Incubation of the reaction was for 15min at 37C, followed by incubation at 56C for a further 15min. Another unit of cip was added and the incubation steps repeated. This method

was used successfully to remove 5' phosphate groups from both 'blunted' and 'sticky-ended' DNA fragments.

(iii) End-filling reactions

DNA fragments with 5' overhangs were blunt-ended following the method described in the BRL cloning manual (see Appendix for Suppliers).

(iv) Ligations.

DNA was ligated with T4 DNA ligase following the instructions described in the BRL cloning manual.

2.1.10 TRANSFORMATION OF E.COLI.

(i) Preparation of bacteria.

Exponentially growing E.coli were made competent for transformation by two washes in calcium chloride solutions. The cells were first pelleted at 3,000rpm for 5min at 4C, gently resuspended in 0.5 volumes 50mM ice-cold $CaCl_2$, and incubated on ice for 30min. The cells were then pelleted at 2,500rpm for 5min at 4C, resuspended in 0.05 volumes 50mM $CaCl_2$, and left on ice for a minimum of 30min before transformation.

(ii) Transformation of E.coli with plasmid DNA.

Competent cells were divided into 200µl aliquots on ice, 10ng of plasmid DNA was added, and the tubes vortexed briefly to help the DNA adhere to the surface of the cells. The tubes were incubated on ice for 45min, heat shocked at 42C for 90s, and allowed to recover on ice for 5min, before the addition of 300µl L-Broth. The tubes were then incubated at 37C for 60min to enable the transformed cells to begin expression of the plasmid-borne antibiotic resistance gene. The cells were then pelleted in a microfuge for 30s and resuspended in 160µl fresh L-Broth. 10µl, 50µl and 100µl were then spread onto L-ampicillin plates and incubated, inverted, overnight at 37C.

(iii) Transformation of E.coli with ss/RF M13 DNA.

200ul of competent cells were mixed with 1ng M13 DNA,

vortexed briefly, and incubated on ice for 30min. After heat-shocking for 45s, the cells were returned to ice for a further 5min, before plating out in 3ml molten B-top agar (8g Bacto-tryptone, 5g NaCl, 6g Bacto-agar per litre) containing 10µl 10mM IPTG, 25 µl 40mg/ml X-Gal (in dimethylsulphonide) and 300µl of a culture of exponentially growing E.coli, on glucose/minimal media plates. The plates were incubated inverted overnight at 37C.

2.2.11 SEQUENCING OF DNA.

(i) Preparation of DNA templates.

When screening putative mutants after site-directed mutagenesis ssM13 preparations were prepared and used directly as templates in sequencing reactions. To prepare caesium gradient-purified plasmid DNA for sequencing, the DNA strands were separated by incubating 2µg of DNA, diluted to 18ul in TE, with 2µl 2M NaOH for 5min at room temperature. The base was neutralised and the DNA precipitated by the simultaneous addition of 8µl 5M ammonium acetate pH7.5 and 100µl of -20C absolute ethanol. The tube was left on dry ice for 30min, the DNA pelleted for 10min in a microfuge, washed in absolute ethanol, and dried in vacuo. It was resuspended in 2µl H_2O and used immediately as a template.

To screen a number of bacterial colonies for a particular plasmid-borne sequence, the alkaline lysis mini-prep. method of Maniatis et al, 1989, was used with the following modifications. Cells from 5ml of an overnight culture were lysed instead of 1.5ml. Following the centrifugation step to pellet the chromosomal DNA, the plasmid DNA was ethanol precipitated from the supernatant and resuspended in TE. Boiled RNase A was added to a final concentration of 9 units/ml, and the solution incubated at 37C for 30min. The solution was made 150mM NaCl, the aqueous phase phenol/chloroform extracted

twice, extracted once with chloroform:IAA (24:1), and twice with ether. 0.2 volumes of 3M sodium acetate pH6 was added and the DNA precipitated with 0.6 volumes isopropanol. The DNA pellet was resuspended in 50µl TE. An aliquot was restricted and quantitated by gel analysis. 3-4µg of the mini-prep. DNA was then denatured as described for caesium-gradient purified DNA (above).

(ii) Sequencing reactions.

The synthetic oligonucleotides used as sequencing primers in this study, are shown in Table 2.1.11

All sequencing of DNA was performed using methods based on the dideoxynucleotide method of Sanger et al (1977). Early experiments followed the method of Bankier et al (1986), where Klenow was used to extend off the sequencing primer. However, the more reliable and versitile Sequenase™ kit (United States Biochemical), which uses a modified form of T7 polymerase, was later used successfully to sequence M13 and plasmid DNA templates. The method used was exactly as described in the protocol provided with the Sequenase™ kit.

The products of sequencing reactions were resolved on 6% polyacrylamide gels (prepared and run as described in Section 2.1.5). Immediately before loading the samples on the gel, 4ul formamide dye (95% formamide, 20mM EDTA, 0.1% w/v bromophenol blue, 0.1% w/v xylene cyanol FF) was added to each dideoxynucleotide-terminated reaction and they were incubated at 80C for 5min. Typically, loading 2-4µl of each reaction was sufficient to enable the sequence to be read after overnight exposure to Fuji RX X-ray film.

TABLE 2.1.11 **SEQUENCING PRIMERS**

Name (bp)	Sequence (5' - 3')	Reference No.
CL581 (16)	AAA CCG GGA GGA AAT A	-102 to -87
CL585 (17)	A GAT CCT AGC GTA ATT A	600 to 616
CL586 (18)	G CAA AGA CGT AAT GGT TC	696 to 713
CL587 (17)	T GAT GTT AGG GAT GGA A	900 to 916
CL588 (17)	GA TCT AAT GGA AAG TGT	1010 to 1026
NPT2 (20)	CAT CAG AGC AGC CGA TTG TC	1650 to 1631
UNIVERSAL PRIMER (17)	GTA AAA CGA CGG CCA GT	6310 to 6294

The reference points refer to the numbering of the published sequences of preproricin (Lamb *et al*, 1985) for the CL-labelled primers, neomycin (Beck *et al*, 1982) for the NPT2 primer, and M13mp18 (Norrander *et al*, 1983) for the Universal primer. The CL-labelled ricin primers annael to the strand complementary to the published sequence, the NPT2 primer annaels to the published strand, and the M13 Universal primer annaels to the + strand.

2.1.12 SITE-DIRECTED MUTAGENESIS (SDM).

Two methods of SDM were used and both are based on the method of oligonucleotide-directed, site-specific mutagenesis described by Zoller and Smith (1982). A list of the oligonucleotides used in this study are given in Table 2.1.12. All mutagenic reactions were carried out in M13-derived vectors. Consequently, the region of DNA to be mutated was first cloned into a suitable M13 vector and ssDNA was prepared using the methods described in the sections above. The modifications to the original SDM protocol of Zoller and Smith which are described below, were designed to increase the recovery of mutant colonies.

(i) In vitro synthesis of 'homoduplex mutant' DNA.

This method, first described by Nakamaye and Eckstein (1986), uses strand-specific selection to eliminate the non-mutant strand sequence in vitro, so that a homoduplex mutant DNA species can be transformed into E.coli. The E.coli DNA mismatch repair mechanism therefore does not recognise the mutation making the theoretical recovery of mutant progeny very high.

The mutant oligonucleotide is first annealed to the template in the usual way. Nakamaye and Eckstein (1986) found that the incorporation of thionucleotide dCTPalphaS into the mutant strand during second strand synthesis makes this strand resistant to digestion with NciI. The non-mutant strand can therefore be selectively 'nicked' by digestion with NciI. Limited exonuclease digestion is then used to remove the non-mutant region of DNA before re-initiating second strand synthesis. This time the mutant strand is acting as DNA template and the net result of this is the production of homoduplex mutant DNA. This avoids the need for screening large numbers of progeny and the identification of mutants can

TABLE 2.1.12 OLIGONUCLEOTIDES USED AS MUTAGENIC PRIMERS.

Name (bp)	Sequence (5' - 3')	Source
	*	
MW01 (21)	T ATA AGG CCA CTG GTA CCA AA	ICI
	** *** **	
MW02 (29)	A CTG GTA CCA AGG GGC AGT GCT GAT GTT T	ICI
	*** *** **	
MW03 (28)	T TTG CTT ATA GAA GGC CGG GTA CCA AAT	ICI
	* * *	
MW04 (22)	TTC AAT CAC TCG AGA CGA TCG A	WARWICK
	*** **	
MW05 (23)	ATA ACA ACA GTC GAC CCA AAC AA	WARWICK
	* * *	
OAB1049 (28)	TTG TTC AAT CAC TCG AGA CGA TCG ACG G	MILANO
	*** **	
OAB 1179 (29)	AGG ATA ACA ACA GTC GAC CCA AAC AAT AC	MILANO

OAB1049 & OAB1179 were gifts from Dr. Luca Benatti of Farmitalia, Milano. (* = mismatched base with template.)
MW01, MW02 and MW03 were gifts from Dr.M.E.Edge of ICI Pharmaceuticals.

usually be achieved by directly sequencing DNA prepared from 6-12 progenic plaques.

The reagents necessary for this method of SDM were available from Amersham ('Oligo-directed *in vitro* mutagenesis system, Version 2' code: RPN.1532) and the protocol supplied with the reagents was followed exactly.

(ii) DNA Repair-deficient strain of *E.coli*.

In this method, described by Kramer *et al* (1984), heteroduplex is first formed by annealing the mutagenic primer to the ssDNA template and synthesising the second strand *in vitro*. Kramer *et al* (1984) found that transformation of the heteroduplex mutant into *mutL* strains of *E.coli*, which are deficient in their DNA mismatch repair mechanisms, increased the recovery of mutant progeny. To minimise the introduction of opportunistic mutations, the transformants are plated out with exponentially growing wild-type cells, which form a lawn into which the M13-derived progeny transfect upon release from their transformed *mutL* host. The *E.coli* strains *71:18* and *71:18mutL* were used for this work (Table 2.1.1).

The mutagenic oligonuleotide was firstly phosphorylated at its 5' end by adjusting its concentration to 50D units/ml in H_2O, mixing 2.5µl with 3µl 10xKinase buffer (1M Tris-HCl pH8.0, 100mM $MgCl_2$, 70mM DTT, 10mM ATP) 24 µl H_2O, 2 units T4 polynucleotide kinase, and incubating at 37C for 15min. The reaction was stopped by incubation at 70C for 10min and the kinased oligonucleotide stored at -20C.

The oligonucleotide was annealed to the template by drying down 6.2µl of the phosphorylated oligonucleotide (10pmoles) in a 1.5ml eppendorf, resuspending it in 3.5µl H_2O, adding 0.5µg template (0.5mg/ml) and 0.5µl 10xTM (0.1M Tris-HCl, 0.1M $MgCl_2$), and floating the eppendorf in a test-tube of water at 80C. The test-tube was left to cool to room temperature on the bench. Meanwhile, an extension mix was made

up by mixing the following reagents together on ice:

Extension mix (6ul)	
10x TM	1 ul
5mM ATP	1 ul
0.1M DTT	1 ul
5mM dATP,dCTP,dGTP,dTTP	1 ul
H_2O	2 ul
TOTAL =	6 ul

The cooled eppendorf containing the annealing reaction was spun briefly to collect the condensate, and then 3µl extension mix, 3 units of T4 DNA ligase, 0.8 units of Klenow were added and the extension reaction incubated at 16C overnight. The mix was diluted to 100ul with TE and 1ul, 10µl and 20µl used to transform competent 71:18mutL cells, as described in Section 2.1.10., using 300µl of a culture of exponentially growing wild-type 71:18 cells to create a lawn.

To identify mutant progeny, 1.5ml cultures were prepared from 100 white plaques and the cultures grown as for the preparation of ssM13 DNA. 50µl of each supernatant from these cultures, and an equivalent amount of ssM13-derived template (control), were loaded onto a nitrocellulose filter (Hybond C) using a dot-blot apparatus and the filter dried *in vacuo* at 80C for 2 hours. The remainder of each supernatant was stored at 4C.

The filter was probed with radio-labelled mutant oligonucleotide, which was made by phosphorylating 15pmoles of oligonucleotide as described above, except that the 10xKinase buffer was depleted of ATP and 30uCi [$-^{32}P$]ATP was included in the reaction mix. The kinased oligonucleotide was then diluted to 3ml with 6xSSC (made as a 20x stock, *viz.*: 263g NaCl, 132.3g trisodium citrate, dissolved in 1.5l H_2O and adjusted to pH7.0 with concentrated NaOH), filtered through a 0.45µm filtation

unit, the unit washed through with a further 1ml of 6xSSC, and the probe stored at -20C before use. The filter was first pre-hybridised by incubating it for 60min at 67C in 10ml 6xSSC/0.2% w/v SDS/10xDenhardts solution (made as a 100x stock, *viz.*; 2% w/v BSA, 2% w/v Ficoll, 2% w/v polyvinylpyrrolidone). The filter was then rinsed for 1min in 50ml 6xSSC and hybridised with the probe at 67C for 30min. The hybridisation was completed by cooling to room temperature over 30min. The probe solution was then discarded and the filter washed 3 times in 3 x 100ml of 6xSSC for 5min at room temperature. The filter was then drained of excess moisture and monitored using a Geiger counter. The filter was then covered in clingfilm and exposed to Fuji RX X-ray film at -20C for 1 hour using 2 intensifier screens to amplify the signal.

The mutagenic oligonucleotides used in this study were all over 20 nucleotides in length and therefore the dissociation temperature of probe from mutated template (Td) could not be predicted precisely using the Wallace Rules (Suggs *et al*, 1981). For this reason, after exposure to film, the filter was washed in 6x SSC for 5min at temperature increments of 5^{o}C and the filter monitored for radioactivity after each wash. When the measured number of cpm had fallen to approximately half of the value recorded at room temperature, the wet filter was again exposed to X-ray film. The mutant progeny were taken to be those which still had probe bound at the temperature where the probe had been dissociated from the template controls. ssDNA was then prepared from the aliquots of these clones that had been stored at 4C. The presence of the mutated bases was then assessed by sequencing before preparing RF M13 DNA and cloning the mutated fragment back into the original vector.

2.2 IN VITRO EXPRESSION OF PREPRORICIN DNA.

2.2.1 SP6/T7 IN VITRO TRANSCRIPTIONS.

DNA to be expressed in vitro was first cloned into either pSP64T (expression driven from the SP6 viral promoter: see Appendix for map) or a pGEM vector (for expression driven from the T7 viral promoter). Then 10ug of caesium-gradient purified DNA was linearised with a suitable restriction enzyme for which the construct had a unique recognition site located downstream of the preproricin gene. The DNA was phenol/chloroform extracted twice, ethanol precipitated and resuspended to a final concentration of 1mg/ml.

The method of transcription used was as described by Krieg and Melton (1986). A reaction was set up as follows:

SP6/T7 Transcription reaction

linearised DNA (1mg/ml)	2.0 µl
H_2O	2.5 µl
transcription PREMIX	12.0 µl
RNasin	0.5 µl
5mM M-G(5')ppp(5')G (CAP)	1.0 ul
T7/SP6 polymerase	2.0 µl
TOTAL =	20.0 µl

Transcription PREMIX

10xsalts (20mM spermidine, 400mM HEPES KOH pH7.5, 60mM magnesium acetate)	1000 µl
10 mg/ml BSA	100 µl
500mM DTT	100 µl
50mM rUTP, rATP, rCTP	100 µl
5mM rGTP	200 µl
H_2O	4500 µl
TOTAL =	6.0 ml

Stored in 60µl aliquots at -70C.

The transription reaction was incubated at 40C for 60min with 1µl 8mM rGTP (in 20mM HEPES KOH pH7.6) being added after 30min. The RNA synthesised was then stored at -70C for periods of up to 6 months.

To estimate the amount of RNA synthesised in a reaction, 10µCi of ^{3}H-UTP was first dried down in an eppendorf using a Speedivac. apparatus. The reaction mix was then assembled in this tube. After incubation of the transcription reaction, 2 x 1µl aliquots were spotted onto Whatman DE81 paper and allowed to dry. The incorporated counts were estimated on one filter after five, two minute washes in 200ml of 0.15M Na_2HPO_4, followed by two washes in H_2O and two in methanol. The dried, washed filter was then compared to the untreated filter in a scintillation counter using aqueous/non-aqueous scintillant. The ratio of cpm from the washed:unwashed filters (% incorporation) was multiplied by 14.4 (µg of NTPs added in the PREMIX) to give the total amount of RNA synthesised.

To visualise the transcript, 0.25µCi ^{32}P-UTP was included in the transcription reaction. After the 1 hour incubation, 5ul was run on a formaldehyde/agarose gel as described in Section 2.1.6.

2.2.2 WHEAT GERM LYSATE TRANSLATIONS.

Transcripts synthesised *in vitro* were translated in wheat germ lysate using the method of Anderson *et al* (1983). A frozen stock of lysate was first prepared from 12g of commercially available wheat germ exactly as described by Anderson *et al*, and stored in 1ml aliquots under liquid nitrogen. A typical translation reaction was set up on ice by mixing the components listed below in the order shown;

Wheat germ translation reaction	
wheat germ PREMIX	2.35 µl
H_2O	4.40 µl
RNA# (0.5-1.0µg)	1.00 µl
^{35}S-Methionine (15 µCi)	1.00 µl
wheat germ lysate*	3.75 µl
TOTAL =	12.0 µl

= transcription mix containing between 0.5-1.0µg of mRNA

* = the defrosted lysate was centrifuged for 4min before use.

wheat germ PREMIX	
1M HEPES KOH pH7.6	350 µl
0.1M ATP	250 µl
0.4M creatine phosphate (CP)	500 µl
10 mg/ml CP kinase	100 µl
1.5 mg/ml spermine pH7.0	500 µl
0.1M DTT	425 µl
2mM GTP, Mg^{++} #	250 µl
5mM each amino acid mix	125 µl
1M potassium acetate	2100 µl
H_2O	100 µl
TOTAL	4700 µl

= made as a 40mM GTP/magnesium acetate stock, adjusted to pH7

The PREMIX was aliquoted and stored at -70C.

The translation reaction was incubated at 28C for 60min. In order to give an indication of the efficiency of translation, 1 µl of the translation mix was spotted onto Whatman No.1. filter paper and left to dry. The filter was then washed twice for 5min in 100ml ice cold 10% TCA, boiled in 100ml 5% TCA for 15min, rinsed in H_2O, and finally washed in absolute ethanol before being left to dry. The dried filter was counted in a scintillation counter using aqueous/non-aqueous scintillant and the cpm compared to a filter spotted with 1 ul

of a control translation (no exogenous RNA) and treated in the same way.

Translation products were resolved by SDS-PAGE and visualised by fluorography (see Section 2.3.1).

2.2.3 RABBIT RETICULOCTYE LYSATE TRANSLATIONS

Rabbit reticulocyte lysate was obtained from Promega and used exactly as directed. Typically, 15μCi ^{35}S-Methionine, 20ul lysate and 1μl of an *in vitro* synthesised transcript was added to a translation reaction. The transcript was generally added directly from the transcription reaction and not from precipitated nucleic acid. Incubation was for 60min at 30C. The efficiency of translation was determined by measuring TCA-precipitable counts as described for wheat germ lysate translations. Translation products were resolved by SDS-PAGE and visualised by fluorography.

2.2.4 *XENOPUS* EGG CELL-FREE TRANSLATIONS

An egg cell-free extract from *Xenopus* was available in the laboratory, prepared using a method obtained from the laboratory of Professor A.Colman (University of Birmingham), and stored under liquid nitrogen. A translation reaction was set up as follows:

egg extract	100.0μl
125mM spermidine	1.0μl
350mM creatine phosphate	2.5μl
reticulocyte lysate S-100	10.0μl
^{35}S-Methionine	100.0μCi

This was split into 30μl aliquots and 1μg mRNA added to each aliquot. Incubations were at 21C for 2 hours.

Since the system described was not dependent on exogenous RNA the proricin synthesised in the lysate was first immunoprecipitated before SDS-PAGE analysis. The translation

was diluted to 200µl with Immunoprecipitation buffer (100mM Tris pH7.8, 1% w/v Triton X-100, 1% w/v deoxycholate, 0.5% w/v SDS, 5mM $MgCl_2$, 100mM KCl, 1mM methionine, 1mM PMSF: pH8.2), 4µg of rabbit anti-B chain antisera was then added plus 50µl of a 1:1 slurry of protein A Sepharose in immunoprecipitation buffer and the mix was incubated at room temperature on a clinostat for 2 hours. The beads were pelleted and washed three times in immunoprecipitation buffer. The immunoprecipitated protein was finally released from the beads by boiling them in 1xProtein Sample buffer (Section 2.3) immediately prior to SDS-PAGE analysis.

The B chain domain of proricin synthesised in this way was also assessed for lectin (sugar-binding) activity by diluting the translation in Oocyte Homogenisation (OH) buffer (20mM Tris pH7.6, 1% Triton X-100, 0.1M NaCl, 1mM PMSF) and then passing it down a SeLectin^{TM} column. The matrix of the column consists of lactose immobilised onto polyacrylamide beads. Protein bound to the immobilised lactose could be eluted in 100mM galactose/OH buffer. Bound and un-bound fractions were immunoprecitated as above, prior to SDS-PAGE analysis.

Immunoprecipitated proricin made in this translation system was also incubated with endo-N-acetylglucosaminidase H (endoH) in order to determine the extent of N-linked glycosylation. The protein was first removed from the beads by boiling for 5min in 100mM TrisHCl pH8/1%(w/v)SDS/1%(v/v)B-mercaptoethanol. The supernatant from this was diluted in 9 volumes 0.1M sodium citrate buffer pH5.5, 5 milliunits of endoH was added and the reaction incubated at 37C overnight. The protein in this solution was then TCA precipitated (see next section) prior to analysis by SDS-PAGE.

2.3 ANALYTICAL PROTEIN TECHNIQUES.

2.3.1 SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS.

Soluble proteins were resolved by electrophoresing aliquots, under denaturing conditions, through polyacrylamide using the standard technique of SDS-polyacrylamide gel electrophoresis (SDS-PAGE) described by Laemmli (1970). Proteins were first boiled in 1xProtein Sample buffer (see below) for 2min immediately before loading into the wells of a gel.

2xProtein Sample Buffer (REDUCING)	
0.5M Tris-HCl pH6.8	2.5 ml
SDS	0.4 g
Glycerol	4.0 ml
B-mercaptoethanol #	1.0 ml
Bromophenol blue	0.02 g
H_2O to	10 ml

ommited when making 2xNON-REDUCING sample buffer.

2.3.2 SILVER STAINING OF SDS-POLYACRYLAMIDE GELS.

Gels were silver stained according to the method of Merril et al (1981).

2.3.3 COOMASSIE BLUE STAINING OF SDS-POLYACYLAMIDE GELS.

After electrophoresis, gels were immersed in Staining solution (0.25% w/v Coomassie brilliant blue R, 50% methanol, 7% acetic acid) for 1 hour. Stained protein bands were then visualised by repeated washes in Destain solution (40% methanol, 10% acetic acid) to reduce the background.

2.3.4 WESTERN BLOTTING OF SDS-PAGE RESOLVED PROTEINS.

Proteins resolved on a gel were transferred to a nitrocellulose filter (Hybond-N), in transfer buffer (20mM

Tris-HCl, 200mM Glycine, 20% methanol: pH8.3) using a Biorad Transblot apparatus, for 2 hours at 60V. Pre-stained protein markers were run on the gel to give an indication of the efficiency of transfer. After blotting, the gel was silver stained as described in Section 2.3.2. The nitrocellulose filter was washed for 30min in 3x 100ml Blocking buffer (PBS, 0.1% Tween 20, 5% w/v Marvel) in order to block non-specific binding of the antibody to the filter, before incubating it overnight at 4C, in fresh Blocking buffer containing antisera prepared from animals immunised against ricin. Laboratory stocks of antisera used in this study to probe for proricin included anti-ricin A chain and anti-ricin B chain antibodies raised in rabbits. A 95% pure IgG fraction of antisera taken from sheep immunised against recombinant ricin A chain, a gift from J.Chaddock and G.Legname (Warwick), was also used.

Antibody - antigen signals were amplified using the Amersham biotinylated-proteinA/streptavidin-horseradish peroxidase system (RPN. 1231), when using antisera raised in rabbit. Dilutions of 3.3 in 1000 were used with both the biotinylated-protein A conjugate (diluted in PBS, 0.1% Tween 20) and streptavidin-horseradish peroxidase conjugate (diluted in PBS/0.1% Tween 20). The signal was visualised by washing the filter in PBS and developing for 10-15min in Developing solution (30 mg 4-chloro-1-naphthol dissolved in 10ml methanol, and then mixed with 50ml PBS and 60 ul hydrogen peroxide). After development, the filter was finally rinsed in H_2O and left to dry.

For filters probed with sheep IgG anti-ricin antibodies, the signal was amplified by probing with a second antibody/enzyme conjugate (anti-sheep IgG/alkaline phosphatase, Sigma A7789) and the filter developed using the Promega alkaline phosphatase developing kit (Promega AP system W3930).

2.3.5 STORAGE OF POLYACYLAMIDE GELS.

Polyacrylamide gels, once fixed, were dried down onto 3MM paper at 80C under vacuum, using a Biorad heated gel drier apparatus. Once dried, they were stored flat at room temperature.

2.3.6 FLUOROGRAPHY AND AUTORADIOGRAPHY OF RADIO-LABELLED PROTEINS AFTER SDS-PAGE.

After electrophoresis, resolved protein bands were first fixed by washing the gel for 30min in Destain solution, before immersion of the gel in Amplify (Amersham) for 10-15min. The gel was then dried down at 60C on a vacuum gel drier and exposed to Fuji RX X-Ray film in a light-proof cassette, at -80C. Film was developed in Kodak LX24 and fixed in Kodax Unifix.

2.3.7 TRICHLOROACETIC ACID (TCA) PRECIPITATION OF PROTEIN SAMPLES.

Where it was necessary to reduce protein sample volumes before loading onto polyacrylamide gels, the sample was made 10% v/v TCA and the protein precipitated on ice for 30min. BSA was used as a carrier (to a final concentration of 100µg/ml) when the concentration of protein was known to be very low. The protein was pelleted by centrifugation in a microfuge for 10min. The pellet was washed in 100% acetone three times to remove residual TCA, and resuspended in 1xProtein sample buffer. The sample was then boiled and loaded as above.

2.4 TRANSIENT EXPRESSION IN TOBACCO LEAF PROTOPLASTS.

2.4.1 INTRODUCTION.

The method of protoplast transformation described was based on the PEG-mediated method of protoplast transformation as described by Saul *et al* (1988). All manipulations were performed under sterile conditions in a flow hood. All solutions were autoclaved or filter sterilised and all utensils were flamed with ethanol immediately before use.

2.4.2 GROWTH AND MAINTENANCE OF PLANTS.

(i) Sterilisation of seeds.

Seeds of *Nicotiana petit Havana* SR1 were washed for 1min in 75% ethanol. They were then washed in 1% active hypochlorite for 2.5min. The residual hypochlorite was removed by washing the seeds twice in sterile distilled water, before soaking them in 0.5mg/ml GA3. The sterilised, prepared seeds were germinated on 0.5xMS medium (Flow labs) containing 5g/l sucrose, 0.8% agar.

(ii) Maintenance of plants.

Seedlings germinated from sterilised seeds were transferred to 0.5xMS medium, containing 30g/l sucrose and 0.8% agar. Subsequently, cuttings were taken from healthy plants and re-potted in the above medium. All seedlings and plants were grown at 26C and in 800 LUX of artificial, fluorescent light provided by a GROWLUX element.

2.4.3 PREPARATION OF DNA.

Caesium chloride gradient-purified plasmid DNA was ethanol precipitated and the pellet washed three times in 70% ethanol under the sterile hood. The eppendorf was sealed during each wash and the tube gently inverted, to sterilise the inner surface. The excess ethanol from the final wash was drawn off

and the pellet allowed to dry for 2 hours in the sterile hood. The dried DNA was resuspended to 3mg/ml in H_2O.

2.4.4 ISOLATION OF PROTOPLASTS

(i) Enzyme digestion.

Twenty healthy leaves were taken and their midribs excised using a sharp scalpel, after briefly soaking the leaves in enzyme solution (1.2% w/v cellulase, 0.4% w/v maceroenzyme-R10 in K3 medium: filter sterilised, stored at -20C in dark). The leaves were then transferred to a petri dish containing 25mls enzyme solution, and were cut into 0.5cm^2 squares. The petri dish was evacuated for 3-4min in a bell jar connected to a vacuum line, in order to remove the air from within the leaf structure, before returning the petri dish to atmospheric pressure. The petri dish was finally sealed with parafilm and incubated in the dark at 25C overnight.

(ii) Separation of protoplasts from undigested leaf matter.

The petri dish was agitated by hand to release the protoplasts from the leaf skeletal structure, and then allowed to stand for a further 30min at room temperature. The digested solution was gently taken up using a wide-bore pipette and filtered through a 125micron stainless steel filter. The petri dish and filter were rinsed through with 0.5 volumes of Sucrose buffer solution (0.6M sucrose, 0.5% w/v MES: pH5.6 with KOH, autoclaved). The filtered protoplast suspension was then transfered to 15ml sterile plastic tubes and 1ml W5 (154mM NaCl, 125mM $CaCl_2$, 5mM KCl, 5mM glucose: pH 6.0 with KOH, autoclaved) was layered on top of each aliquot. The tubes were centrifuged at 600rpm for 10min at room temperature in a clinical centrifuge. The intact protoplasts collected at the sucrose/W5 interface and were transfered to fresh tubes using a wide-bore pipette. These protoplasts were washed twice in 10ml W5 (in order to dilute out any remaining digestive enzymes) by

resuspending the cells in W5 and pelleting at 600rpm for 5min. The final protoplast pellet was resuspended in 2-4ml W5 and stored at 4C for 30min-3 hours, before transformation.

2.4.5 ESTIMATION OF VIABLE PROTOPLASTS. (Larkin, 1976)

Two 50ul aliquots of the washed protoplast suspension were mixed with an equal volume of diluted fluoroscein diacetate (FDA was stored as a 5mg/ml stock solution in acetone: it was diluted 4:1000 with K3 before use). The viable protoplasts were estimated using a haemocytometer slide mounted on a fluorescent microscope (viable cells fluoresce green). An average value from the two aliquots counted in this way was taken, and the protoplast concentration of the original suspension calculated.

2.4.6 PROTOPLAST TRANSFORMATION.

The protoplasts were pelleted as above, and resuspended in Mannitol/Magnesium (0.5M mannitol, 15mM $MgCl_2$, 0.1% w/v MES: pH 5.6 with KOH, autoclaved and stored at 4C) to a concentration of $1.6x10^6$ viable cells/ml. They were immediately heat-shocked at 45C for 5min, before being left on the bench to cool to room temperature.

Meanwhile, the DNA was aliquoted into a sterile 15ml plastic tube. A typical transformation reaction contained, 30µg sheared herring sperm DNA (used as a carrier and at 1mg/ml), 9µg GUS-encoded plasmid DNA (used as a cytoplasmic control), and 21µg plasmid DNA of interest. To this tube was added 300µl of cooled, competent protoplast suspension, and the tube agitated gently to mix the DNA and protoplast solutions. To mediate the transformation, 300µl of PEG solution (40% w/v PEG 4000, 0.4M mannitol, 0.1M $Ca(NO_3)_2$: pH 8.0 with KOH, autoclaved and stored at 4C) was added dropwise, the tube was again agitated, and then allowed to stand for 30min at room

temperature.

The PEG was diluted tenfold by the successive additions of 1,2, and 7ml of W5 at 3min intervals, mixing each time. The protoplasts were then pelleted at 500rpm for 5min, and resuspended in 0.5ml K3 (3.78 g/l B5 salts, 5mM $CaCl_2$, 12mM NH_4NO_3, 0.4M sucrose, 2mM xylose, 1 mg/l 6-BAP, 1mg/l NAA: pH5.6 with KOH, filter sterilised and stored in the dark at -20C). The surviving protoplasts floated on this medium. The tubes were incubated in the dark at 25C for 36 hours.

2.4.7 ISOLATION AND PREPARATION OF PROTOPLAST AND MEDIA FRACTIONS.

The methods below were based on the methods described by Denecke et al (1989 & 1990). It was not necessary to carry out further manipulations under sterile conditions.

When the tubes were examined after 36 hours' incubation, three distinct phases were visible: an upper protoplast phase, a media phase, with any lysed, non-viable protoplasts forming the lowest phase which lined the bottom of the tubes. If perturbation of the tubes occurred during their removal from the incubator, the suspension was clarified by centrifugation at 500rpm for 10min as above. 300µl of the media phase was transferred to an eppendorf and processed, as described in (ii) below. All solutions and fractions from now on were kept on ice to minimise proteolysis.

(i) Protoplast processing.

The protoplasts were washed in 10ml W5, to dilute any media present, and were pelleted at 500rpm for 5min as above. The W5 was drawn off, the protoplasts resuspended in 400µl (50mM Tris-Cl pH7.5, 1mM PMSF, 2mM DTT, 0.3 mg/ml BSA) and the suspension sonnicated for 10 seconds at an amplitude of 6microns. The insoluble proteins were pelleted at 14,000rpm for 10min at 4C. In early experiments, the soluble protein extract

was used directly in the enzyme assays. However, in later experiments the method was modified and the following steps were included in the processing of the protoplast fractions.

The soluble protein extract was made 60% saturated with respect to ammonium sulphate, and the proteins precipitated on ice for 60min. The precipitated protein was pelleted in a microfuge for 10min at 4C. This protein pellet was resuspended in Enzyme buffer (50mM Tris-HCl pH7.0, 2mM DTT) and stored at 4C overnight, or frozen at -70C over the weekend.

(ii) Media processing.

In early experiments the media fraction was used directly in enzyme assays. However, it became necessary to include the following steps prior to enzyme analysis, in order to obtain reliable results.

Soluble protein, extracted from untransformed tobacco leaf tissue and dissolved in extraction buffer, was added to the media fraction to a final concentration of 100µg/ml. The solution was made 60% saturation with respect to ammonium sulphate and the proteins precipitated on ice for 60min. The precipitated was pelleted for 10min in a microfuge at 4C, and resuspended in 1ml Enzyme buffer. This solution was dialysed by three repeated concentration/dilution steps in an Amicon 'Centricon 10' filter apparatus. This dialysed solution was stored at 4C overnight, or at -70C over the weekend.

2.4.8 ENZYMATIC ASSAYING OF PROTOPLAST AND MEDIA FRACTIONS.

(i) GUS assay (Jefferson, 1987).

An aliquot of each fraction was diluted to 360ul in Enzyme buffer, made 0.1% w/v Triton X-100, and equilibrated to 37C. To this was added 40ul 10mM methylumbelliferryl glucuronide (10mM MUG was made fresh in Enzyme buffer, 0.1% w/v Triton X-100) and the reaction allowed to proceed at 37C. Time points were taken at suitable times between 0 and 3 hours, viz;

100µl of reaction mix was added to 900µl of 0.2M $NaHCO_3$ (this stopped the enzyme reaction). The product, MU, released at a given time, was estimated by measuring the fluorescent emission of a time point at 455nm during excitation at 365nm, in a Jasco FP-550 spectrofluorometer. These values were then plotted on a graph of relative fluorescence vs time, to show that the fluorescence observed increased linearly with time. These readings were then converted to pmoles product released by reference to a standard curve of relative fluorescence vs [MU]. As all readings recorded were 'relative' fluorescence, a fresh standard curve was prepared for each assay.

(ii) nptII assay (M^{C}Donnell *et al*, 1987)

The following solutions were prepared;

5xReaction mix
335mM Tris-Cl
210mM $MgCl_2$
2M NH_4Cl

Titrated to pH7.1 with maleic acid

1xAssay mixture(made fresh)

ATP (10mM stock)	5 µl
^{32}P-ATP (10mCi/ml stock)	1.5 µl
neomycin (22mM stock)	7 µl
NaF (1M stock)	50 µl
1x reaction buffer	4.936 ml
TOTAL	5 ml

To start the reactions, 30µl assay mixture and 30µl of each fraction were mixed and incubated at 37C for 30min. The insoluble proteins were pelleted in a microfuge for 5min at 4C. Duplicate aliquots of 20µl supernatant were spotted onto Whatman P81 paper, which had been pre-soaked in 20mM ATP/100mM Na_2PO_4 and dried in an oven. The filter was washed for 2min in

500ml 10mM phosphate buffer pH7.5 at 80C, washed three times in 500ml 10mM phosphate buffer pH7.5 at room temperature and was then left to dry. Finally, half of the filter was exposed to X-ray film at -70C for 1-7 days. From the second half of the filter, the duplicate spots were suspended in aqueous/non-aqueous scintillant and counted.

2.5 RNA N-GLYCOSIDASE ACTIVITY OF PRORICIN

In this study, proricin was synthesised by injection of preproricin mRNA into Xenopus oocytes. These injections were performed by Prof. A.Colman (University of Birmingham). The proricin was extracted from the oocytes and analysed by SDS-PAGE by Dr. P.Richardson (University of Warwick). The details of these procedures have been described elsewhere (Richardson et al, 1989). All analyses into the activity of proricin species isolated from these extracts however, were performed solely by this author.

The assay used to determine the RNA-dependent N-glycosidase activity of ricin A chain and described below is based on the method of Endo et al (1987).

2.5.1 PREPARATION OF ANILINE REAGENT.

Aniline was distilled twice and stored at -20C in the dark until required. 1ml of this redistilled aniline was mixed with 7ml H_2O and 0.5ml glacial acetic acid in a universal that had been washed in chromic acid to inactivate any RNase activity. A glass pH probe was then used and the solution brought to pH4.5 with acetic acid. The aniline reagent was then made to a final volume of 11.0ml and could be stored at 4C in the dark for periods of up to two months.

2.5.2 INCUBATION OF RABBIT RETICULOCYTE LYSATE RIBOSOMES WITH RIP.

An equal volume of Promega rabbit reticulocyte lysate (non-nucleased batch) and preproricin sample (typically a translation mix of 30µl) were mixed and incubated at 37C for 30min. The ribosomal proteins were then denatured by the addition of SDS to a final concentration of 1% w/v, and the ribosomal RNA (rRNA) extracted in the aqueous phase by two phenol/chloroform extractions. The RNA was ethanol

precipitated, the pellet washed once in 70% ethanol/50mM NaCl, once in 70% ethanol, dried *in vacuo*, and resuspended in 10µl H_2O. 1µl was diluted to 500ul in H_2O and the absorbance measured at 260 and 280 nm. A protein-free rRNA sample gave an A_{260}:A_{280} ratio of 1.8. The concentration of RNA was calculated from the A_{260} as follows:

$$[\text{RNA}]\ (\text{ug/ml}) = \frac{A_{260} \times \text{extinction coefficient (40)} \times 500}{1000}$$

2.5.3 ANILINE TREATMENT OF MODIFIED rRNA.

2-5ug of extracted RNA was added to 20ul aniline reagent and the solution incubated at 60C for 2min in the dark. The solution was made 2.5M with respect to ammonium acetate and 2.5 volumes -20C absolute ethanol added; the solution mixed and placed on dry ice for 30min in the dark. The RNA was pelleted for 10min in a microfuge at 4C, washed first in 70% ethanol/50mM NaCl followed by 70% ethanol, and then dried *in vacuo*. The dried RNA was resuspended in 20ul RNA sample buffer (60% v/v formamide, 0.1%TPE: made fresh) and incubated at 65C for 5min. 2µl of RNA loading buffer was added and the sample loaded onto a 50% formamide, 1.5% agarose gel.

2.5.4 50% FORMAMIDE/1.2% AGAROSE GELS.

Commercial formamide solution was deionised by stirring it with BDH Amberlight Monobed Resin MB-3 until the pH had fallen to pH7. This was then filtered through Whatman No. 1 paper and stored in the dark at 4C for periods up to 2 months. The gel was made by first boiling 1.2g agarose in 50ml 0.1xTPE (made as a 1x stock, *viz*: 36mM Tris-HCl, 30mM NaH_2PO_4, 2mM EDTA). Once dissolved, the agarose was cooled to 60C and 50ml deionised formamide was added. The gel was then poured and allowed to set at 4C.

The gel was run in 0.1xTPE (not submerged) at a constant current of 20mA for 2-3 hours. To visualise resolved RNA bands, the gel was stained by immersion in a solution of 2µg/ml ethidium bromide before placing it onto a UV transilluminator.

2.6 ISOLATION OF PROTEIN BODY ENDOPROTEASE ACTIVITY AND ITS ASSAY

All work with castor bean seed tissue was performed in a toxin-designated area and all utensils, plant material and samples were detoxified by soaking for 24 hours in 0.1M NaOH before discarding.

2.6.1 GROWTH OF PLANTS

Plants were grown from seed in a greenhouse, under tungsten lighting, with a 12 hour photoperiod. Flowering took place 6-8 months after germination, and was induced by keeping the plants pot-bound and starving them of water.

2.6.2 PREPARATION OF EXTRACTS CONTAINING PRORICIN-PROCESSING ACTIVITY FROM DEVELOPING SEEDS.

Enzyme activity was recovered from developing seeds following the method of Harley and Lord (1985). 10g of dehulled castor bean seeds harvested from the fruiting bodies of greenhouse-grown plants at a post-testa stage of development, were homogenised in 10mls 100mM sodium phosphate buffer pH7.0 using a mortar and pestle. The homogenate was passed through four layers of muslin, centrifuged at 16,000rpm, and the supernatant dialysed against 500 volumes of 0.5x citrate phosphate buffer (15.4mM sodium citrate, 19.3mM Na_2HPO_4: pH4.0: from Gomori, 1955) for 4 hours at 4C. It was stored frozen at -20C.

2.6.3 PREPARATION OF EXTRACTS CONTAINING PRORICIN-PROCESSING ACTIVITY FROM THE PROTEIN BODIES OF DRY SEEDS.

The preparation of protein bodies from dry seed tissue (Mettler and Beevers, 1979) was performed, at room temperature, by homogenisation of dehulled castor seeds in 100% glycerol (1ml per g tissue) in an atomix, straining of the homogenate

through four layers of muslin, and pelleting of the protein bodies at 8,500rpm for 15min. The protein bodies were resuspended in glycerol and repelleted as above.

The protein bodies were lysed by resuspending the pellet in ice-cold 25mM sodium phosphate buffer pH7.0. All further manipulations were performed at 4C. The crystalloid and membrane fractions were pelleted by centrifugation for 20min at 11,500rpm, before dialysing the supernatant for 4 hours in 3x 5 litres of 0.5x citrate phosphate buffer pH4. Protein precipitated during dialysis was pelleted at 10,000rpm for 10min. The soluble protein was drawn off and stored at -20C.

2.6.4 REMOVAL OF LECTINS FROM ENZYME SAMPLES USING AFFINITY CHROMATOGRAPHY ON SEPHAROSE 6B.

Sepharose 6B (Pharmacia) that had been pre-washed in 1M proprionic acid was available in the laboratory. The acid treatment exposes sugar residues on the bead matrix to which ricin and RCAI will bind. 0.5ml of crude enzyme sample was first spun in a Beckman TL100 at 50,000rpm for 30min to pellet any aggregates. The supernatant was then passed down a 1ml column of acid-washed Sepharose, equilibrated with 0.5x citrate phosphate buffer, and the unbound proteins eluted. The lectins could be released from the column with 100mM lactose.

2.6.5 AMMONIUM SULPHATE PRECIPITATION OF EXTRACTS FROM DRIED SEEDS.

Soluble protein fractions from dried seeds were brought to 80% saturation with ammonium sulphate on ice and the protein allowed to precipitate on ice for 60min. The precipitated protein was pelleted at 15,000rpm for 10min and resuspended in 0.5ml 0.5x citrate phosphate buffer pH4. The ammonium sulphate was dialysed overnight against 10,000 volumes of buffer.

2.6.6 PREPARATION OF RADIOLABELLED PRORICIN FROM DEVELOPING SEED TISSUE.

Proricin was radiolabelled and extracted from developing seed tissue according to the method of Butterworth and Lord (1983). Developing seeds, harvested at a post-testa stage (stage E : see Roberts and Lord, 1981a), were dehulled, divided into endosperm halves and the cotelydonary tissue removed with a scalpel. 20 endosperm halves were placed on moistened filter paper, layered with 340µCi ^{35}S-Methionine and incubated at 25C for 90min. The label was then chased with 400µl 0.25M unlabelled methionine for 60min at 25C.

The tissue was then homogenised to a fine slurry with a razor in 1.5ml ice-cold grinding buffer (150mM Tricine pH7.5, 1mM EDTA, 10mM KCl, 1mM $MgCl_2$, 100mM lactose, 12% w/w sucrose). The homogenised tissue was strained through four layers of muslin before loading onto a stepped 20% w/w / 30% w/w sucrose gradient (the sucrose was dissolved in 1mM EDTA pH7.5) and centrifuging in a Beckman SW28 rotor for 2.5 hours, at 23,000rpm. The rotor was brought down without the use of the brake and the microsomal fraction, visible as a milky band at the 20% / 30% interface, removed using a syringe needle.

The radiolabelled proricin was immunoprecipitated from the microsomal fraction using rabbit anti-ricin A chain antibodies available in the laboratory by following the method of Roberts and Lord (1981b). The microsomal membranes were dispersed by mixing in equal volume of SPA1 buffer (1% v/v NP40, 10mM Tris-HCl pH7.5, 150mM NaCl, 2mM EDTA, 100mM lactose, 40µg/ml PMSF), and the proricin bound to antibody by incubation with 4µl antisera at room temperature for 60min. Protein A-Sepharose was added (50µl of a 1:1 slurry in SPA1) and the mixture rotated on a clinostat at room temperature for 30min. The protein A-Sepharose - antibody - proricin complex was pelleted in a microfuge for 30 seconds, the supernatant

discarded, and the pelleted washed three times in SPA2 (as SPA1 but with 0.2% NP40), twice in SPA3 (as SPA2 but with 0.5M NaCl), and once in 10mM Tris-HCl pH7.5. The bound proricin was then used as a substrate in an assay (see 2.4.6 below). The proricin / processed products were released from the beads by boiling for 5min in 1xProtein sample buffer, pelleting the beads and loading the supernatant on a SDS-polyacrylamide gel.

2.6.7 SYNTHESIS OF RADIOLABELLED PRORICIN FROM cDNA.

Proricin cDNA was available in the laboratory cloned into an SP6 expression vector. The clone used for *in vitro* expression of proricin was pSP64Tproricin (May, 1988). This construct contains a 1.94 kb fragment of proricin cDNA cloned into the *Bgl II* site of pSP64T (Krieg and Melton, 1984). The DNA encoded for a proricin molecule which resembled the primary translational product of ricin, preproricin, but was lacking the first 11 amino acid residues. Translation was therefore initialised at the ATG at position -72 of the published sequence (Lamb *et al*, 1985). The DNA was transcribed and translated in wheat germ lysate as described in section 2.3. Standardly, 1-2 µl of translation products were used in an assay as substrate.

2.6.8 ASSAY FOR PRORICIN-PROCESSING ACTIVITY.

Proricin substrate was incubated with enzyme fractions at 25C overnight. The products were resolved by SDS-PAGE and visualised by fluorography.

2.7 EXPRESSION OF PRORICIN IN E.coli.

2.7.1 INTRODUCTION

All expression work of proricin in E.coli was performed under HSE Category IV conditions.

The vector, pINIIIompA2proricin, a gift from Dr. Khalid Hussain (Warwick), contains proricin cDNA fused, distal to and in frame with, DNA encoding the 22 amino acid ompA signal peptide, in a pINIII vector (see Appendix for map). This was transformed into E.coli strain JA221 (see Table 2.1.1).

2.7.2 MAINTENANCE AND EXPRESSION OF JA221:pINIIIompA2proricin

Colonies from this transformation were maintained on JA221-minimal, 1.5% agar plates containing 100µg/ml ampicillin (JA221-minimal media was made by supplementing glucose/minimal media with 2µg/ml thiamine, 20ug/ml tryptophan, and 20µg/ml leucine). Overnight cultures were set up by inoculating 10ml JA221-minimal media with a single colony from a plate and incubating at 37C and 200rpm. Expression of the ompA:proricin chimaera was achieved by inoculating 100ml of expression media (JA221-minimal media made 1% w/v Caseamino acids) with 0.05 volumes of an overnight culture, incubating for 2-3 hours (exponential growth) at 37C and 200rpm, and then inducing the expression with 1mM IPTG. The cultures were grown for a further 3 hours before preparing total cell, periplasm, and shocked cell fractions.

2.7.3 PREPARATION AND ANALYSIS OF FRACTIONS PREPARED FROM EXPRESSION CULTURES OF TRANSFORMED JA221.

To prepare total cell fractions, cells from 1.5ml of culture were pelleted, resuspended to a concentration of 20 OD units/ml in ice-cold Sonication buffer (100mM Tris pH8.5, 5mM EDTA), and sonicated at 4microns for 2-5min. The insoluble

material was pelleted in a microfuge for 5min, the supernatant was removed and stored at 4C.

To prepare fractions of periplasmic protein, osmotic shock was used. Cells from 1.5ml of culture were pelleted, resuspended in 500µl 10mM Tris-HCl pH7.1/5mM NaCl and incubated on ice for 5min. The cells were then pelleted and resuspended in 500µl 33mM Tris-HCl pH7.1/0.4mM EDTA/20% w/v sucrose and incubated at room temperature for 10min. The cells were again pelleted and resuspended rapidly and violently in 500µl ice-cold 0.1mM $MgCl_2$. After a 10min incubation on ice, the shocked cells were pelleted for 5min and the supernatant, containing periplasmic proteins, was removed and stored at 4C. The shocked cells were then resuspended to 20 OD units/ml in Sonication buffer, sonicated and pelleted as for total cell fractions, and the supernatant stored at 4C. All fractions were analysed by SDS-PAGE, using 12% acrylamide, and the resolved proteins Western blotted using anti-ricin antisera.

2.7.4 BETA-LACTAMASE ASSAY.

In order to monitor the efficiency of recovery of periplasmic proteins, the total cellular/periplasmic/shocked cell fractions were assayed for the activity of the periplasmically-localised enzyme, beta-lactamase, using the method described by O'Callaghan *et al* (1972). 20µl of a freshly prepared solution of nitrocephin (4mg/ml in DMSO) was mixed with 780µl of each fraction (diluted in 0.1M phosphate buffer pH7.0) and the reaction followed spectrophotometrically by measuring the absorbance of the product at 500nm and at an incubation temperature of 30C, over 1min. After correcting for dilution factors, the relative B-lactamase activities of the three fractions could be compared.

2.8 FACTOR Xa AND THROMBIN ASSAYS

Proteins translated *in vitro* were digested with 0.1 volumes Factor Xa or Thrombin in one of the following buffers:

1xTEN (P.Esnouf, Oxford)= 10mM Tris pH7.4, 1mM EDTA, 140mM NaCl (also 1.25mM $CaCl_2$ for thrombin digestions)

1x XaBuffer (P.Esnouf, Oxford)= 100mM Tris pH7.6, 0.05% Triton-X-100, 10mM $CaCl_2$, 0.05% deoxycholate

The Factor Xa enzyme was a gift from P.Esnouf, Radcliffe Hospital, Oxford. Thrombin was commercially available from Sigma. Reactions were carried out at 26C for 1 hour (unless otherwise stated).

CHAPTER 3

CONSTRUCTION OF PREPRORICIN/NEOMYCIN PHOSPHOTRANSFERASE CHIMERIC GENES IN A VECTOR SUITABLE FOR THEIR EXPRESSION IN PLANTS.

CONTENTS

3.1 INTRODUCTION.

As described in Chapter 1, ricin is initially synthesised as a preproprotein precursor with an N-terminal signal peptide that directs the protein into the lumen of the ER. However, as was also described, all proteins segregated into the lumen of the ER require some sort of positive sorting information to retain them within the endomembrane system and prevent secretion via the default pathway (Denecke *et al*, 1990). For instance, resident soluble ER-lumenal proteins are selectively retained from the bulk flow by possession of a carboxy-terminal KDEL sequence (Pelham, 1989). The presence of an oligosaccharide (mannose-6-phosphate) signal on lysosomal enzymes in animal cells which targets them to lysosomes has also been well documented (see Kornfield and Melham, 1989). It is also now known that the information responsible for the correct targeting of a yeast vacuolar enzyme, CPY, to the vacuole is held within a contiguous stretch of four amino acid residues, QRPL (Valls *et al*, 1990). It has been argued that during plant seed development the bulk of protein transport through the secretory pathway is to the protein bodies and therefore one might expect this route to be the default pathway, with positive sorting being required to target proteins to the cell surface (a proposal considered by Bowles and Pappin, 1988). However, when Dorel *et al* (1989) expressed an ER signal peptide:cytosolic pea albumin chimera in transgenic tobacco they could not find accumulation of the albumin in the protein bodies. This example appears to contradict the proposal that the default secretory pathway in developing seed tissue is to the protein bodies and supports the hypothesis that, as with all other eukaryotic tissue examined to date, the secretion pathway represents the default pathway. It is therefore proposed here that preproricin must also contain a signal that targets newly-synthesised ricin to

the protein bodies in the developing castor bean seed.

The fact that ricin is a glycoprotein has also been discussed. However, in contrast to the lysosomal targeting signal of animal proteins, the protein body targeting signal used by the ricin precursor is not an oligosaccharide. Non-glycosylated forms of ricin synthesised in the presence of tunicamycin, still accumulate within protein bodies (Lord, 1985b). It is therefore reasonable to postulate that the signal for targeting proricin to the protein bodies is held within the amino acid structure of the polypeptide.

Preproricin has 35 amino acid residues at its N-terminus that are not present in the mature protein (referred to in future as the 'N-terminal presequence'). It has been considered for some time that not all 35aa residues function as an ER signal peptide (Lynne Roberts, personal communication). If this is the case then there must be an intermediate form of ricin present in the ER lumen that has an N-terminal propeptide (described in future as the N-propeptide to distinguish it from the 12 amino acid peptide that links the A and B chains in the precursor forms of ricin). However, efforts in the past to define the exact site of signal peptidase cleavage in preproricin by using microsequencing of *in vitro* synthesised and segregated preproricin have proven unsuccessful. This was because, in the *in vitro* assays used, preproricin was inefficiently imported in microsomes and the quantities obtained were too low for purification and analysis (Lynne Roberts, personal communication).

Many other vacuolar/protein body storage proteins are synthesised as preproproteins - that is they contain regions, in addition to the ER-signal peptide, that are only present in intermediate forms of the protein. Furthermore, it has been shown in some cases that these propeptides are removed after the protein has reached its final organellar destination.

Therefore, it is conceivable that propeptides, which have no role to play in the final structure of the mature protein, may play a role during the transport of the protein to its organelle. Such functions might include;

(i) to maintain the protein in a transport-competent state. It has been shown that misfolded proteins are degraded or sequestered in the ER (Klausner et al, 1990).

(ii) to maintain the protein in an altered conformation from its mature form - for instance in the case of enzymes, to prevent them from acting on potential substrates with which it may be co-transported at some stage whilst en route to its final destination, by either sterically blocking the active site or maintaining the enzyme in an inactive configuration.

(iii) to target the protein to the correct organelle; i.e. the N-propeptide may contain a targeting signal which could be proteolytically removed from the mature protein after reaching its final destination.

This third role for an N-propeptide has been found in Saccharomyces cerevisae. The N-propeptide of yeast vacuolar enzymes proteinase A and CPY contain information which is necessary for correct targeting of the enzymes to the vacuole. In the case of CPY, the targeting signal has been identified as the tetrapeptide QRPL (Valls et al, 1990). Although this sequence does not appear in the N-terminal region of preproricin, it is possible that the targeting machinery has diverged during evolution such that a different signal exists in plant storage protein precursors responsible for the targeting of these proteins to the protein bodies.

3.2 AIMS.

The aims of the work discussed in the next two chapters were to attempt to more clearly define the ER-signal peptide

cleavage site and to see whether the N-terminus of preproricin contains a protein body targeting signal.

3.3 APPROACH.

The N-terminal region of preproricin was analysed to find likely sites of signal peptidase cleavage. Two constructs were then designed to create, in each case, a chimeric gene encoding a protein that would contain the N-terminal region of preproricin followed by the complete amino acid sequence of the marker enzyme, neomycin phosphotransferase (npt2). One construct, termed pPARTIAL, was designed to encode the first 23aa residues of preproricin fused in front of npt2. The other construct, called pFULL, was designed to encode a preproricin/npt2 hybrid which contained all 35aa residues of the preproricin N-terminal region fused in front of the npt2 enzyme. In both DNA constructs the gene was flanked on its 5' side by a constitutive plant promoter and on its 3' side by a plant transcription terminator.

The chimeric gene from the pPARTIAL construct was first expressed *in vitro* in order to determine whether or not the hybrid protein contained a functional ER signal peptide.

As will be discussed in the next chapter, the pPARTIAL and pFULL constructs were then used to transform tobacco protoplasts.

3.4 COMPUTER ANALYSIS OF THE N-TERMINAL REGION OF PREPRORICIN.

The algorithm for predicting the site of signal peptide cleavage described by von Heijne (von Heijne, 1986) was used to identify likely sites of signal peptide cleavage in preproricin. The algorithm assigns a value to amino acid residues greater than 12 residues away from the N-terminal methionine residue. A residue assigned a positive value is more likely to be the site of cleavage than a residue with a

negative value. Unfortunately the algorithm is complicated and fifteen individual probabilities must be summed before it is possible to assign a value for a particular residue in the sequence. Therefore, in order to enable many protein sequences to be screened accurately and quickly, a computer programme was designed that was based on von Heijne's algorithm (see Appendix).

Figure 3.4 shows a graphic representation of the Von Heijne algorithm when applied to the N-terminal region of preproricin. Five residues of the 35 had positive values. The algorithm identified the residues, glycine^{22} and threonine^{30}, as the most likely sites of signal peptide cleavage with values of +4.8 and +5.4 respectively. The algorithm also confirmed the suspicion that, with a value of -9.33, asparagine^{35} was highly unlikely to be the site of signal peptide cleavage.

This data was strong indirect evidence for the presence of an N-propeptide region in proricin.

3.5 CLONING STRATEGIES.

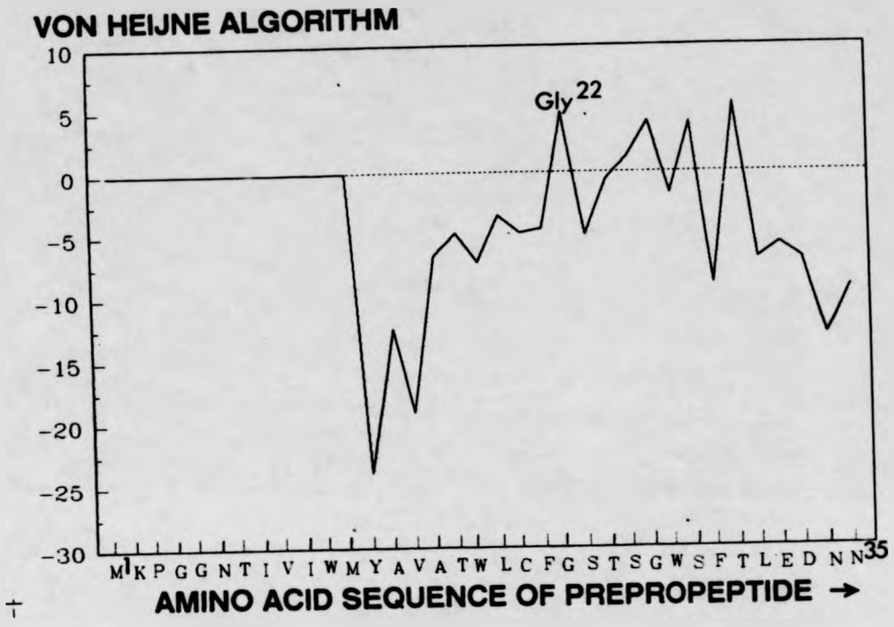
[N.B. all nucleotide reference numbers for preproricin refer to those printed in the published sequence of Lamb *et al* (1985), an amended copy of which can be found in the Appendix.]

3.5.1 SOURCE AND CHARACTERISATION OF VECTORS.

The construct ppRCL617, containing the entire preproricin coding sequence (Lamb *et al*, 1985 & Richardson *et al*, 1989: see Appendix for map) was available at Warwick. This was used as the source of the preproricin DNA. ppPCL617 has a unique *Xho*I restriction site immediately upstream to the 5' end of the preproricin gene. This preproricin cDNA also has three *Bam*HI sites at positions -41, 852 and 1551. The site at -41 lies within codon Gly^{22}. The computer analysis described in the section above identified this glycine residue as being a likely

FIGURE 3.4 VON HEIJNE PLOT OF N-TERMINAL REGION OF PREPRORICIN

Graphic representation of the von Heijne algorithm when applied to the N-terminal region of preproricin. This algorithm is used to predict the likely sites of signal peptidase cleavage in the N-terminal regions of proteins that are co-translationally translocated into the lumen of the ER. Sites of cleavage are predicted to lie on the carboxyl side of residues that have a high (positive) value. Von Heijne reported that this algorithm has 80% predictive accuracy (von Heijne, 1986). The algorithmic values for preproricin (shown opposite) were obtained by entering the first 37 amino acid residues into a computer programme designed by this author (see Appendix). Residues glycine^{22} and threonine^{30} have high values and can be considered as likely signal peptidase cleavage sites. This plot also shows that asparagine^{35} (the residue immediately before the first residue of the mature A chain) is highly unlikely to be the site of signal peptidase cleavage. This supports the hypothesis that the intracellularly transported form of ricin (proricin) contains an N-terminal propeptide.



site of signal peptide cleavage. It was therefore decided to utilise this BamHI site in order to create a construct encoding a fusion protein between an N-terminal fragment of preproricin and the marker enzyme npt2, using standard molecular biology procedures (described in Section 3.5.3). This fusion protein would then be used to check that the first 22 N-terminal residues of preproricin were sufficient to co-translationally transport a foreign protein across the ER membrane and subsequently, to determine whether the ER-segregated protein is secreted (by default) in tobacco protoplasts.

In order to determine the effects of fusing the complete N-terminal region of preproricin (i.e. the first 35aa residues) to the N-terminus of npt2 on its intracellular transport, it was necessary to mutate the preproricin DNA to introduce a unique restriction site at the junction of the N-terminal region and the mature A chain region. The cloning procedures followed to achieve this are described below in Section 3.5.4.

The plant system expression vector, pCaMVNEO (Fromm et al, 1986), was a gift from Dr. Aldo Ceriotti, Milan. It contains the 35S-cauliflower mosaic virus promoter (35S-CaMV), the coding region of the the neomycin phosphotransferase gene (npt2) from transposon Tn5 and the nopaline synthase transcription terminator sequence (nos). This vector was used as the expression vector for the plant transformation experiments described in the next chapter and as a source of the npt2 gene. However, in order to use this npt2 gene for the construction of ricin:npt2 chimeric genes it was necessary to mutate the translation initiating codon ATG, to prevent initiation of translation at this site in the fusion proteins.

The first step in this process was to determine the exact DNA sequence around the N-terminus of the npt2 gene. Based on the sequence data for npt2 published by Beck et al (1982), a sequencing oligonucleotide (npt2 primer) was made that was

complementary to the published strand and annealed approximately 100 bases downstream of the npt2 initiation codon. Its sequence and exact annealing positions are shown in Table 2.1.12. A large scale plasmid preparation of pCaMVNEO was made and an aliquot was plasmid sequenced using the npt2 primer. Restriction mapping was also performed on the vector. The vector, restriction map, and DNA sequence around the 5' terminus of the npt2 gene are shown in figure 3.5.1.

Study of the restriction map revealed that there were no XhoI restriction sites. It was therefore decided to introduce simultaneously a unique XhoI site at the 5' end of the npt2 gene whilst changing the npt2 initiation (ATG) codon. Close inspection of the sequence showed that a 3 base mismatch mutation was required to introduce the XhoI recognition sequence, CTCGAG, such that it also altered the initiating ATG (methionine) to GTG (valine);

i.e. (DNA sequence in pCaMVNEO)...TTTCGC **ATG...>npt2 gene**

* * *

(same region after mutagenesis)...TCTCGA GTG...>npt2 gene

The next section describes the cloning procedures followed to achieve this.

3.5.2 CONSTRUCTION OF pCaMVNEOΔX.

(for cloning strategy, see figure 3.5.2.1)

A 1.6Kb HindIII - BglII fragment from pCaMVNEO was cloned into the HindIII - BamHI sites of M13mp19 (Messing, 1986). Single-stranded M13 preparations from white plaques selected on IPTG/X-gal plates were sequenced using the npt2 primer to select M13 clones containing a single copy of the npt2 fragment. These were named M13mp19npt2 and were used as a template for site-directed mutagenesis.

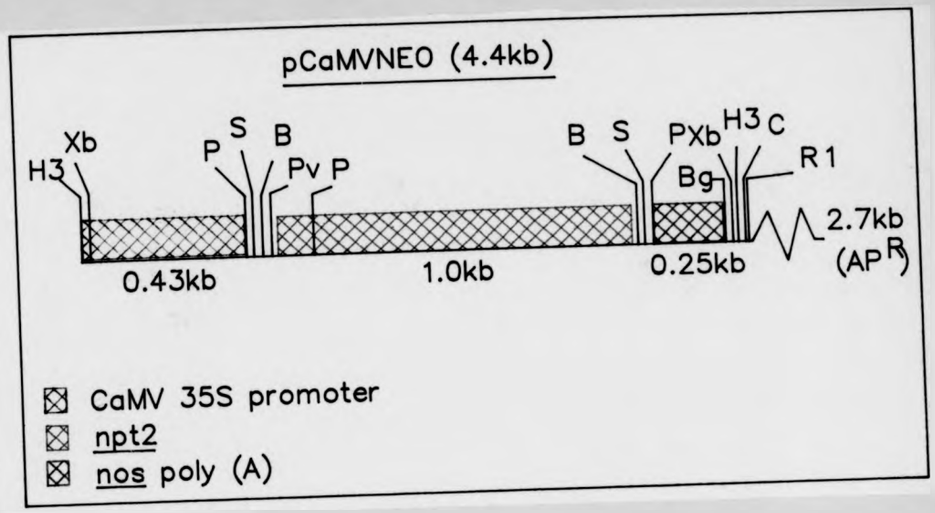
An oligonucleotide, MW04 (see Table 2.1.12), was designed and synthesised on an Applied Biosystems 380B DNA synthesiser

FIGURE 3.5.1 RESTRICTION MAP AND PARTIAL DNA SEQUENCE OF pCaMVNEO

A/ Restriction map of the plant expression vector pCaMVNEO (Fromm *et al*, 1986), assembled by deduction from the paper of Fromm *et al* and by restriction mapping carried out by this author. In this vector, the gene for neomycin phosphotransferase (npt2) from the transposon Tn5 (see Beck *et al*, 1982) is situated between the constitutive plant promoter from the cauliflower mosaic virus (35S-CaMV) and the nopaline synthase transcription terminator region (*nos* poly A). This vector contains the B-lactamase gene (AP^R), used to confer ampicillin resistance to bacterial colonies during cloning. Of relevance is the absence of any *Xho*I restriction sites in this vector.

B/ DNA sequence of the vector pCaMVNEO in the region around the 5' terminus of the npt2 gene, sequenced using the npt2 primer as described in the text. The *npt2* sequence, the npt2 primer and its site of annealing on the npt2 gene are shown in red.

A



B

TTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGAC

GTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTAT

P S B

ATAAGGAAGTTCATTTCATTTGGAGAGGA CGACCTGCAGGTCGACGGATC

Pv 1550

CGTCGATCGTTTCGCATG ATT GAA CAA GAT GGA TTG CAC GCA

1600

GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA (TGA CTG GGC

TAT GAC TGG GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC

C TGT TAG CCG ACG AGA CTA C

3' 5'

NPT2 SEQUENCE IN RED
(Nucleotide nos. refer to Beck et al, 1982)
CaMV 35S IN BLUE
LINKER IN pCaMVNEO IN BLACK
Bracketed sequence from Beck et al (1982)

FIGURE 3.5.2.1 CLONING STRATEGY FOR pCaMVNEO△X

Step 1: Ligation of 1.6Kb HindIII - BglII fragment from pCaMVNEO into HindIII & BamHI -restricted M13mp19, to produce the construct M13mp19npt2. Single-stranded DNA was prepared and used as a template for site-directed mutagenesis.

Step 2: Site-directed mutagenesis of M13mp19npt2 to mutate simultaneously the initiating codon ATG to GTG and introduce a XhoI restriction site at the 5' end of the npt2 gene. Putative mutants were screened and the clone D7 was found to contain the correct mutation.

Step 3: Ligation of 1.0Kb BamHI fragment (containing the mutated sequence) into the 3.4Kb BamHI vector fragment from wild type pCaMVNEO to produce the construct, pCaMVNEO△X.

KEY.

Restriction sites:

B	BamHI
Bg	BglII
H3	HindIII
R1	EcoRI
X	XhoI

Ap^r B-lactamase gene (confers resistance to Ampicillin)

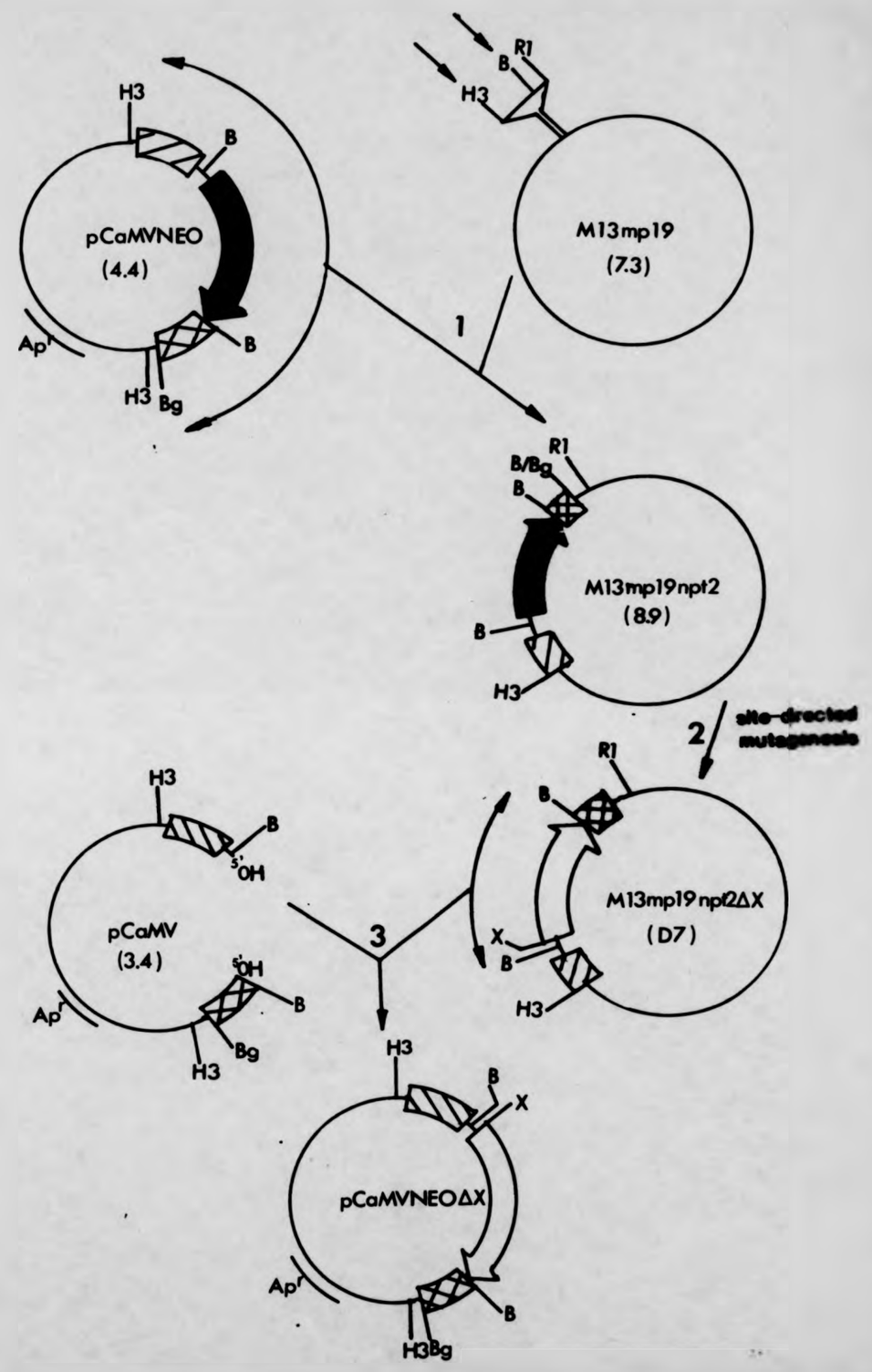
35S CaMV promoter

neomycin phosphotransferase (npt2) gene

mutated npt2 gene (initiation codon ATG - GTG)

nopaline synthase transcription termination sequence

Sizes of plasmids shown in brackets (in Kb).



apparatus, that would enable mutagenesis of the initiating ATG codon of *npt2* and create a *Xho*I restriction site. This oligonucleotide was annealed to the template and oligonucleotide-directed, site-specific mutagenesis based on the method of Nakamaye and Eckstein (1986), was used to attempt to generate the desired mutants. This method is a modification of the oligonucleotide-directed, site-specific mutagenesis protocol described by Zoller and Smith (1982). Nakamaye and Eckstein (1986) described how the incorporation of thionucleotides into the mutant strand enabled the *in vitro* restriction and exonuclease digestion of the non-mutant strand. This sufficiently increases the recovery of mutants to make feasible the successful screening of a small number of putative mutant progeny. However, when this method was used to try to generate an *npt2 Xho*I mutant, sequencing analysis of DNA from 12 putative mutants revealed that all twelve were wild-type sequences.

It was then checked that the MW04 oligonucleotide had annealed correctly to the 5' end of the npt2 gene. The Sequenase™ kit was used to sequence from the 3' end of MW04 after the latter had been first annealed to the template. No sequences were visible in any tracks where 0.5pmols (1x the suggested amount of sequencing primer), 5pmols (10x) and 50pmols were annealed to the template. When 500pmols of MW04 was annealed to the template (1000x), a sequence was visible and corresponded to a region of M13mp19 that had no significant sequence homology to the primer used. It was therefore concluded that the oligonucleotide had a defective sequence and a new oligonucleotide was synthesised. This second oligonucleotide, named OAB1049 (see Table 2.1.12), was 6 nucleotides longer than MW04 in order to improve its specificity for annealing to the template. This oligonucleotide was used as a sequencing primer and it gave a readable sequence

at the suggested primer concentration that corresponded to the expected region of the template. This confirmed that OAB1049 annealed uniquely to the 5' region of the npt2 gene.

Mutagenesis was then repeated using the method described above and with OAB1049 as the mutagenic oligonucleotide. However, all 33 putative mutants which were subsequently sequenced were found to contain wild-type sequences in the region of the desired mutation.

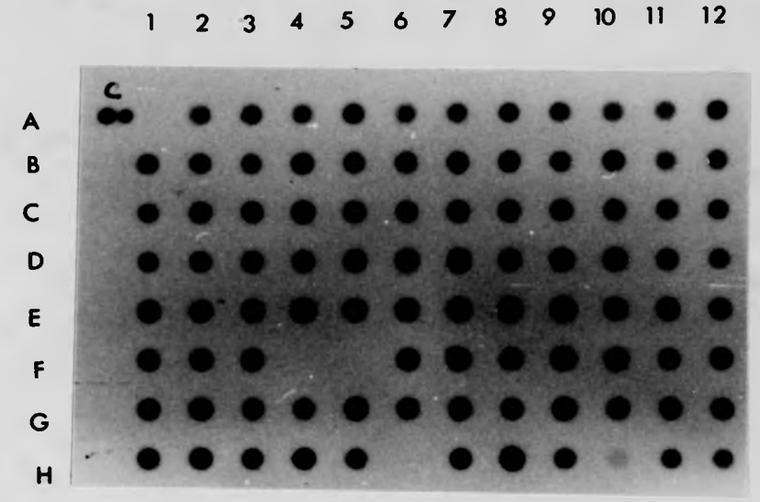
An alternative method of site-directed mutagenesis was then used in an attempt to generate mutant progeny (Kramer, Kramer and Fritz, 1984). This procedure (described in detail in Section 2.1.12) has a single *in vitro* 'second-strand synthesis' step before transformation into a *mutL* strain of *E.coli* (a strain deficient in DNA repair mechanisms, used to increase the mutation frequency). It was hoped that by reducing the number of *in vitro* mutagenic manipulations an increased recovery of mutants would be possible. After following this protocol, DNA from 96 putative mutant plaques was hybridised onto nitrocellulose and probed with ^{32}P-labelled OAB1049. Figure 3.5.2.2 shows the results of this probing after washing the filters at room temperature (Fig.3.5.2.2A) and then at 65C (Fig.3.5.2.2.B). From this, 15 positives were identified giving a mutagenic frequency of 15.6%. Single-stranded DNA was prepared from seven of these, A3,B1,B8,B10,C3,C9,D7. The npt2 primer was used to sequence each clone in the region of the mutation. Clone A3 had a wild-type sequence; B1 was a deletion mutant; but the sequenced clones B8-D7 all contained a *Xho*I restriction site in the desired position (Figure 3.5.2.3).

RF DNA was then prepared from clone D7. A 1Kb *Bam*HI fragment, containing the mutated npt2 gene, was ligated to a 3.4Kb *Bam*HI gel-isolated vector fragment from pCaMVNEO. Competent cells were transformed with this ligation mix and plasmid DNA prepared from colonies selected on ampicillin. The

FIGURE 3.5.2.2 HYBRIDISATION SCREENING OF PUTATIVE XHOI MUTANTS USING THE ^{32}P-LABELLED MUTAGENIC OLIGONUCLEOTIDE, OAB1049.

Autoradiographs of a nitrocellulose filter onto which was fixed 50ul aliquots of supernatant, containing ssM13 DNA, from 96 putative mutant progeny. The filter was probed with 15pmoles of the ^{32}P-labelled OAB1049 mutagenic oligonucleotide and subsequently washed in 6xSSC at the temperatures shown. Two 100ng aliquots of the ssM13mp19Xnpt2 DNA (the template for the mutagenesis) were used as controls (C) to show the temperature above which the probe would no longer remain bound to the wild-type (non-mutated) sequence. Note the selective binding of the probe to certain clones at the higher temperature, but not to the control DNA.

20C



65C

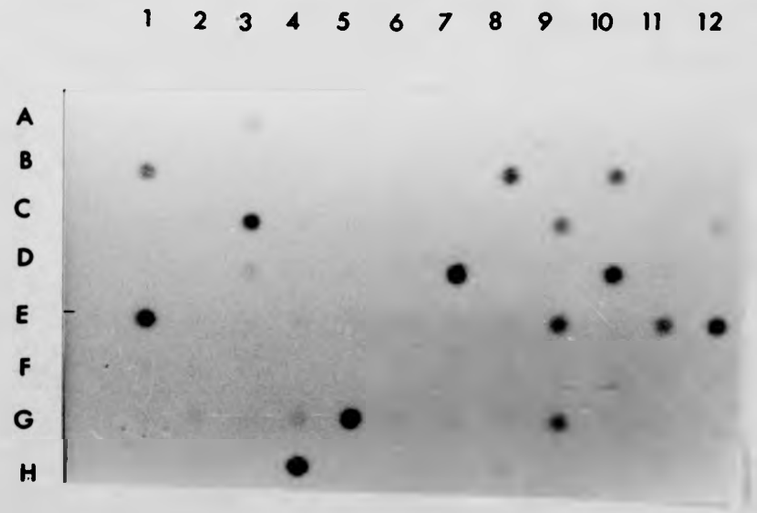
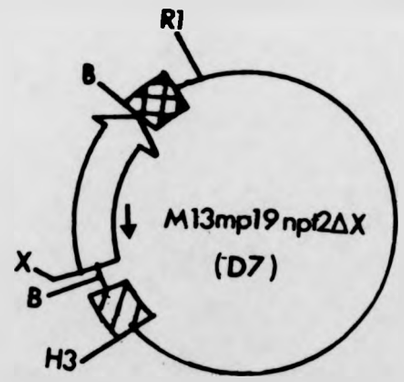
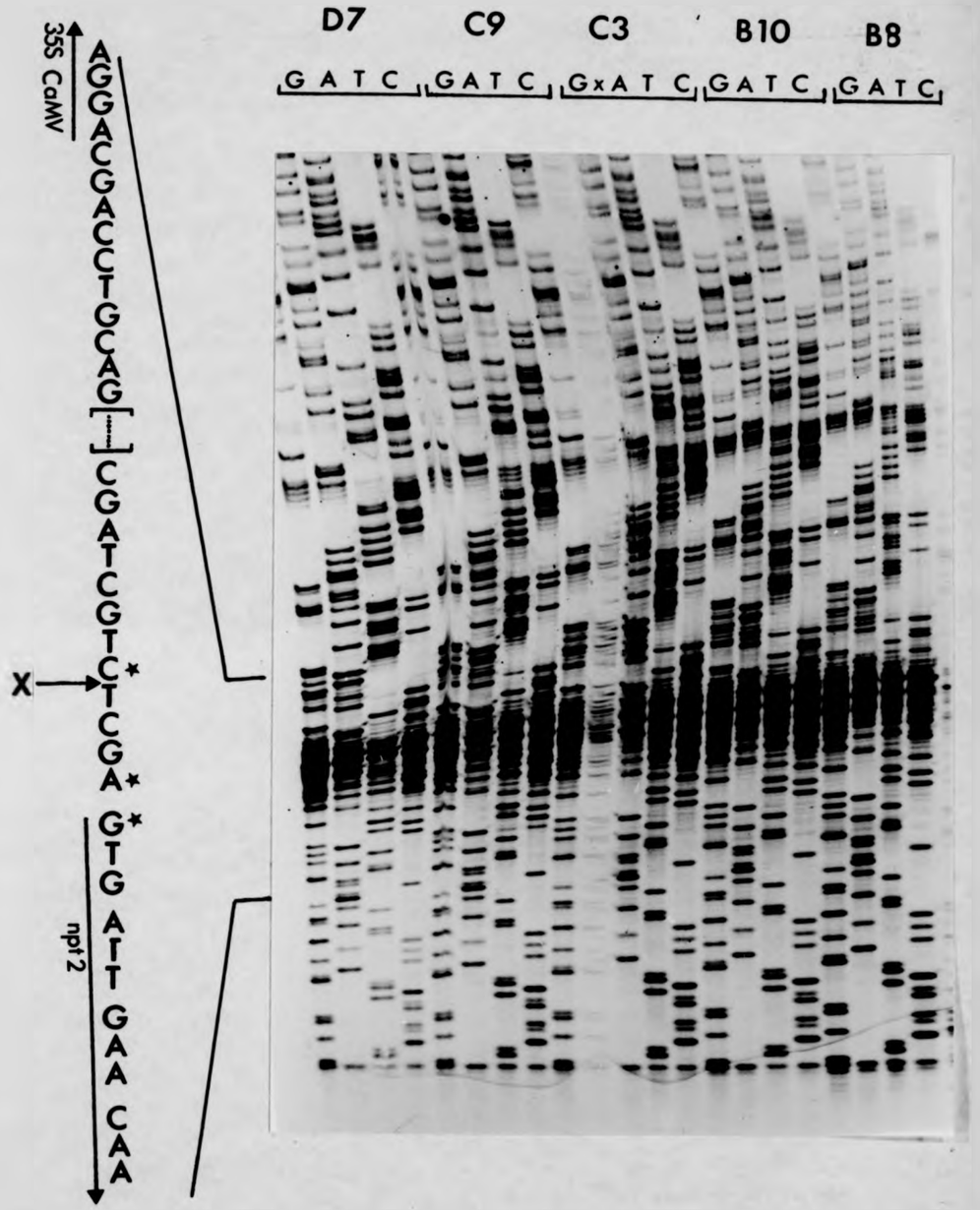


FIGURE 3.5.2.3 DNA SEQUENCE ANALYSIS OF PUTATIVE XHOI MUTANT CLONES B8, B10, C3, C9 AND D7.

Single-stranded M13 DNA was prepared from these clones after they had been identified as putative mutants by hybridisation screening (Figure 3.5.22). 2ug of each ssDNA was then sequenced using the npt2 primer and the Sequenase™ kit. All clones shown contained the 3 base mismatch mutation (asterisked) which confirmed them as XhoI mutants. The introduced restriction site is arrowed and labelled 'X'. Note the compression present in all the tracks. The significance of this compression with respect to the failure of the Amersham *in vitro* mutagenesis kit to generate XhoI mutants is discussed in the text.



35S-CaMV
npt2
nos poly A
→ npt2 primer and direction of sequencing.



orientation of the fragment was determined by restriction digest analysis (Figure 3.5.2.4). One clone, named pCaMVNEOΔX, contained the correctly mutated npt2 gene in the desired orientation with respect to the 35S-CaMV promoter. The orientation of the fragment in this clone was finally checked by plasmid sequencing, using the npt2 primer.

3.5.3 CONSTRUCTION OF pPARTIAL.

(for cloning strategy, see Figure 3.5.3.1)

A 71bp XhoI - BamHI fragment from the vector ppRCL617, which contains the full-length coding sequence of preproricin, was end-filled and gel-isolated. This DNA fragment encodes the first 23 N-terminal amino acid residues of preproricin. It was cloned into the XhoI restriction site of the mutated vector, pCaMVNEOΔX. This was accomplished by first cutting the vector with XhoI before end-filling the 3'-recessed 'sticky-ends'. The vector was then phosphatased and the 71bp ricin fragment ligated. Mini-preparations of plasmid DNA from transformed E.coli were sequenced to identify a pCaMVNEOΔX clone that contained a single, correctly-orientated 71bp fragment. Once identified, a large scale DNA plasmid preparation was made and the construct named pPARTIAL. It was finally sequenced once more using the npt2 primer.

3.5.4 CONSTRUCTION OF pFULL.

(for cloning strategy, see Figure 3.5.4.1)

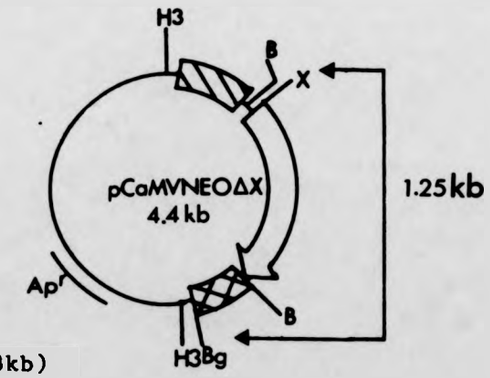
As there were no convenient restriction sites in the preproricin cDNA in the region encoding the junction between the N-terminal region and the first amino acid residue of the mature A chain, it was decided to use site-directed mutagenesis to introduce a SalI restriction site at position +1 in the DNA.

A 702 bp XhoI - BglII fragment from ppRCL617 was cloned into the XhoI/BamHI sites of the M13mp18-derived vector,

FIGURE 3.5.2.4 RESTRICTION DIGEST ANALYSIS OF pCaMVNEOΔX SHOWING THE CORRECT ORIENTATION OF THE 1KB BAMHI FRAGMENT (CONTAINING THE MUTATED NPT2 GENE).

1ug of pCaMVNEOΔX DNA was digested with 10 units each of XhoI and BglII for 1 hour at 37C. DNA bands were then resolved by electrophoresis through a 1%agarose (1xTBE buffer) gel containing 0.5ug/ml EtBr. Bands were visualised on a UV-transilluminator.

1. 1ug 'wild-type' pCaMVNEO digested with XhoI and BglII. As there is no XhoI restriction site the plasmid is linearised.
2. 1ug of a pCaMVNEOΔX clone in which the 1kb BamHI fragment was inserted in the wrong orientation. Digestion with XhoI and BglII therefore releases a small fragment of approximately 250bp (predicted size).
3. 1ug of pCaMVNEOΔX digested with XhoI and BglII. The predicted release of a 1.25kb fragment is seen (see plasmid map of pCaMVNEO X below).
M. Lambda DNA digested with EcoRI and HindIII. Sizes (in bp) are shown.



35S CaMV (0.43kb)
npt2 (1.0kb)
nos poly A (0.25kb)

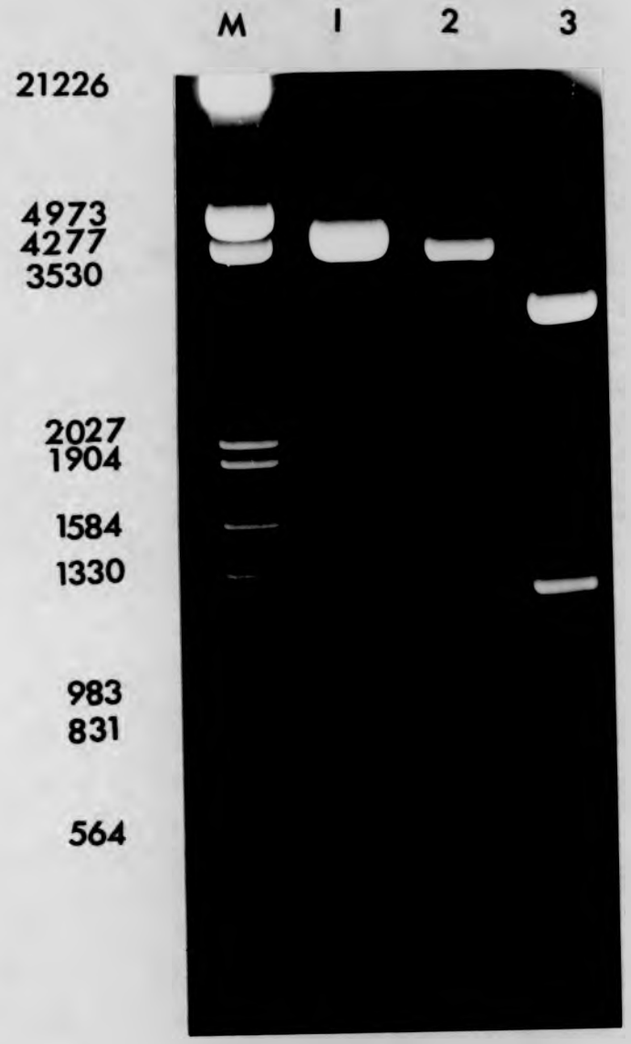


FIGURE 3.5.3.1 CLONING STRATEGY FOR pPARTIAL CONSTRUCT

Step 1: The 71bp XhoI - BamHI fragment from ppRCL617 (containing the first 23 codons of preproricin) was end-filled, gel isolated and ligated into pCaMVNEOΔX, after it had first been linearised with XhoI and end-filled. The ligated DNA was transformed into E.coli TG1 cells and DNA from ampicillin-resistant colonies was sequenced to identify a clone which contained the 71bp preproricin fragment inserted (in the correct orientation) in front of the npt2 gene such that the reading frame was conserved (i.e. in-frame). This construct was termed 'pPARTIAL'.

KEY.

Restriction sites:	B	BamHI
	Bg	BglII
	H3	HindIII
	R1	EcoRI
	X	XhoI

Ap^r	B-lactamase gene (confers resistance to Ampicillin)
▨	35S CaMV promoter
➡	neomycin phosphotransferase (npt2) gene
⇨	mutated npt2 gene (initiating codon ATG - GTG)
▩	nopaline synthase transcription termination sequence
S	region encoding preproricin N-terminal presequence
A	ricin A chain coding sequence
L	linker region coding sequence of preproricin
B	ricin B chain coding sequence

Sizes of plasmids shown in brackets (in Kb).
Nucleotide reference numbers refer to the published cDNA sequence of preproricin (Lamb et al, 1985).

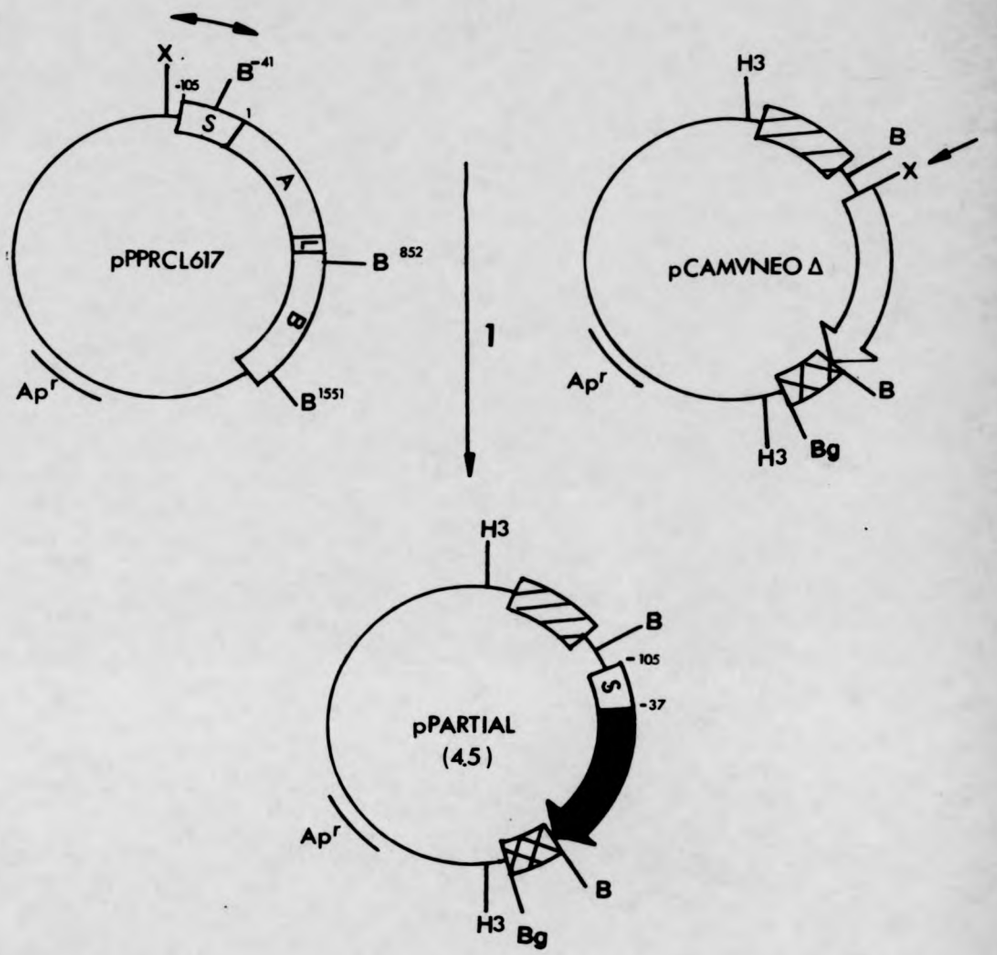


FIGURE 3.5.4.1 CLONING STRATEGY FOR pFULL CONSTRUCT

Step 1: The 702bp XhoI - BglII fragment from ppRCL617 (containing the preproricin 5' end encoding the entire N-terminal presequence) was cloned into the multiple-cloning site of M13mp18X, after first restricting the M13 vector with XhoI and BamHI. Single-stranded DNA was prepared and used as a template for site-directed mutagenesis.

Step 2: Site-directed mutagenesis was used to introduce a SalI restriction site immediately 3' to the last codon of the preproricin 35aa residue N-terminal region. The DNA from putative mutants was sequenced and the clone F10 identified as containing the correct mutation. RF DNA was prepared from this clone.

Step 3: The 124bp SalI fragment from clone F10 (containing the complete preproricin N-terminal region coding sequence) was then cloned into the XhoI restriction site of pCaMVNEOΔX. A clone containing the preproricin N-terminal region coding sequence, fused in frame with and in the correct orientation with respect to the npt2 coding sequence was identified by restriction digest analysis and plasmid sequencing. This clone was termed 'pFULL'.

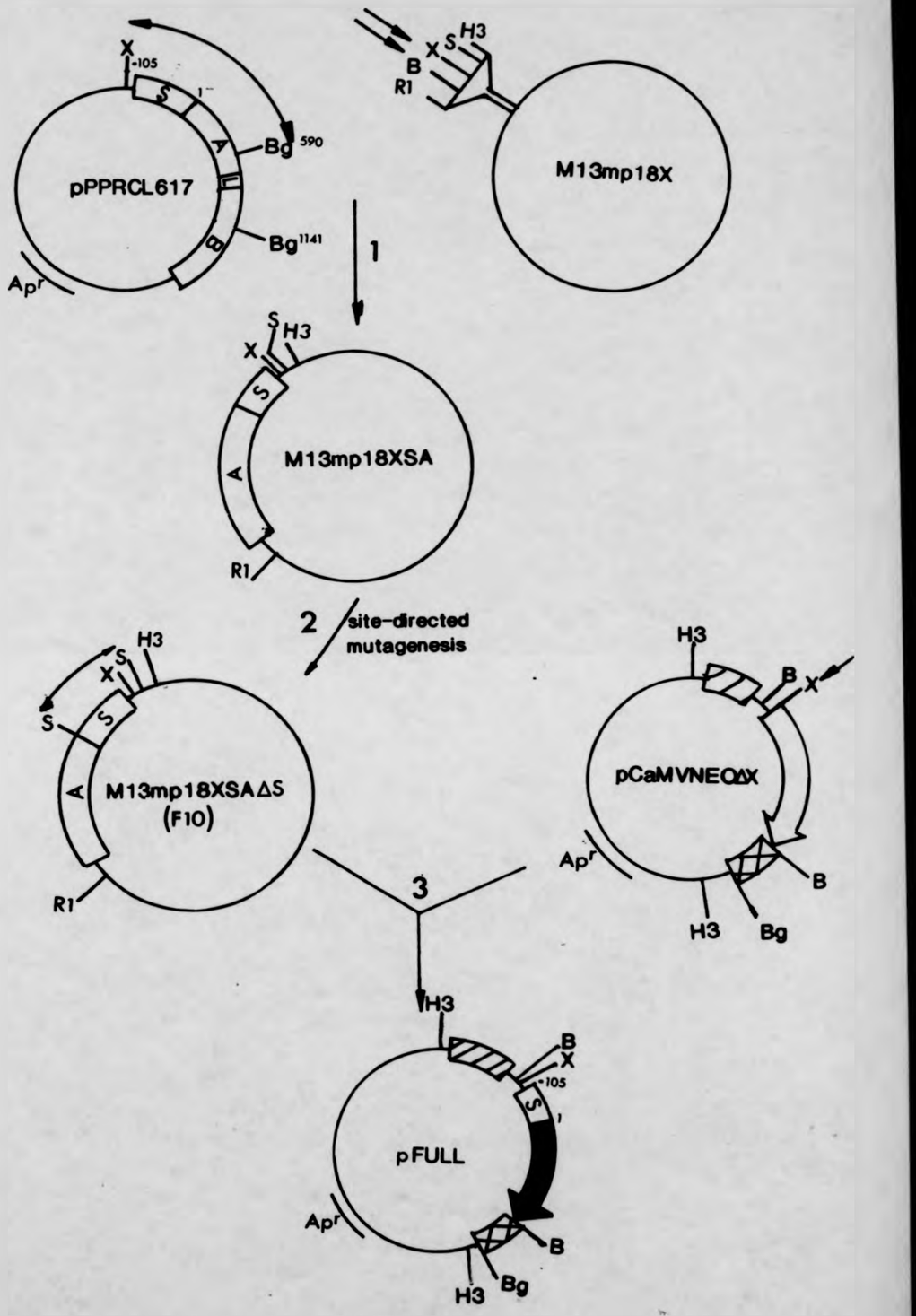
KEY.

Restriction sites:

B	BamHI
Bg	BglII
H3	HindIII
R1	EcoRI
S	SalI
X	XhoI

Ap^r	B-lactamase gene (confers resistance to ampicillin)
▨	35S CaMV promoter
➡	neomycin phosphotransferase (npt2) gene
⇨	mutated npt2 gene (initiating ATG - GTG)
▩	nopaline synthase transcription termination sequence
S	region encoding preproricin N-terminal presequence
A	ricin A chain coding sequence
L	linker region coding sequence of preproricin
B	ricin B chain coding sequence

Sizes of plasmids shown in brackets (in Kb).
Nucleotide reference numbers refer to the published cDNA sequence of preproricin (Lamb et al, 1985).



M13mp18X (Turner, 1991). This vector contains a unique XhoI restriction site in the M13 polylinker. Single-stranded DNA preparations were sequenced using the M13 universal primer (Table 2.1.12) and a clone containing the correctly-inserted preproricin fragment was used as a template for mutagenesis.

A mutagenic oligonucleotide, MW05 (Table 2.1.12), was designed to introduce a five base mismatch mutation. This oligonucleotide was made on the same machine as oligonucleotide MW04, which had been shown to be defective (above). Therefore, MW05 was first used as a primer in a sequencing reaction and this confirmed that it annealed to the template in the correct place. MW05 was then used in conjunction with the method for site-specific mutagenesis described by Nakamaye and Eckstein (1986) to try to create the desired mutation. However, all putative mutants screened by sequencing were found to be wild-type. When this method was repeated it again proved unsuccessful at generating mutants. The ability of the mutagenic oligonucleotide to prime sequencing of the template illustrated that it was annealing in the correct place and the 3' end of the oligonucleotide was attached to the template. However, in light of the problems encountered with the mutagenic oligonucleotide, MW04 (discussed in Section 3.5.2 above), synthesised at the same time as MW05, it was thought likely that MW05 was also defective, possibly at its 5' end which would affect the ligation reaction during the second-strand synthesis stage of mutagenesis and affect the frequency of mutants. Therefore, it was decided to re-synthesise a mutagenic oligonucleotide.

This second oligonucleotide, OAB1179 (Table 2.1.12), was used in conjunction with the method of mutagenesis described by Kramer, Kramer and Fritz (1984), which had proven successful in the creation of pCaMVNEOΔX (above). DNA from 96 putative mutants was probed with ^{32}P-labelled OAB1179 and the results

are shown in Figure 3.5.4.2. From this, 10 positive clones were identified which gave a mutation frequency of 10.4%. Single-stranded DNA was prepared from six of these positive clones (B4, C1, C5, E9, E12, F10) and sequenced using the universal primer. Clone B4 contained a wild-type sequence. Clones C1-E12 had poor sequence in the area of the desired mutation and it was not possible to determine if all 4 bases had been correctly mutated. However, clone F10 possessed the correct sequence (Figure 3.5.4.3).

RF DNA was then prepared from clone F10 and a 124bp SalI fragment, containing the entire coding region for the preproricin N-terminal region, was cloned into the XhoI site of pCaMVNEOΔX. Mini plasmid DNA preparations were sequenced using the npt2 primer to identify a pCaMVNEOΔX clone with the correctly-inserted and orientated ricin fragment. A large scale plasmid preparation was made from this and the construct was termed pFULL.

3.6 COMPUTER ANALYSIS OF pPARTIAL and pFULL CONSTRUCTS.

The predicted N-terminal amino acid sequences and von Heijne plots of the ricin:npt2 chimeras encoded by the pPARTIAL and pFULL constructs are shown in Figure 3.6. It is relevant to note that, in the construction of both chimeras, two 'extra' codons were created at the junction between the ricin and npt2 coding sequence. They encode the 'foreign' residues arginine (positions 24 and 36 for the pPARTIAL and pFULL gene products respectively) and valine (positions 25 and 37).

The von Heijne algorithm gave a strong, predicted signal peptide cleavage site after Gly^{22} for the pPARTIAL chimeric gene product. This means that the environment around this site in the protein is favourable for signal peptidase action. When the predicted N-terminal sequence of the pFULL gene product was

FIGURE 3.5.4.2 HYBRIDISATION SCREENING OF PUTATIVE SALI MUTANTS USING THE ^{32}P-LABELLED MUTAGENIC OLIGONUCLEOTIDE, OAB1179.

Autoradiographs of a nitrocellulose filter onto which was fixed 50ul aliquots of supernatant, containing ssM13 DNA, from 96 putative mutant progeny. The filter was probed with 15pmoles of the ^{32}P-labelled OAB1179 mutagenic oligonucleotide and subsequently washed in 6xSSC at the temperatures shown. Two 100ng aliquots of the ssM13mp18signal DNA (the template for the mutagenesis) were used as controls (C) to show the temperature above which the probe would no longer remain bound to the wild-type (non-mutated) sequence. Note the selective binding of the probe to certain clones at the higher temperature, but not to the control DNA.

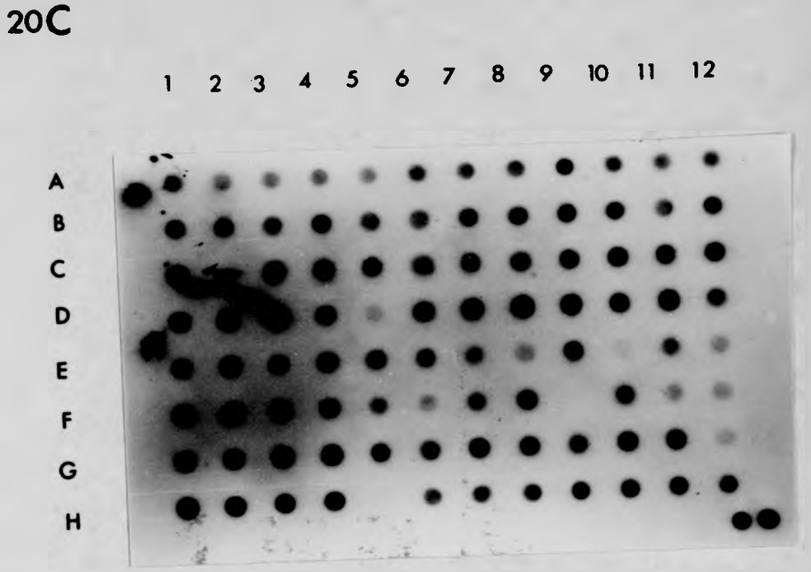
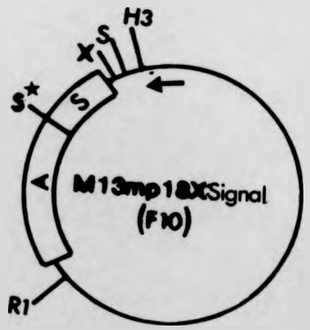


FIGURE 3.5.4.3 DNA SEQUENCE ANALYSIS OF PUTATIVE SALI MUTANT CLONES B4 & F10.

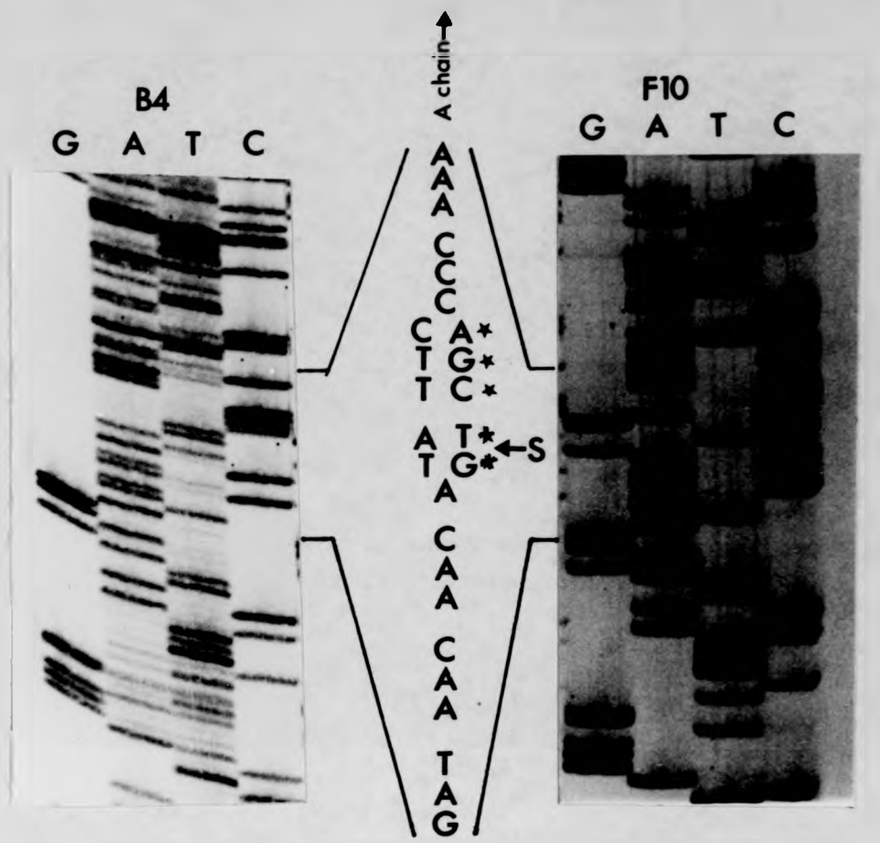
Single-stranded M13DNA was prepared from these clones after they had been identified as putative mutants by hybridisation screening (Figure 3.5.43). 2ug of this DNA was sequenced using the M13 universal primer and the Sequenase™ kit. An autoradiograph of the sequencing gel is shown giving the relevant regions in clones B4 and F10. The clone B4 appeared as a non-mutant (i.e. wild-type sequence) in this region, but clone F10 contained the desired 5-base mismatch mutation (asterisked), which introduced a SalI restriction site at the position shown (arrowed and labelled 'S'). Below the photographs is the 5' DNA coding sequence of preproricin, amended from Lamb et al (1985)(see Appendix for the complete DNA coding sequence). The region shown in the photographs above is highlighted.



S N-terminal region coding sequence of preproricin

A ricin A chain coding sequence

→ M13 universal primer and direction of sequencing



5' -105 -100

---CTCGAGG ATG AAA CCG GGA AGA AAT ACT ATT GTA ATA TGG

Met Lys Pro Gly Gly Asn Thr Ile Val Ile Trp

(--- --- --- --- --- --- --- --- --- --- ---

ATG TAT GCA GTG GCA ACA TGG CTT TGT TTT GGA TCC ACC TCA GGG TGG TCT TTC AGA TTA GAG
Met Tyr Ala Val Ala Thr Trp Leu Cys Phe Gly Ser Thr Ser Gly Trp Ser Phe Thr Leu Glu
--- ---

GAT AAC AAC ATA TTC CCC AAA CAA TAC CCA ATT ATA AAC TTT ACC ACA GGG GGT GCC ACT GTG CAA AGC TAC ACA AAC TTT ATC AGA GCT
Asp Asn Asn Ile Phe Pro Lys Gln Tyr Pro Ile Ile Asn Phe Thr Thr Ala Gly Ala Thr Val Gln Ser Tyr Thr Asn Phe Ile Arg Ala
--- --- ---

GTT CGC GGT CGT TTA ACA ACT GGA GCT GAT GTG AGA CAT GAT ATA CCA GTG TTG CCA AAC AGA GTT GGT TTG CCT ATA AAC CAA CGG TTT
Val Arg Gly Arg Leu Thr Thr Gly Ala Asp Val Arg His Asp Ile Pro Val Leu Pro Asn Arg Val Gly Leu Pro Ile Asn Gln Arg Phe
Glu

ATT TTA GTT GAA CTC TCA AAT CAT GCA GAG CTT TCT GTT ACA TTA GCC CTG GAT GTC ACC AAT GCA TAT GTG GTC GGC TAC CGT GCT GGA
Ile Leu Val Glu Leu Ser Asn His Ala Glu Leu Ser Val Thr Leu Ala Leu Asp Val Thr Asn Ala Tyr Val Val Gly Tyr Arg Ala Gly
Gln Ser

FIGURE 3.6 DNA SEQUENCE AND VON HEIJNE ANALYSIS OF PREDICTED N-TERMINAL REGIONS OF FUSION PROTEINS ENOCODED BY THE CONSTRUCTS pPARTIAL AND pFULL.

A/ DNA sequences and deduced amino acid sequences of the N-terminal regions encoded by the chimeric genes in pPARTIAL and pFULL. The 'extra' codons created as a result of cloning manipulations are highlighted by dashed lines. The mutated initiating codon of the npt2 gene (ATG to GTG) is asterisked.

B/ Graphic representation of the von Heijne algorithm when applied to the predicted N-terminal sequences shown in **A**. The residues are shown along the x-axis. The first 23 residues are invariant (shown by dashes). In both cases the residue glycine[22] remains a good candidate for the site of signal peptidase cleavage. In the predicted N-terminal sequence for the pFULL chimeric gene product, threonine[30] also has a high value.

A

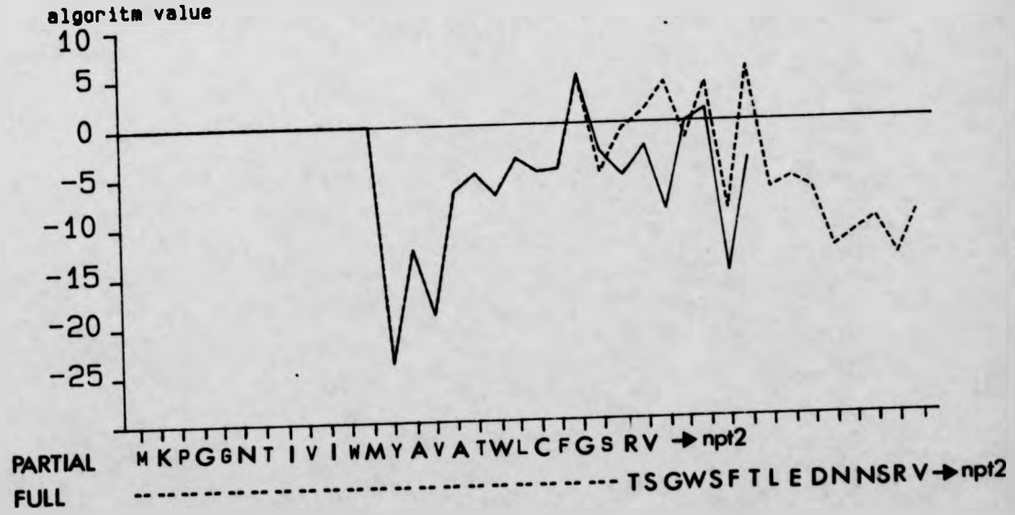
PARTIAL

```
                                                    30                                  60
ATG AAA CCG GGA GGA AAT ACT ATT GTA ATA TGG ATG TAT GCA GTG GCA ACA TGG CTT TGT
Met Lys Pro Gly Gly Asn Thr Ile Val Ile Trp Met Tyr Ala Val Ala Thr Trp Leu Cys

                        ★
TTT GGA TCT CGA GTG →npt2
Phe Gly Ser Arg Val
```

FULL

```
                                                    30                                  60
ATG AAA CCG GGA GGA AAT ACT ATT GTA ATA TGG ATG TAT GCA GTG GCA ACA TGG CTT TGT
Met Lys Pro Gly Gly Asn Thr Ile Val Ile Trp Met Tyr Ala Val Ala Thr Trp Leu Cys

                                                    90                      ★
TTT GGA TCC ACC TCA GGG TGG TCT TTC ACA TTA GAG GAT AAC AAC AGT CGA GTG →npt2
Phe Gly Ser Thr Ser Gly Trp Ser Phe Thr Leu Glu Asp Asn Asn Ser Arg Val
```

B



analysed, it showed a similar profile to that of the preproricin N-terminus (compare with Figure 3.2).

3.7 IN VITRO EXPRESSION OF THE CHIMERIC GENE FROM THE pPARTIAL CONSTRUCT.

The ability of the partial preproricin N-terminal sequence cloned into the pPARTIAL construct to act as a functional ER signal peptide was then assessed using an *in vitro* expression system. The 1Kb BamHI chimeric gene fragment from the pPARTIAL construct was first cloned into the BglII site of the SP6 expression vector, pSP64T (Krieg and Melton,1984) and the new construct named pSP64Tpartial. The 1Kb BamHI fragment from pCaMVNEO, containing the wild-type npt2 gene was also cloned into pSP64T (named pSP64Tnpt2) to be used as a control.

Transcripts from pSP64Tnpt2 and from pSP64Tpartial were translated in rabbit reticulocyte lysates in the presence or absence of dog pancreatic microsomes. The results are shown in Figure 3.7. In both cases, when translated in the absence of microsomes a single labelled protein product was observed (lanes 1 and 3). The protein expressed from pSP64Tpartial was estimated to be 30kD in size, which was about 2kD larger than the npt2 gene translated from pSP64Tnpt2 mRNA. This 'wild-type' npt2 gene product was similar in size when translated in the presence of dog microsomes (lane 2). However, when pSP64Tpartial mRNA was translated in the presence of dog microsomes (lane 4), a second band (of similar size to npt2) was observed in addition to the full-length fusion protein (compare lanes 1,3 and 4). It is proposed that this lower band represented protein which had been co-translationally translocated across the microsomal membrane and been processed by signal peptidase. As there are no tripartite (Asn-X-Ser/Thr) sequences in the fusion protein, no asparagine-linked

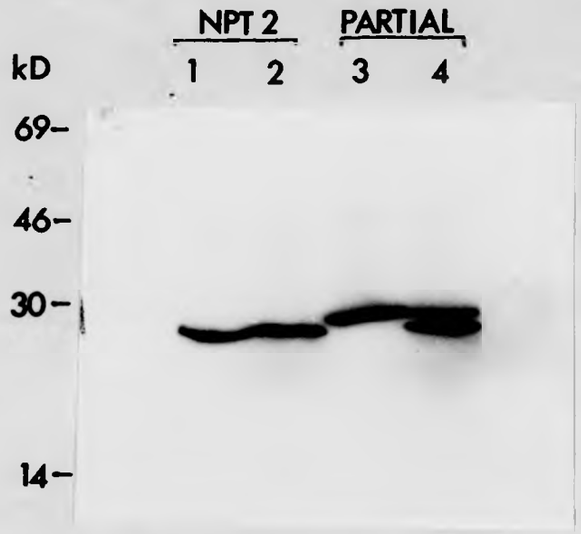


FIGURE 3.7 IN VITRO EXPRESSION OF THE CHIMERIC GENE FROM pPARTIAL.

The npt2 gene from pCaMVNEO and the chimeric partial ricin presequence/npt2 gene from pPARTIAL were cloned into the *in vitro* expression vector, pPSP64T, and transcribed and translated in rabbit reticulocyte lysate using the methods described in Chapter 2. The translated products were resolved by reducing SDS-PAGE and the bands visualised by fluorography. Lanes 1 and 3 show the single protein products from the *in vitro* expression of npt2 and pPARTIAL genes respectively. In the translations shown in lanes 2 and 4, 1ul of dog pancreatic microsomes had been added at the start of each incubation.

glycosylation would be expected to occur. The segregated fusion protein, after cleavage of the N-terminal signal, was therefore smaller in size than the non-segregated fusion protein. Furthermore, the site of signal peptidase cleavage in the fusion protein must be at/close to the site of cleavage predicted by the von Heijne algorithm - namely to the carboxyl side of Gly^{22} - as the segregated protein was of similar size to wild-type npt2. It was therefore concluded that the 22 residues at the N-terminus of preproricin was able to function as a signal peptide and mediated the transport of the pPARTIAL-encoded fusion protein across the ER membrane.

3.8 DISCUSSION.

The algorithm of von Heijne was applied to the N-terminal residues of preproricin and a computer programme was designed to help do this. The algorithm identified the residues $glycine^{22}$ and $threonine^{30}$ in the N-terminal region of preproricin as being the most likely sites for signal peptide cleavage. On the other hand, the last residue of the 35aa residue N-terminal presequence, $asparagine^{35}$, was given a negative value which according to von Heijne (1985) renders it a very poor candidate for signal peptide cleavage. This is, therefore, strong indirect evidence to support the idea that once translocated into the lumen of the ER, proricin has an N-terminal propeptide of 5-13 residues.

Two constructs suitable for expression in plant systems were then successfully constructed. These contained chimeric genes between preproricin cDNA and the gene for the marker enzyme npt2. In the process of their construction, two site-directed mutagenesis steps were necessary and were carried out successfully. However, the method of mutagenesis described by Nakamaye and Eckstein (1986) failed to generate any mutant progeny. In the construction of the mutated vector pCaMVNEOΔX,

the failure of this method to introduce a 3-base mismatch mutation was attributed to the region of DNA near to the site of mutation which showed compressions when sequenced (see Figure 3.5.2.3). This is characteristic of secondary structure formation in the DNA and this may have hindered one or more of the *in vitro* manipulations involved in this mutagenesis protocol. Based on this observation, another mutagenesis protocol (Kramer, Kramer and Fritz, 1984) was followed which involved fewer *in vitro* manipulations. This protocol used a strain of *E.coli* which is deficient in its DNA repair mechanisms and so increases the viability of heteroduplex plasmid DNA. This second method proved successful in generating both of the desired mutants with a mutation frequency of 10-15%. Because this is lower than the mutation frequencies claimed possible with the method for mutagenesis of Nakamaye and Eckstein (1986), an additional hybridisation screening step was necessary to identify putative mutants before DNA sequencing was used to confirm the presence of the desired mutation.

Of the two constructs designed, one, termed pPARTIAL, encoded a protein chimera with 23aa residues from the N-terminus of preproricin fused onto npt2. This protein therefore contained the N-terminal presequence of preproricin up to and including glycine^{22}, one of the two residues identified by the von Heijne algorithm as being a likely site of signal peptide cleavage. The second construct named pFULL, encoded a protein with the complete preproricin N-terminal presequence (35 residues) fused to npt2. The von Heijne algorithm was applied to the predicted N-terminal regions of both gene products and both were found to have residues near their N-termini which were good candidates for being the sites of signal peptide cleavage.

Since both constructs pPARTIAL and pFULL were designed

for expression in plant protoplasts in order to study plant protein targeting, it was necessary to check that the fusion protein encoded by pPARTIAL contained a functional ER signal peptide. The ability of the first 23 amino acid residues of preproricin to mediate the transport of a nascent ricin/npt2 polypeptide chain across the ER-membrane was therefore assessed by cloning the chimeric gene from the pPARTIAL construct into the *in vitro* expression vector pSP64T, and then expressing the gene *in vitro* in the presence or absence of dog pancreatic microsomes. The results revealed that in the presence of microsomes about 50% of the fusion protein was processed to a size similar to wild-type npt2 (Figure 3.7, lanes 1 & 4). As expected, there was no such processing observed when wild-type npt2 was translated in the presence of microsomes (Figure 3.7, lane 2). This suggests that the preproricin-derived portion of the fusion protein mediated this processing response. Since the processing of the chimaera resulted in a protein product which was of similar size to wild-type npt2, it was concluded that cleavage of the signal peptide by the ER signal peptidase occurred at or very near to the carboxyl side of the residue glycine^{22}. This had been identified by von Heijne's algorithm as being a good candidate for the site of signal peptide cleavage in preproricin. However, whether glycine^{22} is the actual site of signal peptide cleavage in preproricin cannot be unequivocally determined from this data.

CHAPTER 4

TRANSIENT EXPRESSION OF pPARTIAL AND pFULL IN TOBACCO LEAF PROTOPLASTS

CONTENTS

4.1 INTRODUCTION.

The construction of the plasmids pPARTIAL and pFULL have been described in the last chapter. This chapter will deal with the expression of these ricin/npt2 chimeric genes in tobacco leaf protoplasts. The chapter will begin with a short section which explains the reasons for the selection of npt2 as a suitable marker enzyme and the transient expression in protoplasts as an appropriate system for expression of the ricin/npt2 chimeras. The results will then be presented, along with the conclusions that have been drawn from them.

It should be noted that all of the work presented in this chapter was performed at the Istituto Biosintesi Vegetali, Milan, under the supervision of Dr. Aldo Ceriotti.

4.2 CHOICE OF MARKER ENZYME AND EXPRESSION SYSTEM.

The use of recombinant DNA techniques for studying protein targeting in eukaryotic cells is well documented and was reviewed by Garoff (1985). B-glucuronidase (GUS), neomycin phosphotransferase (npt2), chloramphenicolacetyltransferase (CAT), and phosphinothricin acetyltransferase (PAT) have all been used successfully as reporter genes in plants. Sensitive assays have been described for each reporter gene system (Jefferson *et al*, 1987 - GUS: McDonnell *et al*, 1987 - npt2: Gormen *et al*, 1982 - CAT: De Block *et al*, 1987 - PAT). They have been used for both studies in transgenic plants and rapid analysis using transient expression in protoplasts.

The regeneration of transgenic plants after transformation with a foreign gene is a useful system for studying the tissue-specific expression of foreign plant promoters. It is also necessary if accumulation of relatively large quantities of the foreign protein is desirable. Boutry *et al* (1987) have also used transgenic plants to study intracellular targeting of a reporter gene (npt2) to the

mitochondria after tagging with a mitochondrial presequence.

However, the high sensitivity of the marker enzyme assays and the relatively rapid analysis capable with transient expression studies have made this alternative an attractive system for the rapid screening of many different constructs, in preference to the time consuming process of regenerating transgenic plants. For example, Werr and Lorz (1986) described a transient expression system for analysing plant promoters in Graminaceae protoplasts using npt2 as the marker gene. Topfer *et al* (1988) compared the relative activities of GUS, CAT and npt2 in transient expression studies using tobacco leaf protoplasts. Their results showed that the order of relative sensitivity for the three reporter genes under these conditions was CAT < npt2 < GUS.

Denecke *et al* (1990) used transient expression in tobacco leaf protoplasts to study the plant secretion pathway. They compared the relative rates of secretion of GUS, npt2 and PAT by fusing an ER signal peptide sequence to each gene and then transforming protoplasts. After a 36 hour incubation they analysed the cell and media fractions for enzyme activity. In each case the marker enzyme was targeted to the lumen of the ER and was then secreted into the media. However, the relative rates of secretion of GUS, npt2 and PAT into the media differed considerably. They found that the order was GUS < npt2 < PAT. GUS is a tetrameric protein and its assembly in the ER was thought to retard its rate of secretion. GUS also has the disadvantage of containing two potential N-glycosylation sites which, when glycosylated, result in a complete loss of enzyme activity. These considerations meant that GUS was not a good candidate for making ricin/marker enzyme fusions to be targeted to the endomembrane system. However, the observations of Topfer *et al* (1988: above) illustrates its high sensitivity and make it an excellent cytosolic marker.

On the other hand, npt2 has a significantly high sensitivity in transient expression systems (Topfer et al, 1988: above), an acceptable rate of secretion from the ER lumen (Denecke et al, 1990: above) and it is able to tolerate relatively long N-terminal extensions without appreciable loss of enzyme activity (Denecke, personal comm.). This made it the most attractive candidate for this work and, therefore, the constructs pPARTIAL and pFULL were designed and made as described in Chapter 3.

The site of action of the neomycin phosphotransferase enzyme encoded by the npt2 gene from transposon Tn5 (Beck et al, 1982) on its substrate, neomycin, is shown in figure 4.2. The vector pCaMVNEO, used as a building block for the ricin/npt2 chimeras, encodes a copy of this gene. Details of its enzyme kinetics can be found in a review by Davies and Smith (1978).

4.3 ISOLATION AND TRANSFORMATION OF PROTOPLASTS.

4.3.1 ISOLATION OF PROTOPLASTS

The protocol described in Section 2.7 was used for the asceptic isolation of SR1 protoplasts from the leaves of N.tabaccum petit Havana var. Early attempts to isolate intact protoplasts proved unsuccessful. This failure was attributed to the high light intensity under which the plants were grown. It was thought to be sufficient to cause a significant accumulation of starch grains in the chloroplasts which ruptured the protoplasts during centrifugation steps. The light intensity was therefore adjusted to 700 Lux, the illumination being provided by a single GROWLUX tungsten element, and this enabled an average yield of 5×10^6 viable protoplasts from 20 healthy leaves (See Figure 4.3.1).

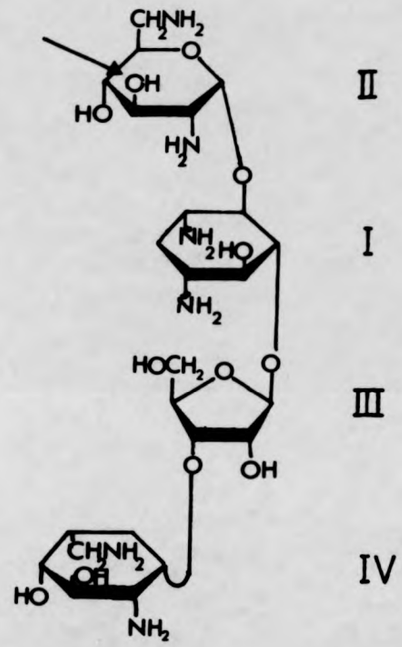


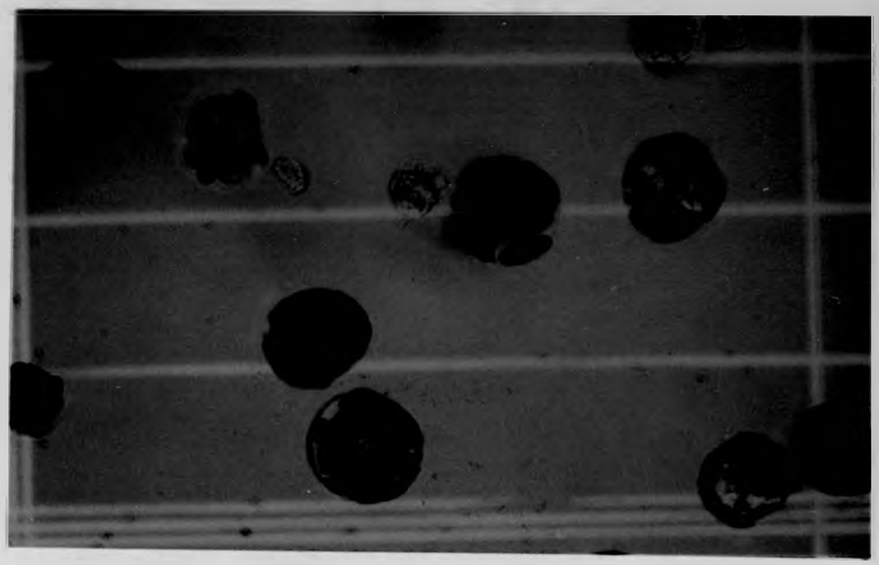
FIGURE 4.2 SITE OF ACTION OF NPT2 ON NEOMYCIN.

The npt2 gene product from Tn5 (see Beck et al, 1982) is classified as a 3'-O-phosphotransferase (3'APH). It catalyses the transfer of the gamma phosphate group from ATP to the arrowed hydroxyl group on carbon ring II of neomycin. This forms the basis of a sensitive assay for npt2 activity (McDonnell et al, 1987).

FIGURE 4.3 IMMUNOFLUORESCENCE OF VIABLE PROTOPLASTS

The photographs opposite show how viable protoplasts can be estimated by fluorescein diacetate (FDA) staining. 50µl of a 4:1000 dilution of FDA (5mg/ml stock in acetone) in K3 media was mixed with 50µl of the isolated protoplast suspension (in W5) using a wide bore pipette. Photograph **A** shows the protoplasts under transmitted light: photograph **B** is the view seen under fluorescent light. The dye passes freely across the plasma membrane but only in living protoplasts are the molecules cleaved by the action of esterases (Larkin, 1976). Cleavage results in the releease of fluorescein which is no longer able to pass back across the membrane. The fluorescein therefore accumulates within intact cells and can be excited by ultraviolet light. By carrying out this operation on a haemocytometer slide, the numbers of viable protoplasts can be estimated.

A



50µ

B

4.3.2 PEG/CALCIUM NITRATE-MEDIATED PROTOPLAST TRANSFORMATION.

Intact, isolated protoplasts were transformed with the npt2-encoded vector of interest. During PEG treatment the protoplasts were seen to form small clumps. This was thought to be a result of the effects of PEG on the cell membranes. As the PEG solution was first diluted and then finally removed from the protoplast suspension, the protoplasts became re-dispersed into single cells.

Approximately 90% protoplast survival was observed after this PEG treatment. This was estimated visually as being the proportion of protoplasts that floated in the K3 expression medium into which the protoplasts were placed after transformation. In this medium, which is 0.4M with respect to sucrose, only intact protoplasts will float (Saul et al, 1988).

Transformed protoplasts were stable in this medium when incubated in the dark, at 26C and for 36 hours.

4.4 TRANSFORMATION OF PROTOPLASTS WITH pCaMVNEO, pPARTIAL AND pFULL.

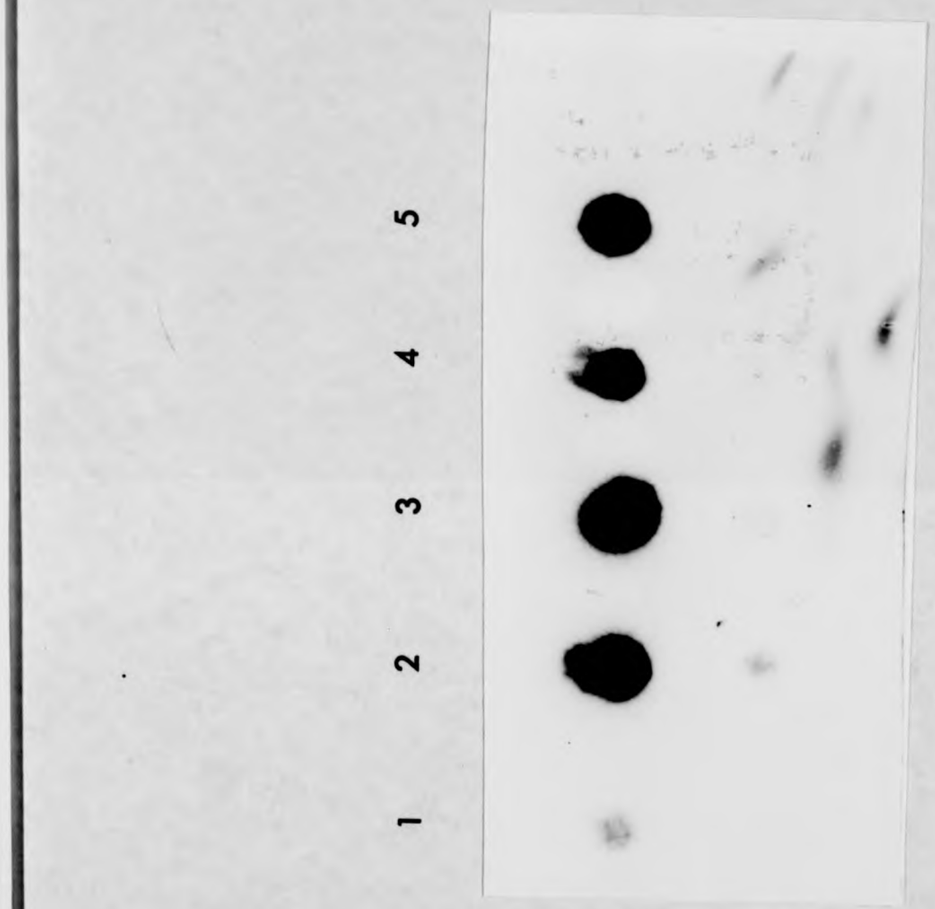
Viable protoplasts were transformed with 21ug of sterile, covalently-closed, circular npt2-encoded plasmid DNA. After 36 hours' incubation, the protoplasts and media were analysed separately for npt2 activity. Figure 4.4 shows the result of the npt2 activities in protoplast and media fractions. No endogenous npt2 activity was seen in protoplast or media fractions from untransformed protoplasts. This showed that there was no measurable background npt2 activity in tobacco protoplasts when treated in this way.

On the other hand, when protoplasts were transformed with pCaMVNEO, pPARTIAL or pFULL, high levels of npt2 activity were observed in the cell fractions. From this it was concluded that a soluble, functional npt2 enzyme was being expressed from each of the constructs.

FIGURE 4.4 **NPT2 ACTIVITIES IN CELL AND MEDIA FRACTIONS OF PROTOPLASTS, 36 HOURS AFTER TRANSFORMATION WITH pCaMVNEO, pPARTIAL and pFULL.**

The dot blot assay of McDonnell et al (1987) was used to determine the npt2 activities in cell (protoplast) and media fractions (see Methods). Under the assay conditions, npt2 activity present in a fraction catalyses the transfer of ^{32}P-labelled gamma phosphate groups from ATP to neomycin. At the end of the assay, the soluble protein was spotted onto Whatman P-81 paper and the un-incorporated ^{32}P-ATP washed away. The filter was then dried and any bound, radiolabelled neomycin was identified by autoradiography. Fractions containing npt2 activity appear as black spots.

Lane:	npt2-encoded plasmid (21ug)
1	no DNA
2	pCaMVNEO
3	pCaMVNEO
4	pPARTIAL
5	pFULL



en in any of the media fractions.
for protoplasts transformed with
ıpt2 was being expressed in the
ent in the media could only have
and cytosolic release. However,
:pected for the fusion proteins
ıRTIAL and pFULL constructs. One
ılts is that once expressed, the
he cytosol. However, the results
of the pPARTIAL construct (shown
t this fusion protein contains a
In addition, other workers have
cid residue N-terminal presequence
: in the pFULL construct) can also
f fusion proteins across the ER
1988). Therefore it is reasonable
ıts transformed with pPARTIAL or
n protein is co-translationally
mbrane.

on that the fusion proteins are
there must be another reason for
ty in the medium. One possible
gregated products are misfolded in
men and are therefore retained in
mple, association with the binding
, though unlikely, interpretation
which is normally expressed in the
f redundant targeting or retention
signal which interacts with a receptor in the secretory pathway and so retards it from bulk flow secretion. Finally, the preproricin-derived residues on the segregated fusion proteins expressed from pPARTIAL and pFULL may contain a targeting or retention signal which interacts with some factor along the

No npt2 activity was seen in any of the media fractions. This result was as expected for protoplasts transformed with pCaMVNEO. In these cells, npt2 was being expressed in the cytosol and any activity present in the media could only have arisen as a result of lysis and cytosolic release. However, this result was not as expected for the fusion proteins expressed from both the pPARTIAL and pFULL constructs. One interpretation of these results is that once expressed, the fusion proteins remain in the cytosol. However, the results from the *in vitro* expression of the pPARTIAL construct (shown in Figure 3.7) indicate that this fusion protein contains a functional signal peptide. In addition, other workers have shown that the full 35amino acid residue N-terminal presequence of preproricin (as is present in the pFULL construct) can also mediate the translocation of fusion proteins across the ER membrane (Richardson *et al*, 1988). Therefore it is reasonable to assume that in protoplasts transformed with pPARTIAL or pFULL, the expressed fusion protein is co-translationally translocated across the ER-membrane.

Based on the assumption that the fusion proteins are segregated in the ER lumen, there must be another reason for the absence of npt2 activity in the medium. One possible interpretation is that the segregated products are misfolded in the environment of the ER lumen and are therefore retained in this compartment by, for example, association with the binding protein, BiP. An alternative, though unlikely, interpretation of the results is that npt2, which is normally expressed in the cytosol, contains some form of redundant targeting or retention signal which interacts with a receptor in the secretory pathway and so retards it from bulk flow secretion. Finally, the preproricin-derived residues on the segregated fusion proteins expressed from pPARTIAL and pFULL may contain a targeting or retention signal which interacts with some factor along the

secretory pathway and prevents secretion of the proteins into the medium.

Denecke *et al* (1990) have shown that when npt2 is targeted to ER lumen in tobacco protoplasts by fusion to a signal peptide, it was efficiently secreted into the medium. From this data it can be concluded that 'wild-type' npt2 folds in the ER in such a way that there is no permanent association with BiP and that also the protein contains no retention/targeting signals, since it is secreted by bulk flow into the medium. From the data presented in the previous chapter were the pPARTIAL construct was expressed *in vitro* (see Figure 3.7), it was concluded that the ER-segregated form of the fusion protein was similar in size to wild-type npt2. If cleavage of the signal peptide occurs at the same site during expression of the pPARTIAL construct in protoplasts, then the segregated form must closely resemble wild-type npt2. This would suggest that it folds correctly in the ER lumen and is not retained or targeted away from the bulk flow secretion pathway. Therefore, it was expected that npt2 activity would be found in the medium from protoplasts transformed with pPARTIAL. The only remaining alternatives as to why this was not observed were considered to be, either the preproricin signal peptide does not function correctly in this plant expression system and therefore the fusion proteins remain localised in the cytosol, or that secreted npt2 is inactivated in the media. However, the existing data was insufficient to enable discrimination between these alternatives

In light of this apparently anomalous result with expression of the gene from pPARTIAL, it was impossible to make any further conclusions from the data obtained with the pFULL construct; such as the possibility of preproricin-borne signal-mediated retention of npt2 activity in the cells of pFULL-transformed protoplasts. A control, known to produce a secreted

protein in this system, was clearly desirable. Furthermore, it was deemed necessary to modify the protocol and approach both to test the hypothesis of npt2 inhibition in the media fractions (mentioned above) and to monitor the amount of cell damage. These are described in the section below.

4.5 PROTOCOL MODIFICATIONS AND FURTHER TRANSIENT EXPRESSION STUDIES.

4.5.1 AMENDED APPROACH.

In an effort to get a better understanding of the destination of the ricin/npt2 chimeras expressed in this transient expression system, the following modifications to the original method were applied.

(a) The plasmid, pDE307 (a gift from Jurgen Denecke: for map see Denecke *et al*, 1990), was obtained. It encodes a fusion protein between a functional ER signal peptide from the pathogenesis-related protein 1b (Cornelissen *et al*, 1986) and npt2. Denecke and co-workers showed that when this plasmid was transformed into tobacco leaf protoplasts, approximately 50% of the expressed npt2 activity was found in the media 36 hours after transformation. They concluded that the signal peptide had directed the npt2 into the lumen of the ER from where it was secreted by default into the medium. This plasmid was therefore used to provide a control for default secretion of npt2 from the ER.

(b) All protoplasts were co-transformed with pDP33GUS and the npt2-encoded plasmid of interest. The plasmid pDP33GUS (a gift from Aldo Ceriotti) contains the GUS gene located between a 35S CaMV promoter and a nopaline synthase transcription terminator sequence. Protoplasts transformed with this plasmid should therefore express GUS in the cytosol. This was used as a sensitive cytosolic marker. By analysing aliquots of the cell and media fractions for GUS activity (as well as npt2

activity), it was hoped that the amount of contamination of the media with cytosol could be determined and this in turn would assist interpretation of any npt2 activity found in the media fractions. This will be explained in more detail with actual data obtained, in Section 4.5.2 below.

(c) Following the transformation and incubation of protoplasts, the soluble proteins in the cell and media fractions were precipitated with ammonium sulphate (see Methods 2.4.7). The protein was then resuspended in enzyme reaction buffer and the salt was dialysed away prior to assaying each fraction for GUS and npt2 activity. This step was an attempt to remove any enzyme-inhibiting factors from the cell/media fractions (see Denecke *et al*, 1989: also, personal communication, Jurgen Denecke). It was hoped that this would reduce the likelihood of there being inhibition of npt2 activity in the media fractions (as was a possibility considered in the last experiment described above).

(d) Soluble protein from the leaves of a pCaMVNEO-transformed transgenic tobacco plant was used as a source of npt2. Controls could therefore be set up using appropriate amounts of this protein, diluted in cell and media buffers. These controls were processed in an identical fashion to the fractions from transformed protoplasts and were then assayed for npt2 activity. By doing this it was possible to check the relative recoveries of npt2 activity between cell and media fractions after they had been processed. In the third experiment described below, 6 milliunits of commercially-obtained GUS (Sigma) was also added to these controls before precipitation and dialysis, in order to check the relative recovery of GUS activity in the same way.

Using these modifications, three further transformation experiments were performed. The three experiments will be referred to in chronological order as Experiments 1, 2, and 3.

In each experiment, the following transformations were done:

1. 21µg pCaMVNEO
2. 21µg pDE307
3. 21µg pPARTIAL
4. 21µg pFULL
5. no DNA

In Experiment 1, protoplasts were heat-shocked at 45C for 5min immediately prior to co-transformation with 7µg pDP33GUS and the relevant npt2-encoded plasmid indicated above.

Experiment 2 was as above, but without the heat-shock step.

In Experiment 3 the protoplasts were not heat-shocked and only 1.3µg pDP33GUS was co-transformed with each of the npt2-encoded plasmids.

4.5.2 RESULTS FROM GUS ASSAYS.

After 36 hours' incubation, cell and media fractions were prepared from transformed protoplasts and aliquots assayed for GUS/npt2 activity. The results of the GUS assays are shown in Figures 4.5.2.1 and 4.5.2.2. Figure 4.5.2.1 shows how the raw data from these GUS assays was converted to give the final values for GUS activity, expressed in the preferred units of nmoles product released/min (Jefferson, 1987) and displayed graphically in Figure 4.5.2.2.

For each sample four time points were taken. The samples from the control protoplasts which had been PEG-treated in the absence of pDP33GUS (labelled 'no DNA') had very low levels of relative fluorescence and these values did not increase with time. From this it was concluded that extracts from 'no DNA' PEG-treated protoplasts contained negligible levels of GUS activity and there was also a negligible level of non-specific MUG hydrolysis under the assay conditions used.

FIGURE 4.5.21 **CALCULATION OF GUS ACTIVITY IN CELL AND MEDIA FRACTIONS.**

The graphs opposite show a plot of relative fluorescence versus time for the GUS assays performed in Experiment 3. The following calculations were performed to convert these relative fluorescence values to the prefered units for GUS activity; nmoles product/min/fraction (Jefferson, 1987). As an example, the calculation for the 'pCaMVNEO'-transformed protoplast (cell) sample is shown in **bold type.**

1. Calibrate the spectrofluorometer with a series dilution of $10^{-2}$to 10^{-6}mM MU:

[MU](mM)	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
rel. fluor.	**24.0**	**3.15**	**0.315**	**0.032**	**0**

therefore rel. fluor. [MU] between 0.032 and 3.15

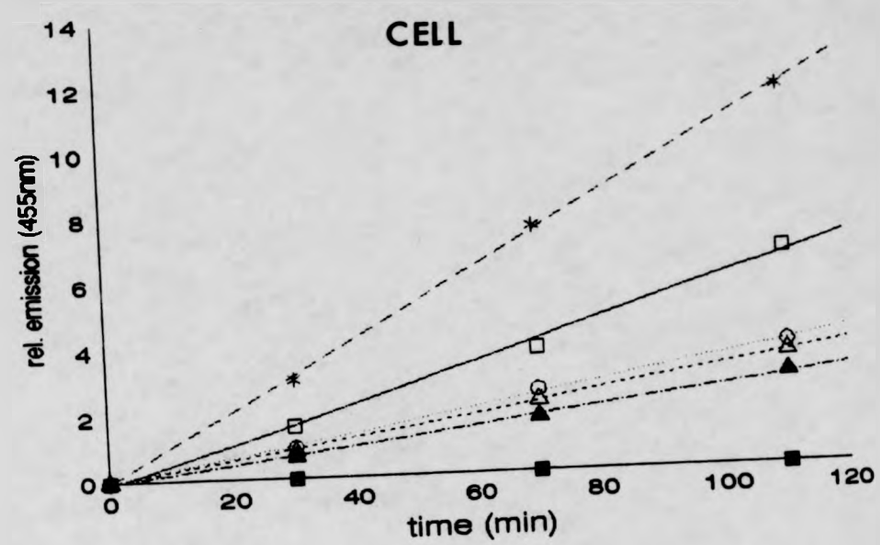
2. Read each time point of sample (diluting with 0.2M $NaHCO_3$ if necessary to get rel. fluor. betwwen 0.032 & 3.15):

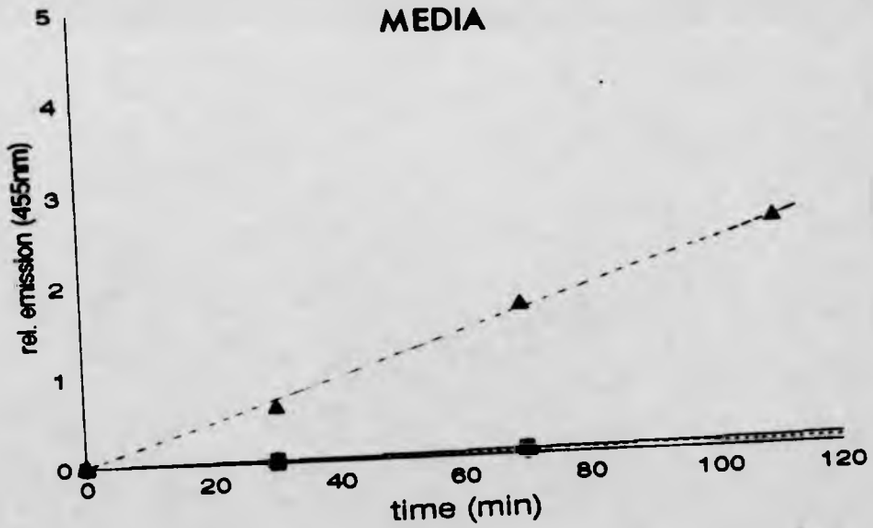
time (min)	0	30	71	110
pCaMVNEO(rel. fluor.)	**0.025**	**1.635**	**0.380***	**0.665***

(* = 10x dilution)

3. Plot graph of rel. fluor. versus time to show linear relationship (taking into account dilutions made above):
see open squares on graphs opposite.
4. Read off fluorescence at 60min:
= 3.24units [4]
5. As 1ml of 10^{-3}MU (rel. fluor. =3.15) contains 1nmole, calculate nmoles MU released per hour for sample:
[4] x 1/3.15 = 1.03nmoles MU/hour [5]
6. Only 1/4 of the total reaction was stopped at each time point (100ul from 400ul):
[5] x 4 = 4.11nmoles MU/hour/assay [6]
7. Correct for activity per minute:
[6] x 1/60 = 0.07nmoles MU/min/assay [7]
8. Correct for total activity in cell/media:
[7] x 416(volume of total soluble protein from cell fraction) / 65(volume of above used in GUS assay)

= 0.44nmoles MU produced/min/cell fraction





KEY: □ pCaMVNEO ○ pPARTIAL ■ no DNA
△ pDE307 * pFULL ▲ npt2/GUS control

FIGURE 4.5.22 GUS ACTIVITIES IN CELL AND MEDIA FRACTIONS OF PROTOPLASTS, 36 HOURS AFTER TRANSFORMATION WITH pDP33GUS.

3×10^5 viable protoplasts were transformed with the plasmids indicated below and the genes expressed for 36 hours by incubation of the protoplasts at 26C in K3 medium in the dark. Protoplast (cell) and media fractions were then prepared as described in Methods (2.4.7) and assayed for GUS activity according to Jefferson (1987). This assay follows the GUS-catalysed release of methylumbelliferone (MU) from 4-methyl-umbelliferyl glucuronide (MUG) spectrofluorometrically. The amount of MU present at different time points was determined by diluting aliquots of the reaction in 9 volumes 0.2M $NaHCO_3$, measurement of the relative fluorescence of this solution (at 455nm when excited at 365nm) and reference to a standard curve of relative fluorescence versus [MU] (not shown). These values were corrected for assay volume versus total cell/media volume and the GUS activities were then expressed as nmoles MU released/min/total fraction. An example of such a calculation is shown in Figure 4.5.21

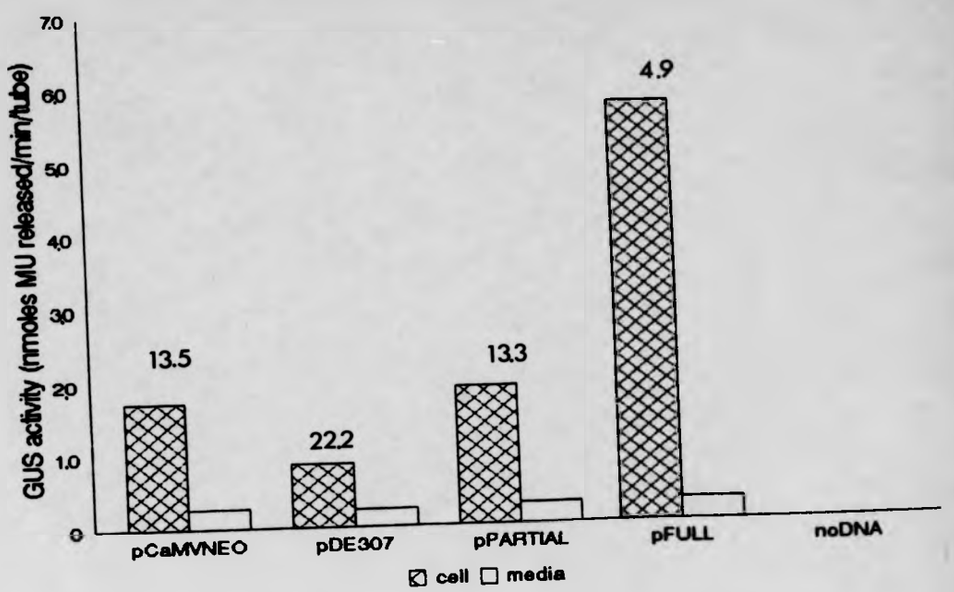
Graph A: Protoplasts were heat-shocked for 5min at 45C immediately before transformation with 7.0µg pDP33GUS.
Graph B: No heat shock: transformation with 7.0µg pDP33GUS.
Graph C: No heat shock: transformation with 1.3µg pDP33GUS.

In each case (A,B,C) protoplasts were co-transformed with 21.0µg of the npt2-encoded plasmid indicated along the X-axis. "No DNA" indicates that protoplasts were PEG-treated in the absence of any npt2/GUS-encoded plasmid DNA. "npt2" (GRAPH C only) indicates control tubes that contained 0.1mg soluble protein from leaves of transgenic tobacco and 6mU of commercially-obtained GUS.
Values above each bar cluster represent % cell lysis;

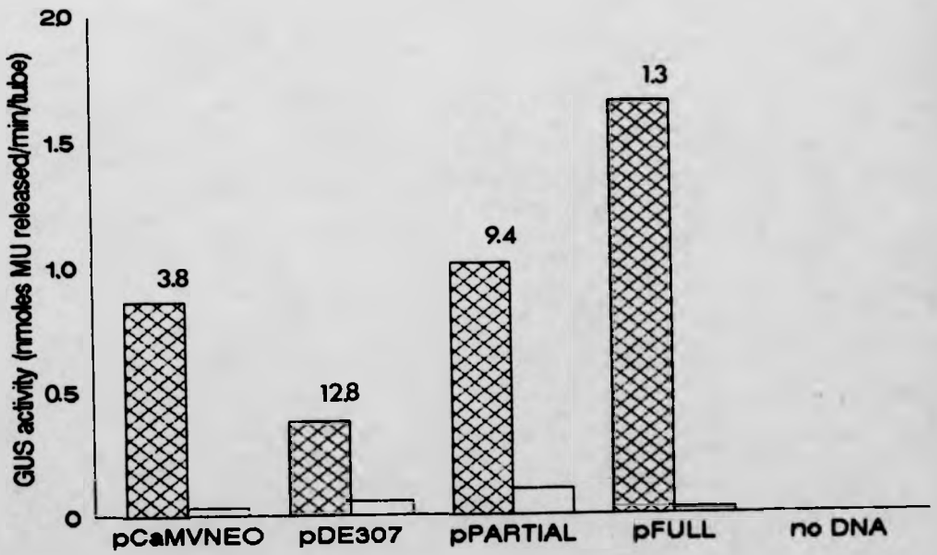
i.e. $\dfrac{\text{GUS activity in media}}{\text{Total GUS activity in cell + media}} \times 100$

Graph D: Comparison of the total GUS activities (cell+media) from the three experiments (expressed as nmoles MU released/min/ug pDP33GUS)

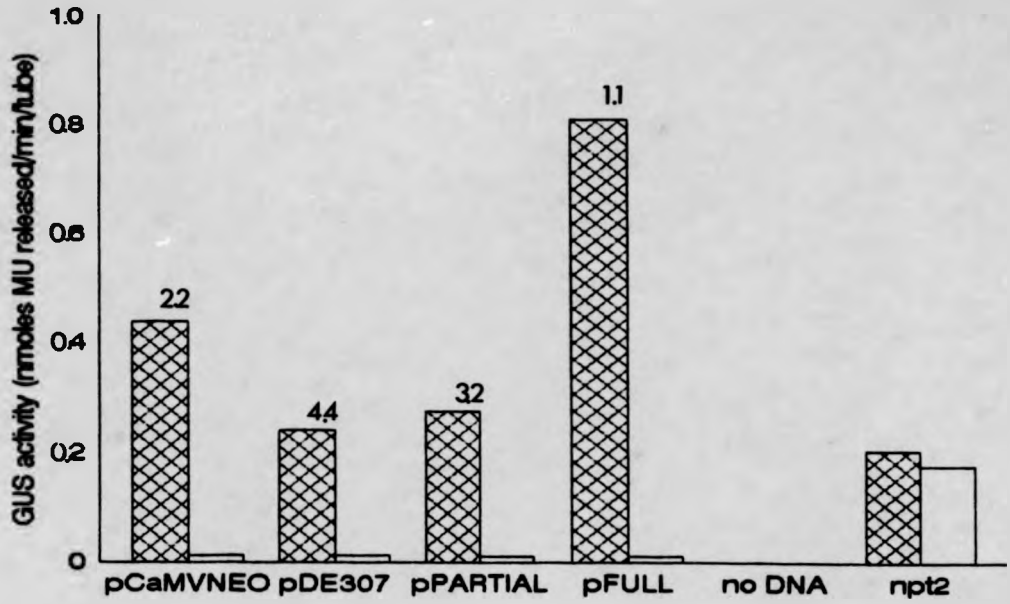
A



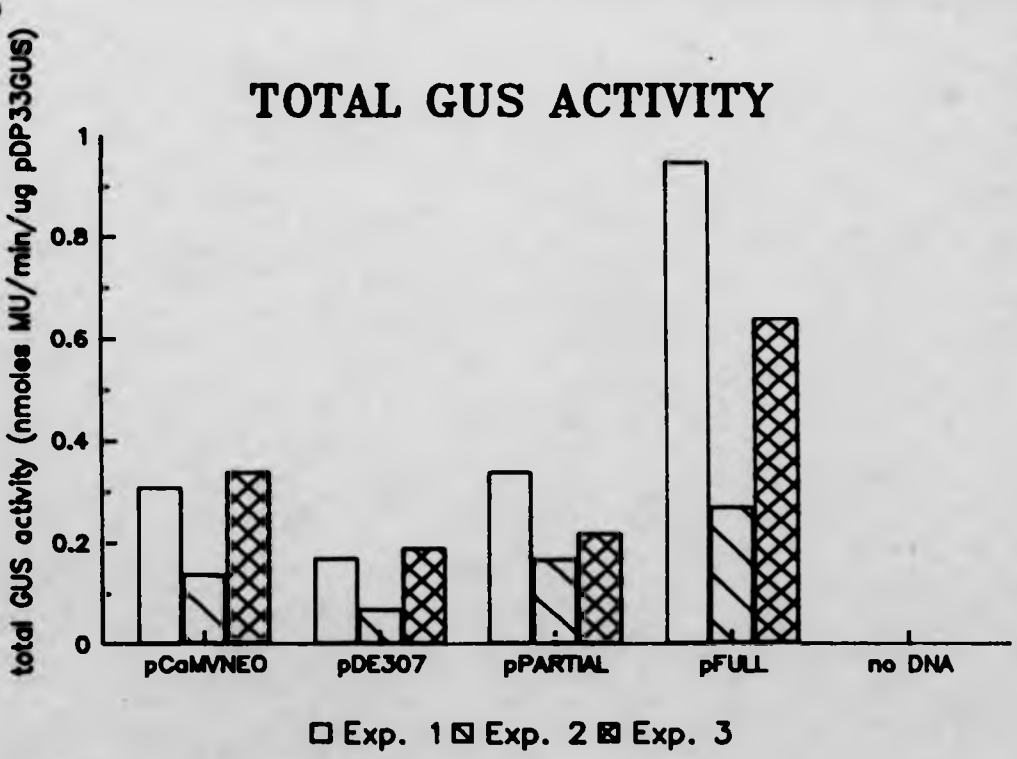
B



C


GUS activity (nmoles MU released/min/tube)
1.0
0.8
0.6
0.4
0.2
0
2.2
4.4
3.2
1.1
pCaMVNEO
pDE307
pPARTIAL
pFULL
no DNA
npt2

D


TOTAL GUS ACTIVITY
total GUS activity (nmoles MU/min/ug pDP33GUS)
1
0.8
0.6
0.4
0.2
0
pCaMVNEO
pDE307
pPARTIAL
pFULL
no DNA
□ Exp. 1 ▨ Exp. 2 ▩ Exp. 3

However, with all cell extracts from pDP33GUS-transformed protoplasts a detectable level of fluorescence was observed. By plotting the relative fluorescence of these time points versus time (see graphs in Figure 4.5.2.1) it was also possible to confirm that the observed fluorescence in extracts from pDP33GUS-transformed protoplasts increased linearly with time during the assay. From these results it was concluded that the conversion of MUG to MU, in samples from pDP33GUS-transformed protoplasts, was GUS-catalysed.

Jefferson (1987) has reported that the PEG treatment of protoplasts can sometimes result in the production of phenolic compounds which give a high background fluorescence in the GUS assay. This problem was not encountered in these experiments, however, and, as stated in the paragraph above, by following a time course it was possible to distinguish between background fluorescence and GUS-catalysed hydrolysis of the substrate, MUG.

Figure 4.5.2.2 shows the processed results for the GUS assays from all three experiments expressed in graphical form. Although the values for GUS activity shown in these graphs may appear low, they are deceptive as the units used mask the high sensitivity of the assay. For example, the background GUS activity in the 'no DNA' controls was calculated to be less than 1×10^{-4}nmoles MU/min/tube, whereas even the lowest value for GUS activity recorded in pDP33-transformed protoplasts (see Figure 4.5.2.2, graph B) was 4000 times greater than this value (0.4nmoles MU/min/tube). It was therefore concluded that, in every case, transformation of the protoplast samples had been successful and each was expressing significant levels of functional GUS enzyme.

These results shown in Figure 4.5.2.2 also revealed that within each experiment the GUS activity varied between samples, despite the fact that each sample had been transformed with an

equal amount of pDP33GUS. Furthermore, these variations were not random, as is shown in Graph D, where the total GUS activities (cell+media) were standardised to take into account the fact that each protoplast sample in Experiment 3 was transformed with 1.3µg of pDP33GUS (as opposed to 7µg in Experiments 1 & 2). This graph shows that in each experiment the order of GUS activity was as follows:

GUS activity pDE307 < pCaMVNEO/pPARTIAL < pFULL.

As the only variable factor between transformations was the npt2 plasmid with which pDP33GUS was co-transformed, it must be concluded that the npt2 plasmid has some effect on the efficiency of transformation or level of expression of the pDP33GUS plasmid. Further discussion of this observation must be left until after the results of the npt2 assays have been presented below.

As was discussed in Section 4.5.1, the reason for co-transformation of the protoplasts with pDP33GUS in addition to the npt2 plasmid of interest was to use the former as a cytosolic marker in the hope of calculating the level of contamination of the media fractions with cytosol. It was hoped that this, in turn, would make the results from the npt2 assay more meaningful. In graphs A-C of Figure 4.5.2.2 the level of cytosolic contamination, expressed as percentage cell lysis, is shown above the bars. The justification for using this value as a valid indication of cell lysis, was that all GUS activity observed in the media fractions could only have originated from within protoplasts transformed with pDP33GUS and expressing the enzyme.

Using this data to compare the results obtained in Experiments 1 and 2, reveals that although the heat-shock treatment, performed prior to transformation in Experiment 1 only, increased the total average level of expression of GUS in protoplast samples from the first experiment (3.1 versus

1.4nmoles MU/min/tube respectively), this treatment also increased the average amount of cell lysis (13.5% versus 6.8% respectively). The increased transformation observed in heat-shocked versus non-heat-shocked protoplasts has been reported before (Saul et al, 1988). However, in this instance it is desirable to minimise cell lysis in order to maximise the relevance of the results from the npt2 assays. Therefore, it was decided to also omit the heat-shock step in Experiment 3.

As a final check to show that there was a comparable recovery of GUS activity from cell and media fractions after processing the samples, two controls (see lanes labelled 'npt2' in Figure 4.5.2.1 and Graph C of Figure 4.5.2.2) were set up which consisted of 100µg of soluble leaf protein in the relevant extraction buffer and 'spiked' with 6mU of commercially-available GUS (Sigma). These controls were processed at the same time and in an identical way to all other cell and media fractions. Graph C in Figure 4.5.2.2 shows that a very similar activity was recovered. When these values were compared to a time course, also set up using 6mU of GUS activity but which had not been processed (0.24nmol MU/min/tube: not shown on graph), it was calculated that 83% of the original GUS activity had been recovered after cell fraction processing and 75% after media processing. As only one set of controls were done, however, these values were not thought to be sufficiently statistically valid to allow empirical adjustment of the values for the cell and media fractions of pDP33GUS-transformed samples (i.e. to take into account that only 83%(cell) or 75%(media) of the original GUS activity in fractions was detected after processing. On a more positive note, what these results do show is that the huge differences observed between cell and media fractions were not simply artefacts of the extraction and processing steps but

were due to real differences in GUS activity between the cell and media fractions.

4.5.3 RESULTS FROM NPT2 ASSAYS

The fractions from the three experiments were then assayed for npt2 activity using the same dot blot method of McDonnell et al (1987) as was used in the initial transformation experiment (Section 4.4). The results of Experiments 1 and 2 are shown in Figure 4.5.3.1.

The controls used to compare recovery of npt2 activity between cell and media processing (see C5-C8), where soluble protein from leaves of pCaMVNEO-transformed transgenic tobacco was extracted and processed in a similar way to either cell or media protoplast fractions, revealed that npt2 enzyme activity could now be recovered from both fractions. This meant that by following the amendments to the protocol as described in Section 4.5.1 above, it should now be possible to detect any npt2 activity in the media fractions of protoplasts transformed with npt2 constructs.

In the cell fractions of protoplasts PEG-treated in the absence of any plasmid DNA ('no DNA': see lanes A5 & B2) spots of diffuse radioactivity, surrounded by a darker 'halo', were observed. By comparison to the results from the previous transient expression experiment (Figure 4.4), where npt2-specific activity in cell fractions appeared as dark spots centred around the point of sample application to the filter (^{32}P-neomycin phosphate only diffuses slowly before becoming bound), it was concluded that the diffuse spots seen in these controls was probably due to non-neomycin-bound ^{32}P-phosphate (by remaining unassociated from neomycin the radiolabel can diffuse further from the point of application before becoming bound to the filter). For some inexplicable reason this unbound phosphate was not not washed off the filter completely during

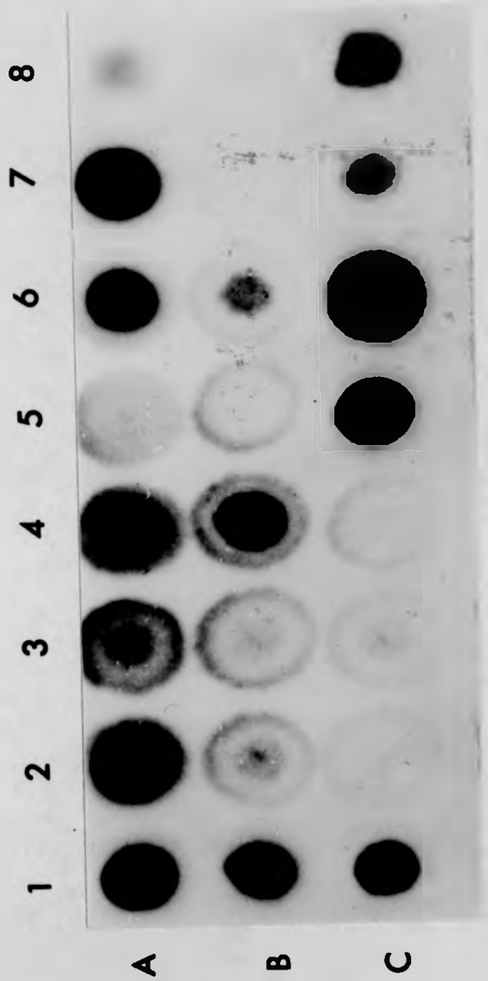
FIGURE 4.5.31 **NPT2 ACTIVITY IN CELL AND MEDIA FRACTIONS OF PROTOPLASTS 36 HOURS AFTER TRANSFORMATION: EXPERIMENTS 1 & 2.**

The same assaying procedure was used as described in Figure 4.4. Protoplasts were heat-shocked for 5 min. at 45 C immediately before transformation in Experiment 1. No heat-shock step was performed in Experiment 2.

EXPERIMENT 1 (LANE:)	npt2-encoded plasmid (21ug)	EXPERIMENT 2 (LANE:)
A1	pCaMVNEO - cell fraction	A6
A2	pDE307 - " "	A7
A3	pPARTIAL - " "	A8
A4	pFULL - " "	B1
A5	no DNA - " "	B2
B3	pCaMVNEO - media	B8
B4	pDE307 - "	C1
B5	pPARTIAL - "	C2
B6	pFULL - "	C3
B7	no DNA - "	C4
C5	npt2 control - cell fraction	C7
C6	" " - media	C8

0.1mg of soluble protein from the leaves of pCaMVNEO-transformed transgenic tobacco was used as a source of enzyme activity for the npt2 controls (C5-C8).

All protoplast samples were co-transformed with 7.0ug pDP33GUS.



the rinsing steps and so was detected on the autoradiograph. In addition to this non-specific background radioactivity, the 'no DNA' cell fraction (Lane B2) showed a low level of npt2-like activity which suggested that there might be a low level of endogenous activity in non-transformed protoplasts.

In cell fractions of protoplasts transformed with the npt2-encoded plasmids, pCaMVNEO, pDE307 and pFULL on the other hand, dark spots of npt2-specific activity were observed (Experiment 1, A1-A4: Experiment 2, A6-B1). From this it was concluded that all protoplasts transformed with these plasmids were expressing a functional npt2 enzyme. However, in the cell fractions of protoplasts transformed with pPARTIAL, only a very low level of activity above background was visible (Lanes A3 and A8). This result was in direct contradiction with the result from the previous transient expression experiment where a high level of npt2 activity had been observed in the cell fraction of protoplasts transformed with the same preparation of pPARTIAL DNA (compare Lanes A3 & A8 with Lane 4 in Figure 4.4).

No npt2-specific activity was observed in media fractions of the 'no DNA' controls. In the media fractions of protoplasts transformed with pCaMVNEO (lanes B3 & B8), no significant activity was seen. However, in the protoplasts transformed with pDE307, npt2 activity was observed in the medium (Lanes B4 & C1). As the there were negligible levels of npt2 in the cell fractions of pPARTIAL-transformed protoplast it was of no surprise that no npt2 activity was detected in the media fractions either. Some npt2 activity was observed in the media fraction of pFULL-transformed protoplasts from Experiment 1 (Lane B6) but not in the corresponding fraction from Experiment 2 (Lane C2).

Taken in isolation, these results suggested that in the protoplasts transformed with pDE307 and, to a lesser extent

those transformed with pFULL, the expressed npt2 derivative was being secreted into the media. However, there is another possible reason for the detection of npt2 activity in the media. In these cases the media may have been contaminated with cytosol containing npt2 activity. To check whether this was likely, it was possible to refer to the results of the relevant GUS assay described above. In Experiment 1 it was estimated from the GUS assay that there had been 22% protoplast lysis, which might explain why a significant proportion of the npt2 activity was found in the media. However, the pCaMVNEO-transformed protoplasts were estimated to have suffered 14% lysis in the same experiment and yet there was no significant npt2 activity seen in the media fraction (Lane B3). These results were unclear and therefore it was decided in Experiment 3 to omit the heat-shock treatment prior to transformation and to exercise great care when separating the cell and media fractions in order to minimise the contamination

The results from the third experiment are shown in Figure 4.5.3.2. The npt2 controls (Lane 6) showed that the relative recovery of npt2 activity after cell and media processing was similar. The fractions from protoplasts treated with PEG in the absence of plasmid DNA showed no significant endogenous npt2 activity (Lane 5: no DNA). Conversely, npt2 activity was observed in the cell fractions from protoplasts transformed with npt2-encoded plasmids (Lanes 1-4). Although there was a large amount of npt2 activity in cell fractions of protoplasts transformed with pCaMVNEO, no activity was observed in the media fraction above background. However, in protoplasts transformed with pDE307, 12.5% of the activity was observed in the media (as estimated by densitometry). As in Experiments 1 and 2, there was only a low level of npt2 activity in the cell fraction of protoplasts transformed with pPARTIAL. No npt2 activity was observed in the media fraction from these

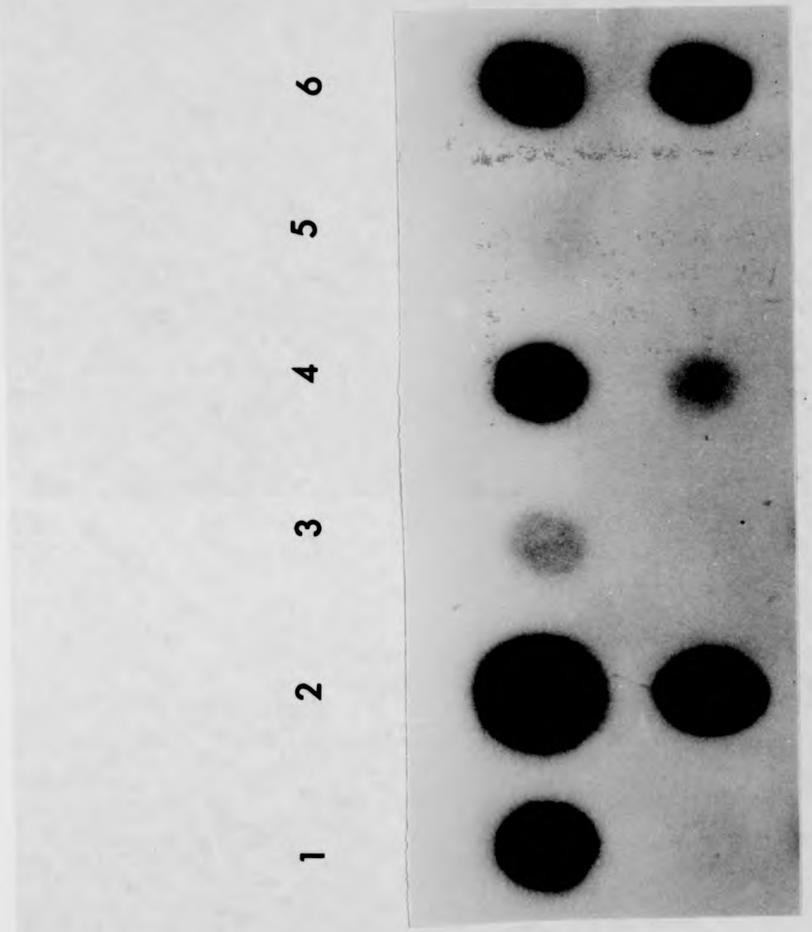
FIGURE 4.5.32 **NPT2 ACTIVITIES IN CELL AND MEDIA FRACTIONS OF PROTOLASTS, 36 HOURS AFTER TRANSFORMATION: EXPERIMENT 3.**

The same assaying procedure was used as described in Figure 4.4. Protoplasts were NOT heat-shocked before transformation.

LANE:	npt2-encoded plasmid (21µg)
1	pCaMVNEO
2	pDE307
3	pPARTIAL
4	pFULL
5	no DNA
6	npt2 control

0.1mg of soluble protein from the leaves of pCaMVNEO-transformed transgenic tobacco was used as a source of enzyme activity for the npt2 controls (C5-C8).

All protoplast samples were also co-transformed with 1.3µg pDP33GUS.



protoplasts. From the autoradiograph in Figure 4.5.32 it was also estimated by densitometry that 7% of the npt2 activity in protoplasts transformed with pFULL was in the media fraction.

The results of the GUS assays on cell and media fractions in Experiment 3 showed that in protoplasts transformed with pDE307 or pFULL, 4.4% and 1.1% cell lysis occurred respectively. This meant that the cytosolic contamination of media could not account for the levels of npt2 activity found in these two fractions. It was therefore concluded that the npt2 activity in the media fractions of protoplasts transformed with pDE307 and pFULL was present as a result of secretion of the enzyme from the cell.

As mentioned above, the amount of GUS activity expressed in the protoplasts transformed in all three experiments was apparently linked to the npt2-encoded plasmid with which pDP33GUS was co-transformed. The possibility that contaminating factors present in the DNA preparations of the npt2 plasmids affected the expression of the pDP33GUS plasmid was thought unlikely as pDE307-transformed protoplasts that consistently showed the lowest GUS activities, expressed the highest levels of npt2 activity. If contamination was the determining factor with regards to the level of expression, then one would expect low GUS activities to be mirrored by low npt2 levels and vice versa. Furthermore, all plasmids were purified on caesium chloride gradients and checked spectrophotometrically for protein contamination. Another possible explanation for this discrepency would be competition between the GUS-encoding gene and npt2-encoding gene for the same cellular RNA and protein expression machinery. Tobacco leaf mesophyll cells are primarily photosynthetic organs. This means that, unlike other heterologous expression systems, they do not have high turnover rates of protein synthesis. It is conceivable therefore that in protoplasts co-expressing two genes with a highly efficient

constitutive 35S-CaMV promoter, there would be competition for the same cell transcription and translation machinery. If this were the case then one would expect to see an inversely proportional relationship between the levels of GUS and npt2 in each protoplast sample. This was observed for the pDE307-transformed protoplasts which consistently had the lowest GUS activities but the highest npt2 activities of any samples tested. Conversely, the pFULL-transformed protoplasts had high GUS activities but low npt2 activities. However, the extracts from pPARTIAL-transformed protoplasts did not fit this pattern as they had low GUS and npt2 levels. Therefore, there appeared to be no single satisfactory explanation for the patterns in GUS activities observed.

4.6 DISCUSSION.

Stable protoplasts were produced using the methods described. The PEG-mediated method for protoplast transformation described by Saul *et al* (1989) enabled the transformation of protoplasts with npt2 and GUS-encoded plasmids. Protoplasts transformed with these plasmids and incubated for 36 hours, expressed detectable levels of the plasmid-encoded enzymes above background. Although heat-shock treatment of the protoplasts prior to transformation increased the efficiency of transformation and lead to a higher level of cytosolic expression, it also lead to a higher degree of cell lysis which was undesirable.

Both constructs, pPARTIAL and pFULL, encoded functional npt2 enzymes. The enzyme can therefore tolerate the preproricin extensions fused to its N-terminus without loss of activity upon removal of these extensions (in part or wholly). However, in the space of time between the initial experiment and the next set of three experiments, the transformation efficiency of the preparation of the pPARTIAL construct dropped

dramatically. The reason for this anomaly is unknown as both the original preparations of pPARTIAL and pFULL constructs were stored together at -20C.

Partial purification of the media fractions, by ammonium sulphate precipitation, was necessary in order to detect npt2 activity present in these fractions. It seems therefore that there are factors in the crude media which inhibit npt2 activity. This observation is in agreement with the results from other groups who have used npt2 for studying the secretory pathway in transient expression systems (Denecke, personal communication).

Co-transformation of pDP33GUS with the different npt2-encoded plasmids resulted in a similar expression pattern of the cytosolic marker, GUS, in each of the three experiments. However, there was no obvious correlation between the level of expression of the GUS gene and the level of expression of the npt2 gene; i.e. high GUS levels were not consistently balanced by low npt2 levels (and vice versa). Therefore this variation in the level of expression of the GUS gene when co-transformed with different plasmids did not seem to be due to a competition for the cellular translation machinery within the same cells. It was also not thought to be solely as a result of inhibitory factors being present within the npt2-encoded plasmid preparations. No single reason for this observation was obvious to this author.

The results of the GUS and npt2 enzyme assays from Experiment 3 would suggest that the npt2 expressed in protoplasts transformed with pDE307 and pFULL was being exported to the media at a similar rate. The appearance of npt2 activity in the media did not co-incide with a significant contamination of the media with the co-expressed cytosolic marker, GUS. It is therefore, reasonable to suggest that this export was occurring from intact protoplasts. As both genes are

known to encode proteins with functional ER-signal peptides, it is likely that this enzyme activity is being exported from the ER lumen by the default pathway. This data would suggest therefore that the 35 residue N-prepropeptide of preproricin does not contain a contiguous stretch of amino acids representing a targeting or retention signal.

Unfortunately, the limited time of study in Milan did not allow time for more experiments.

Recently, a number of investigations into vacuolar/protein body targeting in plants have been reported. As has been described earlier, Chrispeels and co-workers have decided to use a yeast expression system to study the vacuolar targeting of the bean lectin, PHA. They showed that when PHA was expressed in transgenic yeast it was correctly targeted to the vacuole (Tague and Chrispeels, 1987). This suggested that the mechanisms and signals involved in vacuolar/protein body targeting were similar between yeast and higher plants. They then constructed hybrids between N-terminal regions of PHA and the yeast secretory enzyme, invertase. When expressed in yeast they found that the hybrid was targeted to the vacuole (Tague et al, 1990). They concluded that the N-terminal region of PHA contained a vacuolar targeting signal which re-directed the invertase in the hybrid protein to the vacuole. They identified a region in the N-terminal region of PHA, LQR, which resembles the yeast vacuolar targeting signal, QRPL, and they suggested that this may represent a plant vacuolar targeting signal (Chrispeels and Tague, 1990). In order to test this hypothesis they expressed the PHA/invertase hybrid in transgenic tobacco, but found that it was secreted (Chrispeels, 1991). Therefore, it seems that the components of the vacuolar targeting mechanism in yeast and higher plants are not compatible. Studying plant vacuolar proteins in yeast systems may reveal more about the yeast vacuolar targeting machinery

than the signals in the plant protein which are involved in the targeting of the protein to the plant vacuole.

On the other hand Raikhel and co-workers have expressed barley lectin precursor in transgenic tobacco and have shown that it was correctly targeted to the vacuoles (Wilkins *et al*, 1990). The lectin precursor contains a negatively-charged, 14 amino acid residue C-terminal propeptide which is co-translationally N-glycosylated. After comparison to the targeting of animal proteins to the lysosome, which involves a mannose-6-phosphate signal/receptor pathway, they hypothesised that this propeptide-borne glycan may be involved in the targeting of barley lectin to the vacuole. They used site-directed mutagenesis to remove the N-glycosylation signal in the propeptide and expressed this mutant in transgenic tobacco. However, this was glycosylation-minus mutant was also targeted to the vacuole (Wilkins *et al*, 1990). On the other hand, when this propeptide was removed completely the propeptide-minus mutant was secreted from the tobacco cells (Bednarek *et al*, 1990). It therefore seems that the C-terminal propeptide of barley lectin is involved in the targeting of the protein to the vacuole, but the glycan moiety does not form part of this signal.

More recently, Nakamura and co-workers have shown that the N-terminal propeptide of the sporamin precursor, the vacuolar-localised storage protein from the tuberous roots of the sweet potato, is involved in vacuolar targeting. When they expressed the cDNA for the sporamin precursor (Hattori *et al*, 1985) in transgenic tobacco, they found that it was correctly targeted to the vacuole (Matsuoka *et al*, 1990). However, when they used site-directed mutagenesis to remove a highly-charged 16 amino acid residue N-terminal region, identified by *in vitro* studies as being a propeptide (Hattori *et al*, 1987), and expressed it in suspension-cultured tobacco, they found that

this propeptide-minus mutant was secreted (Matsuoka and Nakamura, 1991). One observation made by Matsuoka et al (1990) which suggests that there are limits to the use of transgenic tobacco expression for the study of targeting and processing of foreign plant proteins, was the proteolytic processing of the wild-type sporamin precursor in the vacuoles of transgenic tobacco to a nearly-mature size which was longer than the sporamin isolated from sweet potato by three residues.

Thus the groups of Raikhel and Nakamura have shown an involvement of propeptides in vacuolar targeting. However, they have yet to show that these regions of vacuolar protein precursors contain all the information necessary for such targeting. Work is now under way using recombinant DNA techniques and studies in transgenic tobacco to determine whether the propeptides of these proteins can act autonomously and mediate the transport of normally non-vacuolar proteins to the vacuole.

A detailed computer analysis of plant storage protein sequences has recently identified regions of homology which may be involved in vacuolar/protein body targeting (Sebastiani et al, 1991). These regions of homology lie within the mature protein sequence and it suggested that different regions may interact forming part of a 3-D signal patch, only when the protein has folded into its tertiary or quaternary structure. It remains to be proved whether or not these regions represent the elusive signal.

CHAPTER 5

BIOLOGICAL ACTIVITIES OF RECOMBINANT PRORICIN.

CONTENTS:

5.1 INTRODUCTION.

It has long been known that ricin enzymatically inhibits protein synthesis (reviewed by Olsnes and Pihl, 1982). The inhibitory effect is attributable to the A subunit, or effectomer, whose structure and mode of action seems to resemble other proteins that have been collectively termed ribosome inactivating proteins (RIPs). Cell-free translation studies have been used in the past to study the kinetics of RIP action. For example, Olsnes *et al* (1976) compared the polyphenylalanine-synthesising activity of untreated and ricin-treated rabbit reticulocyte ribosomes in a cell-free lysate. They performed competition studies with elongation factor preparations and calculated the affinity constant (Km) and rate of reaction for the A chain action on the ribosomes. However, although ricin strongly inhibited the translating ability of the reticulocyte ribosomes, the significance of their conclusions as to the true kinetics of ricin inhibition is questionable since they used crude elongation factor preparations and artificially high (non-physiological) magnesium salt concentrations in their assays.

Using the same cell-free translation system, Olsnes *et al* (1976) also showed that non-reduced 'whole' ricin had relatively no inhibitory effect on the translating-ability of ribosomes. They concluded that the close association of the B chain when disulphide bonded with the A chain either blocked the active site or maintained the A chain in an altered (inactive) conformation. However, there have been no studies on the precursor forms of ricin.

More recently, the precise molecular mechanism of A chain action on eukaryotic ribosomes has been elucidated. Endo and Tsurugi (1987) showed that the A chain is an RNA-specific N-glycosidase. Furthermore, Endo *et al* (1987) showed that other RIPs share the same mode of action; that is, they depurinate a

specific adenine residue in eukaryotic 26S or 28S ribosomal RNA (rRNA). This particular adenine (A^{4324} in rat 28S rRNA) lies in a highly-conserved loop in the 26S or 28S rRNA component of the 60S ribosomal subunit, which may explain why ricin and other RIPs are active against all eukaryotic ribosomes. The ricin-catalysed depurination at this site makes the phosphodiester bond susceptible to mild base hydrolysis and in this way Endo et al (1987) used acetic aniline to design a sensitive assay for following the depurination of eukaryotic ribosomes. Osborn and Hartley (1990) have since used this assay to show that both translation initiation and elongation steps are inhibited by the action of A chain on ribosomes.

The data of Olsnes et al (1976) discussed above would suggest that proricin should have no inhibitory effect on eukaryotic ribosomes as there is a close association between A and B chains. However, as was discussed above, the assay system used by Olsnes and co-workers is not necessarily the best way to study the inhibitory effects of RIPs. The aniline assay now provides a sensitive alternative since it allows ricin catalysed inactivation of ribosomes to be done under physiological conditions (e.g. in an in vitro translation) before analysing the modification of the rRNA via aniline hydrolysis post-catalysis.

5.2 AIMS AND APPROACH.

The aim was to determine if recombinant proricin possessed any ribosome inactivating ability. The most direct way to study this would be to translate a proricin clone in vitro in a rabbit reticulocyte lysate and subsequently to examine the extracted rRNA for proricin-catalysed depurination. This approach has been used by May et al (1989) who showed that recombinant ricin A chain synthesised in rabbit reticulocyte lysate is biologically active on the 28S rRNA. This suggests

that in this *in vitro* expression system the A chain can fold into an active conformation. However, initial attempts to translate transcripts of proricin in reticulocyte lysates appeared to fail (M.May; unpublished data). Very little protein was synthesised and the ribosomes appeared modified, suggesting that proricin possessed RNA-specific N-glycosidase activity. However, on long exposures of the protein gels it was found that there were many labelled protein bands below the full-length proricin band. These could have arisen from premature termination events during translation of the proricin transcripts or from post-translational proteolysis of the full-length proricin polypeptide. In either case, it was thought that some of these bands may possess ribosome inactivating properties. For this reason it was not possible to inequivocally attribute the ribosome-inactivating activity observed to the full-length proricin polypeptide. Therefore, another system for the expression of recombinant proricin was necessary.

The expression of genes by injection of transcripts into *Xenopus* oocytes provides a system for recombinant proteins, synthesised with signal peptides, to be segregated in the lumen of the ER. In the context of this study, should proricin contain any ribosome-inactivating activity, the sequestering of the protein in the ER lumen would segregate it from the ribosome population and so enable further translation of proricin transcripts. A further advantage of ER-segregation is that the proricin will be co-translationally N-glycosylated and disulphide bonded. Richardson *et al* (1988) have demonstrated that ricin B chain synthesised and segregated in this system is stable and possesses lectin activity. Therefore, it will be possible to determine whether the recombinant proricin made in this system possesses any lectin activity.

In this work the oocyte injections were performed by

Professor A.Colman (Birmingham): the protein fractionation and SDS-PAGE analysis was performed by Dr. P.Richardson (Warwick). The results from this study have been published (Richardson et al, 1989).

5.3 RESULTS

5.3.1 SOURCE OF RECOMBINANT PRORICIN.

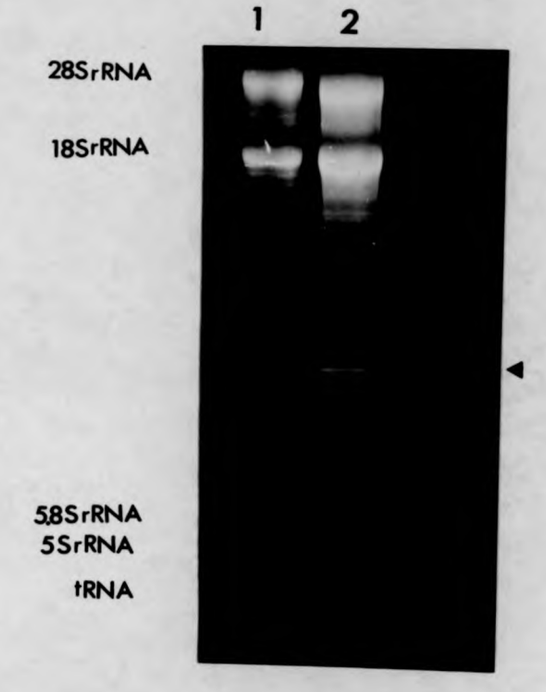
The complete preproricin clone, ppRCL617 (see Appendix B), was available at Warwick in the in vitro expression vector, pGEM1. Preproricin transcripts expressed from this construct were injected into oocytes (30-50ng mRNA per oocyte) and newly synthesised proteins were labelled by incubation of the oocytes in a medium containing ^{35}S-methionine (as described by Richardson et al, 1988). A crude homogenate from these oocytes was then used as the source of recombinant proricin.

5.3.2 ANILINE ASSAY.

Protein from the oocytes was extracted in 1x oocyte homogenisation buffer. It was therefore necessary to confirm that depurination of ribosomes was not inhibited in this buffer. The acetic aniline reagent was first prepared as described in Chapter 2 and this was then used to treat extracted rabbit reticulocyte rRNA after pre-incubation of the ribosomes in 1x homogenisation buffer in the absence or presence of purified recombinant ricin A chain. The results are shown in Figure 5.3.2. An extra band was detected in the RNA sample where ribosomes had been preincubated in the presence of A chain. This represents the diagnostic 390 nucleotide fragment cleaved from the modified 28S RNA by aniline. It was concluded that the RNA-specific N-glycosidase activity of A chain was not inhibited in the oocyte homogenisation buffer and therefore

FIGURE 5.3.2 **RICIN-CATALYSED DEPURINATION OF RABBIT RIBOSOMES IN OOCYTE HOMOGENISATION BUFFER.**

Denaturing gel electrophoresis of 4µg of aniline-treated rRNA, extracted from 30µl of rabbit reticulocyte lysate (Promega) that had been pre-incubated, at 37C for 30min, in 1x oocyte homogenisation buffer in the absence (Lane 1) or presence (Lane 2) of 10^{-8}M purified recombinant ricin A chain. Note the extra band (arrowed) in the track containing A chain-treated rRNA. This represents the 390 nucleotide fragment, released by aniline hydrolysis from the 28S rRNA component, which is diagnostic of ricin-catalysed depurination. RNA bands were stained in ethidium bromide solution and visualised on a UV-transilluminator.



this assay could be used to study the activity of recombinant proricin made in oocytes.

5.3.3 RESULTS FROM THE FIRST EXPERIMENT.

Results from immunoprecipitation and SDS-PAGE analysis of the crude oocyte homogenate revealed that expression of preproricin in oocytes produced a disulphide bonded, N-glycosylated proricin product of the expected size (see Figure 5.3.3 A). However, a small population of immunoreactive bands of lower molecular weight were observed under reducing conditions (Lane 3). This suggested that a proportion of the proricin synthesised was either prematurely terminated (leading to the release of partial-length proricin species into the ER-lumen) or, that a proportion of the full-length proricin polypeptide was somehow endoproteolytically cleaved in the oocyte. It was likely that some of these lower molecular weight species would have RNA N-glycosidase activity and therefore it was decided to pass the crude homogenate down a G-75 Sephadex column under reducing conditions in an attempt to separate the full-length proricin from these other species before assaying the oocyte homogenate samples for depurination activity.

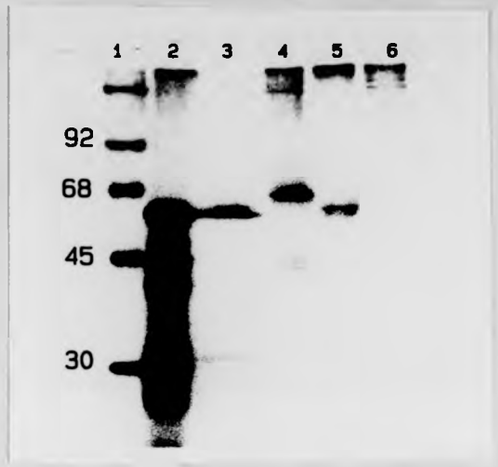
Figure 5.3.3 B shows the result when 36 x 1ml fractions eluted from a G75 Sephadex column were examined for depurination activity. Two peaks of depurination activity were seen, in fractions 13-16 and fractions 34-37. When aliquots of each fraction were immunoprecipitated and examined by SDS-PAGE (not shown) proricin was detected in fractions 13-16 which would suggest that recombinant proricin made in oocytes has RNA N-glycosidase activity. However, faint low molecular weight bands were also seen in these tracks which may have represented biologically-active A chain or partial-proricin species. The presence of these low molecular weight species in the proricin

FIGURE 5.3.3 **RNA N-GLYCOSIDASE ACTIVITY OF CRUDE HOMOGENATE FROM OOCYTES INJECTED WITH PREPRORICIN mRNA.**

A. SDS-PAGE analysis of protein immunoprecipitated from oocytes injected with preproricin mRNA (SDS-PAGE performed by Dr. P.Richardson).

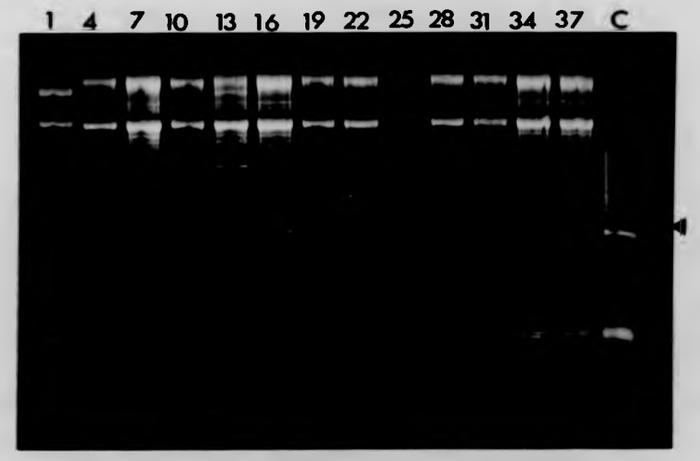
B. 1ml of crude homogenate (from 25 oocytes) was loaded onto a 1 x 30cm G-75 Sephadex column equilabrated with 1x homogenisation buffer containing 50mM DTT. 36 x 1ml fractions were collected and 30µl aliquots of every third fraction was incubated, at 37C for 30min, with an equal volume of rabbit reticulocyte lysate. RNA was then extracted and 4µg of this was analysed for ricin-catalysed depurination using the aniline assay. Note the diagnostic 390 nucleotide band present in fractions 13-16 and 34-37, indicating ricin modification. Control track (C) contains 4µg of aniline treated RNA extracted from reticulocyte lysate after pre-incubation in 1x oocyte homogenate containing 10^{-8}M biochemically-pure ricin A chain.

A



Synthesis of recombinant proricin. The in vitro transcript encoding preproricin was translated in a wheat germ cell-free system or microinjected into *Xenopus* oocytes in the presence of [^{35}S]methionine. Proricin was recovered from oocyte homogenates by immunoprecipitation using anti-ricin B chain antibodies, analysed by SDS-polyacrylamide gel elecrophoresis and visualized by fluorography. Lanes: (1) molecular mass markers (values indicated on the left, in kDa), (2) wheat germ product, (3) oocyte product after treatment with endoH, (4) oocyte product prepared under normal (reducing) conditions, (5) unreduced oocyte product, (6) control oocytes.

B



fractions was thought to be due to the fact that crude homogenate was loaded onto the column. This may have lead to the formation of aggregates between these proteins and other oocyte proteins which eluted from the column, in or close to, the void volume with the 66kD full-length proricin. A proportion of the low molecular weight species must have been retarded by the Sephadex beads and separated from the proricin however, to explain the second peak of activity in the fractions 34-37.

5.3.4 RESULTS FROM THE SECOND EXPERIMENT

The protocol was therefore modified in an effort to purify the proricin component before reduction and size fractionation on the Sephadex G-75 column. Proricin was first affinity-purified from crude homogenate by passage down a SeLectin^{TM} column (immobilised lactose matrix). This procedure allowed testing of any lectin activity of the precursor. Proricin was retained on the column by its association with the immobilised lactose indicating that this lectin precursor contained a biochemically-active B chain moiety. After washing the column through to remove all unbound oocyte proteins, the proricin could be eluted in buffer containing 100mM galactose. Using SDS-PAGE analysis, it was found that this method allowed single-step purification of proricin from the crude homogenate and indicated that the ricin precursor is an active lectin (Figure 5.3.4.1).

After reduction of this affinity-purified proricin fraction, it was passed down a G-75 Sephadex column and fractions analysed as in the previous experiment. The results are shown in Figure 5.3.4.2. They show that a peak of depurination activity was observed in fractions 31-34, presumably corresponding to the biologically-active low molecular weight species observed in the previous experiment.

FIGURE 5.3.4.1 AFFINITY PURIFICATION OF RECOMBINANT PRORICIN

A crude homogenate from preproricin RNA-injected oocytes was diluted in oocyte homogenisation buffer and passed three times down a 1.0ml SeLectin^{TM} column (lanes 2-4). The column was then washed three times with oocyte homogenisation buffer (lanes 5-7) before eluting the bound proricin with 100mM galactose (lanes 8-10). The bound proricin is eluted in the first fraction (lane 8). (Performed by Dr. P.Richardson, Warwick.)

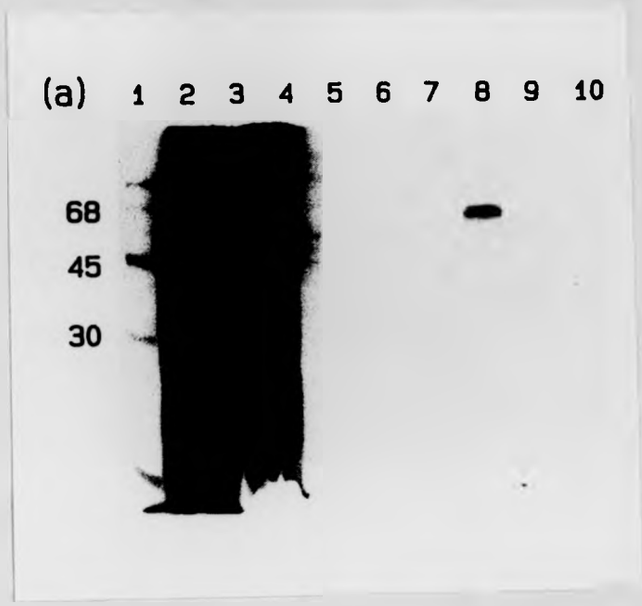
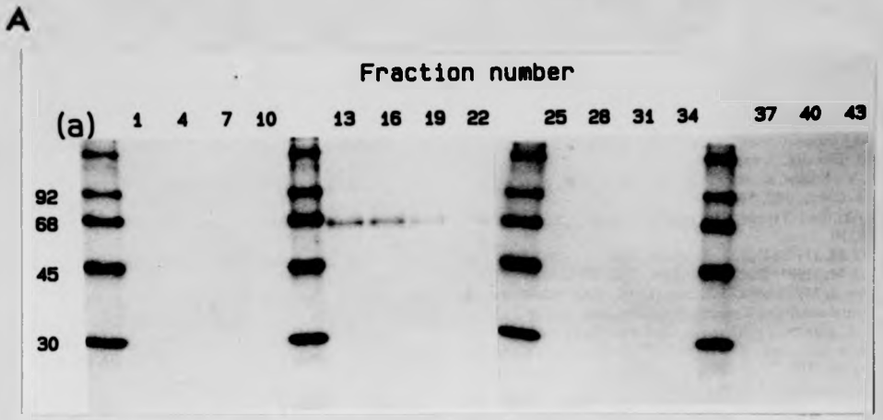
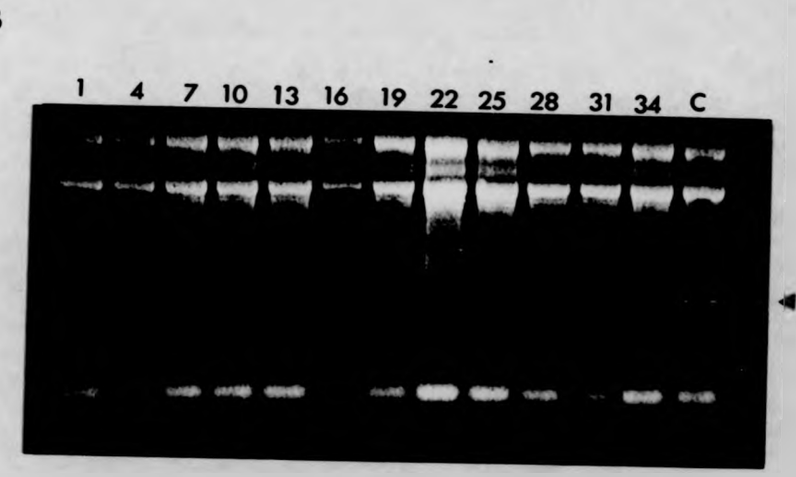


FIGURE 5.3.4.2 **RNA N-GLYCOSIDASE ACTIVITY OF AFFININTY-PURIFIED RECOMBINANT PRORICIN.**

A. SDS-PAGE analysis of G-75 fractions after immunoprecipitation. 0.9volumes of each fraction eluted from the column was immunoprecipitated with anti-ricin B chain antibodies to identify fractions containing recombinant proricin. Note an immunoreactive band of the expected size in fractions 13-19. (SDS-PAGE performed by Dr. P.Richardson).

B. 0.1volumes of each fraction was incubated with an equal volume of rabbit reticulocyte lysate for 30min at 37C. RNA was then extracted and 4μg was analysed for ricin-catalysed depurination using the aniline assay. Note the diagnostic bands present in fractions 31-34, but not in the fractions identified as containing the recombinant proricin (above). Control lane (C) was prepared as in figure 5.3.3.





These polypeptides must have had an active lectin-binding component to be retained on the SeLectin™ column, but this was possibly released from the A chain component by reduction of the disulphide bond prior to gel filtration.

On the other hand, in fractions 13-19 containing the full-length proricin, no depurination of the reticulocyte ribosomes was observed. From this it was concluded that full-length recombinant proricin did not possess RNA-specific N-glycosidase activity.

5.4 DISCUSSION

Soluble, disulphide bonded, N-glycosylated and ER lumen-segregated proricin was synthesised in Xenopus oocytes. Crude homogenate from oocytes injected with preproricin transcripts were shown to have RNA-specific N-glycosidase activity, but this activity was later attributed to the likely contamination of the proricin fractions with lower molecular weight species. These lower molecular weight proteins were immunoreactive, suggesting that they were related to the full-length proricin species. They may represent prematurely-terminated and segregated forms of proricin which had depurinating activity. The low molecular weight species were most clearly visible when resolved by SDS-PAGE under reducing conditions. This suggested that these species may have arisen by endoproteolytic cleavage of proricin in the oocytes which lead to the formation of disulphide bonded heterodimers. These low molecular weight proteins could be separated from the full-length recombinant proricin species by gel filtration under reducing conditions, but only after first purifying the proricin by affinity chromatography on an immobilised lactose column.

This purification step also indicated that the oocyte-synthesised proricin had a biologically-active B chain

moiety. It can be inferred that the A chain moiety folded correctly also since the immunoreactive, lower molecular species described above, were shown to have RNA N-glycosidase activity. However, the results from this study showed that the full-length recombinant proricin synthesised in oocytes does not depurinate 28S rRNA from rabbit ribosomes and therefore does not possess any RNA N-glycosidase activity.

CHAPTER 6

IN VITRO PROCESSING OF RECOMBINANT PRORICIN USING A SOLUBLE MATRIX FRACTION FROM DRY CASTOR BEAN SEEDS.

Contents:

2.1 INTRODUCTION.

As with many other plant seed storage proteins, ricin is initially synthesised as a preproprotein and is co-translationally translocated into the ER lumen where it becomes glycosylated and loses its N-terminal signal peptide. It is then transported in this intermediate form, via the Golgi (where it receives modifications to its oligosaccharide side chains), to the protein bodies (Lord, 1985ab). It is only upon reaching this final sub-cellular organelle that the single chain proricin intermediate is processed into its mature, heterodimeric form. As was discussed earlier, processing is believed to involve the removal of an N-propeptide as well as the removal of the 12 amino acid residue linker. The carboxy-terminal residue of both the putative N-propeptide and the linker is an asparagine.

The protein body-localised proricin processing activity has been demonstrated *in vitro* by Harley and Lord (1985). They first used a total soluble protein extract from developing castor seeds to process ^{35}S-labelled immunoprecipitated proricin, isolated from the endosperm of developing seeds, into A chain and B chain-sized fragments. They found that complete conversion of proricin to ricin subunits required an overnight incubation, and that the pH optimum for the reaction was pH4.6 (consistent with the pH inside plant vacuoles/protein bodies). They also prepared a protein body-enriched fraction from dry castor bean seeds which processed the immunoprecipitated proricin in a similar way. Extracts from both the developing seed and the dry seed also processed the precursor of the closely related castor bean lectin, RCA.

They concluded that this proricin processing activity was localised in the protein bodies, processing was slow, had a pH optimum of 4.6, and did not have a unique

substrate specificity.

In vitro processing of an 11S-globulin precursor from pumpkin, using a vacuolar matrix fraction from developing pumpkin cotyledons, has also been demonstrated (Nishimura and Nishimura, 1987). Nishimura and Nishimura (1987) suggest that the enzyme(s) involved in the cleavage of the pumpkin proglobulin polypeptide is an asparagine-specific thiol protease with an acidic pH optimum of 4. They also reported that a protein body matrix fraction from dry castor bean seeds processed the proglobulin from pumpkin into mature-sized globulin fragments.

Comparison of the predicted protein sequences for seed storage protein cDNAs suggests that asparagine plays a prominent role in precursor processing sites. This has led to the hypothesis that the enzyme(s) involved in the maturation of most (but not all) storage protein precursors is an asparagine-specific endoprotease (Lord and Robinson, 1986).

At the time when this work was initiated nothing was known about the asparagine-specific proteases. It was therefore deemed desirable to purify and fully characterise a storage protein precursor processing enzyme. As dry castor bean seeds have been shown previously to contain proricin processing activity, they provide a source of enzyme. Also, proricin clones are available in *in vitro* expression vectors, providing a means of synthesising a substrate for an enzyme assay. It was therefore decided to design an *in vitro* assay to enable purification of the proricin processing enzyme(s).

6.2 RESULTS.

6.2.1 SOURCE OF PRORICIN SUBSTRATE.

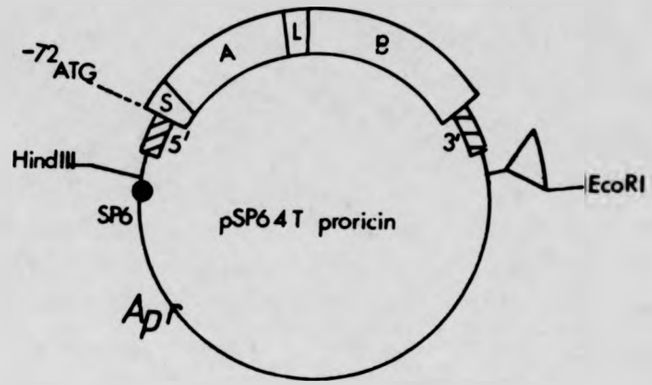
The construct, pSP64Tproricin (M.May, PhD thesis, Warwick), was used as the source of proricin cDNA (see Figure 6.2.1). This construct contains the proricin clone

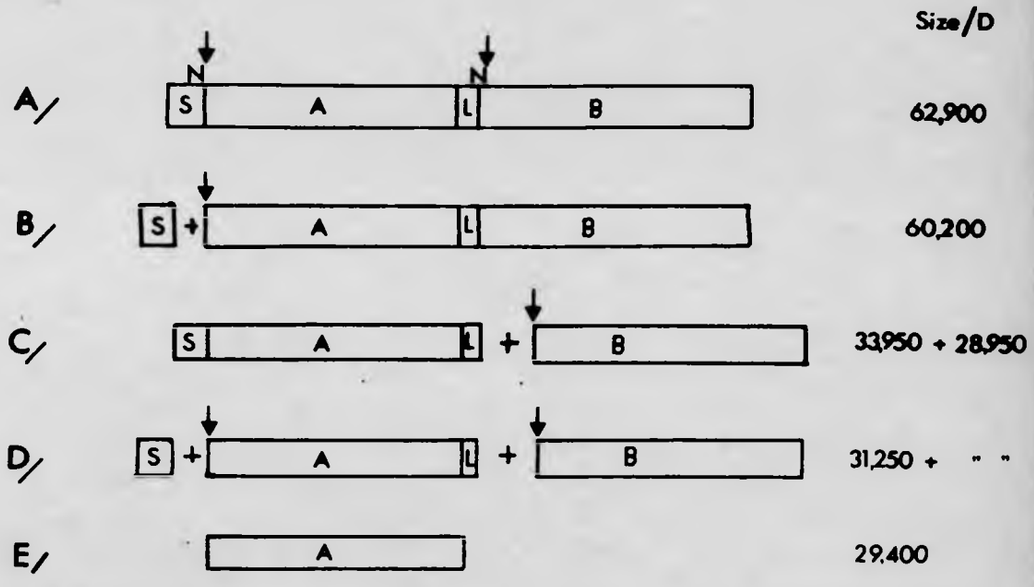
FIGURE 6.2.1 **SOURCE OF PRORICIN SUBSTRATE AND PREDICTED SITES OF ASPARAGINE-LINKED PROCESSING.**

The construct, pSP64Tproricin (M.May, PhD thesis, Warwick), contains the incomplete cDNA clone pRCL617delta4, cloned into the expression vector pSP64T at the BglII site. The proricin clone in this construct is therefore flanked by by 5' and 3' untranslated regions of B-globin DNA (hashed boxes). The construct pSP64Tproricin was first linearised by digestion with EcoRI, before *in vitro* transcription and translation. Translation is initiated in the proricin clone at position -72 (Lamb *et al*, 1985: Appendix A). Translation of this clone in a wheat germ lysate produces the non-glycosylated 62.9kD proricin form shown in A. Asparagine-linked proteolytic processing of this polypeptide at one or both of the sites arrowed, is predicted to yield the various sized fragments shown in B, C and D. The production of a mature A chain (shown in E) would require additional enzyme activities to remove the linker region at the C-terminus of A chain-linker fragments.

KEY.

S = (partial) N-terminal presequence.
A = A chain
L = linker
B = B chain





pRCL617delta4, inserted into the *in vitro* expression vector, pSP64T (Krieg and Melton, 1984). Expression of the gene is driven from the SP6 promoter, and translation is initiated from the ATG codon at position -72 (see Appendix A). The protein product encoded by this gene therefore lacks the first 11 N-terminal residues of the preproricin presequence.

6.2.2 TRANSCRIPTION OF THE PRORICIN GENE.

Proricin transcripts were synthesised *in vitro* from pSP64Tproricin, after first linearising the construct by digestion with *Eco*RI. Transcripts were quantitated by the addition of ^{3}H-UTP to the transcription reaction mix, and then by comparing the radioactivity incorporated into RNA to the total radioactivity present at the end of the incubation. Typically, 4-6µg of proricin mRNA was synthesised from a 20µl transcription reaction (containing 2µg of linearised pSP64Tproricin and 2units of SP6 RNA polymerase).

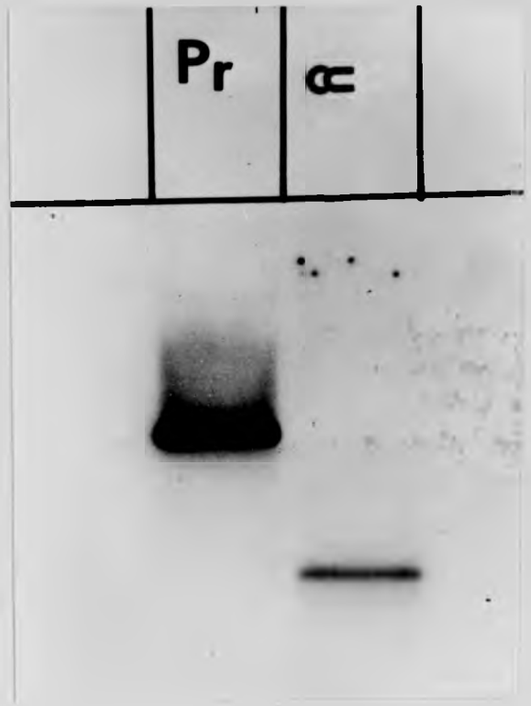
Proricin transcripts were also visualised by synthesis in the presence of ^{32}P-UTP and subsequent resolution of the labelled mRNA by denaturing agarose gel electrophoresis (Figure 6.2.2). The proricin mRNA appeared as a single band, which indicated that the proricin transcript synthesised *in vitro* from pSP64Tproricin existed as a single species.

6.2.3 OPTIMISATION OF WHEAT GERM LYSATE.

A wheat germ lysate and reaction Premix was prepared as described by Anderson *et al* (1983). As part of the protocol, low molecular weight inhibitors are removed from the lysate by gel filtration. The low molecular weight inhibitors appear as a bright yellow band which is retarded on the column relative to the active lysate fraction. The latter elutes as a pale brown band in the void volume and was collected in 1ml fractions.

FIGURE 6.2.2 **DENATURING AGAROSE GEL ELECTROPHORESIS OF PRORICIN TRANSCRIPTS.**

5µl from a 20µl transcription reaction (containing 2ug linearised pSP64Tproricin and 0.25µCi ^{32}P-UTP), denatured and resolved on a 50% formaldehyde/1.5% agarose (1xMOPS) gel (Pr). A 650 ribonucleotide transcript, synthesised from a yeast preproalpha-factor clone pGEM2alpha36 (a gift from J.Rothblat, EMBL, Heidleberg, Germany: see Rothblat *et al*, 1987), was run as a control (α).The gel was run at 40mA for 3 hours before drying under vacuum and visualising bands by autoradiography.



Aliquots of these fractions were then assayed for their ability to translate transcripts in vitro. Figure 6.2.3A shows the incorporation of ^{35}S-methionine into TCA-precipitable protein, when the aliquots were used to translate 1ug of pea polyA$^+$ mRNA. The highest incorporation of radioactivity into protein was observed in translations containing fractions 20 and 25. It was concluded therefore, that fractions 20-25 were the most active. These were stored under liquid nitrogen and used in all subsequent wheat germ translation reactions.

In order to optimise the synthesis of proricin in wheat germ lysate, in vitro translation reactions were set up using various amounts of the proricin transcription reaction mix from 6.2.2 above. Figure 6.2.3B shows the TCA-precipitable counts recovered from these reactions. It was found that 1.0µl of proricin transcription mix gave the highest level of protein synthesis. In the translation reactions containing less than 1.0µl, the amount of transcript is likely to be the limiting factor. However, a drop in protein synthesis was observed in reactions containing 1.5µl and 2.0µl, suggesting an inhibition of translation (probably due to other factors present in the proricin transcription reaction mix). Although the amount of proricin mRNA synthesised in in vitro transcription reactions will vary slightly, altering the slope of the curve from 0-1.0µl and the maximum level of protein synthesis (Figure 6.2.3B), all other factors in the transcription reaction mix (e.g. enzyme, RNasein, Premix) remain constant and therefore would be expected to inhibit a translation reaction to the same degree. It was concluded therefore to use 1.0µl of proricin transcription mix in all future translation reactions.

Translation of 1.0µl of proricin transcript in wheat germ lysate under these optimised conditions gave 200,000-250,000cpm/µl of TCA-precipitable protein.

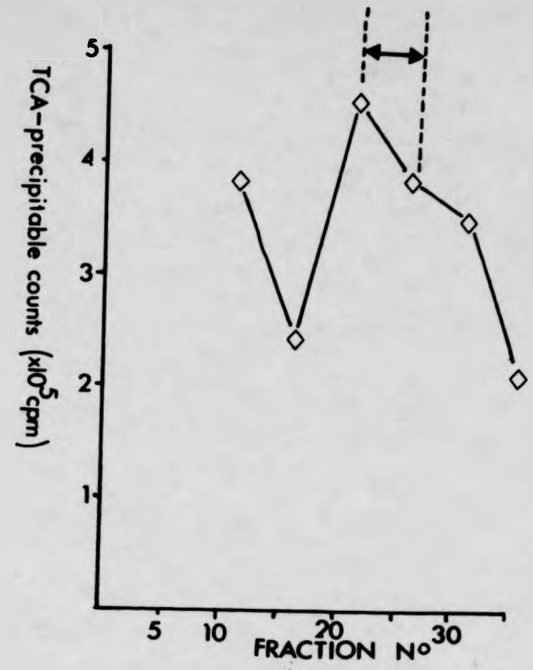
FIGURE 6.2.3 **OPTIMISATION OF WHEAT GERM LYSATE FOR IN VITRO TRANSLATION OF PRORICIN.**

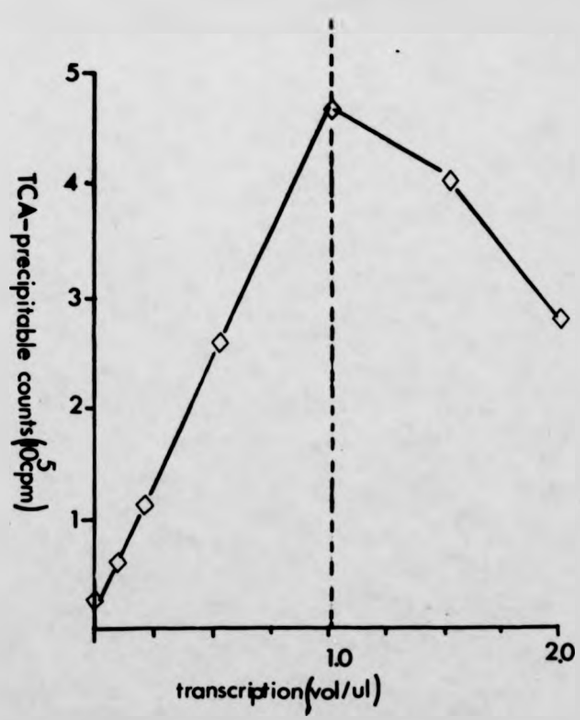
GRAPH A: Identification of active wheat germ lysate fractions eluted from a Sephadex G-75 column.

TCA-precipitable counts from 1/6th of <u>in vitro</u> translation reactions where 1ug of polyA+ RNA from pea (gift from P.Kirwin, Warwick) was translated, in the presence of 15uCi ^{35}S-methionine, using aliquots of wheat germ lysate fractions eluted from a 30cm G-75 Sephadex column. The most active fractions (20-25: arrowed) were then used for all subsequent wheat germ lysate translations.

GRAPH B: Optimisation of proricin transcript in wheat germ lysate.

TCA-precipitable counts from 1/6th of wheat germ translation reactions in which variable amounts of a transcription reaction, containing newly synthesised proricin transcripts, were added. The optimal amount, 1µl, typically contained 200-300ng proricin mRNA (estimated by comparison of incorporated cpm versus total cpm from a transcription reaction containing ^{32}P-UTP).





6.2.4 PROTEIN BODY MATRIX FRACTIONATION.

The isolation of protein bodies from dry castor bean seeds using the non-aqueous method of Yatsu and Jacks (1968) has been used previously to localise the proricin maturation enzyme activity (Harley and Lord, 1985: Nishimura and Nishimura, 1987). Protein bodies were therefore extracted from dry castor beans in this way (as described in Section 2.6.3) and aliquots of the soluble matrix fraction from the isolated protein bodies were analysed using SDS-PAGE.

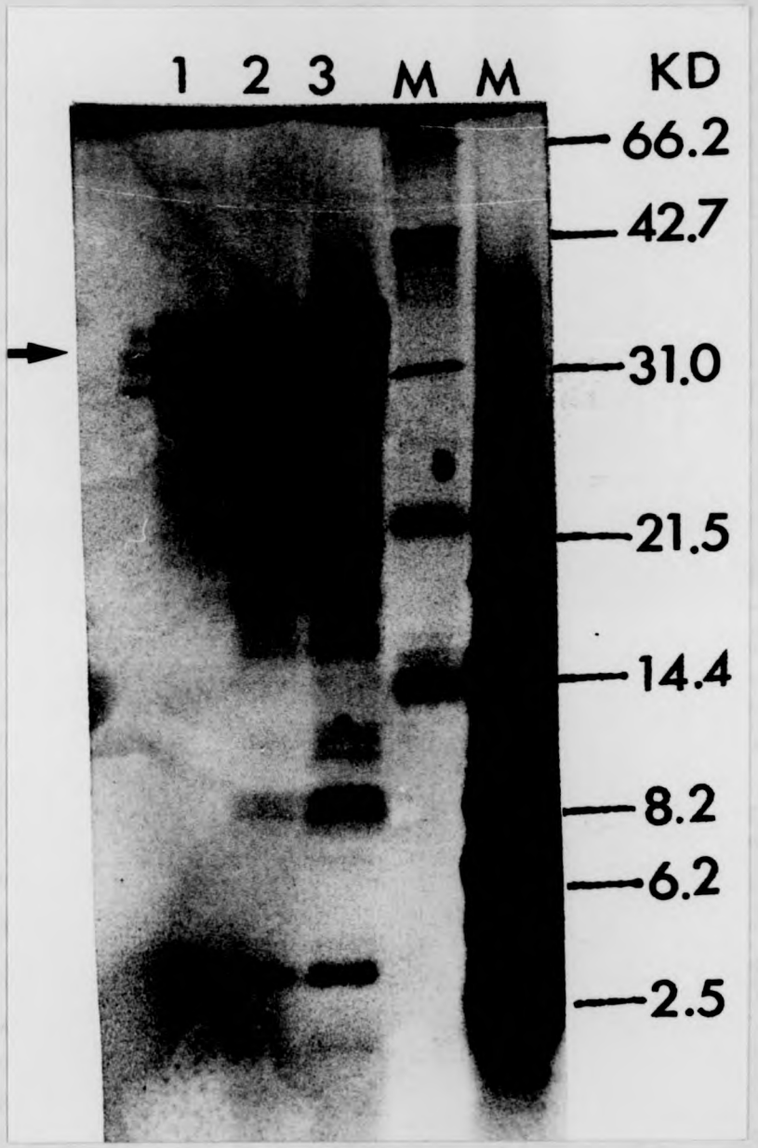
The three major polypeptides identified in the soluble protein body matrix fraction were between 32-37kD in size and correspond to the reduced ricin and RCA A and B subunits. The bands represent (in ascending order of M^r), ricin/RCA A chains, ricin B chain and RCA B chain. All other proteins which were stained using this gel system appeared lower in molecular weight than the lectin subunits. It was estimated that the lectins RCA and ricin accounted for over 80% of the protein visualised in the soluble matrix fraction.

The low molecular weight proteins were further resolved using the modified SDS-PAGE method of Shagger and von Jagor (1987), in which tricine replaces the glycine in the upper reservoir buffer (for details see Irwin, 1989). Figure 6.2.4 shows the protein profile of the protein body soluble matrix fraction resolved using this gel system. This gel reaffirms the estimate that the lectins account for over 80% of the protein in the matrix fraction. The other major ingredients in this fraction, the 2S albumins, were not visualised as they cannot be silver stained (Irwin, 1989).

The extract was initially stored overnight at 4C. Upon storage or dialysis the lectins were slowly precipitated out of solution.

FIGURE 6.2.4 **SDS-PAGE ANALYSIS OF SOLUBLE MATRIX FRACTION FROM PROTEIN BODIES ISOLATED FROM DRY CASTOR BEAN SEEDS.**

1, 5 and 15µg of total soluble protein from the matrix fraction of protein bodies, isolated from dry castor bean seeds, was analysed by SDS-PAGE (16% acrylamide/ tricine buffer: lanes 1, 2 and 3 respectively). Bands were visualised by silver stain. The most prominent bands (arrowed) correspond to the reduced A and B chains of the castor bean lectins. Low molecular weight markers used were; myoglobin fragments I+II (14.4kD); myoglobin fragment I (8.16kD); myoglobin fragment II (6.21kD); myoglobin fragment III (2.51kD). Source was S.D.Irwin, Warwick.



6.2.5 IN VITRO ASSAY FOR PROCESSING ACTIVITY.

Aliquots of the protein body soluble matrix fraction described above were then assayed for their ability to cleave the *in vitro* synthesised proricin substrate. Aliquots were assayed in citrate phosphate buffer at the pH optimum for the processing activity described by Harley and Lord (1985). The incubation temperature (26C) and length of incubation (overnight) required to maximise processing were also as described previously. Proricin and any cleaved products were then analysed by SDS-PAGE analysis as shown in Figure 6.2.5.

When proricin, synthesised *in vitro* in wheat germ lysate, was incubated overnight in citrate phosphate buffer alone, a band of the expected size (63kD) was observed (Figure 6.2.5: lane 1). Labelled polypeptide bands, which appeared below the full-length proricin product in lane 1 in the form of a smear, were also visible in freshly-translated proricin samples (not shown), and were therefore produced during the *in vitro* synthesis reaction, as opposed to being caused by proteolysis during overnight incubation in the citrate phosphate buffer. As the fidelity of the proricin transcript had been confirmed previously (Figure 6.2.2), it was concluded that these lower molecular weight bands were the result of incomplete elongation and premature termination events occurring in the wheat germ lysate during translation of the proricin message.

In samples where proricin was incubated with the soluble protein body matrix fraction, either before (lane 2) or after (lane 4) dialysis (to remove the residual glycerol present after the non-aqueous fractionation of the protein bodies), or after 80% ammonium sulphate precipitation (lane 6), the full-length proricin band was partially digested and bands of 29-31kD were observed. Distortion of the bands was observed, together with precipitation of radioactivity in the wells and

FIGURE 6.2.5 **IN VITRO ASSAY FOR PRORICIN PROCESSING ENZYMES.**

Autoradiograph of *in vitro* synthesised, radiolabelled proricin after overnight incubation with soluble matrix fractions from castor bean protein bodies, and resolution of polypeptides by SDS-PAGE. Aliquots of matrix fractions were assayed before (lane 2) or after (lane 4) dialysis in 3x 1000 volumes citrate phosphate buffer; after 80% ammonium sulphate precipitation, resuspension in citrate phosphate buffer and subsequent dialysis (lane 6); and after 10 fold or 20 fold dilution to reduce the distortion of the lanes by the 30kD lectins (lanes 8 and 10 respectively). Lanes 3, 5, 7, 9 and 11 represent the fractions assayed in 2, 4, 6, 8, and 10 (respectively), except that the former were assayed for activity after storage at -80C. As a control, proricin (P) was incubated overnight in citrate phosphate buffer pH4.6 (lane 1)

In some lanes significant amounts of radiolabel were precipitated in the wells (W) or at the interface (I) betweeen stacking and resolving gel mixes. Approximate protein concentrations of the protein body matrix fractions used in the assay were; 150µg/ml (lanes 2,3,6,7); 120µg/ml (lanes 4,5); 60µg/ml (lanes 8,9) and 6µg/ml (lanes 10,11). Protein was estimated by Bradford protein assay or estimation of protein on stained gels.

stacking/resolving gel interface, caused by the high amounts of endogenous lectin present in the soluble matrix fraction. When assays containing 10 fold and 20 fold dilutions of the soluble matrix fraction were analysed (lanes 8 and 10), only in the former were any faint bands observed.

Aliquots of these fractions remained active after storage at -80C (lanes 3, 5, 7, 9, 11).

6.2.6 REMOVAL OF LECTINS BY AFFINITY CHROMATOGRAPHY.

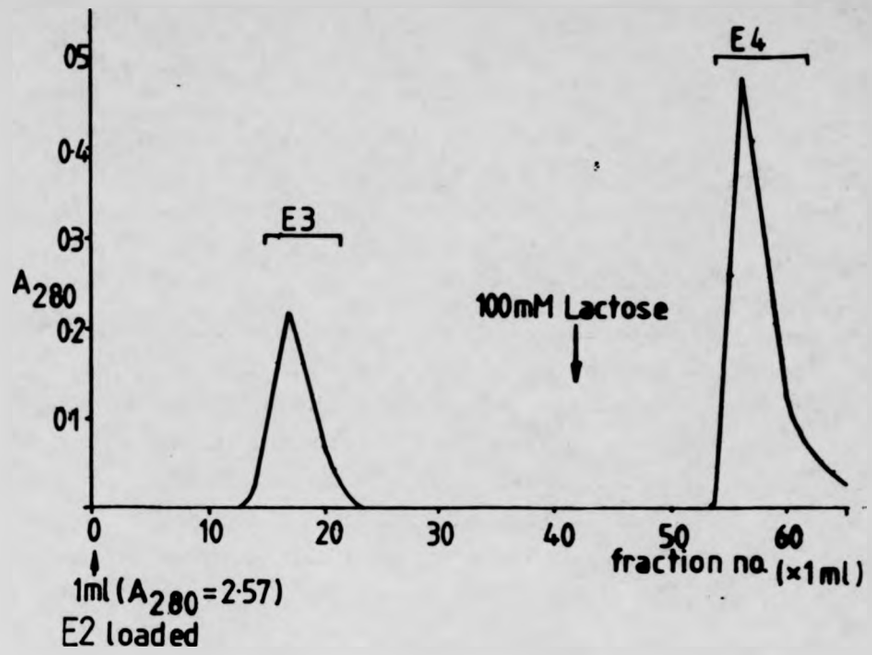
As illustrated in Figures 6.2.4 and 6.2.5, the major protein components of the protein body soluble matrix fraction are the lectins which distorted the lanes in the in vitro processing assay. Overloading of the reduced lectin bands meant that complete digestion of the proricin and accurate size estimation of the digested fragments was not possible. It was therefore desirable to remove the lectins from the soluble matrix fraction.

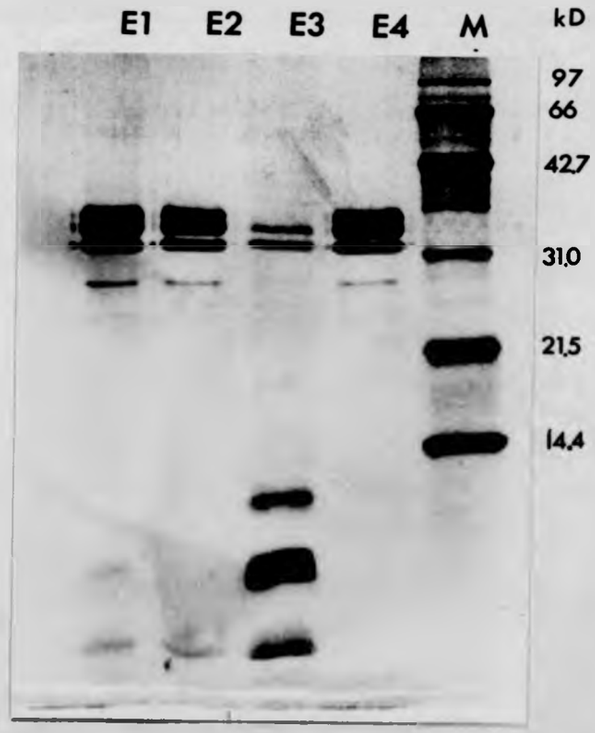
The lectins were removed by affinity chromatography on propionic acid-treated Sepharose 6B. The lectins are selectively retained on the column by interaction with the sugar residues exposed on the gel matrix.. Figure 6.2.6 shows the column profile and SDS-PAGE analysis of the unbound (E3) and lectin (E4) fractions eluted from a Sepharose 6B column, onto which was loaded a dialysed fraction of soluble protein body proteins (E2), prepared from a crude protein body matrix extract (E1). Although most of the lectin components were bound to the column, SDS-PAGE analysis revealed that the unbound fraction still contained a significant proportion of ricin (Figure 6.2.6: E3). The apparent discrepancy between the estimated amount of protein loaded in lane E3 and the amount of protein visualised after silver staining is likely to be due to the fact that the major components of this fraction were the 2S albumins which are not stained by silver.

FIGURE 6.2.6 **REMOVAL OF LECTINS FROM PROTEIN BODY SOLUBLE MATRIX FRACTION BY AFFINITY CHROMATOGRAPHY ON SEPHAROSE-6B**

A/ A_{280} profile of fractions eluted from Sepharose-6B affinity column, loaded with a protein body matrix fraction (E2). Lectins bound to the column were eluted in citrate phosphate buffer containing 100mM lactose (arrowed).

B/ SDS-PAGE analysis of crude protein body matrix fraction before (E1) or after (E2) dialysis to remove glycerol; and of (pooled) fractions eluted from the Sepharose 6-B column in citrate phosphate buffer (E3) or in the same buffer made 100mM lactose (E4). Approximately 100ng of protein was loaded in lanes E1, E2 and E4 (as determined by the Bradford protein assay). 1µg protein was loaded in lane E3. Resolved polypeptide bands were visualised by silver stain.





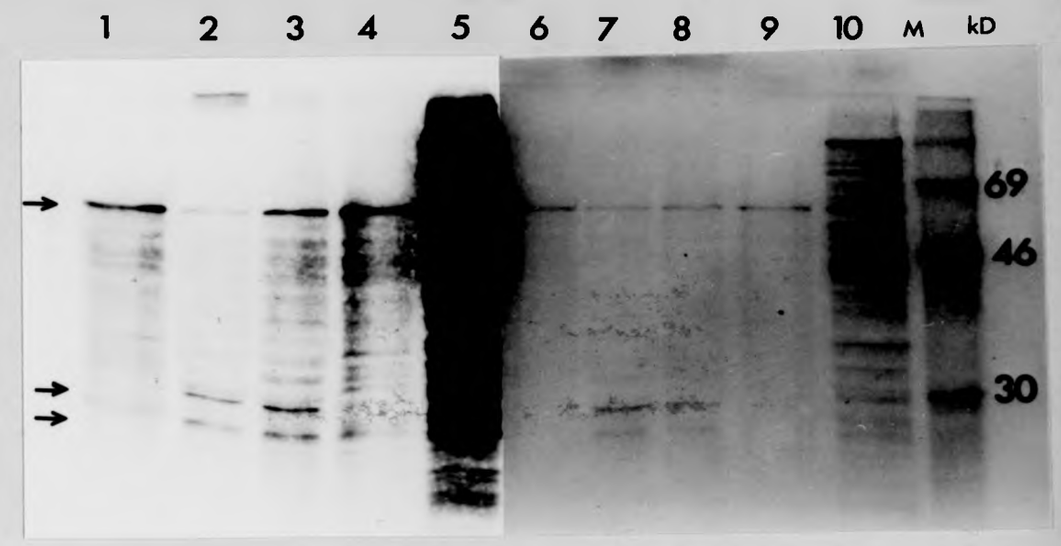
6.2.7 EFFECTS OF AMINO ACID ANALOGUE INCORPORATION/DTT ON THE PROCESSING OF PRORICIN *IN VITRO*.

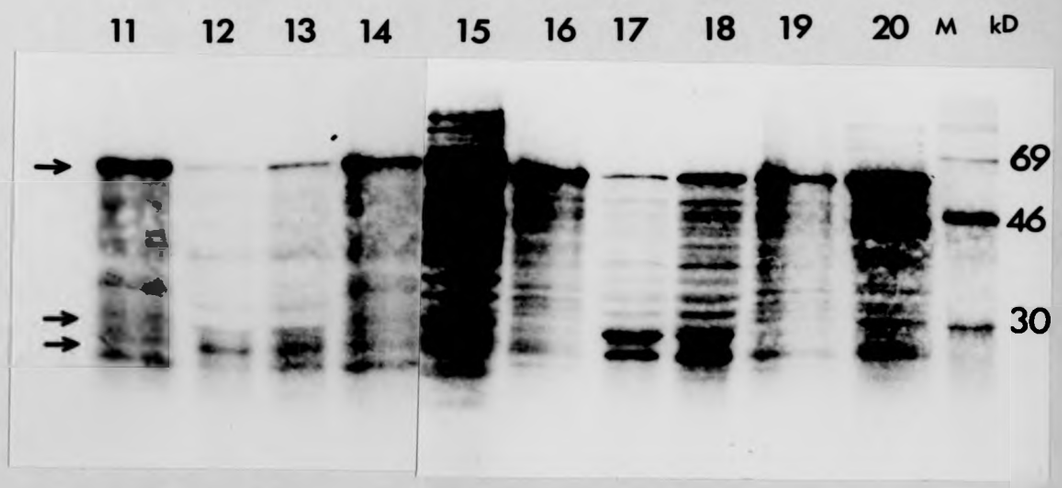
The fractions E1-E4 described in the section above were assayed for proricin processing activity by incubation with *in vitro* synthesised proricin (Figure 6.2.7: lanes 1-5). The fractions were also assayed using proricin substrates containing either the amino acid analogues azetidine or canavanine (azetidine = proline analogue; canavanine = arginine analogue: lanes 6-10 and 11-15 respectively), or proricin synthesised in 1mM DTT (lanes 16-20). The processing of the analogue-substituted/DTT substrates was compared to the processing of the 'wild-type' proricin substrate. It was thought that proricin synthesised in a wheat germ, lacking the intrachain disulphide bonds normally present when proricin is synthesised and sequestered in the ER lumen of the developing castor bean endosperm, may not fold into a conformation which would allow access of the processing enzyme to the linker region. The enzyme would therefore be unable to cleave the polypeptide backbone to release the A and B chain fragments. It was therefore thought worthwhile to synthesise proricin containing the amino acid analogues canavanine and azetidine, to see whether the polypeptides produced had an altered conformation which would allow access of the enzyme to the linker. Similarly it was hoped that the DTT present would prevent the formation of any 'non-productive' (i.e. not present in native proricin from castor bean) disulphide bonds in the proricin substrate, which may alter its conformation and prevent access of the enzyme to the linker.

As with the previous assay (Figure 6.2.5) partial processing was observed in protein body matrix fractions before and after dialysis (E1 and E2). No processing however, was observed with fraction E3. In lanes containing the lectin peak (E4), the reduced lectin subunits somehow affected the

FIGURE 6.2.7 **EFFECT OF AMINO ACID ANALOGUE-INCORPORATION INTO PRORICIN ON THE PROCESSING OF THE PRORICIN PRECURSOR.**

Proricin transcripts, translated in a wheat germ lysate using a normal 5mM amino acid mix as described in Methods (lanes 1-5), a 5mM amino acid mix in which azetidine replaced proline (lanes 6-10), a 5mM amino acid mix in which canavanine replaced arginine (lanes 11-15), or which contained a normal reaction made 1mM DTT (lanes 16-20), were incubated in citrate phosphate buffer (lanes 1, 6, 11 and 16); or with crude protein body matrix fractions E1 (lanes 2, 7, 12 and 17), dialysed fractions E2 (lanes 3, 8, 13 and 18), unbound column fractions E3 (lanes 4, 9, 14 and 19) and lactose-eluted fractions E4 (lanes 5, 10, 15, 20). Polypeptides and digested fragments were resolved by SDS-PAGE and visualised by fluorography. Estimated protein concentrations were; 90µg/ml (E1); 60µg/ml (E2); 10µg/ml (E3, E4). Protein concentrations estimated by Bradford protein assay.





resolution of radioactive protein and no conclusions could be drawn. However, it was not expected to find any processing activity in this peak.

Incorporation of the amino acid analogues into the substrate or translation in 1mM DTT had no visible effects on its processing. In both cases, the fractions E1 and E2 caused partial cleavage; E3 was inactive and in lanes containing E4, labelled polypeptides were resolved poorly. Furthermore, the level of expression of proricin containing azetidine was very low. It was concluded from this that the folding of the proricin substrate synthesised *in vitro* in the wheat germ lysate could not be altered by incorporation of amino acid analogues to improve its processing by the protein body matrix fractions.

All attempts to concentrate E3 fractions failed to show any processing activity (not shown) and also resulted in a concentration of the endogenous ricin which had not bound to the Sepharose 6B column (see Figure 6.2.6: lane E3).

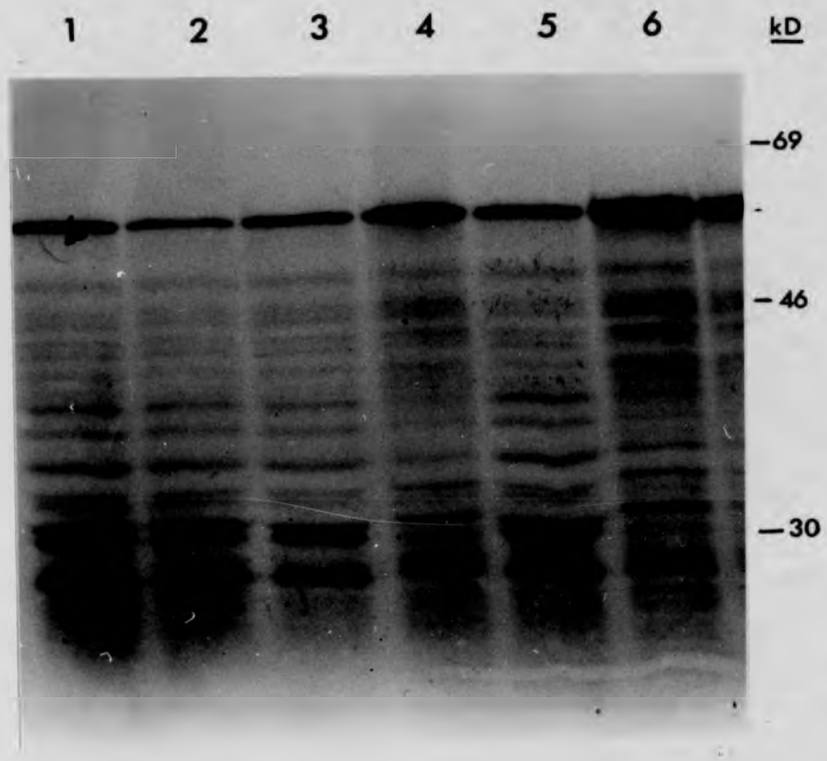
6.2.8 EFFECTS OF EDTA AND PMSF ON ENZYME ACTIVITY.

PMSF is a known inhibitor of serine proteases. EDTA chelates magnesium and calcium ions and therefore specifically inhibits enzymes which require either of these cations for activity. The proricin processing activity described by Harley and Lord (1985) was not affected by either of these chemicals. It was decided to test the activity of the soluble matrix fraction on the *in vitro* synthesised proricin substrate in the presence of these protease inhibitors (Figure 6.2.8).

In this assay a low concentration of E2 was used. This gave a low degree of processing but enabled good resolution of the digested fragments. In the assay EDTA had no visible effect on the processing of the proricin substrate. However, in the reaction containing 5mM PMSF no fragments were

FIGURE 6.2.8 **EFFECT OF EDTA AND PMSF ON PRORICIN PROCESSING ACTIVITY ISOLATED FROM DRY SEEDS.**

Aliquots of a protein body matrix fraction from dry castor bean seeds were assayed for proricin processing activity (as in Figure 6.2.7), after a 20min pre-incubation at 30C in either citrate phosphate buffer pH4.6 (lane 5) or buffer containing 10μM EDTA (lane 1), 100μM EDTA (lane 2), 1mM EDTA, and 5mM PMSF (lane 4). Non-specific proteolysis of proricin was tested by incubation of proricin in citrate phosphate buffer in the absence of castor bean extract (lane 6). Labelled proricin was resolved by SDS-PAGE and visualised by fluorography. Protein concentration of the matrix fraction (60μg/ml) was estimated by Bradford protein assay.



visible. It was concluded that the processing activity observed in these experiments was inhibited by 5mM PMSF, in contradiction to the results of Harley and Lord (1985).

6.2.9 IN VITRO PROCESSING OF PULSE-CHASED AND IMMUNOPRECIPITATED PRORICIN ISOLATED FROM DEVELOPING SEEDS.

In light of the apparently anomalous result above, it was thought possible that the processing activity described in this chapter, was not the same as the activity described by Harley and Lord (1985) who had used pulse-labelled immunoprecipitated proricin from developing seeds as a substrate. Therefore, it was decided to confirm that the soluble matrix fractions used for the experiments described in this chapter did contain an enzyme activity capable of processing proricin isolated from developing seeds.

Endosperm slices from developing castor bean seeds were pulse-chased with ^{35}S-methionine and labelled proricin immunoprecipitated from a microsomal fraction, as described in Methods. This was used as a substrate for assaying the proricin processing activity in protein body soluble matrix fractions before (E2) or after (E3) passage down a Sepharose 6B column. A crude homogenate from developing seeds was also assayed for processing activity.

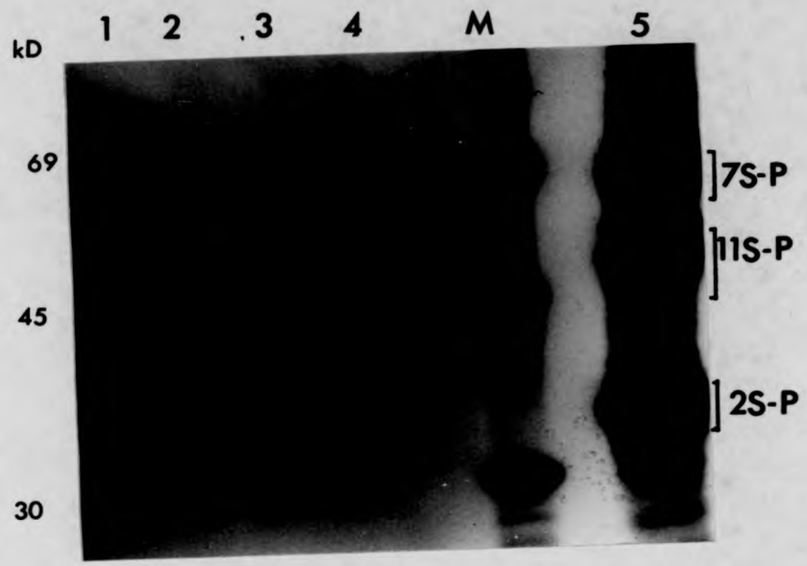
Figure 6.2.9A shows the result from this experiment. The major labelled proteins visible in the microsomal fraction were the precursors to the 7S lectins, 11S globulins and 2S albumins (lane 5). A band of 68kD was immunoprecipitated from the microsomal fraction corresponding to glycosylated proricin (lane 1). A fainter band of approximately 34kD was also visible in the immunoprecipitated fraction. This was present before overnight incubation and therefore did not represent a processed form of the 68kD band, but instead corresponds to the 2S albumin precursor immunoprecipitated by the anti-ricin

FIGURE 6.2.9 **IN VITRO PROCESSING OF PULSE-CHASED, ^{35}S-LABELLED PRORICIN FROM DEVELOPING SEEDS.**

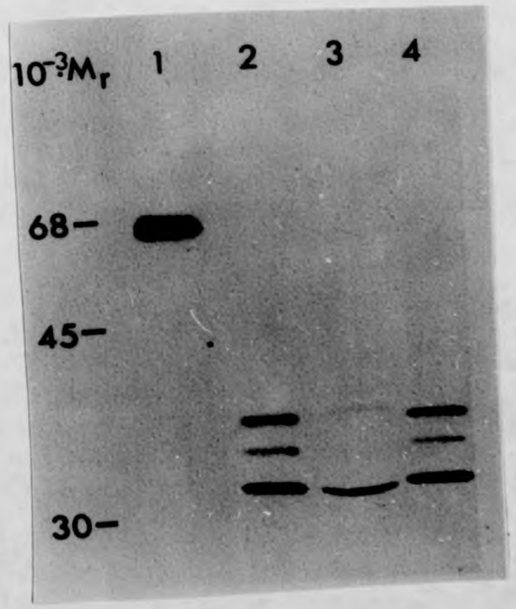
A/ Radiolabelled proricin, immunoprecipitated from the microsomal fraction of ^{35}S-methionine pulse-chased castor bean endosperm (60min pulse with 300μCi ^{35}S-methionine followed by 90min chase with 0.25M methionine) was incubated overnight at 26C with citrate phosphate buffer (lane 1), an extract from developing castor bean seeds (lane 2), and protein body matrix fractions E2 (lane 3) and E3 (lane 4) from dry castor bean seeds. Lane 5 shows the unbound (i.e. protein not immunoprecipitated) protein from the endosperm microsomal fraction, with the prominent bands present corresponding to the precursors of the 7S-lectins (7S-P), the 11S-globulins (11S-P) and the 2S-albumins (2S-P) (Lord, 1985a). Proricin was immunoprecipitated using rabbit anti-ricin antisera. Labelled polypeptides were resolved by SDS-PAGE and visualised by fluorography.

B/ The same experiment performed by Harley and Lord (1985) (lane 1, control; lane 2, incubated with extract from developing seeds; lane 3, incubated with extract from protein bodies of dry seed; lane 4, incubated with extract from dry seeds after removal of the lectins). Proricin and the processed fragments were immunoprecipitate with anti-RCA antisera, which also cross-react with ricin. Reproduced by permission of J.M.Lord.

A/



B/



antisera (compare lanes 1 and 5). Harley and Lord (1985) also reported 2S-albumin precursor contamination in some immunoprecipitated fractions during their study.

Overnight incubation with the fresh and dry seed tissues completely processed aliquots of the immunoprecipitated proricin (lanes 2-4). In the incubation with the fresh seed tissue, two bands of 32kD and 34kD were seen (lane 2). The 32kD band was distorted in the track containing the dry seed preparation E2, due to the high concentration of endogenous lectins. This 'smile' was corrected in the track containing the protein body matrix fraction E3, in which the lectins had been removed. In the tracks 2-4 there is a faint band present of 37kD which may represent RCA B chain, processed from contaminating RCA precursor co-immunoprecipitated with proricin.

These results correspond to those of Harley and Lord (Figure 6.2.9B). Unfortunately, the antibodies used seemed to cross-react with the 2S-albumin and RCA precursors and did not seem to successfully immunoprecipitate all the labelled proricin. However, the processing of the proricin precursor in all tracks containing seed extracts, to bands of 32kD and 34kD, indicate that the fractions E2 and E3 contained proricin processing activity.

6.3 DISCUSSION.

^{35}S-labelled proricin was successfully synthesised *in vitro* using a wheat germ lysate. Once optimised, the lysate synthesised 200,000-250,000cpm/µl of TCA-precipitable radiolabelled protein. A 63kD polypeptide was resolved by SDS-PAGE analysis of the translation reaction, corresponding to the expected size of 62.9kD for non-glycosylated proricin (Figures 6.2.5, lane 1; 6.2.1, A). A smear of bands appeared below the full-length proricin polypeptide caused by incomplete

elongation/premature termination events during the *in vitro* translation in wheat germ lysate. These bands could not be removed and made difficult the analysis of assay results. Immunoprecipitation did not clarify this profile (data not shown) presumably because all of the bands were translated from the proricin transcript and so were detected by the anti-ricin antibodies.

A protein body preparation from dry seeds was made using the non-aqueous method of Yatsu and Jacks (1967). The soluble matrix fraction from the isolated protein bodies appeared to consist of approximately 80% lectins upon SDS-PAGE analysis (Figure 6.2.4). When aliquots of this fraction were assayed for proricin processing activity a disappearance of the 63kD radiolabelled proricin substrate with concomitant appearance of bands at 30kD was observed. It was proposed that this represented specific proricin-processing activity in the soluble protein body extract. Processing was approximately 90% in the assays where protein body matrix fractions containing 150µg/ml protein were used (Figure 6.2.5; lane 2). The activity appeared stable after dialysis to remove residual levels of glycerol, frozen storage at -80C and after ammonium sulphate precipitation (Figure 6.2.5; lanes 2-7). However, the high levels of endogenous lectins present in these samples distorted the radiolabelled protein bands, making accurate size determination impossible. The overloaded lectins also caused significant precipitation of radiolabel in the wells and at the gel interface. It could be argued that the disappearance of the 63kD proricin band observed in the lanes containing the protein body matrix fractions was due to precipitation of the proricin substrate in the well and at the gel interface. However, when a control incubation was 'spiked' with protein body matrix fraction immediately before SDS-PAGE the proricin band was visualised at 63kD, despite the presence of the

radiolabel higher up the gel (not shown). It was therefore concluded that the radiolabel observed in the wells and gel interface was un-incorporated ^{35}S-methionine and did not represent labelled proricin.

When the protein body matrix fraction was diluted to reduce the adverse effects of the endogenous lectins discussed above, partial processing was observed at a protein concentration of approximately 60µg/ml (e.g. Figure 6.2.8, lane 2). At this lower protein concentration 3 bands were resolved after overnight incubation with the proricin substrate. There was a doublet of approximately 31kD and a lower polypeptide band of 29kD. By comparison to the predicted sizes of fragments produced by asparagine-linked proteolysis of proricin (shown in Figure 6.2.1) it was suggested that the lower band represents B chain (predicted size is 28,950kD) and the size of the doublet resembles the predicted size for an A chain-linker polypeptide (31,250kD). There was no visible evidence on any fluorographs of a 60kD polypeptide; the predicted size for an A chain-linker-B chain fragment produced by proteolytic cleavage at the asparagine residue at the C-terminus of the N-terminal presequence.

The protein concentrations of the protein body matrix extracts were 10-50 fold more dilute than those used by Harley and Lord (1985). Precipitation of the radiolabel in the wells of the gel and gross distortion of the lanes occurred at high concentrations and so limited the amounts used in the assay. Attempts to remove these endogenous lectins by affinity chromatography on Sepharose 6B were only partially successful, and ricin subunits were still visualised upon SDS-PAGE analysis of unbound protein fractions eluted from the column. Furthermore, when these fractions were assayed for proricin-processing activity, they failed to cleave the substrate, before or after protein concentration.

The partial processing activity observed in dilute samples of the protein body matrix fraction was not inhibited by the chelating agent EDTA, but was inhibited by 5mM PMSF. This contradicts both the results of Harley and Lord (1985) and Nishimura and Nishimura (1987), who each reported that the processing enzyme activity they had demonstrated *in vitro* was not inhibited by PMSF. The conditions used were not identical as those used by either Harley and Lord (1985) or Nishimura and Nishimura (1987), who both assayed for the processing activity by using a *de novo*-labelled native substrate, instead of the *in vitro*-synthesised substrate described above. One explanation would be therefore, that the protease activity discussed in this chapter is different to the ricin-processing activity reported by Harley and Lord (1985). In other words the digestion of the wheat germ lysate-synthesised proricin substrate was catalysed by a serine protease (sensitive to PMSF inhibition) present in the protein body matrix fraction, and not by the processing enzyme identified by Harley and Lord. This could be due to the fact that the *in vitro* synthesised substrate was not folded in such a way as to allow access of the processing enzyme to the linker region. The attempts to fold the substrate into a different conformation (by incorporating amino acid analogues into the polypeptide) failed to increase the processing observed.

However, the possibility of experimental error cannot be ruled out as this experiment was only repeated once (not shown) and the result clearly appears anomalous.

In retrospect, dry castor bean seeds were not an ideal source for the purification of an asparagine-specific endoprotease. Although it had distinct advantages; namely, a constant supply of material, the ability to isolate protein bodies from dry seed tissue and the potential for the selective removal of the co-compartmentalised lectins, one might expect

that the maturation protease activity in dessicated seeds is significantly lower than in the developing seed, where it is required for processing of the lectin precursors. It may be therefore, that the endoprotease activity recovered from dry seeds demonstrated by Harley and Lord (1985) and shown above only represents a residual level of functional enzyme. Furthermore, homogenisation of large quantities of castor bean seed tissue releases high levels of ricin. Manipulation of these extracts is consequently restricted in order to minimise the risk of exposure to toxin, making this a problematical source of a storage protein processing enzyme.

The substrate used was also far from optimal. SDS-PAGE analysis of the proricin synthesised in wheat germ lysate revealed a smear of polypeptide bands below the 63kD full-length product. As these were all synthesised from the same proricin transcript they were all recognised by anti-ricin antisera. This made identification of the 30kD processed fragments by immunoprecipitation with either anti-A chain or anti-B chain antibodies impossible.

In any future work it may be more prudent to use another seed tissue as a source of processing enzyme, since the hypothesis proposed at the beginning of this chapter is that many different plant storage proteins are processed by a common asparagine-specific endoprotease. Also, recombinant proricin synthesised in Xenopus oocytes would be a preferable substrate since the results presented in the previous chapter have shown that proricin synthesised in this way is glycosylated, disulphide bonded and biologically active suggesting that its conformation closely resembles native ricin.

Most recently, Nishimura and Nishimura (1991) have reported the purification of the proglobulin-processing enzyme from pumpkin. No details of the purification strategy have been

published, but they confirm that the protein is a thiol protease.

CHAPTER 7

CONSTRUCTION OF PRORICIN CLONES WITH LINKERS ENCODING FACTOR Xa AND THROMBIN RECOGNITION SEQUENCES: EXPRESSION IN E.COLI.

CONTENTS:

7.1 INTRODUCTION.

As discussed in the previous chapters, it was initially desirable to cleave proricin *in vitro* to yield biologically-active A and B chains. After the unsuccessful efforts to purify the endogenous endoprotease(s) responsible for proricin maturation in the plant, an alternative approach was designed. It was decided to use molecular biological techniques in order to introduce into the proricin linker region, a recognition sequence for a purified and commercially-available enzyme. It was hoped that such a molecule would also provide a ready source of active subunits and would therefore circumvent problems of producing biologically-active recombinant ricin B chain. However, the low specificity of many commonly available endoproteases made them unsuitable. For example, trypsin cleaves Lys/Arg-X peptide bonds, chymotrypsin cleaves on the carboxyl side of aromatic and other bulky nonpolar residues, and proteinase V8 from *Staphylococcus aureus* cleaves Glu-X bonds. Digestion of proricin with any of these enzymes would yield a host of fragments.

On the other hand, Factor Xa and thrombin are trypsin-like proteases which, under appropriate conditions possess a unique substrate specificity. Both enzymes are blood clotting factors involved in the complicated hemostatic cascade reactions (reviewed by Mann, 1987). They normally exist in the blood as inactive zymogens, Factor X and prothrombin, but upon vascular damage a network of cellular and protease reactions are initiated to localise the hemorrhage. These involve the activation of Factor X (to Factor Xa) and prothrombin (to alpha-thrombin: termed 'thrombin' hereafter). Both enzymes have been purified from various sources. Factor X can be activated *in vitro* using Russell's viper venom (Fujikawa, Legaz and Davie, 1972) and prothrombin is the physiological substrate of Factor Xa (see Mann, 1987).

Their unique substrate specificity is derived from the fact that, although they cleave peptide bonds on the carboxyl side of arginine residues with a similar reaction mechanism as trypsin, the molecular environment around their active sites restricts the access of most amino acid sequences. This means that they have preferred amino acid 'recognition' sequences: L-V-P-R-N-S is the preferred sequence for thrombin (Chang, 1985), and I-E-G-R has been shown to be the Factor Xa recognition sequence at both of its cleavage sites in prothrombin (see Nagai and Thogerson, 1984), with cleavage occurring after the arginine (R) residue in both cases.

The incorporation of thrombin and Factor Xa recognition sequences into hybrid proteins has been reported (Nagai and Thogerson, 1984: Smith and Johnson, 1988). Nagai and Thogerson expressed in E.coli a hybrid between the amino terminal region of lambda cII protein and human beta-globin, which were linked by a Factor Xa recognition sequence. The lambda-derived region enabled high expression of the hybrid and the beta-globin was subsequently released by digestion with Factor Xa. Smith and Johnson reported a method for single-step purification of recombinant eukaryotic proteins by expressing them in E.coli as fusions with glutathione-S-transferase. Once expressed, the hybrid protein is affinity purified by specific binding of the glutathione-S-transferase domain to a column containing a glutathione-bound matrix. By designing the chimeric gene to encode a fusion protein with a Factor Xa or thrombin recognition sequence at the junction between the two protein domains, the eukaryotic protein of interest can be separated from the glutathione-S-transferase domain by protease digestion following affinity purification.

It was therefore decided to design proricin constructs encoding proteins with either a Factor Xa or thrombin recognition sequence within the natural proricin

linker region. These constructs would then be expressed and the linker regions cleaved to produce heterodimeric recombinant proteins, hopefully with the desired biological activities.

Soluble recombinant A chain has been made in E.coli (O'Hare et al, 1988) but efforts to produce a soluble B chain in E.coli have not proved successful. However, B chain is known to be stabilised in solution by the presence of the A chain (whole ricin is stable in aqueous solution). It was therefore hoped that recombinant proricin produced in E.coli would be stable. In order to promote the formation of disulphide bonds in the recombinant protein, a factor important for the stability of B chain and essential for covalent linkage of the A chain and B chain following proteolytic cleavage, it was decided to target proricin to the periplasmic space, which has an oxidising environment better suited to disulphide bond formation than the cytoplasm. Secretion of recombinant proteins into the periplasmic space has been reported (reviewed by Schein, 1989) and reduces the risk of producing insoluble aggregates or inclusion bodies (Mitraki and King, 1989). It also has the advantage that the recombinant protein can be selectively fractionated away from cytosolic E.coli proteins, which is an aid in the subsequent design of a protein purification strategy.

The aims of the work described in this chapter therefore were to design and create mutant proricin clones encoding linkers with Factor Xa and thrombin recognition sequences, and to assess the use of an E.coli periplasmic expression system for the synthesis of soluble, disulphide-bonded recombinant (mutant) proricin.

7.2 CLONING AND MUTAGENESIS.

7.2.1 SOURCE OF VECTOR, PINIIIompA2proricin.

The construct pIIIompA2proricin was a gift from K.Hussain (Warwick). It contains the partial proricin clone pRCL617ox16 (see Appendix A) inserted into the high expression E.coli secretion vector, pINIIIompA2 (Ghrayeb et al, 1984: Appendix C for map), such that it was in-frame with the ompA signal peptide coding sequence. The chimeric gene in pINIIIompA2proricin therefore encodes a fusion protein with a 21 amino acid residue N-terminal ompA-derived signal peptide fused to the 5th residue of the A chain. Expression of the gene is under the control of the lipoprotein/lactose promoter and once expression is induced (by IPTG) the signal peptide should direct the fusion protein to the periplasmic space.

7.2.2 CONSTRUCTION OF CLONES CONTAINING FACTOR Xa AND THROMBIN ENCODING LINKERS.

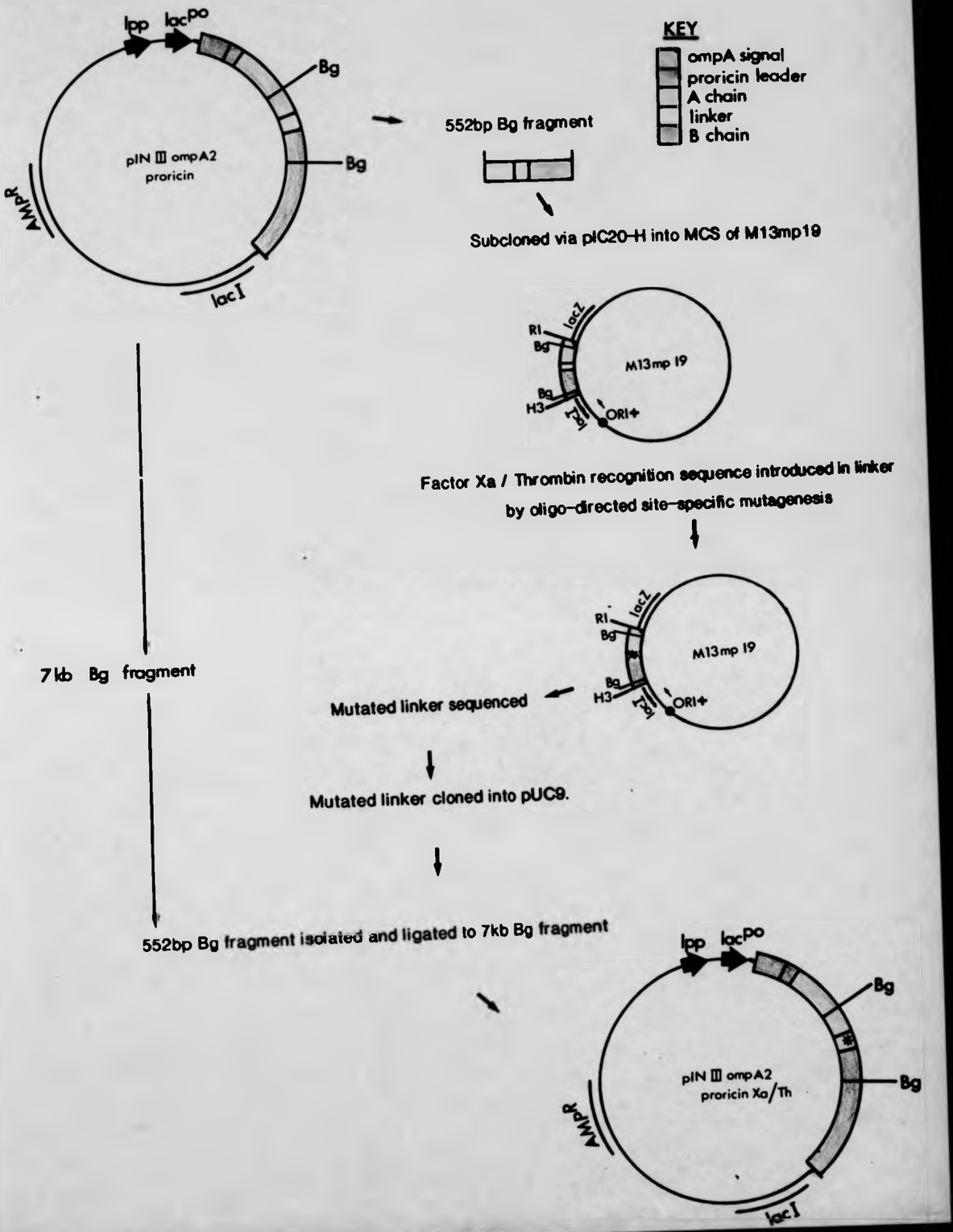
(see Figure 7.2.2.1 for cloning strategy)

(a) Construction of mutagenic template.

A 552bp BglII fragment, containing the proricin linker coding region, was excised from pINIIIompA2proricin and cloned into the unique BglII restriction site of pIC20H (see Appendix C for map). A clone containing a single copy of the 552bp proricin fragment was identified by colour selection on Xgal/IPTG/L-Amp plates and by restriction digest analysis of mini-prep. DNA.

A 600bp HindIII/EcoRI fragment from this clone was then cloned into the polylinker of M13mp19. Single-stranded M13 DNA was prepared from white plaques and sequenced (using the universal primer) to identify a clone in which the DNA strand of the 552bp proricin DNA fragment present in the ssM13 was complementary to the published sequence of Lamb et al (1985:Appendix A). This was the preferred orientation for the

FIGURE 7.2.2.1 CLONING STRATEGY FOR THE CREATION OF PRORICIN LINKERS CONTAINING EITHER A FACTOR Xa OR THROMBIN CLEAVAGE SITE.



ricin fragment in the M13mp19 vector because the ricin sequencing primers (which would be needed at a later stage to identify mutant clones) were the same sense as the published sequence (see Table 2.1.11 for list of primers). Once identified, this clone was used as a template for mutagenesis. In all subsequent mutagenic manipulations, care was taken to minimise the creation of codons which rarely appear in E.coli genes (Edge et al, 1983).

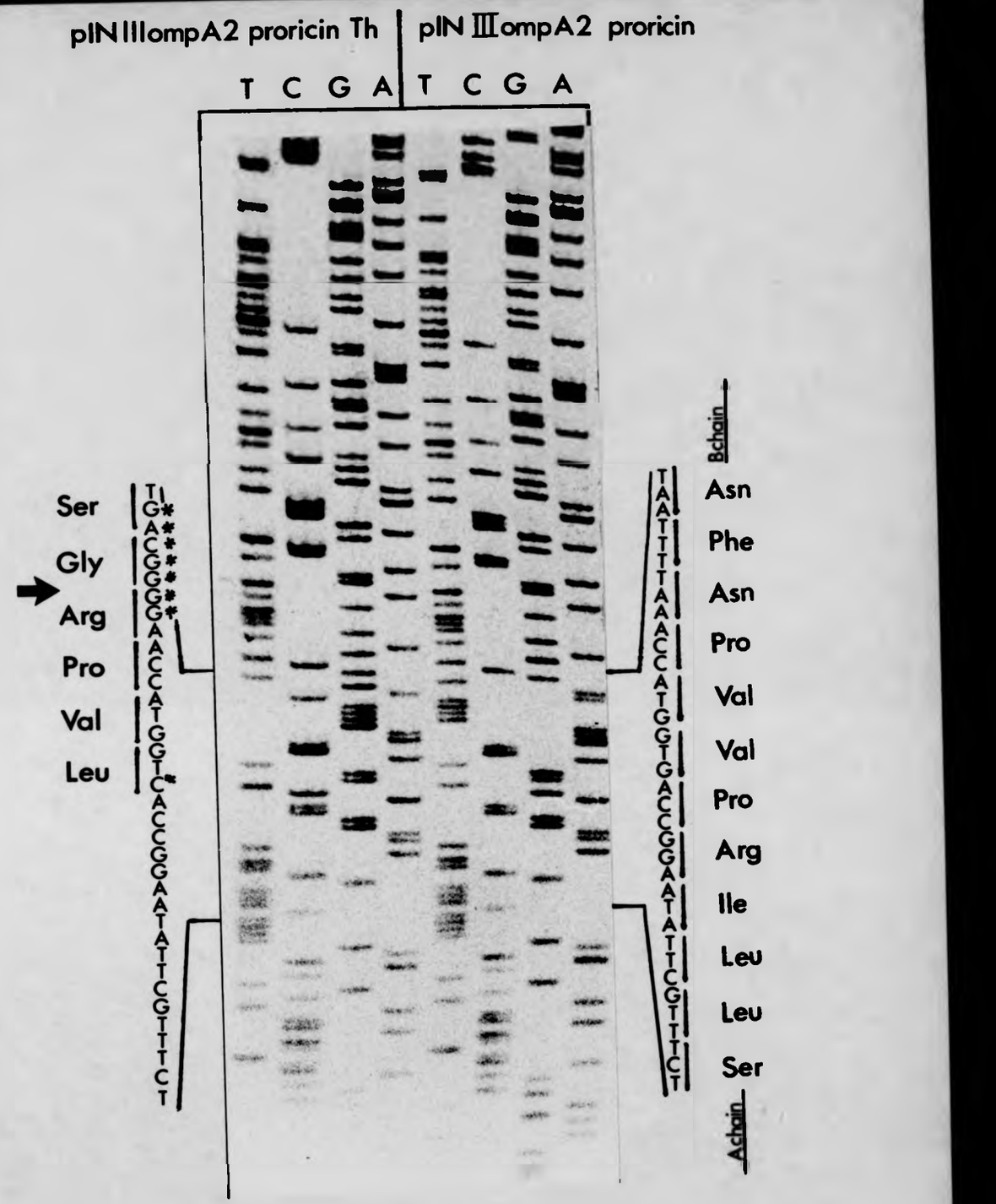
(b) Mutagenesis to create a thrombin recognition sequence in the proricin linker region.

A 21 base mutagenic oligonucleotide, MW01, was designed which would introduce a one-base mismatch mutation in the linker coding region, G^{820} to C^{820} (Table 2.1.12). The method of site-directed mutagenesis described by Nakamaye and Eckstein (1986) was followed and four mutants were identified by sequencing ssM13 preparations from 12 putative mutants, using the ricin sequencing primer CL586 (Table 2.1.11). ssDNA from one of these was then used as template in a second round of mutagenesis with the mutant oligonucleotide MW02 (Table 2.1.12). This second oligonucleotide was designed to introduce a 7 base mismatch mutation, $A^{830}TTTTAA^{836}$ to $G^{830}GGGCAG^{836}$. The method of Nakamaye and Eckstein (1986) was again used to perform the mutagenesis and of six putative mutant progeny sequenced, two were found to contain the correct mutations. The mutant DNA sequence is shown alongside the wild-type sequence in Figure 7.2.2.2. The complete 552bp ricin-derived region from one of these mutants was sequenced using the ricin primers to ensure that no opportunistic mutations had occurred.

This mutated linker will now be referred to as the 'thrombin linker', and all proricin clones containing this linker will be suffixed "Th".

FIGURE 7.2.2.2 **DNA SEQUENCE ANALYSIS OF WILD-TYPE PRORICIN AND THROMBIN MUTANT.**

DNA sequence of pINIIIompA2proricin (wild-type) and pINIIIompA2proricinTh (thrombin mutant) in the linker coding region, showing the 7 base mismatch mutation in the mutant (asterisked). The deduced amino acid sequence of the wild-type linker region is shown on the right and the thrombin recognition sequence, Leu-Val-Pro-Arg-Gly-Ser, encoded by the thrombin mutant is shown on the left with the predicted site of cleavage arrowed.



(c) Mutagenesis to create a Factor Xa recognition sequence in the proricin linker region.

A 28 base mutant oligonucleotide, MW03 (Table 2.1.12), was designed to introduce an 8 base mismatch mutation, $A^{814}GGCCAGT^{821}$ to $G^{814}AAGGCCG^{821}$, and this was used in conjunction with the M13 template from (a) above. The method of Nakamaye and Eckstein (1986) was used and of the 3 putative mutagenic progeny sequenced all were found to contain the correct mutations. Figure 7.2.2.3 shows the sequence of the mutated linker region and the deduced amino acid sequence (with the I-E-G-R Factor Xa recognition sequence) from one of these clones. One mutant was sequenced further to check that no opportunistic mutations had occurred in the 552bp ricin-derived sequence.

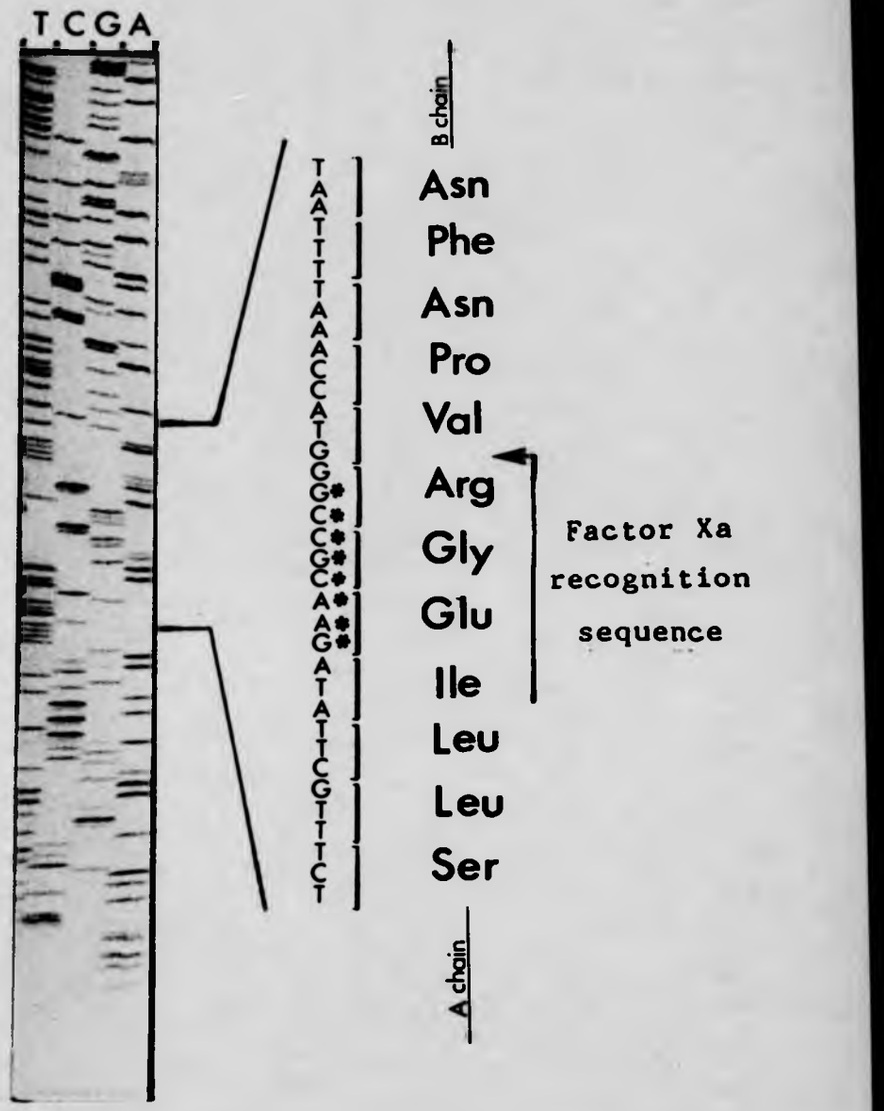
This mutated linker will now be referred to as the 'Factor Xa linker' and all proricin clones containing this linker will be suffixed "Xa".

(d) Cloning of thrombin and Factor Xa linkers back into pINIIompA2proricin.

RF M13 DNA was prepared from the mutated clones in (b) and (c) above. The 600bp EcoRI/HindIII fragment from each was then cloned into the polylinker of pUC9. This step was necessary as M13mp19 contains a BglII restriction site and therefore the 552bp ricin-derived BglII fragment could not have been excised directly from the M13 mutants and cloned easily. In each case, pUC9 clones were screened by plating on colour selection plates. Mini prep DNA prepared from one of these colonies, containing a clone with the mutant linker, was cut with BglII to release the 552bp ricin-derived region. This fragment was gel-isolated and cloned into the large 7kb BglII fragment of pINIIIompA2proricin, which had first been phosphatased to minimise the recovery of religated vector. The orientation of the thrombin/Factor Xa linker in the

FIGURE 7.2.2.3 **DNA SEQUENCE ANALYSIS OF PRORICIN LINKER CODING REGION FROM A FACTOR Xa MUTANT.**

DNA sequence of the linker region from a Factor Xa proricin mutant clone, with the 8 base mismatch mutation asterisked. The deduced amino acid sequence is shown alongside with the Factor Xa recognition sequence, Ile-Glu-Gly-Arg, marked. The predicted site of cleavage is arrowed.



pINIIIompA2proricin vector was determined by digesting mini prep DNA with BamHI and ClaI (not shown). Caesium chloride purified plasmid DNA was prepared from clones containing correctly-orientated mutant linkers and labelled pINIIIompA2proricinXa and pINIIIompA2proricinTh.

7.3 EXPRESSION IN E.COLI.

The constructs pINIIompA2proricin, pINIIIompA2proricinXa, and pINIIIompA2proricinTh were transformed into the E.coli strain JA221 (see Table 2.1.1 for genotype) and transformants were selected by streaking on L-Amp plates. This strain of E.coli was selected for this work because it has a modified cell membrane (lipoprotein minus phenotype) which has been shown to improve the yield of expressed proteins that are targeted to the periplasm (Takagi et al, 1988).

When transformants were grown on this medium, 'patchy' growth was observed which was taken to be due to expression of recombinant proricin from a 'leaky' lpp/lac promoter. Expressed proricin, containing the ompA signal peptide, should be targeted across the cell membrane into the periplasmic space which could weaken the cells during cell growth and division. Therefore, all subsequent transformants were selected and grown on L-Amp plates made 10mM glucose/10mM $MgCl_2$/10mM $MgSO_4$ (glucose is a metabolic repressor of the lac operon and magnesium strengthens cell wall formation). Colonies appeared to grow well on this medium.

Single colonies from these plates were used to inoculate JA221-minimal media (see Methods) which were then grown overnight at 37C/200rpm. These cultures typically reached an $O.D._{600}$ of approximately 1.5units. They were then used to inoculate pre-warmed JA221-expression media (JA221/minimal media supplemented with 1% Caseamino acids). Cultures were

grown to mid-log phase (this took 2-3 hours with a 1/50th dilution of an overnight culture) before inducing expression of proricin with 1mM IPTG. Incubation was continued for a further 3 hours. Figure 7.3.1 shows a typical growth curve for pINIIIompA2proricin-transformed JA221, compared to growth of a culture of untransformed JA221 under similar conditions (with the single exception that ampicillin was not included in the medium). The growth ~~curve~~ curves appear similar for both transformed and untransformed cultures. This would indicate that either there was no expression of recombinant proricin in the transformed cultures or the expression of proricin was not having any detectable effect on the growth of the pINIIIompA2proricin-transformed cells. The growth curves for JA221 cultures transformed with Xa/Th proricin mutants also appeared similar (not shown).

Cell fractions were prepared from cultures before and after induction with 1mM IPTG and analysed for proricin expression. Total cell protein fractions, periplasmic protein fractions (the supernatant obtained after osmotic shock of cells) and cytoplasmic protein fractions (obtained from the cell pellet after osmotic shock) were taken at different time points, resolved by SDS-PAGE and proricin detected by western blotting techniques. Figure 7.3.2A shows the protein profile for such fractions from a JA221 culture expressing proricin. No protein bands were detected in lanes loaded with periplasmic fractions. In the total cell and shocked cell fractions a detectable protein profile was observed for each time point, which was similar to the protein profile obtained from fractions isolated from a culture of untransformed JA221.

The western blot analysis (Figure 7.3.2B) revealed that only fractions from the pINIIIompA2proricin-transformed culture contained any immunoreactive polypeptides. This indicated that proricin was being expressed from the

FIGURE 7.3.1 **GROWTH CURVES FOR pIIIompA2proricin-TRANSFORMED AND UNTRANSFORMED JA221 CULTURES, BEFORE AND AFTER INDUCTION WITH 1mM IPTG.**

JA221 expression media (see Methods) was inoculated with 1/50 volume overnight cultures of pINIIIompa2proricin-transformed JA221 (Pr) and non-transformed JA221 (C). The cultures were incubated at 37C, 200rpm, and grown to mid-log phase before induction with 1mM IPTG (arrowed). Aliquots were taken at 0, 1, 2, and 3 hours after induction and analysed for proricin expression (see Figure 7.3.2).

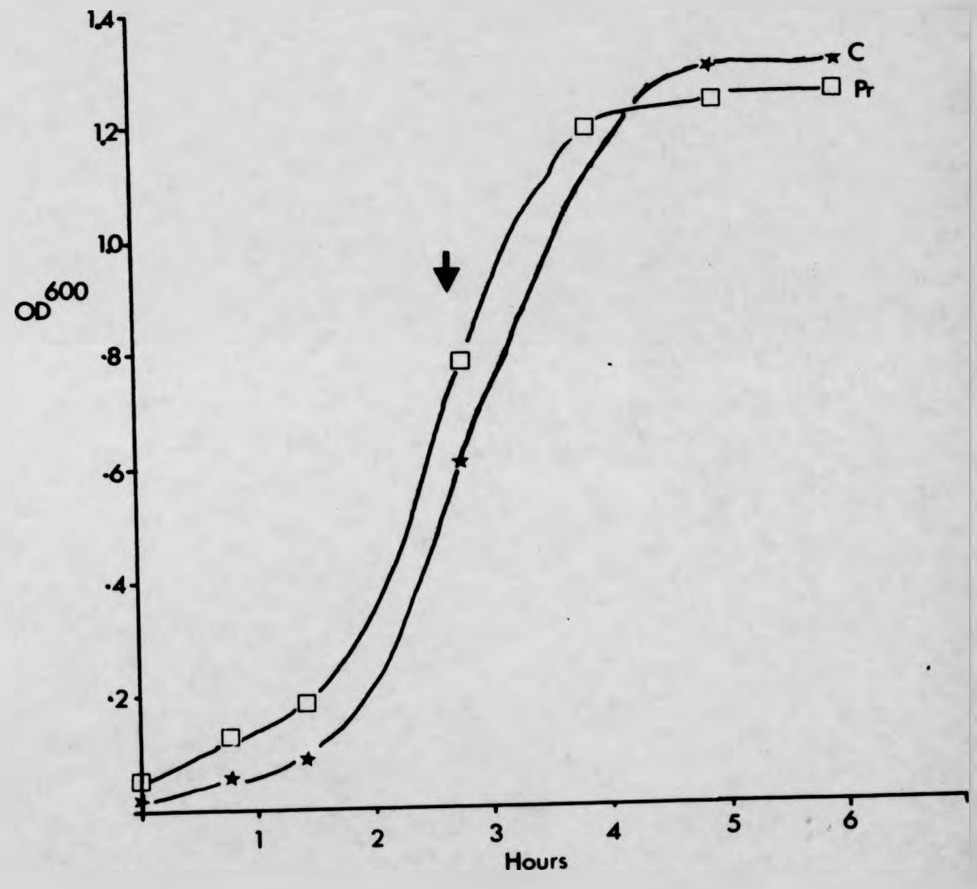
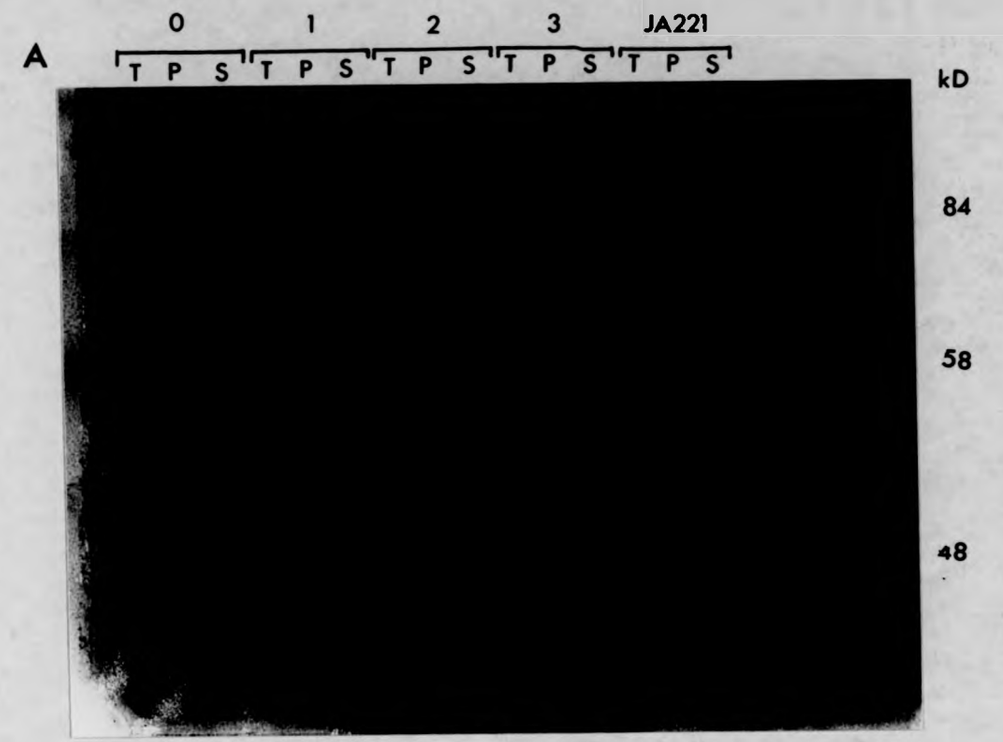


FIGURE 7.3.2 **EXPRESSION OF RECOMBINANT PRORICIN IN pINIIIompA2proricin-TRANSFORMED JA221, BEFORE AND AFTER INDUCTION WITH 1mM IPTG.**

SDS-PAGE (A) and Western Blot (B) analysis of total cell (T), periplasmic (P) and shocked cell (S) fractions from a culture of pINIIIompA2proricin-transformed JA221, 0, 1, 2 and 3 hours after induction with 1mM IPTG. Cultures were grown at 37C, 200rpm to mid-log phase before induction. All fraction volumes were adjusted to represent a cell concentration of 20 O.D.^{600} units/ml (the O.D. of the culture being measured immediately prior to fractionation) and 25ul of each fraction was loaded onto a 10% polyacrylamide gel. Lanes labelled 'JA221' represent fractions from a non-transformed JA221 culture after 3 hours IPTG induction. Protein bands were visualised by Coomassie staining (gel A). The nitrocellulose filter was probed with an IgG fraction of sheep anti-ricin A chain antisera and developed using the alkaline phospatase system. The upper and lower arrows represent the putative unprocessed and processed proricin forms respectively.



pINIIIompA2proricin plasmid. In the total cell and shocked cell fractions a doublet was detected at approximately 58kD, which were the predominant immunoreactive bands and which were taken to be the uncleaved (upper band) and signal peptide-cleaved (lower band) forms of the ompA2/proricin hybrid. Numerous immunoreactive bands were also detected below this doublet, which may represent breakdown products of proricin or premature termination events during translation of the proricin mRNA (similar to that described during the *in vitro* expression of proricin in wheat germ lysate).

No immunoreactive bands were detected in any periplasmic fractions. Since approximately 50% of the proricin was in the ompA-cleaved form in total cell fractions, indicating that 50% of the protein was reaching the periplasm (signal peptidase is localised in the periplasmic space), it was concluded that the method of osmotic shock was not efficiently releasing proteins from the periplasmic space. Another possible interpretation would be that the fusion protein becomes trapped in the cell membrane during translocation. This would allow the ompA signal to be cleaved (assuming this region of the protein was exposed on the periplasmic surface of the cell membrane) but would not allow release of the protein during osmotic shock.

A low level of proricin expression was detected in the total cell and (more clearly visible in Figure 7.3.2) in the shocked cell fractions from cells before IPTG induction. As the media used was a defined media, it was not likely that it contained any *lac* inducer, and therefore it was concluded that the *lac/lpp* promoter was not completely switched off in the absence of IPTG.

The levels of proricin expression in the culture appeared constant over the time period of the experiment (compare total cell fractions after 1, 2 and 3 hours

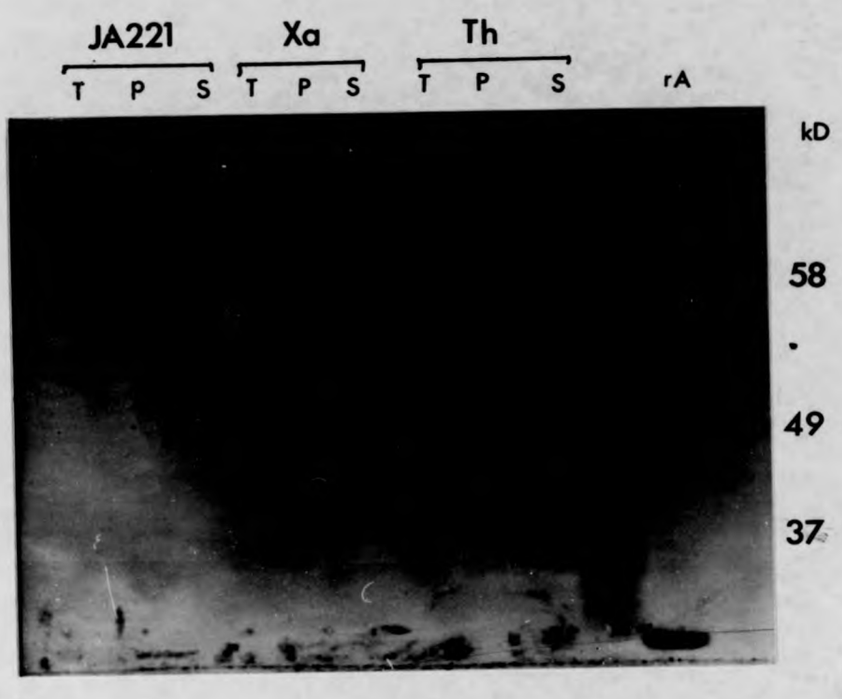
induction). The fractions were all standardised to take account of the different cell densities at the three time points. This result indicates that once induced, the cells reach a steady state level of proricin synthesis. The maximum amount of proricin produced in a culture is therefore achieved when the cells reach stationary phase (2 to 3 hours). On this basis it was deemed sufficient in all future experiments to induce cultures for a 2-3 hour period before fractionation and analysis. A similar expression pattern was observed for JA221 cultures transformed with the Factor Xa and thrombin proricin mutant clones.

In order to increase the resolution of the immunoreactive proricin doublet, fractions were loaded onto a 10% acylamide/0.1% bisacrylamide polyacrylamide gel. Figure 7.3.3 shows the western blot of a gel loaded with fractions from cultures expressing the Factor Xa and thrombin proricin mutants. Two bands of approximately 58kD were clearly resolved and were most prominent in the tracks containing cytoplasmic/membrane protein from shocked cells (S). Surprisingly, the immunoreactive bands detected in the periplasmic fractions (P) were largely in the unprocessed form. It also appeared that the majority of the recombinant protein expressed was not released by osmotic shock and remained within the shocked cells.

In an effort to assess the effectiveness of the osmotic shock treatment to release the periplasmic protein from the cells, fractions from a JA221 culture expressing proricin were also assayed for beta-lactamase activity. The beta-lactamase gene is contained on the pINIIIompA2proricin plasmid and is normally used for antibiotic selection on ampicillin. The protein is periplasmically-localised and can be assayed spectrophotometrically (O'Callaghan et al, 1972). Figure 7.3.4

FIGURE 7.3.3 **RESOLUTION OF PUTATIVE PROCESSED AND UNPROCESSED RECOMBINANT PRORICIN BANDS USING 'LOW BIS' SDS-PAGE.**

Western Blot analysis of total cell (T), periplasmic (P) and shocked cell (S) fractions from pINIIIompA2proricinXa-transformed (Xa) and pINIIIompA2proricinTh-transformed (Th) JA221, incubated at 37C and after 3 hours induction with 1mM IPTG. Protein bands were resolved using 10%acrylamide/0.1%bisacrylamide SDS-PAGE and the amount of each fraction loaded represents protein from 0.5ml culture. 200ng of recombinant A chain (rA) was loaded as a control and the nitrocellulose blot was probed with an IgG fraction of sheep anti-A chain antisera.



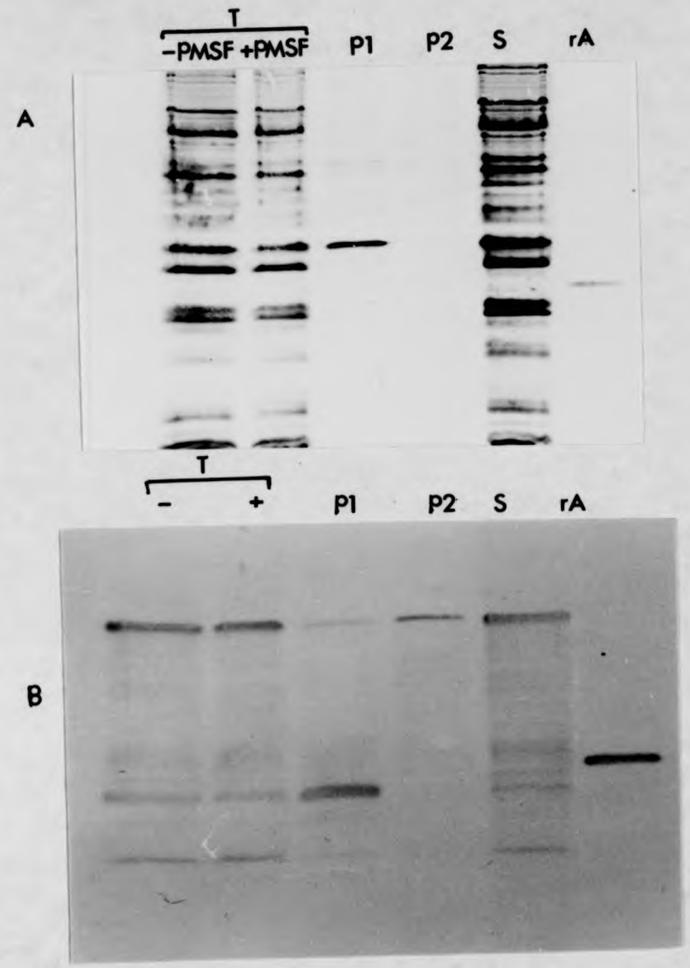
shows the gel profiles, western blot of proricin products and the results of the beta-lactamase assay.

The beta-lactamase activity all appeared to be released from the cells in the first periplasmic fractionation (compare values for total cell and periplasm fractions) with negligible beta-lactamase activity being observed in the shocked cell fraction. The silver-stained gel showed the presence of a single strong polypeptide band of approximately 35kD in the first periplasmic fraction (Gel A; lane P1) which is the predicted size for beta-lactamase. This gel also showed that there was selective release of beta-lactamase over other E.coli proteins (compare total cell fraction to periplasmic fraction). This indicates that the beta-lactamase activity recovered by osmotic shock was not caused by cell lysis and non-selective release of soluble E.coli proteins. There was no corresponding band observed in the second periplasmic fraction (Gel A; compare lanes P1 and P2). It was concluded therefore, that the first osmotic shock treatment efficiently released the periplasmically-localised beta-lactamase.

However, when the fractions were analysed for immunoreactive recombinant proricin (Figure 7.3.4; gel B) very little full-length proricin was released in the first periplasmic fraction. Instead a lower molecular weight immunoreactive polypeptide appeared released in this fraction. More of the periplasmically-localised recombinant proricin was released after the second osmotic shock, and in this fraction there was no evidence of the lower molecular weight immunoreactive bands. However, the majority of the recombinant proricin remained associated with the shocked cell fraction (compare lanes T, P and S of gel B). The same profiles were observed with cultures expressing the Factor Xa proricin mutant (not shown).

FIGURE 7.3.4 **EFFICIENCY OF OSMOTIC SHOCK IN THE RECOVERY OF PERIPLASMICALLY-LOCALISED RECOMBINANT PRORICIN.**

SDS-PAGE (A) and Western Blot (B) analysis of total cell (T), periplasmic (P1, P2) and shocked cell (S) fractions from a culture of pINIIIompA2proricin-transformed JA221 after induction for 2 hours at 37C with 1mM IPTG. Each lane represents protein from 0.2ml culture. An aliquot of the total cell fraction was sonicated in 10mM PMSF to determine whether the bands observed below the full length proricin product were due to proteolysis during fractionation. During the preparation of the periplasm fractions, the cells were shocked twice by successive suspensions in 1/100th volume ice cold 0.1mM $MgCl_2$: lanes P1 and P2 refer to the supernatant from the first and second osmotic shock respectively. 100ng of recombinant A chain was loaded as a control (rA), the nitrocellulose blot probed with an IgG fraction of sheep anti-A chain antisera and the immunoreactive bands detected by the alkaline phosphatase protocol described in Methods. Protein bands in A were visualised by silver stain. Fractions were also assayed for B-lactamase activity (shown in table) using the method of O'Callaghan *et al* (1978). The activity in the culture medium was also determined (M).



Fraction	M	T	P1	P2	S
$\Delta OD^{500}_{min^{-1}}$	0.093	0.317	0.344	ND	0.021

From this data it was concluded that the method of osmotic shock was effective for selective release of some periplasmically-localised proteins, but did not mediate the efficient release of the recombinant proricin from expression cultures of pINIIIompA2proricin-transformed JA221.

Total cell fractions were also prepared in the presence of 10mM PMSF (Figure 7.3.4) to determine whether the low molecular weight immunoreactive bands, seen in fractions from cultures expressing proricin, were caused by serine protease activity released during fractionation. However, there was no detectable difference between the fractions when analysed by SDS-PAGE and Western Blot which suggested that this was not the cause of the bands observed.

The solubility of recombinant proricin (and proricin mutants) recovered in the periplasm fractions was determined by centrifugation at 100,000xg for 30min at 4C. The supernatant after this step was taken to contain soluble protein. When supernatants and pellets were analysed by SDS-PAGE and Western blot analysis however, all immunoreactive proricin was found in the pellet (not shown).

An increase in the levels of expression of soluble recombinant proteins by a reduction of the expression temperature have been reported (O'Hare *et al*, 1988, for expression of recombinant ricin A chain: Hussain *et al*, 1989, for expression of recombinant B chain: also see general review by Schein, 1989). Takagi *et al* (1988) have reported that incubation of induced *E.coli* cultures at 23C increased the level of expression of active substilisin E from a pIIIompA-derived vector. Hussain *et al* (1989) found that ricin B chain expression was increased by overnight expression at 18C. The facilities were not available for routine incubation of cultures at 18C, but one experiment was possible. Figure 7.3.5 shows the results from this experiment.

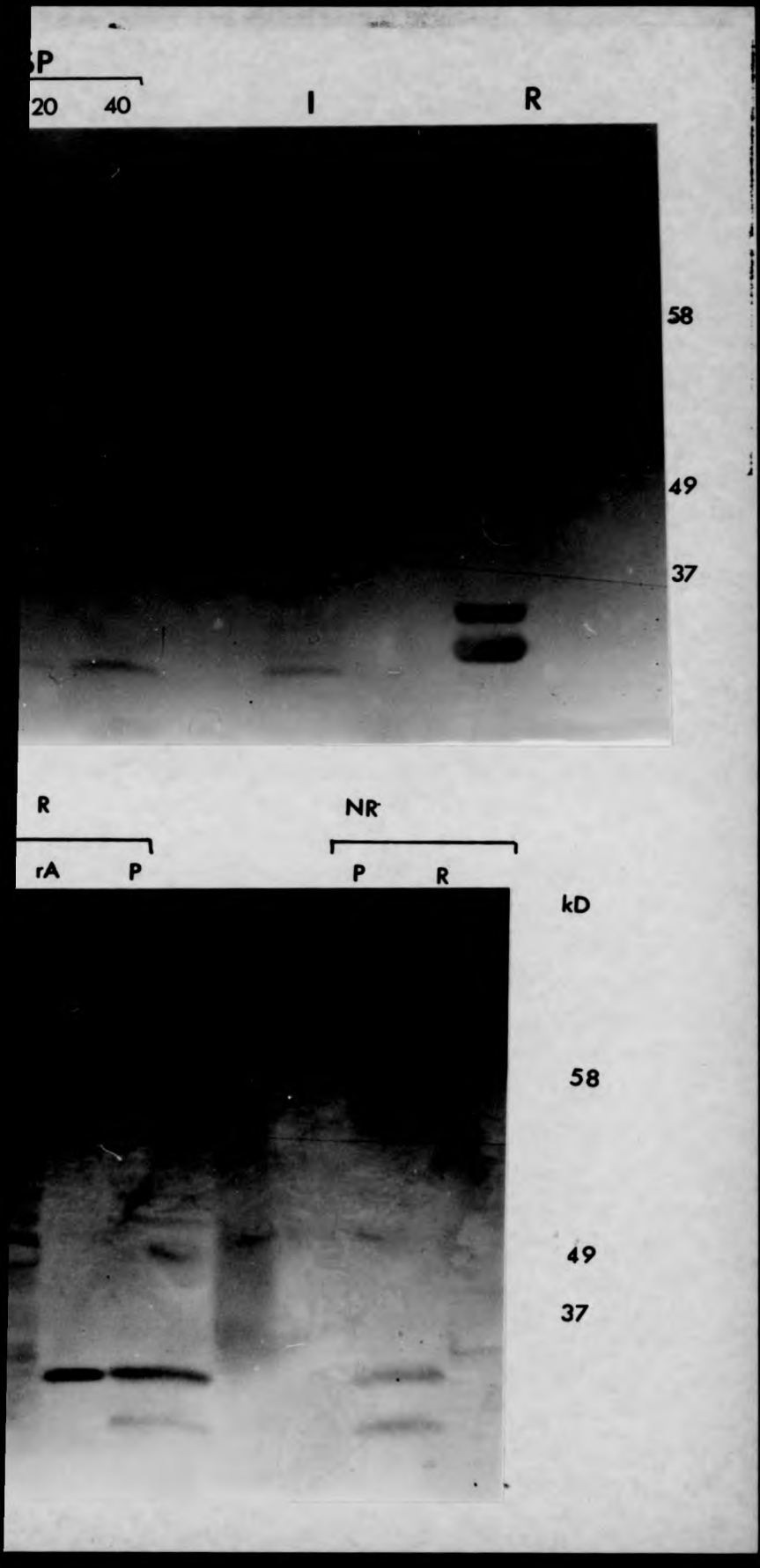
FIGURE 7.3.5

A: EFFECT OF LOWERING THE INCUBATION TEMPERATURE TO 18C ON THE SOLUBILITY OF PERIPLAMSICALLY-LOCALISED RECOMBINANT PRORICIN.

Western Blot analysis of periplasmic fraction isolated from a culture of pINIIIompA2proricin-transformed JA221, after incubation overnight at 18C in the presence of 1mM IPTG. The periplasmic fraction was concentrated 10x fold in an Amicon Centricon filtration unit and then centrifuged at 100,000xg, 30min, 4C, to pellet the insoluble protein. Proteins from 10, 20 and 40µl of the soluble periplasmic fraction (representing soluble periplasmic protein from 10, 20 and 40ml of culture: labelled SP) and all of the insoluble fraction (representing 500ml of culture: lane I) were resolved by SDS-PAGE and blotted onto nitrocellulose. 2µg ricin (Sigma) was run as a control. The blot was probed with an IgG fraction of sheep anti-A chain antisera and developed using the alkaline phosphatase system.

B: EVIDENCE FOR SOME DEGREE OF DISULPHIDE BONDING WITHIN SOLUBLE PERIPLASMICALLY-LOCALISED RECOMBINANT PRORICIN.

Western Blot of soluble periplasmic fraction (P), resolved by SDS-PAGE under reducing (R) and non-reducing (NR) conditions. Protein loaded represents the soluble periplasmic protein recovered from 80mls of the overnight culture (shown in A, above). The protein in lane S represents the shocked cell fraction from 4mls of culture; lane rA contains 500ng recombinant ricin A chain; and lane R is ricin run under non-reducing conditions (NR).


SP
20
40
I
R
58
49
37
R
NR
rA
P
P
R
kD
58
49
37

The western blot analysis showed that approximately 90% of the recombinant proricin expressed at 18C was recovered as soluble material in the periplasm fraction (Figure 7.3.5, blot A: compare lanes SP-40 and I). Furthermore, the majority of the soluble recombinant proricin recovered was in the ompA-cleaved form. Another immunoreactive band of lower molecular weight (approximately 30kD) was also visible in the soluble and insoluble periplasmic fractions.

The migration of the soluble recombinant proricin in the periplasm fraction was also compared under reducing and non-reducing conditions (Figure 3.7.4, blot B). The proricin migrated further under non-reducing conditions (compare lanes P-R and P-NR), indicating that the periplasmic proricin expressed under these conditions has a conformation maintained by disulphide-bonds. It is not possible from this data however, to determine whether the correct pairings of cysteine residues exists in the recombinant product. The effect of disulphide bond formation on the migration of proteins under non-reducing conditions is illustrated by ricin (R), which appears as a single band of approximately 58kD under non-reducing conditions in contrast to its characteristic heterodimeric profile under reducing conditions (compare lane R in blots A and B).

There were two additional bands of lower molecular weight detected by the anti-A chain antisera in the soluble periplasm fraction. The higher band appeared to be exactly the same size as recombinant ricin A chain (compare lanes rA and P: Figure 7.3.5B) and was more intense than the lower band. A possible explanation for the presence of these bands was that they represented processed A and B chain-related polypeptides cleaved from the mature proricin. Such a proteolytic step would have to occur in the periplasm, since cytoplasmic proteolysis would trap B chain, lacking a signal peptide region, in the cytosol. Furthermore, cross-reaction of the sheep anti-A chain

antisera with B chain polypeptides is seen with whole ricin (gel A; lane R). However, the size of the lower band is less than expected for B chain and both immunoreactive bands remain visible in the soluble periplasm fraction under non-reducing conditions suggesting that these immunoreactive polypeptides were not linked by a disulphide bridge. Alternatively, the bands may both represent A chain-like fragments (full-length and partial), produced by premature termination events during translation of proricin mRNA or proteolysis of the recombinant proricin in the cytoplasm. The ompA signal peptide would target both to the periplasm.

Finally, it was estimated that only 0.03mg/l of soluble recombinant proricin was recovered from the periplasm of JA221 cultures transformed with pINIIIompA2proricin and induced overnight at 18C. Over 90% of the recombinant proricin expressed in these cultures still remained trapped in the shocked cell fractions (0.25mg/l).

7.4 DISCUSSION.

Proricin clones with linkers encoding Factor Xa and thrombin recognition sequences, were made successfully. Proricin and Factor Xa/thrombin mutant clones were expressed in E.coli as fusion proteins with the ompA signal peptide, using the periplasmic secretion vector pINIIIompA2. Expression of the fusion protein was inducible with 1 mM IPTG. However, the lpp/lac promoter could not be completely switched off under the conditions used, and low levels of recombinant protein were detected even in the absence of IPTG.

Cultures transformed with pINIIIompA2proricin, pINIIIompA2proricinXa and pINIIIompA2proricinTh all expressed an immunoreactive product of the predicted size for proricin (58kD). Using 'low bisacrylamide' SDS-PAGE it was possible to clearly resolve this 58kD proricin into two bands, which were

taken to represent the signal peptide processed and unprocessed forms of the recombinant fusion protein. All fractions also were shown to contain many immunoreactive bands of lower molecular weight than the 58kD doublet. These polypeptides did not arise by the action of serine-specific proteases during fractionation. Other possible explanations for their presence is that they are caused by premature termination events during translation of the recombinant proricin mRNA or by other classes of proteases.

Efforts to recover the recombinant proricin from the periplasm of expression cultures using osmotic shock only succeeded in recovering approximately 10% of the total immunoreactive protein. This method however, successfully recovered all of the co-expressed and periplasmically-localised beta-lactamase from the same cultures. There are two possible reasons for the poor recovery of recombinant proricin from osmotically-shocked cells. One is that the ompA2/proricin fusion protein cannot easily cross the cell membrane and becomes trapped within the membrane. Schein (1989) reports that the structure of the recombinant protein is important for secretion. Alternatively, the recombinant proricin may become correctly segregated into the periplasm, but it cannot easily be drawn through the cell wall during osmotic shock because, for example, it exists in the form of large insoluble aggregates or simply because it is too bulky to pass through the gaps in the wall. Evidence in support of this was presented in Figure 7.3.4, where the majority of the proricin recovered during osmotic shock was recovered in the second fraction. Furthermore, the small quantity of proricin that was recovered in periplasmic fractions was found to be insoluble.

By lowering the expression temperature from 37C to 18C it was possible to recover a small amount of soluble periplasmically-localised recombinant proricin (0.03mg/l). This

recombinant proricin product was mostly in the processed (ompA-cleaved) form with some evidence of disulphide bonding. Over 90% of the recombinant proricin however, remained within the shocked cell fraction. The optimal total amount of recombinant proricin that was synthesised (0.28mg/l) was from an overnight culture incubated at 18C. Other workers have demonstrated the significance of the expression temperature on the yield of recombinant proteins. In particular, Takagi et al (1988), who were using a similar expression system to target recombinant substilisin E to the periplasm of E.coli, found that expression at 23C increased the yield significantly. They also found that low inducer concentrations improved the yield of substilisin E. However, the effects of different IPTG concentrations on proricin yield were not investigated in this study.

The recombinant proricin recovered in the periplasmic fraction from the culture grown at 18C, also contained significant amounts of two immunoreactive polypeptides that were not linked by a disulphide bridge, but one of which was exactly the same size as recombinant A chain. It is possible that these may represent A chain and B chain-like fragments which are cleaved from the mature recombinant proricin during periplasmic expression.

From the results presented it is clear that the problems with low yield and insolubility of recombinant proricin would make large scale purification impractical. The low molecular weight immunoreactive bands might also make interpretation of any toxicity studies difficult (using the RNA RNA N-glycosidase assay for example) with respect to the biological activity full-length recombinant protein. These problems have been discussed already with regards toxicity studies on recombinant proricin synthesised in Xenopus oocytes (Chapter 5). Although lowering the expression temperature did increase the amount of periplasmically-localised recombinant

protein, it was not possible to perform such experiments routinely at Warwick under the necessary Category IV conditions required by the Health and Safety Executive (HSE). Furthermore, at this lower incubation temperature, two immunoreactive bands were detected. These might mask results of the intended in vitro proteolytic experiments using the Factor Xa and thrombin mutants. It was therefore concluded that this system was not suitable for the expression of recombinant proricin and the Factor Xa/thrombin proricin mutants.

CHAPTER 8

EXPRESSION OF PRORICIN AND PRORICIN MUTANTS IN VITRO USING A XENOPUS EGG CELL-FREE TRANSLATION SYSTEM.

Contents:

8.1 INTRODUCTION.

The low yield of soluble recombinant proricin from E.coli was discussed in Chapter 7. It was concluded that the bacterial expression system was not suitable for expression and analysis of proricin and proricin mutants. The greatest success obtained in expression of recombinant proricin during this study was when preproricin transcripts were injected into Xenopus oocytes (Chapter 5). Using this system it was possible to synthesise a soluble, radiolabelled, immunoreactive product which could be assayed for its biological activities.

Although the technique of oocyte injection requires considerable experience to perform it successfully, an alternative expression system has recently been developed which uses an extract from Xenopus eggs to synthesise recombinant radiolabelled protein in vitro. The egg-free extract is prepared from freshly laid eggs and is stored under liquid nitrogen until use. Immediately before use it is thawed, supplemented (with creatine phosphate, spermidine, tRNA and an S-100 preparation from rabbit reticulocyte lysate) and mixed with the transcript. ^{35}S-methionine is added to radiolabel proteins synthesised during the incubation period and standard immunoprecipitation procedures identify immunoreactive products.

The aims of the work described in this chapter therefore were to assess the uses of this expression system for the synthesis of recombinant proricin and proricin mutants so that the biological activities of the mutants could be analysed before and after proteolytic cleavage. In this way, the accessibility of the new linkers to their proteases, the substrate specificity of the proteolytic reaction and the structure of the A and B subunits (assessed by their biological activities and disulphide bonding) can be ascertained.

8.2 RESULTS.

8.2.1 CLONING OF Xa AND TH MUTANT LINKERS INTO pGEMppRCL617.

The pINIIIompA2proricin (wild-type and mutant) constructs used in the E.coli expression study in the previous chapter, encoded fusion proteins between the bacterial ompA signal peptide and the incomplete proricin clone, pRCL617ox16. This meant that the proricin-derived portion of the chimeric gene in each case lacked a 5'ATG initiation codon. In order to express the Factor Xa/thrombin mutant linkers in a preproricin clone containing a eukaryotic signal peptide, it was necessary to isolate the 552bp BglII fragments from the pINIIIompA2proricinXa and Th mutants and to re-clone these into the large BglII fragment isolated from a full-length preproricin clone. The construct pGEMppRCL617 (Richardson et al, 1989), containing the full ricin presequence and which was used as the expression vector for the oocyte work described in Chapter 5, was therefore cut with BglII, the large fragment phophatased, gel-isolated and ligated with the 552bp BglII Xa and Th mutant linker fragments. Mini prep DNA was prepared from ampicillin-resistant colonies and digested with BamHI to orientate the BglII mutant linker fragment (Figure 8.2.1).

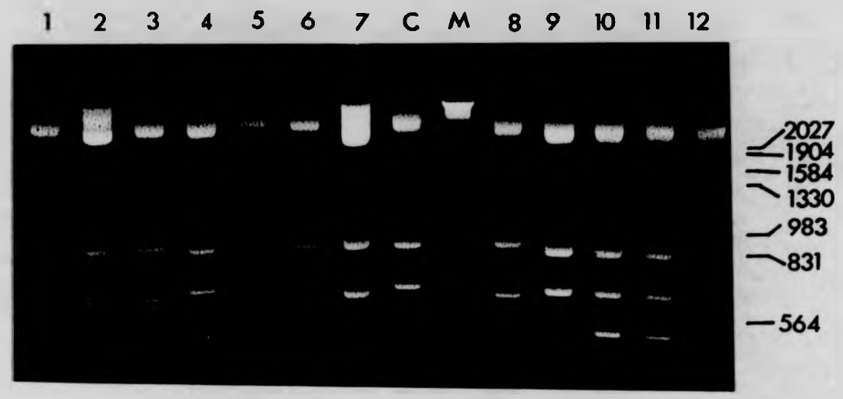
Factor Xa and thrombin preproricin mutants were identified and large scale caesium chloride purified DNA prepared. This DNA was finally sequenced in the region encoding the linker in order to confirm the presence of the substituted mutations in each case.

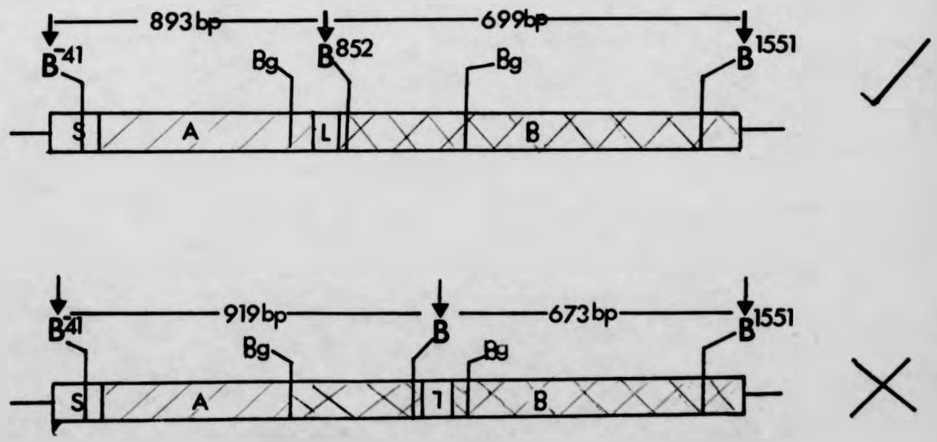
8.2.2 IN VITRO EXPRESSION OF PREPRORICIN AND MUTANT PREPRORICIN CLONES.

The preproricin and mutant preproricin constructs described above were orientated in the pGEM-derived vector such that transcription was initiated from the T7 promoter. The DNA

FIGURE 8.2.1 **ORIENTATION OF THE 552bp BglII FRAGMENT IN pGEMppRCL617Xa and pGEM617PPRCL617Th MUTANTS.**

Restriction digest analysis of mini prep DNA isolated from colonies transformed with pGEMppRCL617, after substituting the 552bp BglII wild type linker fragment with the 552bp BglII Factor Xa and thrombin mutant linker fragments. 1ug DNA was digested with BamHI and fragments were resolved on a 1.2% agarose (1xTBE) gel. Bands were identified by ethidium bromide staining and visualised on a UV transilluminator. Lanes 1-7 contain DNA from Factor Xa mutants; lanes 8-12 contain DNA isolated from thrombin mutants. 1ug wild-type pGEMppRCL617 DNA was also digested with BamHI as a control (C). Lambda markers were used to size the fragments (M): sizes are shown in bp. The predicted sizes of the DNA fragments from preproricin clones with the BglII fragment inserted in the correct orientation (tick) or incorrect orientation (cross) are also shown opposite. Clones in lanes 4, 5, 6, 9 and 12 contain the BglII fragment in the correct orientation. Digestions shown in lanes 1, 10 and 11 have an extra band of approximately 500bp and therefore these clones contain two BglII fragments.





was first linearised and then each construct was transcribed using T7 RNA polymerase. Incorporation of ^{3}H-UTP into RNA standardly gave estimated amounts of RNA synthesised of between 6-8µg per 20µl transcription reaction.

Transcripts encoding preproricin and preproricin mutants were then translated in rabbit reticulocyte lysate and wheat germ lysate systems (see Figure 8.2.2). All preproricin and mutant clones encoded a full-length polypeptide of 64kD when translated in wheat germ lysate, confirming that the substituted 'Factor Xa' and 'thrombin' fragments had not destroyed the reading frame or introduced any opportunistic termination codons. The control (preB), consisting of the ricin presequence fused to the B chain, also encoded a polypeptide of the expected size (32kD) in wheat germ. When translated in rabbit reticulocyte lysate a preB chain band of 32kD was observed. However, when preproricin transcripts or transcripts encoding the preproricin mutants were translated in rabbit reticulocyte lysate, no bands were observed in any tracks. This result agrees with the results of May (1988) who reported that translation of proricin transcripts in reticulocyte lysate was toxic: this toxicity is likely to be due to the synthesis of small amounts of active A chain-like fragments from the preproricin transcript or results from the fortuitous cleavage of the preproricin product to yield a proportion of active A chain fragments capable of inactivating the rabbit ribosomes. Wheat germ ribosomes are 1000 fold less sensitive than mammalian (and other non-plant) ribosomes to ricin and therefore are capable of translating active A chain polypeptides (Rupert Osborn, PhD thesis, Warwick)

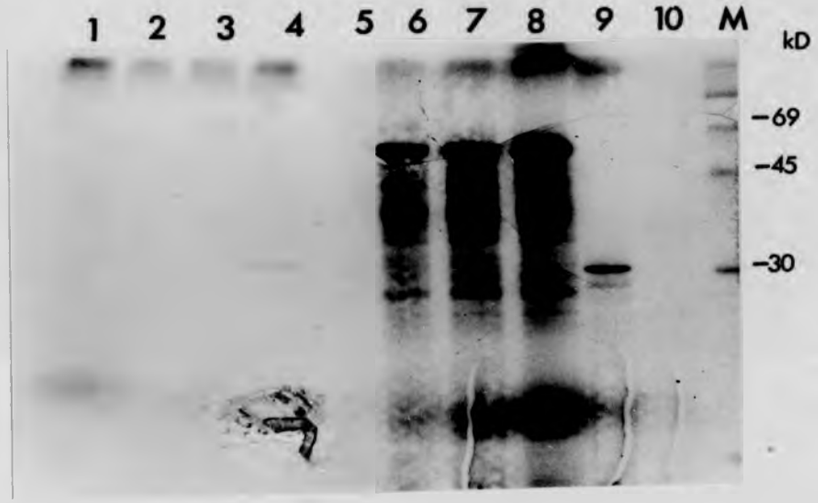
This result might suggest that expression of proricin in any animal system, including Xenopus sp., would be unsuccessful, as the release of active A chain fragments during translation of the proricin transcript would arrest further

FIGURE 8.2.2 **IN VITRO EXPRESSION OF FACTOR Xa AND THROMBIN PREPRORICIN MUTANTS.**

Preproricin transcripts and transcripts encoding Factor Xa and thrombin preproricin mutants were translated in rabbit reticulocyte lysate (lanes 1-5) and wheat germ lysate (lanes 6-10) in the presence of ^{35}S-Methionine. Labelled protein was resolved by SDS-PAGE and visualised by fluorography.

Lanes 1 and 6: preproricin
2 and 7: Factor Xa preproricin mutant
3 and 8: Thrombin preproricin mutant
4 and 9: translation of transcript preB*
5 and 10:no RNA control translation

* - gift from Dr.Richard Wales (Warwick).



expression. However, the results presented in Chapter 5 clearly show that expression is successful in systems where the active A chain fragments can be segregated away from the ribosomes: in the Xenopus oocytes it was assumed that the preproricin signal peptide mediated the translocation of any active polypeptide species across the ER membrane, where they became sequestered in the ER lumen. It was hoped therefore, that any active fragments synthesised in the Xenopus egg cell-free would be targeted and sequestered into microsomes in a similar way, thus enabling further preproricin/mutant expression.

8.2.3 PROTEOLYTIC CLEAVAGE OF IN VITRO SYNTHESISED PREPRORICIN AND PREPRORICIN MUTANTS.

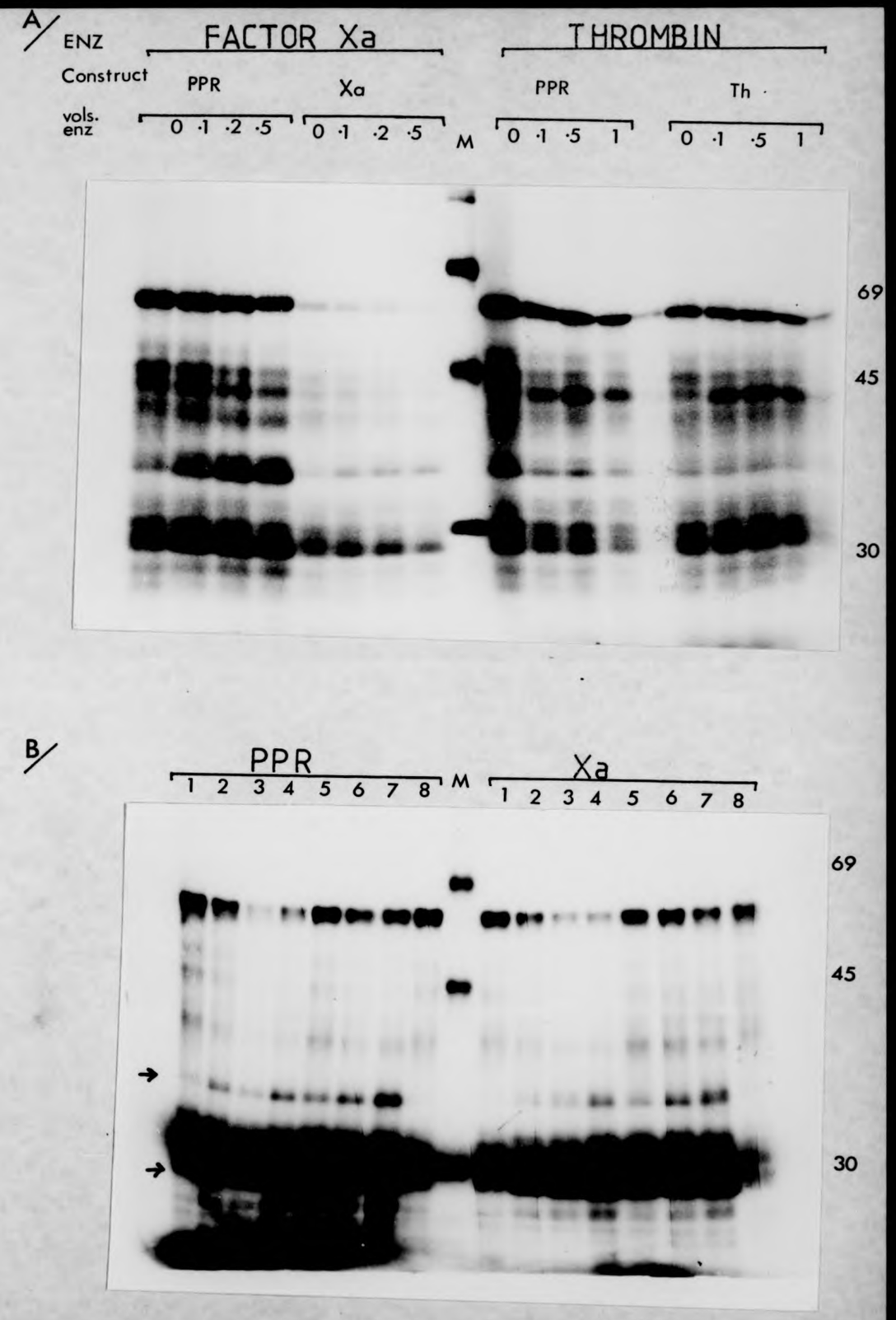
Radiolabelled preproricin and mutant preproricin synthesised in vitro in wheat germ lysate were then incubated with various quantities of Factor Xa and thrombin enzymes, to determine whether specific cleavage could be demonstrated in vitro. The activity of the Factor Xa enzyme was first checked by incubation of an aliquot with its physiological substrate prothrombin, followed by SDS-PAGE analysis. This confirmed that the enzyme was active. It was not possible to check the activity of the thrombin enzyme however, as SDS-PAGE analysis revealed that the commercially-available fibrinogen (the physiological substrate of thrombin) was heavily contaminated with processed fragments. The thrombin enzyme was therefore purchased, stored and used exactly as described by Smith and Johnson (1988).

Figure 8.2.3A shows the result of incubating preproricin, preproricinXa and preproricinTh with increasing amounts of the appropriate proteolytic enzyme. When preproricin and preproricinXa were digested with Factor Xa, two bands appeared which matched the sizes of fragments predicted for the specific cleavage of the linker (i.e. 34.5kD and 29.4kD). This

FIGURE 8.2.3 **PROTEOLYTIC CLEAVAGE OF IN VITRO SYNTHESISED PREPRORICIN MUTANTS.**

A/ ^{35}S-Methionine labelled Factor Xa (Xa) and thrombin (Th) preproricin mutants, synthesised in vitro in a wheat germ lysate, were incubated with different amounts of the appropriate enzyme for 60min at 26C in 1x TEN buffer. Wild type preproricin (ppr) was used as a control. Fragments were resolved by SDS-PAGE and visualised by fluorography.

B/ ^{35}S-Methionine labelled preproricin and Factor Xa preproricin mutant, synthesised in vitro in wheat germ lysate, were digested with 0.1 vol (lanes 2-4) or 0.05 vol (lanes 5-7) Factor Xa at 26C in 1x TEN buffer, for 10min (lanes 2 and 5), 30min (lanes 3 and 6), or 60min (lanes 4 and 7). Control digestions were performed in 1x TEN for 0min (lane 1) or 60min (lane 8)in the absence of Factor Xa in the absence of Factor Xa. Arrows refer to proteolytically-cleaved products.



was as expected for the mutant containing the Factor Xa recognition sequence. However, the wild-type control should not have been cleaved in the linker region. The predicted amino acid sequence of wild-type preproricin was then inspected around the linker region (see Appendix A), to identify arginine residues which may be cleaved 'non-specifically' by Factor Xa. These are listed below with the predicted sizes of the fragments produced by cleavage at each position.

Arg^{274}-Val#	34,500 + 29,500
Arg^{258}-Cys	32,500 + 31,600
Arg^{272}-Pro	34,300 + 29,700
Arg^{291}-Ile	36,800 + 27,300
Arg^{295}-Asn	37,300 + 26,800

(# = predicted site of Factor Xa cleavage in Xa mutant)

Cleavage at only one of these sites would result in fragments of the sizes seen in Figure 8.2.3A. However, trypsin-like proteases do not usually cleave peptide bonds where the carboxyl residue is proline. It is therefore unlikely that non-specific cleavage occurs at this position. Furthermore, cysteine, proline and asparagine residues are not found on the carboxyl side of known Factor Xa cleavage sites (see Nagai and Thogerson, 1984). Wearne (1990) has reported non-specific Factor Xa cleavage of Lys-X peptide bonds in fusion proteins. However, there are no lysine residues within 30 amino acid residues of the linker in either direction.

Factor Xa digestion of the wild-type and mutant was then followed over a time course to determine whether the specific cleavage at the Factor Xa recognition site in the mutant and the non-specific cleavage of the wild-type could be resolved (Figure 8.2.3B). However, there was no distinguishable difference between the rates of appearance of the fragments in either the wild-type or the mutant. This observation therefore remains unexplained.

When thrombin digestions were studied, no specific cleavage of the wild-type or thrombin mutant was observed at any concentration of the enzyme (Figure 8.2.3A). In tracks were relatively large amounts of the enzyme were used, deformation of the 64kD preproricin (or mutant preproricin) band was seen due to the high amounts of bovine serum albumin (66kD) present in the Sigma thrombin preparation.

In the highly reducing, membrane-free wheat germ *in vitro* expression system the proricin synthesised is unglycosylated, and has no disulphide bonding. It is likely therefore that its conformation is unlike native proricin from the plant. The linker region in the wheat germ-synthesised products may not be accessible to the protease enzymes. This may explain why the results obtained above were not as expected and why expression in *Xenopus* egg cell-free extract, where the preproricin should be glycosylated, disulphide-bonded and segregated into ER-derived microsomes, was desirable.

8.2.4 CLONING OF PREPRORICIN AND PREPRORICIN MUTANTS INTO pSP64T.

When preproricin transcripts, synthesised *in vitro* from pGEMPPRCL617, were translated in the *Xenopus* egg-free translation system, immunoprecipitated recombinant proricin was only visible as faint bands on autoradiographs after long film exposures (> one month at -70C). The fluorograph showing the optimisation of the egg cell-free translation system for the preproricin transcript (Figure 8.2.5A) is an example of the long exposures necessary to visualise bands when using pGEMppRCL617-derived transcripts (1 month exposure).

It was therefore decided necessary to re-clone the preproricin and preproricin mutant genes from the pGEM-derived constructs into the high expression transcription vector, pSP64T (Krieg and Melton, 1984: Appendix C for map). This

vector was introduced in Chapter 6, where the pSP64T-derived construct, pSP64Tproricin, was used to synthesise proricin *in vitro*. The vector has 5' and 3' untranslated regions of *Xenopus* beta-globin DNA that have been shown to increase the level of expression of genes cloned into the *Bgl*II site situated between them (Krieg and Melton, 1984).

In each case a 1.7kb *Xho*I / *Eco*RI fragment, containing the wild-type/mutant preproricin clone, was isolated from its pGEM-derived construct and the 5' and 3' cohesive termini blunted using the Klenow fragment of DNA polymerase. This fragment was then cloned into pSP64T, after *Bgl*II restriction and blunt-ending of the vector. Mini prep DNA was prepared from ampicillin-resistant colonies and a clone was selected which contained a single copy of the preproricin gene in the correct orientation for SP6 transcription. The restriction digest of a pSP64Tpreproricin clone is shown in Figure 8.2.4.

Finally, large scale caesium chloride purified plasmid preparations of each clone were prepared and plasmid sequenced to confirm the presence of the correct (wild-type or mutant) linker sequence.

8.2.5 TRANSLATION OF TRANSCRIPTS ENCODING PREPRORICIN AND PREPRORICIN MUTANTS IN *XENOPUS* EGG-FREE EXTRACT.

Transcripts were prepared from SP6 *in vitro* transcription of (*Eco*RI) linearised pSP64Tpreproricin/mutant constructs. These were then translated in egg cell-free extracts and, in each case, a radiolabelled recombinant product of the predicted size was identified after immunoprecipitation, SDS-PAGE and 1 week exposure to film (Figure 8.2.5). From this it was concluded that all constructs expressed a full-length polypeptide which could be visualised after relatively short exposure to film.

FIGURE 8.2.4 **CORRECT ORIENTATION OF PREPRORICIN CLONE PPRCL617, IN pSP64T.**

Restriction digest analysis of pSP64Tpreproricin (lanes 1-4) and pSP64T (lanes 5-8). 1µg of DNA was digested with HindIII (lanes 2 and 6), HindIII and ClaI (lanes 3 and 7), or EcoRI (lanes 4 and 8). Fragments were resolved on a 1% agarose (1x TBE) gel, identified by ethidium bromide staining and visualised on a UV-transilluminator. 1ug of uncut DNA was run as a control (lanes 1 and 5).

FIGURE 8.2.5 **SYNTHESIS OF PRORICIN AND FACTOR Xa/THROMBIN PRORICIN MUTANTS IN XENOPUS EGG CELL-FREE TRANSLATION SYSTEM.**

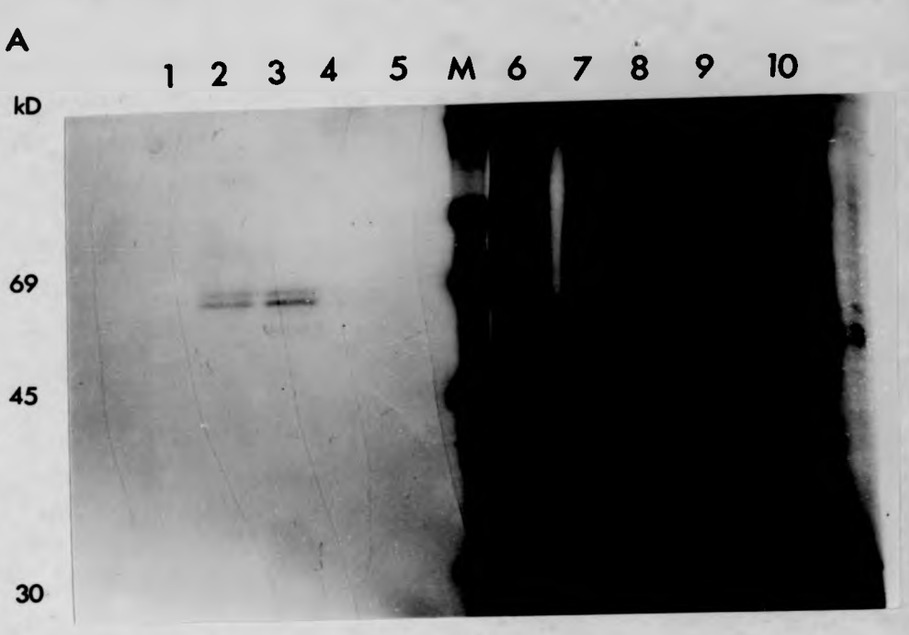
A/ Optimisation of system for transcript.

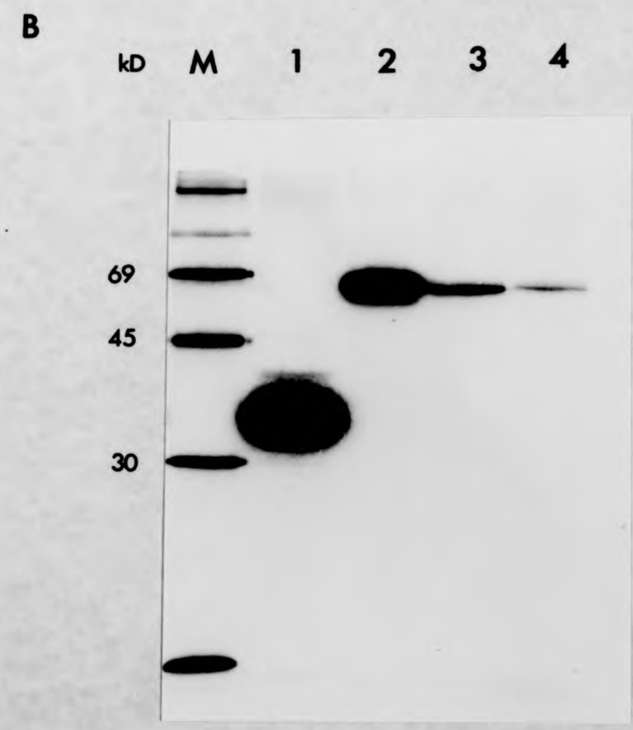
Preproricin transcript was translated in 30µl egg cell-free extract for 2 hours at 21C in the presence of 25µCi ^{35}S-methionine. Recombinant proricin was immunoprecipitated using rabbit anti-B chain antisera. Immunoprecipitated (lanes 1-5) and non-immunmoprecipitated (lanes 6-10) proteins were resolved by SDS-PAGE and visualised by fluorography.

0.2µg transcript: lanes 1 and 6
0.5µg " " : " 2 " 7
1.0µg " " : " 3 " 8
2.0µg " " : " 4 " 9
no RNA : " 5 " 10

B/ Egg cell-free translation of preproricin, Factor Xa and thrombin preproricin mutant transcripts.

1ug of transcripts, synthesised <u>in vitro</u> and encoding preB (lane 1), preproricin (lane 2), Factor Xa mutant preproricin (lane 3) and thrombin mutant preproricin (lane 4), were translated for 2 hours at 21C in 30µl egg extract in the presence of 25µCi ^{35}S-methionine. Proricin was immunoprecipitated using rabbit anti-B chain antisera; bands were resolved by SDS-PAGE and visualised by fluorography.





Although the same amount of RNA was added to each translation reaction, the levels of expression of the wild-type, the mutants and the preB control appeared different (compare intensities of bands in lanes 1,2,3 and 4 in Figure 8.2.5B). Since the 1μg of the preB transcript (1.0kb) contains almost twice as many molecules as the same amount of preproricin RNA (1.8kb) and also, as the rate of synthesis for a 30kD protein should be twice as fast as the the rate of synthesis for a 60kD protein, it is possible to account for the difference between the intensities of the preB and proricin bands. However, the order of expression, PPR > Xa > Th, could not be easily explained. In all further experiments, using different DNA and RNA preparations, this order of expression was maintained.

Another observation made when translating these RNAs in egg cell-free extracts, was that two immunoreactive bands of similar molecular weight were resolved (seen clearly in Figure 8.2.5A). This could either represent signal peptide-processed and unprocessed forms of preproricin or two different glycosylated forms of microsome-segregated proricin. This was investigated in more detail in the next section.

8.2.6 GLYCOSYLATION AND LECTIN ACTIVITY OF PRORICIN AND PRORICIN MUTANTS.

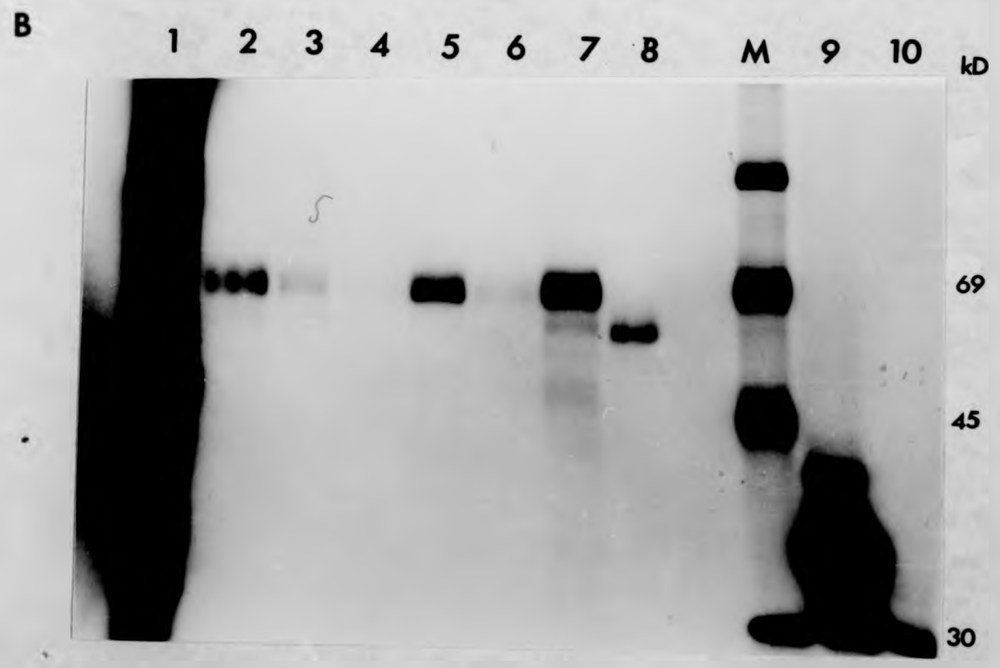
The preproricin and preproricin mutant transcripts encode proteins with a signal peptide which, in the case of wild-type preproricin expressed in Xenopus oocytes (Chapter 5), mediates its co-translational translocation across the ER-membrane where asparagine-linked glycosylation takes place, on the glycosylation sites in the A and B chains. The signal peptide is cleaved and the glycosylated proricin is sequestered in the ER-lumen. In this form, the oocyte-synthesised recombinant proricin was shown to possess lectin activity.

In order to determine whether the egg cell-free translation system co-translationally glycosylates the preproricin and preproricin mutants, one aliquot of the immunoprecipitated protein in each case, was digested with N-acetylglucosaminidaseH (endoH). This enzyme specifically removes all but the N-acetylglucosamine moiety from asparagine-linked core oligosaccharide side chains on glycoproteins such as those sequestered in the ER lumen. When the endoH-treated immunoprecipitated protein was analysed by SDS-PAGE and compared to untreated protein, a shift was observed (see lanes 7 and 8: Figure 8.2.6). As was observed in section 8.2.5 above, the immunoprecipitated protein translated from the preproricin (mutant) transcript could be resolved into two bands. However, in endoH-treated tracks, a single band was observed. This band was lower in molecular weight than the bands in the untreated sample, indicating that the latter represented different glycosylated forms of proricin and not, as was suggested above, processed versus unprocessed forms of preproricin. The wild-type, Factor Xa and thrombin mutants appear to be co-translationally translocated across the ER-membrane and N-glycosylated in this system. Protease protection experiments would have been necessary to confirm this.

In order to test the lectin activity of the proricin and proricin mutants synthesised in the egg cell-free extract, the translation mix was diluted in oocyte homogenisation buffer (to solubilise the microsomal membranes) before loading onto a SeLectin^{TM} column. This column contains lactose immobilised onto polyacrylamide beads and was used for the single-step purification of recombinant proricin synthesised in oocytes (Chapter 5). Unbound proteins were eluted and the column washed thoroughly in buffer, before eluting any bound protein with 100mM lactose.

FIGURE 8.2.6 **LECTIN ACTIVITY AND GLYCOSYLATION OF RECOMBINANT PRORICIN AND PRORICIN MUTANTS SYNTHESISED IN XENOPUS EGG CELL-FREE EXTRACT.**

Labelled protein from a 60µl egg-free translation reaction, containing 2µg transcript and 50µCi ^{35}S-methionine, was diluted in oocyte homogenisation buffer and passed three times down a 2ml SeLectin^{TM} column. The column was then washed with 3x 0.5ml oocyte homogenisation buffer before eluting the bound protein with 2x 0.5ml 100mM lactose. Proricin (gel A), Factor Xa mutant proricin (gel B), and thrombin mutant proricin (gel C) was immunoprecipitated from each fraction using rabbit anti-B chain antisera, resolved by SDS-PAGE and visualised by fluorography (lanes 1= 10% of total unbound protein, before immunoprecipitation; 2= 90% of total unbound protein, after immunoprecipitation; 3= 1st fraction of wash; 4= 2nd fraction of wash; 5= 1st fraction of 100mM lactose wash; 6= 2nd fraction of lactose wash). Also, immunoprecipitated recombinant proricin from a 30µl egg-free translation was divided into two equal aliquots and half was incubated overnight with 2.5mU endoH before loading on the gel (lanes 7=untreated; 8=endoH treated). An aliquot of immunoprecipitated recombinant B chain was also endoH treated as a control (gel B: lanes 9=untreated; 10=endoH treated).



C

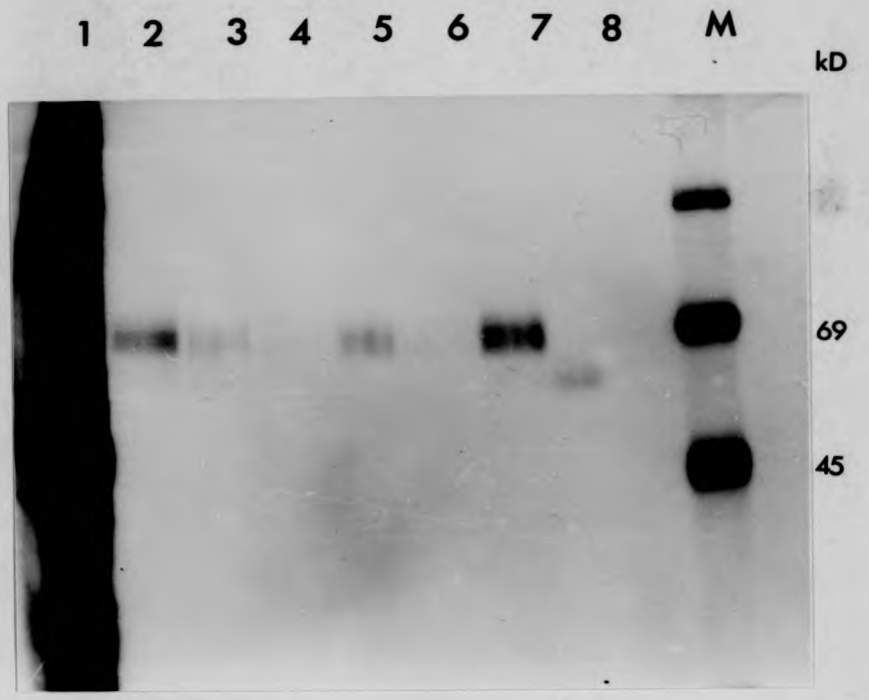


Figure 8.2.6 shows the total unbound protein (lane1) and the proricin immunoprecipitated from each of the fractions eluted from the column (lanes 2-6). In each case, approximately 50% of the total immunoprecipitated (mutant) proricin did bind to the column and was eluted in the 100mM lactose. It was therefore concluded that 50% of the proricin, Factor Xa mutant and thrombin mutant synthesised in the egg-free extract possessed a functional B chain with sugar-binding (lectin) properties.

8.2.7 SPECIFIC PROTEOLYTIC CLEAVAGE OF RECOMBINANT PRORICIN Xa MUTANT.

Sufficient amounts of the radiolabelled recombinant proricin Xa mutant were synthesised in the egg cell-free translation system to enable a digestion of the immunoprecipitated protein with Factor Xa. Therefore, both wild-type and Xa mutant were digested for 60min with various amounts of Factor Xa. Figure 8.2.7 shows that the mutant Xa proricin was specifically cleaved into two fragments, one of which was the same size as glycosylated B chain (Figure 8.2.7: lane 9). Some non-specific cleavage was identified in both wild-type and mutant tracks, but the two strongest bands in the Xa mutant tracks were unique to this mutant, of the predicted size for ricin A and B chains, and were therefore caused by specific cleavage at the Factor Xa recognition site in the mutated linker region.

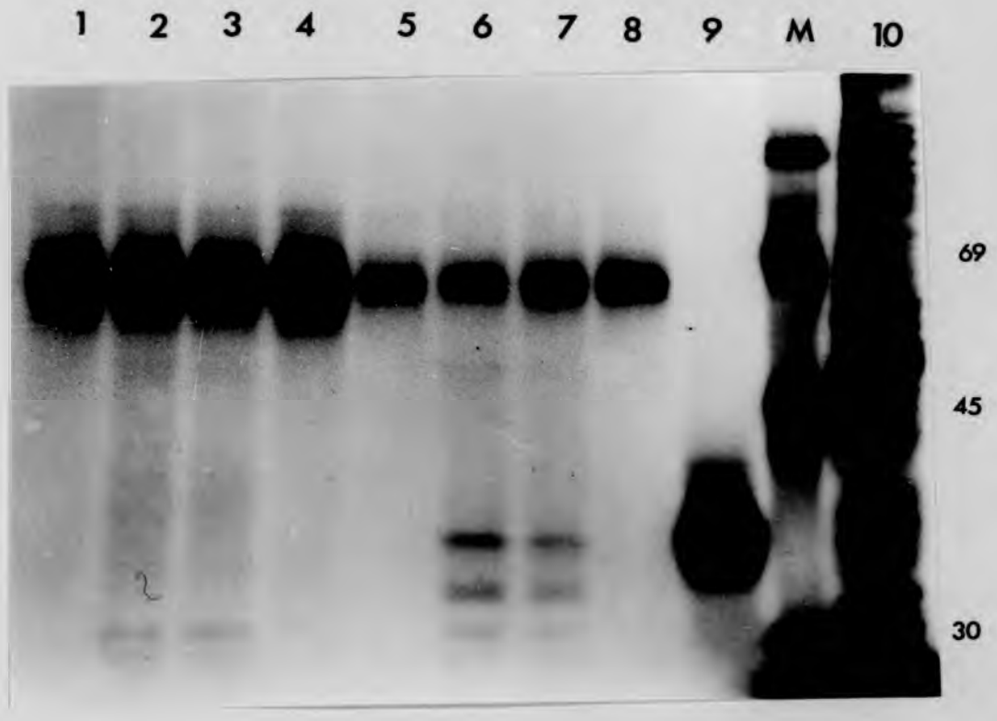
Time did not allow for any further investigations.

8.3 DISCUSSION.

The Factor Xa and thrombin mutant linkers were cloned successfully into a full-length preproricin clone, and wild-type/Xa/Th preproricin were all expressed *in vitro*. Translation in rabbit reticulocyte lysate was not successful,

FIGURE 8.2.7 **SPECIFIC CLEAVAGE OF IMMUNOPRECIPITATED RECOMBINANT FACTOR Xa MUTANT PRORICIN.**

Immunoprecipitated recombinant proricin (lanes 1-4) and Factor Xa mutant proricin (lanes 5-8), synthesised in a Xenopus egg-free system, was incubated with 0.1x volumes (lanes 2 and 6), 0.2x volumes (lanes 3 and 7) or 0.5x volumes (lanes 4 and 8) Factor Xa, for 60min. at 26C in 1x Factor Xa buffer. Immunoprecipitated protein was also incubated as above but in the absence of any enzyme (lanes 4 and 8). Fragments were resolved by SDS-PAGE and were visualised by fluorography. Immunoprecipitated recombinant ricin B chain synthesised in the same system (lane 9) and an aliquot of non-immunoprecipitated proteins from the factor Xa translation (lane 10) was also loaded.



presumably due to the release of active A chain-like fragments during the translation of the preproricin transcripts, which inactivated the rabbit ribosomes. Translation in wheat germ lysate was successful however, as wheat germ ribosomes are known to be 1000x fold more resistant to RIPs than mammalian ribosomes. In the wheat germ lysate wild-type and mutant preproricin polypeptides, all identical in size, were synthesised. This confirmed that the cloning and mutagenesis steps had not introduced any opportunistic stop codons in the mutants.

When the preproricin mutants were incubated *in vitro* with the appropriate endoprotease (Factor Xa or thrombin), only in the preproricin Xa mutant was any proteolysis observed. Bands of the predicted size for cleavage of the linker region (34kD and 29kD) were observed. However, Factor Xa cleavage of the wild-type preproricin was also observed under the same conditions which yielded fragments indistinguishable in size from the fragments released from the Factor Xa digestion of the Xa mutant. It was suggested that non-specific Factor Xa cleavage of the wild-type preproricin was caused because the conformation of the wheat germ synthesised product is likely to differ to the native precursor, possibly occluding access of the enzyme to the linker and exposing other regions of the protein. It was hoped that translation of preproricin mRNA in an expression system containing ER-derived membranes, such as the *Xenopus* egg cell-free extract, would produce a proricin protein that was folded into a more native-like conformation.

Expression of the wild-type preproricin transcript in a *Xenopus* egg cell-free translation system produced a labelled polypeptide of the expected size for glycosylated proricin. However, the levels of expression were very low. It is possible that this was due to the release of active A chain moieties which inactivated the *Xenopus* ribosomes. However,

other workers have reported dramatic increases in gene expression in Xenopus systems by re-cloning of genes from pGEM-derived vectors into pSP64T (Krieg and Melton, 1984). This vector contains the 5' and 3' untranslated regions from Xenopus beta-globin, which are thought to stabilise transcripts of foreign genes cloned into the vector. It was therefore decided to re-clone the wild-type/Xa/Th preproricin genes from pGEMppRCL617 into pSP64T.

The wild-type and mutant preproricin clones were successfully re-cloned into pSP64T. They were plasmid sequenced to confirm the sequences of the linkers in each construct. Transcripts were synthesised in vitro and translated successfully in the egg cell-free extract. As predicted, the proricin immunoprecipitated from this experiment could be visualised after 1-7 days exposure, as opposed to the one month exposures required for visualisation of the immunoprecipitated products encoded by pGEM-derived transcripts. No direct comparisons were made between the levels of proricin expression from the two types of transcript, however.

In the egg cell-free system experiments the bands visualised for the wild-type and mutant proricin polypeptides were identical in size, but not intensity. The immunoprecipitated Xa preproricin mutant was more intense than the Th preproricin mutant, but was less intense than the wild-type. However, both the wild-type and the mutants were N-glycosylated and possessed B chain lectin activity. This is evidence that the amino acid sequence of the Xa and Th mutant linkers did not disrupt the functional folding of the B chain.

Only in the case of the Xa preproricin mutant was expression sufficient to enable a preliminary investigation of the specific Factor Xa cleavage of the mutant linker. The mutant was cleaved into two fragments of approximately 34kD and 32kD by Factor Xa digestion with two out of the three enzyme

volumes used. In the Factor Xa digestions where fragments were visualised, the 34kD fragment appeared the same size as B chain translated from the preB transcript in the same egg cell-free system. At the highest enzyme concentration used (0.5volumes added) no fragments were seen after one hour's incubation. No explanation for this result has been given.

No corresponding fragments of 32kD and 34kD were visualised when the immunoprecipitated wild-type proricin was digested with Factor Xa under the same conditions. From this it was concluded that the fragments were produced by specific cleavage of the Xa proricin mutant. This result also shows that the non-specific cleavage of the wild-type proricin observed in the protease digestions of the wheat germ-synthesised polypeptide, can be resolved from the specific cleavage of the Xa proricin mutant in this system. This lends weight to the argument that the conformation of proricin synthesised is different in the two systems.

The problems encountered with the low levels of expression of the wild-type/mutant preproricin transcripts in this system, requiring time-consuming re-cloning of the clones from pGEM1 into an alternative *in vitro* transcription vector (pSP64T), restricted the endoprotease studies to the single experiment shown in Figure 8.2.7. The Factor Xa specific cleavage of the Xa proricin mutant needs further characterisation. For instance, there is currently no direct evidence to support the assumption that the fragments observed by digestion of the mutant are indeed the A chain and B chain subunits. Furthermore, it is desirable to show that these fragments are disulphide-linked, as is the case in mature, native ricin. Also, the protease digestion was performed on immunoprecipitated protein: the bound antibody molecules may in fact hinder the access of the enzyme to the linker region and so limit the rate of cleavage. It may be preferable to digest

the proricin mutant before immunoprecipitating any of the fragments produced.

A possible strategy for investigating these questions is to pass the crude egg cell-free translations down a SeLectin^{TM} (as in Section 8.2.6). In this way the bound Xa mutant proricin is purified from the other labelled proteins and can be eluted in 1x TEN buffer made 100mM lactose. The purified mutant can then be digested with Factor Xa, before dividing the reaction in three. Two aliquots are reduced by addition of beta-mercaptoethanol before immunoprecipitating with either anti-A chain or anti-B antisera. The immunoprecipitated polypeptides could then be resolved by SDS-PAGE. The third aliquot is immunoprecipitated with a mixture of anti-A chain and anti-B chain antisera and then aliquots of the immunoprecipitated protein resolved by SDS-PAGE under reducing and non-reducing conditions.

Such an approach would identify the fragments produced as A chain-like or B chain-like, since each antiserum should only immunoprecipitate its corresponding fragment. Also, by running immunoprecipitated aliquots under reducing and non-reducing conditions, any fragments which are linked by a disulphide bond will only be resolved from the full-length uncleaved proricin, under reducing conditions.

This approach might also enable an investigation into the thrombin-specific cleavage of the Th proricin mutant synthesised in the egg cell-free system. This mutant was only expressed at very low levels (compare Figure 8.2.6; A and C). However, selective binding to a SeLectin^{TM} column would enable the Th proricin mutant protein synthesised from a number of translations to be concentrated. This would allow further investigations to be performed.

CHAPTER 9

GENERAL DISCUSSION.

Intracellular protein sorting has been an area of intense study for cell biologists in recent years. The yeast Saccharomyces cerevisae has been the model system, with its useful genetics enabling the identification and characterisation of components of the sorting pathway such as BiP, ER retention signals and vacuolar targeting signals. This raised the possibility of looking for a plant vacuolar targeting signal by expression of a plant storage protein in transgenic yeast. However, work by Chrispeels and his colleagues on the targeting of PHA/invertase hybrids in transgenic yeast suggests that the sorting machinery in higher plants has evolved such that the two systems are not compatible. The answer has been to use transgenic tobacco to delineate the elusive plant vacuolar targeting signal.

Two regions of plant proteins have been focussed on. Chrispeels and Tague (1990) suggest that regions within the mature domains of some plant storage proteins which resemble the tetrapeptide QRPL yeast vacuolar targeting signal may be involved in targeting to the vacuole. In support of this, Sabastiani et al (1991) have most recently identified 9 regions of primary sequence homology within the mature domains of the 7S-lectins (one of which was the region identified by Chrispeels and Tague) which, they suggest, may combine in the folded protein to form a 3D signal patch.

On the other hand, the precursor forms of sporamin (Matsuoka and Nakamura, 1991), barley lectin (Bednarek et al, 1990), wheat germ agglutinin (Chrispeels and Raikhel, 1991) and the basic isoforms of B-1,3-glucanases from tobacco (Shinshi et al, 1988) all have propeptides which have been implicated in vacuolar targeting. This would suggest that the vacuolar targeting signal is not held within the mature domain of these proteins. Interestingly, work at Warwick and in Milan on the

type I RIPs dianthin-30 (from carnation leaves) and saporin SO-6 (from the leaves of Saponaris officialis) has recently identified C-terminal propeptides which bear a striking resemblance to the propeptides of barley lectin and tobacco B-1,3-glucanase (Giuseppe Legname and Marco Soria, personal communications). It has yet to be established whether dianthin-30 and saporin SO-6 are localised in the vacuole, but it is tempting to speculate that their C-terminal extensions contain vacuolar targeting information.

Ricin and other type II RIPs have evolved by association of ancestral genes encoding a type I RIP and a lectin-like protein. The homology comparisons of Chrispeels' and Sebastiani's groups on plant storage proteins would suggest that the vacuolar targeting information is held within the mature protein. However, the close homologies of type I RIPs (to which ricin A chain is evolutionarily related) to other vacuolar proteins favours a propeptide-localised signal. Evidence for the presence of an N-terminal propeptide on proricin has been presented in this thesis. Furthermore, 22aa residues of the N-terminal presequence from preproricin were sufficient to function as an ER-signal peptide when fused to npt2 and expressed in vitro. However, the full 35aa residue ricin presequence did not cause retention of npt2 when ricin/npt2 chimeras were expressed transiently in tobacco plant protoplasts. This data would suggest that the ricin presequence alone does not function as a vacuolar targeting signal. Transient expression studies are currently underway at Warwick using presequence-A chain constructs to determine whether the vacuolar targeting signal is held within the A chain domain of proricin.

Plant storage proteins are transported through the endomembrane system in a precursor form, maturation only

occurring after deposition at their final subcellular destination - the protein bodies. The processing of many plant storage protein precursors is thought to involve asparagine-specific proteolytic cleavage. An assay was therefore designed to enable purification of the proricin processing enzyme from castor bean. The results showed that purification of the enzyme using this system was not possible however; the dry castor bean seeds containing low levels of activity and high levels of toxic endogenous ricin (which was co-purified with the enzyme activity in the protein body matrix fraction). The *in vitro* synthesised proricin also was sub-optimal for use as an enzyme substrate. SDS-PAGE analysis showed that it contained radioactive bands covering the whole range of molecular weights up to the full-length proricin polypeptide.

Questions were also raised as to the use unglycosylated, non-covalently folded proricin as a substrate for the processing enzyme. It is unlikely that the conformation of proricin synthesised in this way resembles the conformation of native proricin, which folds in the oxidising environment of the ER lumen. Support for this comes from another area of the thesis. The Factor Xa proricin mutant, synthesised in wheat germ lysate, appeared to be cleaved non-specifically when digested with Factor Xa (both the wild-type and mutant proricin were cleaved: see Figure 8.2.3). However, when the wild-type proricin and Factor Xa mutant were synthesised in *Xenopus* egg cell-free extract, preliminary data showed Factor Xa-specific cleavage of the mutant only. This demonstrates that the synthesis of the same polypeptide in different systems can alter its use as an enzyme substrate.

Nishimura and colleagues have reported the purification of a storage protein processing enzyme (Nishimura and Nishimura, 1991). The publication of a purification

strategy and details of the characteristics of this enzyme are eagerly awaited.

Expression of a recombinant proricin in heterologous systems was hindered by the potent toxicity of ricin. Although recombinant proricin *per se* was shown to possess no RNA N-glycosidase activity when expressed in *Xenopus* oocytes (the B chain was shown, on the other hand, to be biologically active), A chain-sized fragments were produced which were active and depurinated rabbit reticulocyte ribosomes. The inability to express preproricin/proricin transcripts successfully in rabbit reticulocyte lysate has been attributed to the inadvertent synthesis of toxic A chain-like fragments from the (pre)proricin message which inactivate the ribosomes and prevent further translation of the full-length precursor moiety. Finally, the potential for the release of active forms of a recombinant type II RIP, restricted the *E.coli* expression studies of proricin to HSE Category IV conditions. This ultimately prevented a thorough study into the potential of *E.coli* periplasmic expression for production of recombinant proricin and mutant proricin proteins.

The ability of the A chain to stabilise the B chain in aqueous solution made expression of a recombinant, linker-cleavable proricin desirable. Expression of preproricin in *Xenopus* oocytes provides a system whereby the recombinant protein can be N-glycosylated and segregated into the oxidising environment of the ER-lumen. Using expression in a *Xenopus* egg cell-free extract it was possible to synthesise proricin mutants, with linkers containing either a Factor Xa or thrombin recognition sequence. These proteins were glycosylated in a similar way to wild-type proricin and the B chain domains of both mutants possessed lectin activity. Unfortunately, in order to increase the expression of the preproricin and preproricin

mutant clones in this system to a workable level, it was necessary to perform time-consuming cloning steps to transfer the preproricin (mutant) clone from a pGEM-derived vector into the high expression vector pSP64T. However, a preliminary experiment on the recombinant Factor Xa proricin mutant synthesised in the egg cell-free system suggests that specific proteolytic cleavage of the linker is possible. Further experiments are in progress to further characterise these proricin mutants in order to assess their potential use for the synthesis of recombinant toxins.

APPENDIX A

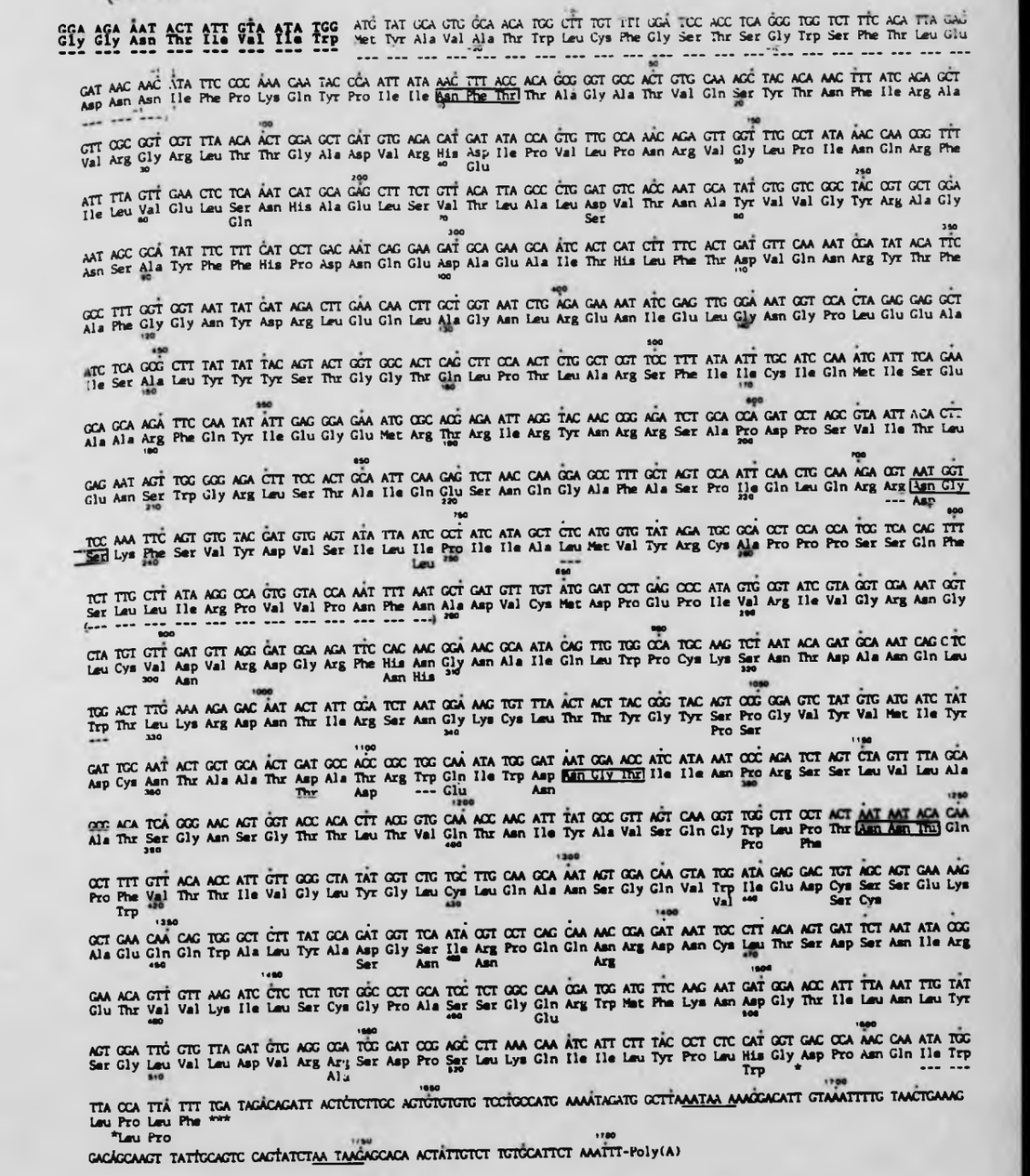
cDNA SEQUENCE AND DEDUCED PROTEIN SEQUENCE OF PREPRORICIN.

(original sequence from Lamb et al, 1985)

The cDNA sequence of preproricin, published by Lamb et al (1985), was found to be an incomplete clone upon comparison of the sequence to that of a preproricin genomic clone (Halling et al, 1985). The cDNA clone lacked a single methionine (initiation) codon at its 5' end.

The amended cDNA sequence, shown opposite, encodes a single polypeptide precursor form of ricin (preproricin). The deduced primary amino acid sequence is shown below the DNA sequence. This precursor consists of a 35 amino acid residue N-terminal region, a 267 amino acid residue A chain region, 12 residue linker peptide and a 241 amino acid residue region corresponding to the mature B chain. The N-terminal region and 12 amino acid linker peptide (both highlighted by dashed lines in parentheses) are not present in the mature protein. Potential sites for Asparagine-linked glycosylation are boxed.

All clones described in the text as "preproricin" contain the full-length gene described here and thus translation is initiated from the ATG codon at position -105. The construct, pSP64Tproricin (used as a source of substrate for the protein body-localised castor bean endoprotease) contains an incomplete 5' region beginning at position -75 and, therefore, expression of this derivative of preproricin uses the ATG codon at position -72 to initiate translation. The construct pINIIIompA2proricin (used for E.coli expression of proricin) contains an incomplete cDNA clone beginning at position +10. It therefore lacks the entire N-terminal region and the first 3 residues of the A chain.



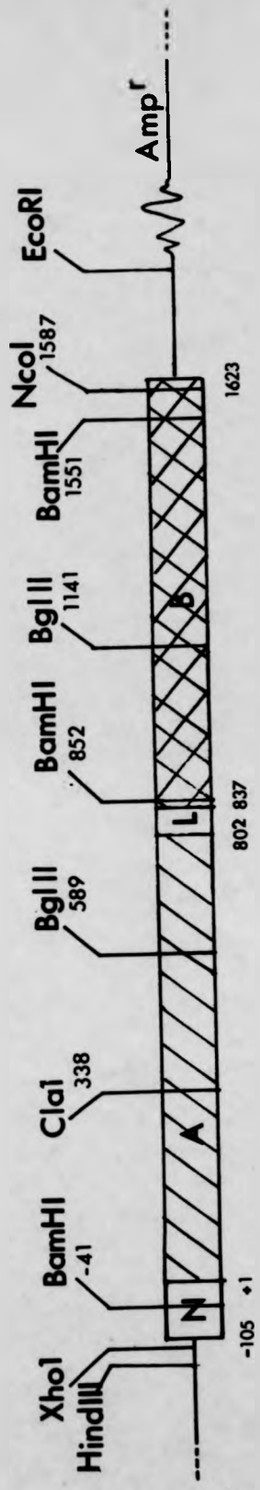
APPENDIX B

RESTRICTION MAP OF pPRCL617

The construct, pPRCL617, contains the amended preproricin cDNA sequence shown in Appendix A (Richardson et al, 1989). This clone was used for the source of preproricin cDNA. The position of restriction sites are shown by numbers which correspond to the nucleotide numbering of Lamb et al (1985) (see Appendix A).

KEY.

N - 35 amino acid residue N-terminal coding region
A - mature A chain coding region
L - 12 amino acid residue coding region
B - mature B chain coding region



APPENDIX C

VECTORS.

A. pSP64T (Krieg and Melton, 1984)

This vector is used for in vitro expression of cloned genes. The gene of interest is cloned into the unique BglII restriction site (in a clockwise orientation with reference to the map opposite). The DNA is then linearised using a suitable restriction enzyme site in the multiple cloning region (MCS) and transcription fired from the SP6 promoter (SP6). The 5' and 3' untranslated regions of Xenopus B-globin mRNA (shaded) stabilise the transcript and result in a high expression of the cloned gene. This can dramatically improve expression of genes in Xenopus oocyte/egg-free extract translation systems. The vector carries the B-lactamase gene for Ampicillin resistance (Amp^R).

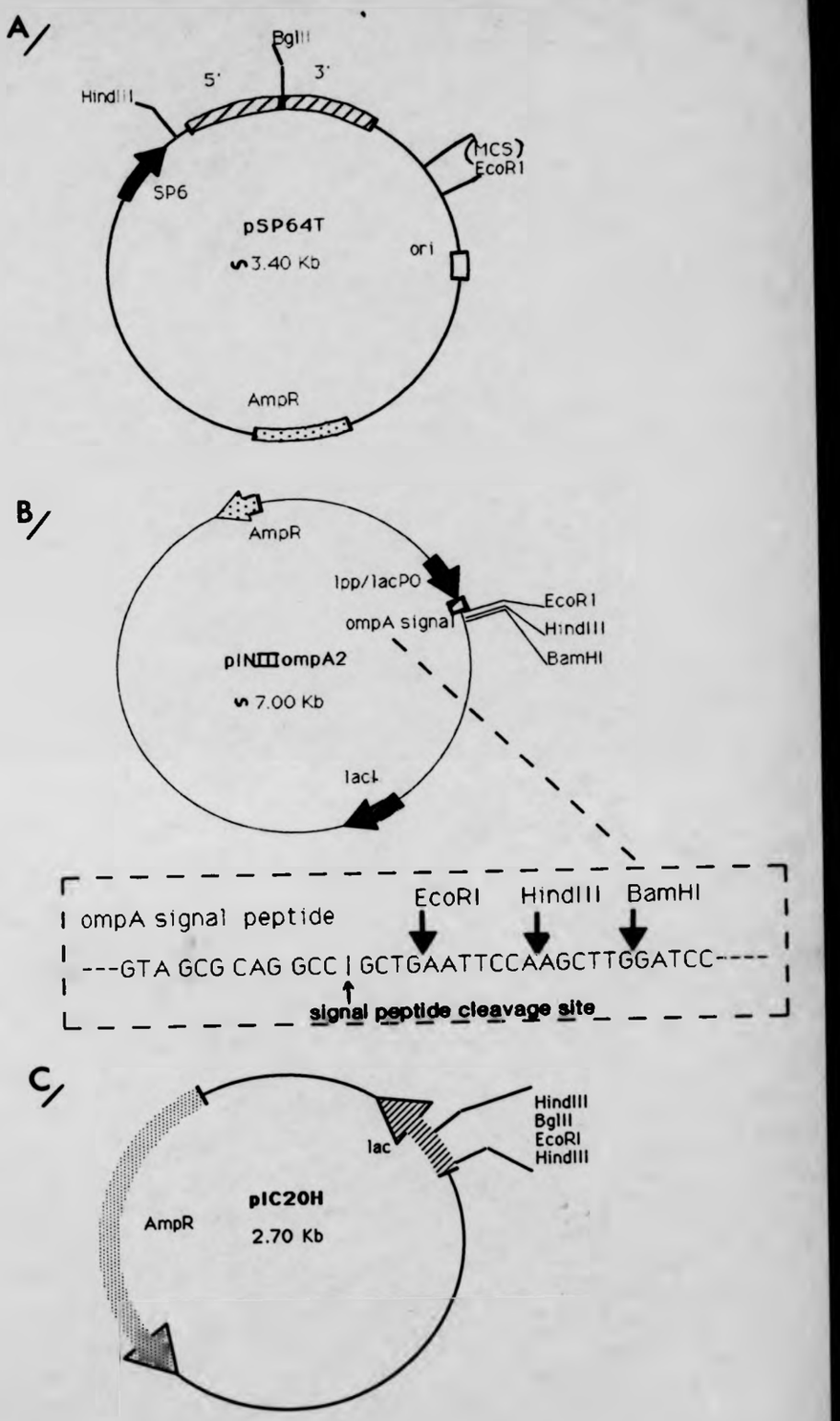
This vector was used to create the constructs 'pSP64Tpreproricin', 'pSP64TXa', and 'pSP64TTh'. May (1988) also used this vector to create 'pSP64Tproricin' which was used by this author as a source of proricin substrate to assay for the protein body-localised castor bean endoprotease.

B. pINIIIompA2 (Ghrayeb et al, 1984)

This vector is used for expression of foreign proteins in E.Coli. It contains a region encoding the signal peptide of the E.coli ompA protein followed by three unique restriction sites. The gene of interest is cloned in-frame with the ompA signal sequence by insertion at one of these three unique restriction sites to create a chimeric gene. The expression of this chimera is under the control of the lipoprotein promoter/lactose operon (lpp/lac PO) and, when expression is induced by IPTG, the fusion protein produced will be secreted into the periplasmic space. Khalid Hussain (Warwick) cloned an incomplete proricin clone, pRCL617ox16, beginning at the fourth codon of the mature A chain (and hence lacking the entire N-terminal region and first 3 residues of A chain) into the EcoRI site to create the construct pINIIIompA2proricin. This construct was used by this author to investigate expression of proricin in E.coli.

C. pIC20H (Marsh et al, 1984)

This vector contains a unique BglII restriction site in the multiple cloning site. This vector was used as part of the cloning strategy for mutation of the proricin linker peptide to create Factor Xa and thrombin recognition sequences.



APPENDIX D

COMPUTER PROGRAMME FOR FINDING LIKELY SITES OF SIGNAL PEPTIDE CLEAVAGE

The computer programme shown opposite was written in BBC Basic language and it was run on a BBC microcomputer. The programme allows the user to enter a sequence of amino acid residues using either the one or three-letter codes, beginning at the N-terminal methionine residue of the protein. Signal peptides described in the literature to date have been between 20 and 35 residues long (von Heijne, 1990) and therefore it is usually sufficient to enter the first 35-40 N-terminal residues. Upon the command [return key], the programme will assign values to each residue entered using the von Heijne algorithm (von Heijne, 1986). The programme calculates this value by reference to the two-dimensional array, z(15,20), which stores a probability table assembled from von Heijne's paper (von Heijne, 1985) and given to this author by Dr Les Dale, University of Birmingham. The final value assigned to each residue is a summation of 15 separate probabilities and values are only assigned to residues from residue^{13} onwards. Residues assigned high positive values are considered good candidates for the sites of signal peptide cleavage, and visa versa.


```
600 DEF PROCCODE
610 FOR N=1 TO 20
620 O$(N)=W$(N)
630 NEXT
635 N=0
640 ENDPROC
700 DEF PROCCHANGE
710 INPUT"WHICH AA DO YOU WISH TO ALTER".H
720 PRINT A$(H)
730 INPUT"NEW RESIDUE",A$(H)
740 INPUT"CHANGE ANOTHER AA (Y/N)"B$
750 IF B$="Y" THEN GOTO 710
760 ENDPROC
1000 DATA A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y
1005 DATA ALA,CYS,ASP,GLU,PHE,GLY,HIS,ILE,LYS,LEU,MET,ASN,PRO,GLN,ARG,SER,THR,V
AL,TRP,TYR
1010 DATA 0.10,-0.11,-0.40,0.04,0.32,0.22,0.22,0.16,0.54,0.04,1.18,-0.88,1.71,0
.22,-0.88
1020 DATA -0.41,0.29,0.69,0.44,0.69,1.13,0.29,0.58,0.11,0.29,1.44,-0.41,0.69,0.
58,-0.41
1030 DATA -2.19,-2.19,-2.19,-2.19,-2.19,-2.19,-2.19,-2.19,-2.30,-1.20,-0.36,-5.08,-0.
58,-5.08,0.12,0.21
1040 DATA -2.30,-2.30,-2.30,-2.30,-2.30,-2.30,-2.30,-2.30,-0.34,-0.11,-5.08,0.84,-5.08,
36,-5.08,0.26,0.34
1050 DATA 0.84,0.47,0.68,0.68,0.07,0.22,1.17,0.84,-0.34,-0.11,-5.08,0.84,-5.08,
2.07,-0.34
1060 DATA -1.11,-1.11,-1.39,-0.70,-1.39,0.07,-1.39,-1.80,0.45,1.03,-0.88, -0.53
.1.17,-0.19,-0.53
1070 DATA -1.22,-1.22,-1.22,-1.22,-1.22,-1.22,-1.22,-1.22,0.39,-1.22,-5.08,0.57
.-5.08,0.16,-0.53
1080 DATA 0.71,0.71,0.08,-0.21,0.40,-0.39,-0.62,0.08,-0.39,-2.00,0.30,-0.39,-5.
08,0.08,-0.06
1090 DATA -2.42,-2.42,-2.42,-2.42,-2.42,-2.42,-2.42,-2.42,-2.42,-1.04,-5.08,-1.
73,-5.08,-0.03,-0.23
1100 DATA 1.77,1.73,1.78,1.88,1.86,1.31,1.67,1.40,-0.19,0.64,-0.41,0.50,-2.49,-
0.41,-1.11
1110 DATA -0.99,0.11,0.95,0.39,-0.99,0.80,-0.30,-0.30,-0.99,-0.99,-5.08,-0.99,-
5.08,-0.99,-0.30
1120 DATA -1.96,-1.96,-1.96,-1.96,-1.96,-1.96,-1.96,-0.86,-0.86,-5.08,0.3
4,-5.08,-0.57,-0.01
1130 DATA-1.31,-2.00,-1.31,-2.00,-2.00,-2.00,-0.62,-2.00,0.08,0.99,0.64,-5.08,-2.00,-
0.90,-2.00,1.09
1140 DATA -1.84,-1.84,-1.84,-1.84,-1.84,-1.84,-0.05,-1.84,-1.84,0.46,0.24,-5.08,1.05,
-0.74,1.10,0.46
1150 DATA -1.34,-2.03,-2.03,-2.03,-2.03,-2.03,-2.03,-2.03,-0.08,-0.64,-5.08,0.6
8,-5.08,0.46,0.17
1160 DATA -0.24,-1.34,-0.35,-0.64,0.13,0.88,0.27,0.34,0.82,-0.04,0.70,0.40,0.56
.0.27,-0.13
1170 DATA -1.58,0.03,0.66,-0.89,-0.66,0.29,-0.33,-0.33,0.21,-0.48,0.56,-0.19,-0
.48,-1.17,0.03
1180 DATA 0.59,0.81,0.30,0.48,0.16,0.30,0.01,0.89,-2.41,0.08,1.06,-1.31,-5.08,-
0.33,0.43
1190 DATA 0.80,0.51,0.51,-0.59,-0.59,0.11,1.20,0.51,-0.59,0.51,-5.08,1.61,-5.08
.0.11,-0.59
1200 DATA -1.72,-1.72,-0.34,-1.72,-1.72,-1.72,-0.62,-1.72,-1.72,-1.03,-5.08,-0.
11,-5.08,-1.72,0.22
```

```
L.
  1 RESTORE
  5 CLS
 10  DIM A$(100)
 11 DIM S(100)
 20 C=0
 30 N=0
 31 DIM O$(20)
 35 FOR O=1 TO 20
 36 READ O$(O)
 37 NEXT
 40 DIM W$(20)
 41 FOR W=1 TO 20
 42 READ W$(W)
 43 NEXT
 50 DIM Z(15,20)
 60 FOR Y=1 TO 20
 70 FOR X=1 TO 15
 75 READ Z(X,Y)
 80 NEXT X
 85 NEXT Y
100 INPUT"ENTER SEQUENCE NAME".S$
110 INPUT"1 OR 3 LETTER AA CODE".C
120 IF C=3 THEN PROCCODE
150 CLS
160 PRINT "NO. OF AA ENTERED = ";N
165 N=N+1
170 INPUT "ENTER A.A OR PRESS RET. TO ANALYSE ";A$(N)
180 IF A$(N)<>""THEN GOTO 150
190 N=N-1
200 M=N
215 PRINT S$
220 FOR N=1 TO M
230 PRINT A$(N);" ";
240 NEXT N
245 PRINT""
250 INPUT "OK? (Y/N)";R$
260 IF R$="N" THEN PROCCHANGE
270  CLS
280 IF R$="N" THEN GOTO 215
300 PRINT " PLEASE WAIT.CALCULATION BEING MADE! "
310 FOR C=13 TO M-2
320 FOR X=1 TO 15
330 N=C+(X-13)
340 PROCAAC
350 S(C)=S(C)+Z(X,Y)
360 NEXT X
370 NEXT C
380 VDU2
385 @%=131594
390 PRINT S$
400 FOR N=13 TO C-1
410 PRINT N;" ";A$(N),S(N)
420 NEXT N
430 VDU3
499 END
500 DEF PROCAAC
510 O=0
511 O=O+1
520 IF A$(N)<>O$(O) THEN GOTO 511
530 Y=O
```

APPENDIX E

SUPPLIERS: NAMES AND ADDRESSES

All chemicals and reagents were obtained form BDH Chemicals Ltd, Poole, Dorset, UK and were of AnalaR grade, except for those listed below.

Amersham International plc, Aylesbury, Bucks., UK.
Radiochemicals; Amplify; biotin-steptavidin blotting kit; Hybond-C nitrocellulose membranes; restriction enzymes; T4 DNA ligase; T4 polynucleotide kinase; T4 DNA polymerase; E.coli polymerase 1 (Klenow fragment); lambda DNA; T7 RNA polymerase.

BDH Chemicals Ltd, Poole, Dorset, UK.
1-Naphthyacetic acid (NAA).

Beckman RIIC Ltd., High Wycombe, Bucks., UK.
Ready-solv EP liquid scintillation fluid and quickseal vials.

Biorad Laboratories Ltd., Watford, Herts., UK.
Low molecular weight protein standards for SDS-PAGE.

Boehringer Mannheim (BCL), Lewes, East Sussex, UK.
Deoxy- and dideoxy-nucleotides; ribonucleotides; calf intestinal alkaline phosphatase (cip); creatine kinase; creatine phosphate; SP6 polymerase; 4-morpholinethansulfonsaire (MES).

BRL (Bethesda Research Laboratories) Ltd., Paisley, UK.
Restriction enzymes; cloning manual

Difco Laboratories, Detroit, USA
Bacto-agar; Bacto-tryptone; yeast extract.

Fisons plc (Services), Crawley, Sussex, UK.
Acrylamide (electrophoresis grade); butan-1-ol; chloroform; diethyl ether; dimethyl sulphoxide; formamide; formaldehyde; glycine; glycerol; hydrogen peroxide; trichloroacetic acid; Urea

Flow Labs, Irvine, Ayrshire, UK.
Gamborg medium (B5 salts), MS medium, gibberelic acid (GA3).

Fluka Chemicals, Glossop, Derbyshire, UK.
N6-Benzyledenine (6-BAP).

Fuji Photo film Co. (UK) Ltd., London, UK
RX medical X-ray film

Gibco BRL, Uxbridge, Middlesex, UK.
Low melting-point agarose (ultra-pure)

Kodak Chemicals Ltd., Kirby, Liverpool, UK
N'N-methylene bis-acrylamide

May and Baker Ltd., Eccles, Manchester, UK
Acetic acid (glacial); ethanol; hydrochloric acid; methanol

Promega Biotec. Ltd., Maidstone, Wisconsin, USA.
RNasin; SP6 RNA polymerase; alkaline phosphatase detection kit; rabbit reticulocyte lysate.

Sigma Chemical Company, Poole, Dorset, UK
Agarose type II medium EEO; Bovine serum albumin (BSA); ethidium bromide; lysozyme; PMSF; ribonuclease A; Tween 20; Triton X 100; Nonidet P40; L-amino acids; EDTA; silver nitrate; Coomassie Brilliant Blue R; spermine; spermidine;

dithreothreitol (DTT); ATP; 5-bromo-4-chloro-3-indolyl-B-D-galactopyranoside (X-Gal); isopropyl-B-D-thio galacturanoside (IPTG); B-glucuronidase (EC 3.2.1.31: GUS); 4-methyl umbelliferryl glucoronide (MUG); 4-methyl umbelliferrylate (MU); thrombin; xylene cyanol; MOPS; HEPES; fluorescein diacetate (FDA); thrombin.

United States Biochemical Corp., Cleveland, Ohio, USA.
Sequenase kit

Whatman Labscales Ltd., Maidstone, Kent, UK.
Filter paper (3MM; No1; P81; DE81).

Yakult Hosha Co Ltd., Tokyo, Japan.
Cellulase "Onozuka RS"; maceroenzyme R10.

BIBLIOGRAPHY.

Akazawa,T. and Hara-Nishimura,I. (1985): Ann. Rev. Psychol., 36, 441-472.

Anderson,C.W., Straus,J.W. and Dudock,B.S. (1983): Methods in Enzymology, 101, 635-644.

Badenoch-Jones,J., Spencer,D., Higgins,T.J.V. and Millard,A. (1981): Planta, 153, 201.

Balint,G.A. (1974): Toxicology, 2, 77-102.

Bankier,A.T., Weston,K.M. and Barrell, B.G. (1986): in, MRC Labs. Sequencing Manual, Hills Road, Cambridge.

Baranski,T.J., Faust,P.L. and Kornfield,S. (1990): Cell, 63, 281-291.

Bassuner,R., Wobus,U. and Rapoport, T.A. (1984): Febs. Letts., 166, 314-320.

Beck,E., Ludwig,G., Auuerswald,E.A., Reiss,B. and Schaller,H. (1982): Gene, 19, 327-336.

Bednarek,S.Y., Wilkins,T.A., Dombrowski,J.E. and Raikhel,N.V. (1990): Plant Cell, 2, 1145-1155.

Bernstein,H.D., Poritz,M.A., Strub,K., Hoben,P.J., Brenner,S. and Walter,P. (1989): Nature,, 340, 482-486.

Biggin,M.D., Gibson,T.J. and Hong,G.F. (1983): P.N.A.S., 80, 3963-3965.

Blobel,G. and Dobberstein,B. (1975): J. Cell Biol., 67, 852-862.

Blobel,G. (1980): P.N.A.S., 77, 1496.

Bolla,J.M., Lazdunski,C., Innouye,M. and Pages, J.M. (1987): Febs. Letts., 224, 213-218.

Bollini,R. and Chrispeels,M.J. (1979): Planta, 146, 487-501.

Bollini,R., Ceriotti,A., Daminati,M.G. and Vitale,A. (1985): Plant Physiol., 65, 15-22.

Boulter,D. (1984): Phil. Trans. R. Soc. London., 304, 323-332.

Boutry,M., Nagy,F., Poulsen,C., Asyagi,K. and Chua, N-H. (1987): Nature, 328, 340-342.

Bowles,D.J., Marcus,S.E., Pappain,D.J.C., Findlay,J.B.C., Eliopoulas,E., Maycox,P.R. and Burgess.J. (1986): J. Cell Biol., 102, 1284-1297.

Bowles,D.J. and Pappin,D.J. (1988): TIBS, 13, 60-66.

Bowles.D.J. (1990): Curr. Opinion in Cell Biol., 2, 673-680.

Brinegar,A.C. and Petersen,D.M. (1982): Plant Physiol., 70, 1767.

Bulleid,N.J. and Freedman,R.B. (1988): Nature, 335, 649-651.

Bustos,M.M., Luckow,V.A., Griffing,L.R., Summers,M.D. and Hall,T.C. (1988): Plant Mol. Biol., 10, 475-488.

Butterworth,A.G. and Lord,J.M. (1983): Eur. J. Biochem., 137, 57-65.

Carrington,D.M., Auffret,A. and Hanke,D.E. (1985): Nature, 313, 64-67.

Chang,J-Y. (1985): Eur. J. Biochem., 151, 217-224.

Chirico,W.J., Waters,M.G. and Blobel,G. (1988): Nature, 332, 805-810.

Chrispeels,M.J., Higgins,T.J.V. and Spencer,D. (1982): J. Cell Biol., 93, 306-313.

Chrispeels,M.J. (1983a): Planta, 157, 454-461.

Chrispeels,M.J. (1983b): Planta, 158, 140-151.

Chrispeels,M.J. and Tague,B.W. (1990): Int. Rev. Cytol., 125, 1-45.

Chrispeels,M.J. and Raikhel,N.V. (1991): Plant Cell, 3, 1-9.

Cornell,D.G., Dluhy,R.A., Briggs,M.S., McKnight,C.J. and Gierash,L.M. (1989): Biochemistry, 28, 2789-2797.

Crouch,M.L., Tenbarge,K.M., Simon,A.E. and Ferl,R. (1983): J. Mol. & Applied Gen., 2, 273-283.

Croy,R.R.D., Gatehouse,J.A., Tyler,M. and Boulter,D. (1980): Biochem. J., 191, 509.

Davies,J. and Smith,D.I. (1978): Ann. Rev. Microbiol., 32, 469-518.

Dean,N. and Pelham,H.R.B. (1990): J. Cell Biol., 111, 369-377.

DeBlock, Botterman,J., Vandewiele,M. et al (1987): EMBO, 6, 2513-2518.

Della-Cioppa,G., Kishore,G.M., Beachy,R.N. and Fraley,R.T. (1987): Plant Physiol., 84, 965-968.

Denecke,J., Gossele,V., Botterman,J. and Cornelissen,M. (1989): Methods in Mol. Cell Biol., 1, 19-27.

Denecke,J., Botterman,J. and Deblare (1990): Plant Cell, 2, 51-59.

Dorel,C, Voelker,T.A., Herman,E.H. and Chrispeels,M.J. (1989): J. Cell Biol., 108, 327-337.

Dretzen,G., Bellard,M., Sassone-Corsi,P. and Chamdon,P. (1981): Analytical Biochem., 112, 295-298.

Edge,M.D. et al (1983): N.A.R., 11, 6419-6435.

Ellis,R.J. and Hemmingsen,S.M. (1989): TIBS, 14, 339-342.

Endo,Y. and Tsurugi,K. (1987): J. Biol. Chem., 262, 8128-8130.

Endo,Y., Mitsui,K., Motizuki,M. and Tsurugi,K. (1987): J. Biol. Chem., 262, 5908-5912.

Etzler,M.E. (1985): Ann. Rev. Plant Physiol., 36, 209-234.

Etzler,M.E. (1986) in, "The Lectins"; Liener,I.E., Sharon,N. and Goldstein,I.J. eds., San Diego Acad. Press.

Faye,L. and Chrispeels,M.J. (1987): Planta, 170, 217-224.

Foxwell,B.M.J., Donovan,T.A., Thorpe,P.E. and Wilson,G. (1985): Biochem. Biophys. Acta., 840, 193-203.

Frankel,A., Welsh,P., Richardson,J. and Robertus,J.D. (1990): Mol. & Cell. Biol., 12, 6257-6263.

Freedman,M.G. (1984): TIBS, 9, 438-441.

Fromm,M.E., Taylor,L.P. and Walbot,V. (1986): Nature, 319, 791-793.

Fujikawa,K., Legaz,M.E. and Davie,E.W. (1972): Biochemistry, 11, 4892-4898.

Fukasawa,J., Hara-Nishimura,I. And Nishimura,M. (1988): Plant Cell Physiol., 29, 339-345.

Garoff,H. (1985): Ann. Rev. Cell Biol., 1, 403-445.

Ghrayeb,J., Kimura,H., Takahara,M., Hsiung,H., Masu,Y. and Inouye,M. (1984): EMBO, 3, 2437-2442.

Gifford,D.J. and Bewley.J.D. (1983): Plant Physiol., 72, 376-381.

Gomori, (1955): Methods in Enz., 1, 138.

Gormen,C.M., Moffat,L.F. and Howard,B.H. (1982): Mol. Cel. Biol., 2, 1044-1051.

Graham,J.S., Pearce,G., Merryweather,J., Titani,K., Ericsson,L. and Ryan,C.A. (1985): J. Biol. Chem., 260, 6555-6560.

Greenwood,J.S. and Chrispeels,M.J. (1985): Planta, 164, 295-302.

Greenwood,J.S. and Chrispeels,M.J. (1985): Plant Physiol., 79, 65-71.

Griffiths,G. and Simons,K. (1986): Science, 243, 438-443.

Gueze,H.J., Slot,J.W., Strous,G., Hasilik,A. and Figura,K.V. (1985): J. Cell Biol., 101, 2253-2262.

Halling,K.C., Halling,A.C., Murray,E.F., Lodin,B.F., Houston,L.L. and Weaver,R.F. (1985): N.A.R., 13, 8019-8033.

Halpin,C. (1990): PhD thesis, Leicester University.

Harley,S.M. and Lord,J.M. (1985): Plant Sci., 41, 111-116.

Harley,S.M. and Beevers,L. (1989): Plant Physiol., 91, 674-678.

Hattori,T., Ichihara,S. and Nakamura,K. (1987): Eur. J. Biochem., 166, 533-538.

von Heijne,G. (1985): J. Mol. Biol., 184, 99-105.

von Heijne,G. (1986): N.A.R., 14, 4683-4690.

von Heijne,G. (1990): Curr. Opinion in Cell Biol., 2, 604-608.

Hemperly,J.J. and Cunningham,B.A. (1983): TIBS, 8, 100-102.

Herman,E.M. and Shannon,L.M. (1984): Planta, 161, 97-104.

Herman,E.M. and Shannon,L.M. (1985): Plant Physiol., 77, 886-890.

Herman,E.M., Shannon,L.M. and Chrispeels,M.J. (1985): Planta, 165, 23-29.

Herman,E.M., Tague,B.W. and Chrispeels,M.J. (1990): Planta, 182, 305-312.

Higgins,T.J.V., Chandler,P.M., Zurawski,G., Bulton,S.C. and Spencer,D. (1983): J. Biol. Chem., 258, 9544-9549.

Higgins,T.J.V., Chrispeels,M.J., Chandler,P.M. and Spencer,D. (1983): J. Biol. Chem., 258, 9544-9552

Higgins,T.J.V. (1984): Ann. Rev. Plant Physiol., 35, 191-221.

Higgins,T.J.V., Chandler,P.M., Randall,P.J., Spencer,D., Beach,L.R., Blogrove,R.J., Kortt,A.A. and Inglis,A.S. (1986): J. Biol. Chem., 261, 11124-11130.

Higgins,T.J.V., Newbigin,E.J., Spencer,D., Llewllyn,D.J. and Craig,S. (1988): Plant Mol. Biol., 11, 683-695.

Hoffman,L.M. and Donaldson,D.D. (1985): EMBO, 4, 883-889.

Hoffman,L.M., Donaldson,D.D., Bookland,R., Rashka,K. and Herman,E.M. (1987): EMBO, 6, 3213-3221.

Holmes,D.S. and Quigley,M. (1981): Anal. Biochem., 114, 193-197.

Hussain,K., Bowler,C., Roberts,L.M. and Lord,J.M. (1989): Febs. Letts., 244, 383-387.

Inohara,N., Shimomura,S., Fukui,T and Futai,M. (1989): P.N.A.S., 86, 3564-3568.

Irwin,S. (1989): PhD. thesis, Warwick University.

Iturriaga,G., Jefferson,R.A. and Bevan,M.W. (1989): Plant Cell, 1, 381-390.

Jefferson,R.A. (1987): Plant Mol. Biol. Reporter, 5, 387-405.

Jefferson,R.A., Kavanagh,A.A. and Bevan,M.W. (1987): EMBO, 6, 3901-3907.

Johnson,L.M., Bankiatis,V.A. and Emr,S.D. (1987): Cell, 48, 875-885.

Karrenbauer,A., Jeckel,D., Just,W., Birk,R., Schmidt,R.R., Rothman,J.E. and Weiland,F.T. (1990): Cell, 63, 259-267.

Klausner,R.D., Lippincott-Schwartz,J. and Bonifacino,J.S. (1990): Ann. Rev. Cell Biol., 6, 403-432.

Klionsky,D.J., Banta,L.M. and Emr,S.D. (1988): Mol. & Cell. Biol., 8, 2105-2116.

Kornfield,R. and Kornfield,S. (1985): Ann. Rev. Biochem., 54, 631-634.

Kornfield,S. and Melham,I. (1989): Ann. Rev. Cell Biol., 5, 483-525.

Kramer,B., Kramer,W. and Fritz,H.J. (1984): Cell, 38, 879-887.

Krieg,P.A. and Melton,D.A. (1984): N.A.R., 12, 7057-7070.

Krieg,P.A. and Melton,D.A. (1986): Methods in Enz., 155, 397-415.

Krieg,U.C., Walter,P. and Johnson,A.E. (1986): P.N.A.S., 83, 8604-8608.

Kries,M., Forde,B.G., Rahman,S., Miflin,B.J. and Shrewry,P.R. (1985): J. Mol. Biol., 183, 499-502.

Laemmli,U.K. (1970): Nature, 227, 680-685.

Lamb,F.I., Roberts,L.M. and Lord,J.M. (1985): Eur. J. Biochem., 148, 265-270.

Larkin,P.J. (1976): Planta, 128, 213.

Larkins,B.A. and Hurkman,W.J. (1978): Plant Physiol., 62, 256-263.

Lewis,M.J. and Pelham,H.R.B. (1990): Nature, 348, 162-163.

Lewis,M.J., Sweet,D.J. and Pelham,H.R.B. (1990): Cell, 61, 1359-1363.

Lin,T.T.-S. and Li,S.S.-L. (1980): Eur. J. Biochem., 105, 453-459.

Lonnerdal,B. and Janson,J.C. (1972): B.B.A., 278, 175-183.

Lord,J.M. (1985a): Eur. J. Biochem., 146, 403-409.

Lord,J.M. (1985b): Eur. J. Biochem., 146, 411-416.

Lord,J.M. and Harley,S.M. (1985): FEBS Letts., 189, 72-76.

Lord,J.M. and Robinson,C. (1986) in; Plant Proteolytic Enzymes, 2, Dalling,M.J. Editor, CRC Press Inc., Florida.

Lord,J.M., Gould,J., Griffiths,D., O'Hare,M., Prior,B., Richardson,P.T. and Roberts,L.M. (1987), in 'Progress in Medical Chemistry', 24, Ellis & West Editors, Elsevier Science, B.V.

McDonnell,R.E., Clark,R.D., Smith,W.A. and Hinchee,M.A. (1987): Plant Mol. Biol. Reporter, 5, 380-386.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982): 'Molecular Cloning: a Lab Manual', Cold Spring Harbour Lab., NY.

Mann,K.G. (1987): TIBS, 12, 229-233.

Marsh,J.L., Erfle,M. and Wykes,E.J. (1984): Gene, 32, 481-485.

Matsuoka,K., Matsumoto,S., Hattori,T., Machida,Y. and Nakamura,K. (1990): J. Biol. Chem., 265, 19750-19757.

Matsuoka,K. and Nakamura,K. (1991): P.N.A.S., 88, 834-838.

May,J.M.J. (1988): PhD thesis, Warwick University.

May,J.M.J., Hartley,M.R., Roberts,L.M., Krieg,P.A., Osborn,R.W. and Lord,J.M. (1989): EMBO, 8, 301-308.

Merril,C.R., Goldman,D., Sedman,S.A. and Ebert,M.H. (1981): Science, 211, 1437-1438.

Messing,J. (1986): Methods in Enz., 101, 20-78.

Mettler,I.J. and Beevers,H. (1979): Plant Physiol., 64, 506-511.

Meyer,D.I., Krause,E. and Dobberstein,B. (1982): Nature, 297, 647-650.

Mitraki,A. and King,J. (1989): Biotechnology, 7, 690-697.

Montfort,W., Villafranca,J.E., Monzingo,A.F., Ernst,S.R., Katzin,B., Rutenber,E., Xuong,N.H., Hamlin,R. and Robertus,J.D. (1987): J. Biol. Chem., 262, 5398-5403.

Munroe,S. and Pelham,H.R.B. (1986): Cell, 46, 291-300.

Munroe,S. and Pelham,H.R.B. (1987): Cell, 48, 899-907.

Nagai,K. and Thogersen (1984): Nature, 309, 810-812.

Nakamaye,K. and Eckstein,F. (1986): N.A.R., 14, 9679-9698.

Nielsen,N.C. (1984): Phil. Trans. R. Soc. Lon., 304, 287-296.

Nishimura,M. and Beevers,H. (1978): Plant Physiol., 62, 44-48.

Nishimura,H-I. and Nishimura,M. (1987): Plant Physiol., 85, 440-445.

Nishimura,H-I., Nishimura,M., Matsubara,H. and Akazawa,T. (1987): Plant Physiol., 70, 699-703.

Nishimura,H-I. and Nishimura,M. (1991): in, 'Molecular Approaches to Compartmentalisation and Metabolic Regulation', Huang,A.H.C., and Taiz,L. (eds), American Soc. of Plant Physiol.

Norrander,J., Kempe,T. and Messing,J. (1983): Gene, 27, 101-106

O'Callaghan et al (1972): Antimicrobial Agents and Chemotherapy, 1, 283-288.

O'Hare,M., Roberts,L.M., Thorpe,P.E., Watson,G.J., Dron,B. and Lord,J.M. (1988): FEBS Letts., 216, 73-78.

Olsnes,S., Fernandez-Puentes,C., Carrasco,L. and Vazquez,D. (1975): Eur. J. Biochem., 60, 281-288.

Olsnes,S. and Pihl,A. (1982) in, Molecular Action of Toxins and Viruses, Cohen & van Heyningen Editors, Elsevier Biomedical Press.

Olsnes,S., Sandvig,K., Refsnes,K. and Pihl,A. (1976): J. Biol. Chem., 257, 3985-3992.

Osborne,T.B. (1924) in, The Vegetable Proteins, Longmans Green, London.

Osborn and Munson (1974): Methods in Enz., 13, 642-653.

Osborn,R.W. and Hartley,M.R. (1990): Eur. J. Biochem., 193, 401-407.

Parker,M.L. (1982): Plant Cell & Environ., 5, 37-43.

Pelham,H.R.B. (1989): Ann. Rev. Cell Biol., 5, 1-24.

Pernollet,J.C. (1978): Phytochemistry, 17, 1473.

Pfeffer,S.R. and Rothman,J.E. (1987): Ann. Rev. Biochem., 56, 829-852.

Pugsley,A.P. (1990): Current Opinion in Cell Biol., 2, 609-616.

Pusztai,A., Clarke,E.M.W., King,T.B. and Stewart,J.C. (1979): J. Sci. Food Agric., 30, 843.

Rapoport,T.A. (1990): TIBS, 15, 355-358.

Richardson,M., Yarwood,A. and Rouge,P. (1987): FEBS Letts., 216, 145.

Richardson,P.T., Gilmartin,P., Colman,A., Roberts,L.M. and Lord,J.M. (1988): Biotechnology, 6, 565-570.

Richardson,P.T., Roberts,L.M., Gould,J.A. and Lord,J.M. (1988b): BBA, 950, 385-394.

Richardson,P.T., Westby,M., Roberts,L.M., Gould,J.A., Colman,A. and Lord,J.M. (1989): FEBS Letts.,255, 15-20.

Richardson,P.T., Wales,R., Woodland,H.R., Roberts,L.M. and Lord,J.M. (1991): J. Biol. Chem., in press.

Roberts,L.M. and Lord,J.M. (1981a): Planta, 152, 420-427.

Roberts,L.M. and Lord,J.M. (1981b): Eur. J. Biochem., 119, 31-41.

Roberts,L.M., Lamb,F.I., Pappain,D.J.C. and Lord,J.M. (1985): J. Biol. Chem., 260, 15682-15686.

Roberts,L.M., Tregear,J. and Lord,J.M. (1990): in, 'Genetically engineered toxins', A.Frankel (ed), Marcel Dekker Inc., New York.

Robinson,D.G. and Depta,H. (1988): Ann. Rev. Plant Physiol., 39, 53-99.

Romish,K., Webb,J., Herz,J., Prehn,S., Frank,R., Vingron,M. and Dobberstein,B. (1989): Nature, 340, 478-482.

Rose,J.K. and Doms,R.W. (1988): Ann. Rev. Cell Biol., 4, 257-288.

Rose,M.D. and Misra,L.D. (1989): Cell, 57, 1211.

Rothman,J.H., Howald,I. and Stevens,T.H. (1989): EMBO, 8, 2057-2065.

Sandvig,K., Olsnes,S., and Pihl,A. (1976): J. Biol. Chem., 251, 3977-3984.

Sandvig,K., Olsnes,S., and Pihl,A. (1978): Eur. J. Biochem., 82, 13.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977): P.N.A.S., 74, 5463-5467.

Saul,M.W., Shillito,R.D. and Negrutiu,I. (1988): Plant Mol. Biol. Manual, A1, 1-16, Kluwer Academic Publishers, Dordrecht.

Schatz,G. (1986): Nature, 321, 108-109.

Schein,C.H. (1989): Biotechnology, 7, 1141.

Schreier,P.H. and Cortese,R. (1979): J. Mol. Biol., 129, 169-172.

Sebastini,F.L., Farrell,L.B., Vasquez,M and Beachy,R.N. (1991): Eur. J. Biochem., 199, 441-450.

Semenza,J.C., Hardwick,K.G., Dean,N. and Pelham,H.R.B. (1990): Cell, 61, 1349-1357.

Sengupta-Gopalan,C., Reichert,N.A., Barker,R.F., Hall,T.C. and Kemp,J.D. (1985): P.N.A.S., 82, 3320-3324.

Shagger and von Jagov (1987): Arch. Biochem., 166, 368-379.

Sharief,F.S. and Li,S.S.L. (1982): J. Biol. Chem., 257, 14753.

Sharon,N. and Lis,H. (1989): (eds) Lectins, Chapman and Hall, London.

Shinshi,H., Wenzler,H., Neuhaus,J.M., Felix,G., Hofsteenge,J. and Meins,F.Jr (1988): P.N.A.S., 85, 5541-5545.

Simon,R., Altschuler,Y., Rubin,R. and Galili,G. (1990): Plant Cell, 2, 941-950.

Smith,D.B. and Johnson,K.S. (1988): Gene, 67, 31-40.

Sonnewald,U., von Schaewan,A. and Willmitzer,L. (1990): Plant Cell, 2, 345-355.

Staswick,P.E., Hermodsen,M.A. and Nielsen,N.C. (1981): J. Biol. Chem., 8752-8755.

Stillmark,H. (1888): Arbeiten Pharmakoligischen Insitutes zu Dorpat, 3, 50.

Sturm,A., Voelker,T.A., Herman,E.M. and Chrispeels,M.J. (1988): Planta, 175, 170-182.

Suggs,S.V., Hirose,T., Miyake,T., Kawashima,E.H., Johnson,M.J., Itakura,K. and Wallace,R.B. (1981): in 'Developmental biology using purified genes', Brown,D. (ed), Academic Press, NY.

Sun,S.M., McLeester,R.C., Bliss,F.A. and Hall,T.C. (1974): J. Biol. Chem., 249, 2118.

Tague,B.W. and Chrispeels,M.J. (1987): J. Cell Biol., 105, 1971-1979.

Tague,B.W., Dickinson,C.D. and Chrispeels,M.J. (1990): Plant Cell, 2, 533-546.

Takagi,H., Morinaga,Y., Tsuchiya,M., Ikeniura,H. and Inouye,M. (1988): Biotechnology, 6, 948-950.

Thanh,V.H. and Shibasaki,K. (1978): J. Agric. Food Chem., 26, 692.

Topfer,R., Prols,M., Schell,J. and Steinbiss,H.H. (1988): Plant Cell Reports, 7, 225-228.

Trowbridge,I.S. (1974): J. Biol. Chem., 249, 6004-6012.

Tully, R.E. and Beevers, H. (1976): Plant Physiol., 58, 710-716.

Turner,M.M. (1991): MSc. thesis, University of Warwick.

Valls,L.A., Hunter,C.P., Rothman,J.H. and Stevens,T.H. (1987): Cell, 48, 887-897.

Valls,L.A., Winther,J.R. and Stevens,T.H. (1990): J. Cell Biol., 111, 361-368.

Van den Bulcke,M., Bauw,G., Castrensana,C., Van Montagu,M., and Vanderkerckhove,J. (1989): P.N.A.S., 86, 2673-2677.

Vaux,D., Tooze,J. and Fuller,S. (1990): Nature, 345, 397.

Villafranca,J.E. and Robertus,J.D. (1977): J. Mol. Biol., 116, 331-335.

Villafranca,J.E. and Robertus,J.D. (1981): J. Biol. Chem., 256, 554-556.

Vitale,A., Sturm,A. and Bollini,R. (1986): Planta, 169, 108-116.

Vitale,A. et al, (1991): Eur. J. Biochem. (in press).

Voelker,T.A., Florkiewicz,R.Z. and Chrispeels,M.J. (1986): Eur. J. Cell Biol., 42, 218-223.

Voelker,T.A., Herman,E.M. and Chrispeels,M.J. (1989): Plant Cell, 1, 95-104.

Wallace,J.C., Galili,G., Kawati,E.E., Cuellar,R.E., Shotwell,M.A. and Larkins,B.A. (1988): Science, 240, 662-664.

Walter,P. (1987): Nature, 328, 763-764.

Walter,P. and Blobel,G. (1980): P.N.A.S., 77, 7112-7116.

Walter,P. and Blobel,G. (1981): J. Cell Biol., 91, 557-561.

Walter,P. and Lingappa,V. (1986): Ann. Rev. Cell Biol., 2, 49-516.

Warren,G. (1987): Nature, 327, 17-18.

Wearne,S.J. (1990): FEBS Letts., 263, 23-26.

Werr,W. and Lorz,H. (1986): Mol. Gen. Genet., 202, 471-475.

Wiech,H., Stuart,R. and Zimmerman,R. (1990): Seminars in Cell Biol., 1, 55-63.

Wieland,F.T., Gleason,M.L., Serafinin,T.A. and Rothman,J.E. (1987): Cell, 50, 289-300.

Wilkins,T.A., Bednarek,S.Y. and Raikhel,N.V. (1990): Plant Cell, 2, 301-313.

Wilson,K.A. (1986): in 'Plant Proteolytic Enzymes', 2, Dalling,M.J. (ed), CRC Press Inc., Florida.

Yarwood,A., Richardson,M., Morphet,B.Y., Westby,M., Pere,D. and Rouge,P. (1988): Phytochemistry, 27, 1719-1721.

Yatsu,L.Y. and Jacks,T.J. (1968): Arch. BBA, 124, 466-471.

Youle,R.J. and Coloumbatti,M. (1987): J. Biol. Chem., 262, 4676-4682.

Youle,R.J. and Huang,A.H.C. (1976): Plant Physiol., 58, 703-709.

Youle,R.J. and Huang,A.H.C. (1978): Plant Physiol., 61,1040-1042.

Youle,R.J. and Huang,A.H.C. (1982): Plant Physiol., 61, 13.

Zagouras,P. and Rose,J.K. (1989): J. Cell. Biol., 109, 2633-2652.

Zimmermann,R. and Meyer,D.I. (1986): TIBS, 11, 512-515.

Zoller and Smith (1982): N.A.R., 10, 6487-6500.