Molecular Analysis of Actinidin

by

Uta M. Präkelt

submitted for the degree
of Doctor of Philosophy

University of Leicester
September 1987

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

5 month-old Chinese Gooseberry plant

This thesis is my own work unless otherwise acknowledged and has at no time been submitted for another degree.

U. Pisacane

I certify that this statement is correct.

Harry Smith

Acknowledgements

I wish to thank my supervisor, Professor Harry Smith, for his advice and support during the course of this work.

I am greatly indebted to Ray McKee for helpful discussions, inspiration and good (if sometimes cryptic) humour. Many thanks also to Sue Adams, Chris Smith, David Walker, and Orietta Fioroni for help and advice, but also for cheerful company.

My thanks also go to Andy Greenland, and Lucy and Ian Eperon for their valuable advice.

I am particularly grateful to Liz Paul for critical reading of the manuscript and helpful suggestions, but also for moral support and many delicious meals.

Abbreviations

A	Adenosine
Ab	Antibody
ATP	Adenosine-5'-triphosphate
Bis-acrylamide	NN'-methylene-bisacrylamide
BSA	Bovine serum albumin
C	Cytidine
cDNA	Complementary DNA
c.p.m.	Counts per minute
CTAB	cetyl triethylammonium bromide
dATP	deoxy adenosine-5'-triphosphate
dCTP	deoxy cytidine-5'-triphosphate
dGTP	deoxy guanosine-5'-triphosphate
dTTP	deoxy thymidine-5'-triphosphate
DDW	Double-distilled water
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid
G	Guanosine
GTP	Guanosine-5'-triphosphate
HEPES	N-2-hydroxyethylpiperazine N'-2-ethane sulphonic acid
HRP	Horseradish peroxidase
IgG	Immunoglobulin G
IPTG	isopropyl-1-thio-β-D-galactoside
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline

PEG	Polyethylene glycol
Poly (A)	Polyadenylic acid
Poly (A+) RNA	polyadenylated RNA
Poly (A-) RNA	RNA lacking poly (A)
PVP	polyvinyl polypyrrolidone
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
SDW	Sterile double-distilled water
T	Thymidine
TCA	Trichloroacetic acid
TE	10mM TRIS-HCl pH 7.5, 1mM EDTA
TEMED	NNN'-N'-tetramethyl ethylene diamine
TRIS	Tris (hydroxymethyl) amino methane
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Contents

		Page
1. Introduction		1
1.1.	Actinidin, the proteinase of Chinese gooseberry	1
1.2.	Discovery of actinidin	2
1.3.	*Actinidia chinensis*, cultivation and characteristics	2
1.4.	Variety and functions of proteinases	3
1.5.	The family of cysteine proteinases	5
1.6.	Structure of actinidin and similarities with other cysteine proteinases	7
1.7.	Mechanism of hydrolysis of actinidin and other cysteine proteinases	9
1.8.	Substrate specificity of actinidin and papain	10
1.9.	Control of cysteine proteinase activity	13
1.9.1.	Synthesis in the form of zymogens	13
1.9.2.	Subcellular location	14
1.10.	Aims and approach	17
2. Materials and Methods		20
2.1.	Plant material	20
2.2.	Growth of plants	20
2.3.	Extraction and purification of actinidin	20
2.4.	Protease assay	22
2.5.	Sodium dodecyl sulphate polyacrylamide gel electrophoresis	22
2.6.	SDS-urea gels	23
2.7.	Raising and testing antibodies to actinidin	24
2.8.	Immunodetection of antigen: immunodiffusion and immunoelectrophoresis	25
2.9.	Enzyme-linked immunosorbent assay	27
2.10.	Conjugation of horseradish peroxidase to actinidin by the periodate method	29
2.11.	Extracts of Chinese gooseberry plant for use in immunoelectrophoresis and ELISA	29
2.12.	RNA extraction and purification	30
2.13.	Quantification of RNA and assessment of purity and intactness	36
2.14.	Fractionation of RNA by oligo(dT) cellulose column chromatography	37

2.15.	*In vitro* translation of RNA	38
2.16.	Immunoprecipitation of *in vitro* translation products	40
2.17.	Peptide mapping	41
2.18.	cDNA synthesis	42
2.19.	C-tailing of double-stranded cDNA	43
2.20.	G-50 sephadex column chromatography	43
2.21.	Denaturing polyacrylamide gel electrophoresis	44
2.22.	Annealing of homopolymer-tailed cDNA with G-tailed pUC9	44
2.23.	Preparation of competent *E.coli* cells and transformation with recombinant plasmid or phage	45
2.24.	Colony hybridisation	46
2.25.	Plasmid DNA "miniprep"	47
2.26.	Large scale plasmid isolation	47
2.27.	Agarose gel electrophoresis of DNA and Southern blotting	48
2.28.	Recovery of DNA fragments from agarose gels	48
2.29.	Northern blot	49
2.30.	Dot blot of DNA and RNA	49
2.31.	DNA sequencing	50
2.32.	Oligo-labelling of DNA fragments	52
2.33.	End-labelling of synthetic oligonucleotide	53
2.34.	Isolation of plant DNA	53
2.35.	Synthetic oligonucleotide	54

3. Identification of actinidin in various parts of the Chinese gooseberry plant		55
3.1.	Introduction	55
3.2.	Results	56
3.2.1.	Growth of seedlings	56
3.2.2.	Extraction and purification of actinidin	56
3.2.3.	Stability of actinidin	60
3.2.4.	Raising and testing antibodies to actinidin	61
3.2.5.	Estimation of actinidin in plant extracts	62
3.3.	Conclusions	68

4. RNA isolation and *in vitro* translation 69
4.1. Introduction 69
4.2. Results 74
4.2.1. RNA purification 74
4.2.2. *In vitro* translation of Kiwi fruit and apex RNA 81
4.2.3. Immunoprecipitation of *in vitro* translation products 82
4.2.4. Peptide mapping 83
4.3. Conclusions 85

5. CDNA cloning and screening of recombinants 89
5.1. Introduction 89
5.2. Results 91
5.2.1. cDNA synthesis 91
5.2.2. C-tailing of double-stranded cDNA 91
5.2.3. Annealing and transformation 92
5.2.4. Screening of recombinant clones 93
5.2.5. Characterisation of positive recombinants 94
5.3. Conclusions 95

6. Sequence analysis of cDNA clones pAC.1 and pAC.7 97
6.1. Introduction 97
6.2. Results 97
6.2.1. Sequencing strategy 97
6.3. Analysis of sequencing results 98
6.4. Detailed comparison of the cDNA coding regions corresponding to mature actinidin with reference also to other cysteine proteinases 100
6.5. Conclusions 102

7. Actinidin gene expression and estimation of gene number 104
7.1. Introduction 104
7.2. Results 105
7.2.1. Fruit growth 105
7.2.2. Northern blot analysis 105
7.2.3. Southern blot 108
7.3. Conclusions 110

8. Discussion 110
8.1. The approach 110
8.2. Structure of DNAs 111
8.3. Sequence analysis 111
8.4. A family of actinidin genes 113
8.5. Synthesis as a larger precursor 115
8.6. Carboxy-terminal extension 118
8.7. Subcellular location of actinidin 119
8.8. Evolutionary considerations 120
8.9. Actinidin gene expression 121
8.10. Function of actinidin 123

9. References 125

10. Appendix 136

1. Introduction

1.1. Actinidin, the proteinase of Chinese gooseberry

Actinidin (EC 3.4.22.14) is the cysteine proteinase of *Actinidia chinensis*, popularly known as "Chinese gooseberry" or "Kiwi fruit". It is the most abundant protein of the ripe fruit, accounting for 50% of soluble protein (McDowall, 1970). It belongs to a class of highly active proteinases, with representatives in many living organisms, which are capable of utilising virtually all proteins as substrates and hydrolysing these to small peptide fragments.

Due to its ample availability and ease of purification, both the structure and catalytic properties of actinidin have been well characterised, as summarised below. However, there is a surprising shortage of information in the literature concerning the biological aspects of this enzyme. Considering the overwhelming abundance of actinidin in Chinese gooseberry fruit, it seems reasonable to assume that it plays an important physiological role. This is supported by the observation that both its structure and function appear to have been highly conserved during evolution (Kamphius *et al.*, 1985). However, as yet nothing is known about the function of actinidin in the plant. Another intriguing question arising from the large quantity of actinidin in fruit is how the plant cell itself is protected against such a highly active proteinase. This could be achieved by the storage of the enzyme in an inactive form, or by its compartmentalisation within subcellular organelles, but this aspect has apparently not yet been investigated. Finally, it is not known whether actinidin is synthesised exclusively in the fruit or whether it is present in other tissues. The accumulation of large amounts of enzyme in fruit must be controlled, and levels of control include transcription, translation, or post-translational processing, but it may also involve the

genomic organisation of actinidin such as the presence of multiple gene copies. The aim of this work was to try to illuminate some of these questions by a molecular approach.

1.2. Discovery of actinidin

It had long been known in New Zealand that "incorporation of Chinese gooseberry fruit into a table jelly prevents its setting" (Arcus, 1959). The proteolytic enzyme responsible was first described by Arcus (1959), who found that it hydrolysed about one fifth of peptide bonds in gelatin under reducing conditions or in the presence of cysteine.
Later the enzyme was isolated in crystalline form (McDowall, 1970), and its relative molecular mass (M_r) was estimated at around 26 kDa (Boland and Hardmann, 1972), although an earlier estimate had been 12.8 kDa (McDowall, 1970). From subsequent sequencing of the protein the M_r was calculated at 23.6 kDa (Carne and Moore, 1978). The enzyme was found to require a free sulphydryl group for activity (McDowall, 1970), and to be activated by thiol compounds such as cysteine and glutathione, as well as reducing agents (Glazer and Smith, 1971). It has a broad pH optimum between pH 5 and 7, although it is active over a much wider range between pH 3 and 10 (Boland and Hardman, 1972). Its isoelectric point (pI) of 3.1 is unusually low (McDowall, 1970). The crystalline preparation of actinidin was later fractionated into two active components, called proteinase A_1 and A_2, by ion exchange chromatography (McDowall, 1973); however, all subsequent studies refer to only one form of the enzyme.

1.3. *Actinidia chinensis*, cultivation and characteristics of the plant

The Chinese gooseberry has only in recent years become a regular feature on the markets of European countries. As the name suggests, *A. chinensis* originates in China, where several forms of this polymorphic species can

be distinguished (Liang and Ferguson, 1986). It was first introduced into New Zealand in 1906, and seed was introduced into Great Britain around 1900 (Veitch, 1904). In New Zealand several types with higher quality fruit were selected from a single variety, and first propagated in 1930 (Schroeder and Fletcher, 1976). Although the plant grows freely in many other countries, it is cultivated as a major fruit crop in New Zealand which is its main exporter (Schroeder and Fletcher, 1976).

A. chinensis is a deciduous vine or twining shrub and is functionally dioecious. Propagation by seed yields highly variable progeny plants and, although this property was exploited in the selection of superior fruiting types, the plants are now propagated by grafting. Pollination of the pistillate flowers is ensured by grafting staminate laterals onto each female vine (Schroeder and Fletcher, 1976).

The fruit is a berry of 5 to 6cm length, which is densely covered with brown hairs. Each fruit may contain up to 1400 seeds (Schmid, 1978). It develops from anthesis to maturity over a period of around 160 days, and growth follows a double sigmoid curve (Hopping, 1976). During stage I (0-57 days) fruit growth is initially by cell division, complete in most parts after 33 days, and afterwards mainly by cell enlargement. The first stage is followed by a period of slow growth (58-76 days), and by another phase of rapid growth until maturity. Seed development is complete after about 110 days.

1.4. Variety and functions of proteinases

Proteinases are enzymes catalysing the hydrolysis of peptide bonds in proteins. They are present in all living organisms and are believed to have arisen early in evolution, since even primitive organisms must have required them for digestion and metabolism of their own proteins (Neurath, 1984).

Proteolytic enzymes have been classified on the basis of their essential catalytic residue involved in peptide bond hydrolysis, and fall into four major classes, the serine, cysteine, aspartate and metallo proteinases (Neurath, 1984). They display a range of specificities, from the cleavage of a single peptide bond in a particular protein to the general hydrolysis of virtually all protein substrates into small peptides. Among the best known proteinases are the non-specific mammalian digestive enzymes, e.g. trypsin, serving to break down ingested proteins into small peptides suitable for digestion. At the other end of the scale are the enzymes of the blood coagulation cascade and the highly specific peptidases involved in the translocation of proteins across membranes, or the activation of peptide hormone precursors by limited proteolysis at a single peptide bond. The cysteine proteinases are among the most active proteinases in mammals, having the lowest substrate specificity. It has been shown that, in animals, lysosomal cysteine proteinases are responsible for the continuous turnover of intracellular proteins, the extent of which greatly exceeds the digestion by pancreatic enzymes (Barrett and Kirschke, 1981). Lysosomal cysteine proteinases have also been implicated in the proteolytic maturation of other lysosomal enzymes (von Figura and Hasilik, 1986).

The requirements of intracellular protein turnover, protein translocation, and enzyme activation by limited proteolysis apply to the plant kingdom too. Proteinases have been identified in lysosomes of meristematic tissues suggesting a role in plant growth (Ryan, 1973). Mounting evidence supports the concept that the very prominent plant vacuole, serving important functions such as the maintenance of cell turgor and the transport and storage of metabolites, can function as a plant lysosome. It contains many of the enzymes associated with lysosomes in animals, including proteases (Ryan, 1973). Intracellular digestion of proteins in

plants assumes a special role during leaf senescence, enabling the plant to recycle organic matter, and is accompanied by rupturing of the vacuole and release of digestive enzymes into the cytosol. In the special case of the insect trapping plant genera, proteases are secreted for extracellular digestion of prey. Proteolysis has been recognized as a key feature in the physiology of seed germination. Many seeds contain considerable quantities of storage proteins which are mobilised by a number of proteinases with varying specificities, some of which have been shown to be under hormonal control (Ryan, 1973).

1.5. The family of cysteine proteinases

The cysteine proteinases are a family of enzymes which contain an essential cysteine residue at their active site. They have been isolated from prokaryotic as well as eukaryotic organisms. The bacterial cysteine proteinases clostripain, from *Clostridium histolyticum* (Mitchell and Harrington, 1971), and streptococcal proteinase (Liu and Alliott, 1971) are secreted into the medium and probably serve to break down proteins into small peptides suitable for assimilation (Lowe, 1976). The mammalian cysteine proteinases, cathepsins B, H and L, are important lysosomal enzymes and are probably the most active proteinases in the body. In man they turn over 200 to 300 g of intracellular protein per day (Barrett and Kirschke, 1981). Recently, several cysteine proteinases were found to be associated with cellular differentiation of the lower eukaryotic organism, the slime mould, *Dictyostelium discoideum* (Pears *et al.*, 1985; Williams *et al.*, 1985). They are thought to play a role in the autodigestion of intracellular proteins observed during cellular aggregation (Pears *et al.*, 1985). Similarly, a cysteine proteinase has been implicated in the histolysis of larval tissues during insect metamorphosis (Kawamura *et al.*, 1984).

In plants, probably the best known cysteine proteinase is papain which, together with the related enzymes chymopapain and papaya peptidases A and B, is abundant in the fruit and latex of the pawpaw tree, *Carica papaya* (Glazer and Smith, 1976). The ability of extracts of the leaves and fruit of this plant to tenderise meat has been exploited for centuries by inhabitants of tropical countries (Ryan, 1973), and papaya latex or the partially-purified enzyme is still used today in the food processing and brewing industries (Godfrey and Reichelt, 1983).

The cysteine proteinases bromelain from pineapple, *Ananas comosus* (Goto *et al.*, 1976), ficin from fig, *Ficus glabrata* (Glazer and Smith, 1976), calotropin from the madar plant, *Calotropis gigantea* (Heinemann *et al.*, 1982), and asclepain from milkweed, *Asclepias syriaca* (Lynn *et al.*, 1979), all occur in the latex and, in the cases of bromelain and ficin, also in the fruits of the plants. Actinidin probably occurs only in the fruit of the Chinese gooseberry as the plant does not appear to possess a laticifer system. Most of the plant cysteine proteinases occur in several distinct forms; for example, the bromelains isolated from fruit and stem have different characteristics, ficin consists of five distinguishable enzymes, and several forms of papain, chymopapain, asclepain, and calotropin have been isolated (Goto *et al.*, 1976; Lowe, 1976; McKee *et al.*, 1986; Lynn *et al.*, 1980; Pal and Sinha, 1980).

The plant cysteine proteinases listed above are all highly abundant enzymes and, similar to papain, have attracted interest for their usefulness in industrial processes; bromelain and ficin have found limited application in the food processing and pharmaceutical industries (Godfrey and Reichelt, 1983), and the potential of actinidin as a meat tenderizer has been considered (Schroeder and Fletcher, 1967).

As is the case with actinidin, no physiological function has been ascribed to any of the other plant cysteine proteinases. However, it has been

speculated that they may be involved in the defence against insect predators and microbial invasion, although such a role has been attributed to the abundant protease inhibitors of many plants (Lowe, 1976 ; Ryan, 1973). Consistent with such a function is the finding that the latex of *C.papaya* and *F.glabrata* contains enzymes with chitinase activity (Glazer and Smith, 1971), and it is interesting that the natural inhibitors of cysteine proteinases, such as leupeptin and antipain, are of microbial origin (Brocklehurst et al., 1981).

Plant cysteine proteinases whose physiological functions are evident have been found in germinating seeds of barley, *Hordeum vulgare*; sorghum, *Sorghum sp.*; corn, *Zea mays*; and French bean, *Phaseolus vulgaris* (Baumgartner and Crispeels, 1977; Nielsen and Liener, 1984). In barley aleurone layers several cysteine proteinases are synthesised *de novo* under hormonal control and clearly serve the function to mobilise storage proteins during germination (Hammerton and Ho, 1986).

1.6. Stucture of actinidin and similarities with other cysteine proteinases

Actinidin and papain are the two best studied of the cysteine proteinases. Their complete amino acid sequences are known (Carne and Moore, 1978; Husain and Lowe, 1969), and their tertiary structures have been determined by X-ray crystallography at 1.7 Å and 1.65 Å for actinidin and papain, respectively (Baker, 1977; 1980; Drenth *et al.*, 1968; Kamphius *et al.*, 1984).

Actinidin has 220 amino acids, and its calculated M_r is 23,660 Da (Carne and Moore, 1978; Baker, 1980). The molecule has a clearly defined double domain structure, with the active site lying at the interface between the domains (Baker, 1980). Domain I is made up of the first half of the polypeptide chain (residues 19 to 115), and is largely helical, with three significant lengths of α-helix, while domain II, consisting of residues

116 to 213, is built around an irregular, twisted β-sheet. The amino-terminal residues (1 to 18) are associated with domain II and the carboxy-terminal with domain I, both crossing over to the other domain and acting as 'straps' in helping to hold the two halves of the molecule together (Baker, 1980). Both domains have hydrophobic cores, but the interface between them is made up of largely polar and charged side chains. There are three disulphide bridges, between residues 22 to 65, 56 to 98, and 156 to 206. An optimal alignment with the 212 amino acids of papain, requiring three insertions in actinidin of one, two, and four residues and one single amino acid deletion, shows identity of 107 out of the 220 residues of actinidin, corresponding to a level of homology of 48%. In spite of this relatively low level of primary sequence homology the polypeptide backbones of the two enzyme are virtually superimposable. The main differences are in the two chain termini, in several regions of insertions and deletions, and in a central region, which connects the two domains in the form of an extended surface loop and shows the lowest sequence conservation (Baker, 1981).

The complete amino acid sequences of cathepsins B and H (Takio *et al.*, 1983) and partial sequences of bromelain (Goto *et al.*, 1980), ficin (Husain and Lowe, 1970) and asclepain (Lynn *et al.*, 1980), as well as the deduced amino acid sequences from cDNA clones of aleurain (Rogers *et al.*, 1985) and two cysteine proteinases from slime mould (Williams *et al.*, 1985; Pears *et al.*, 1985), show similar levels of homology with actinidin and papain. The most conserved residues are those which are also conserved between actinidin and papain and which are known to affect enzyme function and tertiary structure (Kamphius *et al.*, 1984). Although the primary sequence of calotropin has not yet been determined, its tertiary structure strongly resembles that of papain (Heinemann *et al.*, 1982).

The discovery of these considerable homologies between all the eukaryotic cysteine proteinases has led to the conclusion that they are all derived from a common evolutionary ancestor (Neurath, 1984). It appears that the polypeptide folding patterns of these enzymes have been strongly conserved during evolution, whilst allowing considerable sequence divergence to occur. The fact that homologies are not greater between the enzymes from plants than those between animals, plants and lower eukaryotes, suggests that the ancestral proteinase must have existed before the divergence of animals and plants, more than one billion years ago (Schopf, 1978). The bacterial cysteine proteinases show sequence homology with the eukaryotic enzymes around the active site residues only, and are therefore thought to represent examples of convergent evolution of proteins with similar functions (Neurath, 1984).

1.7. Mechanism of hydrolysis of actinidin and other cysteine proteinases

The active site of actinidin lies at the interface between the two halves of the molecule and is made up of residues from both domains. Cys-25 and His-162 are the essential catalytic residues (Baker, 1977). In the absence of substrate, their side chains, facing each other across the active site cleft, interact to form a thiolate/imidazolium ($S^-...ImH^+$) ion pair. Asn-182 is hydrogen-bonded to the catalytic His-162 and its function is probably to give the correct orientation to the imidazole ring of His-162 (Baker, 1981). The catalytic process can be divided into at least three steps; substrate binding, acylation and deacylation (Glazer and Smith, 1971). Boland and Hardman (1973) suggest that a fourth step, isomerisation of the acyl-enzyme intermediate, may be an additional rate-limiting factor. Upon substrate binding, the polar carbonyl group of the peptide bond to be hydrolysed is "attacked" by the nucleophilic sulphur atom of cys-25 and forms a transient tetrahedral thioester

intermediate by a covalent bond, leading to the formation of an acyl-enzyme complex. The peptide bond is cleaved by protonation from His-162 of the nitrogen atom of the scissile bond. Hydrolysis of the acyl-enzyme complex is thought to involve a second tetrahedral intermediate with a water molecule which also protonates the imidazole nitrogen of His-162.

It has been suggested that the dissociation of the ion pair between Cys-25 and His-162 during substrate binding is facilitated by two further ionised groups, Asp-161 in the vicinity of the active site, and a carboxyl-group further removed from the site and thought to be Glu-52 (Willenbrock and Brocklehurst, 1985). However, studies of the tertiary structure appear not to confirm this proposed involvement of the adjacent carboxyl-group of Asp-161, firstly because the side chain faces away from the active site, and secondly because this residue is Gly in cathepsin B and Asn in cathepsin H, two enzymes which follow the same reaction pathway (Kamphius *et al.*, 1985).

1.8. Substrate specificity of actinidin and papain

Studies of the hydrolysis of a variety of substrates by actinidin and papain have revealed a complex pattern of substrate specificity (Lowe, 1976). In a detailed study of papain using diastereoisomers of alanine, Berger and Schechter (1967) demonstrated that the site of hydrolysis of a polypeptide chain is determined by several 'subsites' involved in substrate binding, and the same was subsequently found to apply to actinidin (McDowall, 1973). The susceptibility to hydrolysis of any peptide bond within a polypeptide was found to be affected by the nature of four residues on the amino-terminal and three residues on the carboxy-terminal side of the bond, as illustrated (P= amino acid residue):

NH_3 - P_4 P_3 P_2 P_1 $P_1'P_2'P_3'$ - COOH Substrate

S_4 S_3 S_2 S_1 $S_1'S_2'S_3'$ enzyme subsites

The binding site (S_2) for the side chain of P_2 showed the highest degree of selectivity and therefore is predominant in determining substrate specificity. Amino acid residues binding preferentially at the S_2 subsite of actinidin are Val, Leu and Phe, but not Tyr, which binds to the equivalent subsite of papain. At the S_1 subsite the most frequently bound residues are His, Gly and Arg, whereas papain preferentially binds hydrophobic side chains at this position. Actinidin cleaves a smaller number of polypeptide bonds in the insulin B chain than papain (McDowall, 1973), and also has a much higher K_m for most substrates (Baker *et al.*, 1980). A comparison of the refined tertiary structures of actinidin and papain shows that differences in the specificities may be correlated with differences found between the amino acid side chains lining the "specificity pocket", S_2. This is composed of the following residues in actinidin (equivalent positions in papain in brackets): Tyr-69 (Tyr-67), Ile-70 (Pro-68), Thr-71 (Trp-69), Ala-136 (Val-133), Val-160 (Val-157), Met-213 (Ser-205) and Ser-215 (Phe-202). In papain the side chains of two of these residues (Val-133 and Val-157) interact most closely with the substrate by hydrophobic forces, and are apparently responsible for substrate specificity. In actinidin the residue equivalent to Val-133 is replaced by Ala. Another important difference is the substitution in actinidin of Met-213 for Ser-205, as this makes the pocket noticeably shorter in actinidin than papain, and likely to affect the specificity of the enzyme (Baker *et al.*, 1980). Other residues near the active site of actinidin, possibly involved in substrate binding, are Gln-19, Gly-23,-24,-67,-68, Ala-159, Ala-163, Asn-182, Ala-211 and Thr-212.

At the opposite end of the active site cleft a cluster of charged side chains, consisting of Lys-17, Glu-35, Glu-50 and Lys-181 has been implicated in widening of the cleft during substrate binding (Baker, 1980). These residues are hydrogen-bonded via a network of eight water molecules crossing the interface of the two domains, and although these water molecules may simply have a charge-spreading role they may provide a more flexible structure. In papain, which also has this arrangement of water molecules, a widening of the active site cleft during substrate binding has indeed been demonstrated by inhibitor studies (Baker, 1980).

The family of cysteine proteinases exhibits a range of sequence variation, which has provided some insight into mechanism and specificity of these enzymes by comparative studies. However, differences between enzymes containing only single amino acid substitutions would demonstrate more clearly the function of individual residues. Therefore considerable interest has been expressed by enzymologists in the synthesis of recombinant proteinases engineered by site-directed mutagenesis (Willenbrock and Brocklehurst, 1984). In addition, since papain is used in a number of industrial processes (Godfrey and Reichelt, 1983), and bromelain has found pharmacological application as an anti-inflammatory agent, proteolytic enzymes with altered properties to suit different needs may be of commercial interest (Carotti *et al.*, 1984). Protein engineering of the serine proteinases trypsin (Graf *et al.*, 1987) and subtilisin of *Bacillus subtilis* (Rath and Fletterick, 1987; Pantoliano *et al.*, 1986) has been used to alter characteristics such as specificity, efficiency, and stability of the enzymes.

1.9. Control of cysteine proteinase activity

1.9.1. Synthesis in the form of zymogens

Evidence is accumulating which suggests that the cysteine proteinases may be synthesised in the form of inactive precursors or zymogens, similar to the serine and aspartate proteinases such as trypsin and pepsin (James and Sielecki, 1986; Stroud, 1974). So far, only cathepsin B has been clearly demonstrated to be proteolytically processed from a 43 kDa precursor to the mature 32 kDa enzyme. An intermediate form of 39 kDa has limited enzyme activity and is believed to be responsible for the highly specific proteolytic maturation of insulin (Docherty *et al.*, 1984). The amino acid sequence of part of the precursor peptide has been deduced from a partial cDNA clone (San Segundo *et al.*, 1985). In addition, the cDNA has revealed coding potential for six amino acids beyond the mature carboxy terminus of cathepsin B, and a similar extension was subsequently found by cDNA cloning of human cathepsin B (Fong *et al.*, 1986). Thus, cathepsin B appears to be synthesised as a precursor which must be proteolytically processed at both the amino and carboxy terminus.

Recently, the sequence of a complete cDNA clone coding for papain was reported and revealed coding potential for an extra 133 amino acid residues at the N-terminus of the mature enzyme, suggesting that papain is synthesised in the form of a zymogen (Cohen *et al.*, 1986). However, the precursor has so far not been isolated.

Aleurain from barley and the two cysteine proteinases from slime mould, whose primary sequences are only known from cDNA clones, all contain additional polypeptides N-terminal to the region of homology with actinidin and papain. However, it is not known whether aleurain or the two slime mould proteinases are, in fact, proteolytically processed. Since mature aleurain has an M_r of 37 kDa, similar to that predicted from the complete nucleotide coding sequence, the authors suggest that most of

the prosequence must remain intact. However, the presence of two potential glycosylation sites within the proteinase domain raises the possibility that the enzyme is a glycoprotein whose carbohydrate side chain contributes significantly to the M_r of the mature enzyme. In this case it may well be proteolytically processed at a site equivalent to the mature N-terminus of other cysteine proteinases.

The prosegments from papain, aleurain and the slime mould proteinases are of similar sizes and show considerable sequence homologies. Some regions and individual residues appear to be more highly conserved than others, suggesting structural and possibly functional similarities between all the prosegments.

1.9.2. Subcellular location

1.9.2.1. Protein synthesis and subcellular targetting

Most of our knowledge of the mechanisms involved in the targetting of proteins to the different subcellular locations has been derived from studies on animal cells and was prompted by a large number of lysosomal storage disorders found in man (von Figura and Hasilik, 1986). Proteins destined for secretion from the cell or for storage in subcellular organelles such as the lysosomes, are synthesised on membrane-bound ribosomes and co-translationally translocated into the lumen of the endoplasmic reticulum (ER; von Figura and Hasilik, 1986). This process depends on the presence of a signal peptide, at the amino terminus of the nascent polypeptide chain, which is co-translationally cleaved off. The signal peptide consists of a core of between 8 and 20 hydrophobic amino acid residues, flanked by a positively charged amino terminus and a more hydrophilic carboxy terminus, but functional signal peptides show a wide range of variation (von Heijne, 1985).

Many proteins synthesised in this way are also glycosylated; pre-assembled

oligosaccharides are transferred to Asn residues which are part of the sequence Asn-X-Ser or Asn-X-Thr (von Figura and Hasilik, 1985). From the ER, proteins move to a second membrane system, the Golgi apparatus, where sorting is believed to take place, and where carbohydrate side chains are modified. Those of the lysosomal enzymes are modified by phosphorylation of one or more mannose residues. Mannose-6-phosphate side chains act as signals for lysosomal targetting, and two receptors involved in this process have been identified (von Figura and Hasilik,1986; Dahms *et al.*, 1987). Tertiary structure components of lysosomal proteins are likely to be required for recognition by the modifying enzymes. However, since in the absence of glycosylation significant levels of lysosomal proteins are still correctly located, a separate pathway involving recognition of protein tertiary structure not involving carbohydrate side chains has been implicated in lysosomal targetting (von Figura and Hasilik, 1985). Both amino-terminal and carboxy-terminal peptide sequences, which may be proteolytically cleaved upon transport, have been proposed as the structural determinants (North, 1986; Erickson and Blobel, 1983).

In plant cells, the same subcellular organelles involved in protein synthesis, translocation and sorting can be distinguished as in animals cells. The best-studied compartmentalised proteins are the abundant seed storage proteins (Higgins, 1984). They are synthesised on membrane-bound ribosomes and co-translationally translocated into the ER. The presence of signal peptides with characteristics similar to those of mammalian proteins has been demonstrated by sequence analysis and *in vitro* translation studies of many seed proteins (Higgins, 1984). The similarities are sufficiently great to allow the correct processing of plant precursors by microsomal membranes from dog pancreas and *Xenopus.* Many seed storage proteins are glycosylated. However, the importance of the carbohydrate side chains in subcellular targetting is not clear, since

both pea vicilin and bean phaseolin, synthesised in the presence of the glycosylation inhibitor, tunicamycin, are still correctly located (Higgins, 1984). By contrast, Faye and Chrispeels (1987) recently demonstrated that the transport from the ER to the protein bodies of concanavalin A (ConA), the lectin from jack bean cotyledons, was dependent on the presence of a carbohydrate side chain. The protein is synthesised as a precursor, pro-ConA, and interestingly, its maturation involves the proteolytic removal of a small peptide containing the carbohydrate side chain.

1.9.2.2. Subcellular location of cysteine proteinases

Evidence for subcellular location within lysosomes or secretory vesicles of cysteine proteinases has been provided in the cases of the mammalian cathepsins, the two cysteine proteinases from slime mould, and the plant enzyme aleurain. Consistent with their location, the amino-terminal extensions of aleurain, CP1, and CP2 described above have the characteristics necessary to function as signal peptides.

Since most of the abundant plant cysteine proteinases are present in latex (with the exception of actinidin), as described above, it seems likely that they are secreted enzymes. However, little is known about the mechanism of latex accumulation or indeed about the formation of laticifers, and therefore it is uncertain whether extracellular secretion is involved. The amino-terminal peptide of the papain precursor apparently does not show the characteristics typical of a signal peptide (Cohen *et al.*, 1986). The authors concluded that papain is probably not a secreted enzyme and that latex containing papain may be synthesised in the cytoplasm of cells which subsequently break down to form the laticifer system. However, a comparison of the prepropapain amino-terminal peptide with published signal sequences (von Heijne, 1985) suggests that it may in

fact function as such a signal.

The bromelains and ficins are glycoproteins, which is consistent with, although not proof of, their presence in membrane-bound organelles. It is interesting that the deduced amino acid sequences of aleurain and both slime mould proteinases each show two potential glycosylation sites within their proteinase domains, and that of the papain sequence has two such sites within the propeptide domain.

1.10. Aims and approach

As outlined above, a number of questions arise from knowledge of the characteristics of actinidin and of its abundance in the Chinese gooseberry fruit.

It would be interesting to discover the mechanism by which actinidin activity is controlled, whether the enzyme is present in an inactive form and/or located within membrane-bound organelles.

It seems likely that actinidin is synthesised in the form of a larger precursor, similar to papain, and that this is inactive until proteolytic removal of the additional peptide.

In order to understand the physiological significance of actinidin in the plant, it would be useful to determine whether it is also present in other tissues or whether it is restricted to the fruit. A study of actinidin gene expression during fruit development may provide some clues concerning the function of the enzyme.

A molecular approach to the study of actinidin was chosen for a number of reasons. The analysis of genes at the DNA level, by isolation of cDNA clones and genomic sequences, has been used in a number of cases to verify or supplement biochemical evidence, for example on the structure and post-translational processing of plant seed storage proteins (Croy *et al.*, 1982; Slightom *et al.*, 1985; Walburg and Larkins, 1986). A comparison of

DNA sequences coding for proteins encoded by multigene families and with homologies between different species has allowed an analysis of their evolutionary relationship and has also provided information on the extent of selection affecting evolutionary divergence between genes (Dunsmuir, 1985; Negoro *et al.*, 1985; Schuler *et al.*, 1982). Using cloned DNA as a probe, it has been possible to detect differential expression of individual members of a gene family in different parts of the plant (Dean *et al.*, 1985; Fluhr *et al.*, 1986). Changes in the level of gene expression during the development of, for example, tomato fruit, have been detected by molecular methods such as *in vitro* translation and Northern blotting (Grierson *et al.*, 1985). Finally, similarities between regions of DNA upstream from promoters of related genes, such as those coding for the legumin family of storage proteins, have led to the proposal of gene-specific elements involved in regulation of gene expression (Baeumlein *et al.*, 1986). Regulatory sequences involved in tissue-specific gene expression, light regulation, or transcriptional activation have been identified for the genes of the small subunit of ribulose 1,5-bisphosphate carboxylase (Timko *et al.*, 1985), chalcone synthase (Kaulen *et al.*, 1986) and zein (Maier *et al.*, 1987).
Factors in favour of a molecular approach to the study of actinidin were its abundance in fruit, ease of purification, and the fact that its amino acid sequence is known. This was likely to facilitate the raising of actinidin-specific antibodies and the construction of a synthetic oligonucleotide, to be used in the identification of recombinant clones. It was expected that actinidin-encoding mRNA represents a major fraction of fruit mRNA, and therefore cDNA cloning was chosen as the method to obtain the actinidin gene. By immunological methods it could be shown that only traces of actinidin are present in tissues other than the fruit. Levels of actinidin mRNA in fruit were assessed by *in vitro*

translation of fruit and leaf RNA, followed by immunoprecipitation with actinidin antiserum. A cDNA library was established using mRNA from fruit. Hybridisation of this library with a differential probe representing RNA sequences from leaf and fruit, in combination with a synthetic oligonucleotide, enabled the isolation of ten actinidin cDNA clones. The largest two of these were sequenced and analysed in detail. The insert of one of the clones was used as hybridisation probe in a study of the changes in actinidin gene expression during fruit development.

2. Materials and Methods

2.1. Plant material

a) Ripe Chinese gooseberry fruits, imported from Italy but of unidentified variety, were purchased from the local grocer. These were used both for the extraction of actinidin and as a source of seeds.

b) Unripe but fully mature fruits and leaves for cloning purposes were obtained from Mr. H.A.Baker at the Royal Horticultural Societies Garden at Wisley, Woking, Surrey, and stored frozen in liquid nitrogen or at -80°C. Developing fruits from the same vines were collected at two- to three-weekly intervals and frozen in liquid nitrogen.

c) Peas of the variety Feltham First were grown in vermiculite at 25°C in the dark.

2.2. Growth of plants

Seeds were a by-product from the homogenisation of fruit for enzyme extraction. They were washed, dried and stored at room temperature. For germination, seeds were washed in several changes of tap water for two days and placed on moist filter paper at 20°C in the light. After radicles appeared the seedlings were planted in John Innes compost and grown either at 20°C in 16 hours daylight or in large pots outdoors.

2.3. Extraction and purification of actinidin

Actinidin was purified by the method of Carne and Moore (1978) with slight modifications.

Fruit was homogenised in a blender with buffer containing 10mM $Na_2S_4O_6$ and 1mM EDTA. After removal of debris by centrifugation at 5000 r.p.m. in a Sorvall GSA rotor, the supernatant was filtered through Miracloth to

remove floating debris, and kept on ice. Polyethylene imine (10% (v/v) aqueous, pH 6.8 with HCl) was added to a final concentration of 0.1% (v/v), the solution stirred for 20 minutes at 4°C, and centrifuged at 10,000 r.p.m.. To the clear supernatant, ammonium sulphate which had been ground to a fine powder, was gradually added (over approximately 3 hours) with stirring at 4°C to 70% saturation. The precipitate was pelleted at 8,000 r.p.m., resuspended in 25ml of 100mM potassium phosphate buffer, pH 6.0, and dialysed against distilled water overnight at 6°C. The solution was clarified by adding $CaCl_2$ to 40mM followed by a solution of 5M K_2HPO_4 to 100mM. The resulting precipitate was removed by centrifugation at 10,000 r.p.m. at 4°C. KCl was added to the supernatant to a final concentration of 100mM. An ion exchange column of 1.6 cm diameter and 40 ml bed volume of deaerated DEAE Sephacel (diethylamino ethyl cellulose, Sigma) was equilibrated with 100mM KCl and 100mM potassium phosphate buffer pH 6.0 at 6°C, and calibrated to a flow rate of 20 ml per hour using a Gilson Minipulse pump. The protein solution was applied to the column and washed with 40ml of 100mM KCl in 100mM potassium phosphate buffer pH 6.0. The enzyme was eluted with two litres of a gradient of 0.1M to 1M KCl in phosphate buffer (as above). Fractions of 5ml were collected throughout.

The elution of protein was monitored by measurement of OD_{280} for each fraction. Column fractions containing protease activity were pooled, precipitated with ammonium sulphate as described in "Results", and dialysed against double distilled water (DDW) at 6°C overnight. Protein concentration was determined by measurements of OD_{280} and OD_{260} (Warburg and Christian, 1942). Actinidin was stored frozen at -20°C at a concentration of 10mg/ml.

2.4. Protease assay

Protease activity was estimated by a method based on the measurement of trichloroacetic acid (TCA)-soluble peptides formed from casein in the presence of the enzyme (Murachi, 1970).

2.4.1. Solutions:

A) 1.2% (w/v) casein in 30mM sodium phosphate buffer pH 7.5, made up freshly and boiled for 15 minutes.

B) 150mM L-cysteine in the same buffer.

C) TCA solution: 9g TCA, 15g sodium acetate, 19.5ml glacial acetic acid, made up to 100ml with DDW.

2.4.2. Method: The assays contained 500µl solution A, 20µl solution B, 80µl test solution. Incubation was at 35°C for 10 minutes, followed by addition of 500µl TCA solution and further incubation at 35°C for 30 minutes. Precipitates were pelleted in a Microfuge and OD_{280} of the supernatants were determined. A standard curve using known amounts of papain was included.

2.5. Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Methods for SDS-PAGE of proteins were based on Laemmli (1970) and used protocols described in Hames and Rickwood (1981). Acrylamide gels of 1mm thickness were prepared from the following stock solutions:

A) 50% (w/v) acrylamide containing 1.33% (w/v) bisacrylamide (both from Sigma).

B) 1M tris(hydroxymethyl)aminomethane (Sigma) adjusted to pH 8.6 with HCl (TRIS-HCL).

C) 1.5% (w/v) ammonium persulphate (Sigma) made up freshly.

D) 10% (w/v) sodium dodecyl sulphate (SDS).

The final concentrations were 290mM TRIS-HCl pH 8.6, 0.1% (w/v) SDS,

0.035% (w/v) ammonium persulphate, 10μl N,N,N',N'-tetramethyl-ethylenediamine (TEMED, Sigma) in 15ml solution, and acrylamide varying as shown in "Results". Gels were poured to within 2.5cm from the top, overlayed with 290mM TRIS-HCl pH 8.6, 0.1% (w/v) SDS and allowed to polymerise for 1 hour. The overlay was replaced by a stacking gel containing 5% (w/v) acrylamide, 0.13% (w/v) bisacrylamide, 125mM TRIS-HCl pH 6.8, 0.1% (w/v) ammonium persulphate, 0.1% (w/v) SDS and 10μl TEMED in 10ml gel solution, and a well-forming comb fitted. Protein samples were mixed with an equal volume of cracking buffer containing 62.5mM TRIS pH 6.8, 2mM EDTA, 2% (v/v) β-mercaptoethanol, and boiled for 1 minute. Electrophoresis was in buffer of 25mM TRIS, 1.4% (w/v) glycine and 0.1% (w/v) SDS (kept as a 5X stock) at 45 volts at room temperature overnight or until the dye reached the bottom of the gel.

2.6. SDS-urea gels

Low molecular weight proteins were separated on acrylamide gels containing urea (Swank and Munkres, 1971), using the following stock solutions:

A) 25% (w/v) acrylamide, 2.5% (w/v) bisacrylamide.

B) Gel buffer stock: 1% (w/v) SDS, 1M H_3PO_4, adjusted to pH 5.0 with TRIS.

C) Reservoir buffer: 0.1% (w/v) SDS, 100mM H_3PO_4, adjusted to pH 6.8 with TRIS.

D) Sample buffer: 1% (w/v) SDS, 8M urea, 1% (v/v) β-mercaptoethanol, 10mM H_3PO_4 adjusted to pH 6.8 with TRIS.

The gel was prepared by mixing 15 ml solution A, 3 ml solution B, 14.4g urea, and 0.3 ml 6% (w/v) ammonium persulphate (freshly prepared). The urea was dissolved by warming to 37°C and the volume adjusted to 30 ml with distilled water. After deaeration 20μl TEMED was added and the gel poured immediately. The gel was overlayed with a ten-fold dilution of gel

buffer stock (B) and polymeried for 1 hour. Sample wells were formed with a low concentration gel containing 2.25 ml solution A, 1.5 ml solution B, 7.2 g urea, and 0.15 ml 6% (w/v) ammonium persulphate, made up to 15ml with DDW and 10μl TEMED added. Protein samples were mixed with an equal volume of sample buffer and electrophoresed in reservoir buffer at 100 volts overnight.

All gels were simultaneously stained and fixed in 30% (v/v) industrial methylated spirit (IMS), 12% (v/v) acetic acid, 0.5% (w/v) Kenacid blue and 0.1% (w/v) cupric acetate, at 37°C for 1 hour, and destained in several changes of 25% (v/v) IMS and 10% (v/v) acetic acid. For autoradiography of gels containing [^{3}H]- or [^{35}S]- labelled proteins, the gels were immersed in Amplify (Amersham) for 20 minutes, placed on Whatman 3mm filter paper, covered with cling film and dried on a gel slab drier (Bio Labs.) at 80°C for several hours. Kodak X-ray film (XAR-5) was exposed for up to 4 weeks.

Marker proteins were the high molecular weight kit (30,000 to 200,000 daltons) and the kit for molecular weights 2,510 to 16,950 from Sigma, and the high molecular weight ^{14}C-methylated marker kit (14,300 to 200,000 daltons) from Amersham.

2.7. Raising and testing antibodies to actinidin

Two rabbits were immunised by injection of a total of 2mg native, and SDS-denatured actinidin , respectively. The antigen was prepared in Freund's complete adjuvant (Difco Laboratories, LTD.), as described in Hudson and Hay (1980), and booster injections were given over 6 weeks at two-weekly intervals. Test bleeds were taken before each injection, and after 8 weeks a total bleed was carried out. All manipulations involving animals were performed by staff of the animal house at Leicester

University Medical School.

After clotting overnight at 4°C serum was recovered by centrifugation at 3,000 r.p.m. in an MSE bench centrifuge for 10 minutes, and in the case of test bleeds, the serum was used without further purification. Serum obtained after the final test bleed was centrifuged as above, and precipitated by addition of 0.6 volumes of saturated ammonium sulphate made up in 20mM TRIS-HCl pH 8.0 and 1mM EDTA. After 30 minutes at room temperature the precipitate was spun at 4,000 ×g in an MSE bench centrifuge, redissolved in 0.4 volumes (compared to original volume) of the same buffer. Precipitation with 0.6 volumes of saturated ammonium sulphate was repeated three times. The final precipitate was dissolved in 0.2 volumes of 5mM sodium phosphate pH 8.0 and dialysed against the same buffer in two changes of 5l at 6°C overnight. The concentration of protein was estimated by measurements of OD_{260} and OD_{280} (Warburg and Christian, 1942). Pre-immune serum for use as a control was prepared from uninjected rabbits in the same way.

2.8. Immunodetection of antigen: immunodiffusion and immunoelectrophoresis

2.8.1. Solutions:

A) barbitone buffer:

1) 12.0g 5'5 diethyl barbituric acid, sodium salt dissolved in 800ml DDW.

2) 4.4g 5'5 diethyl barbituric acid dissolved in 150ml DDW at 95°C.

Solutions 1 and 2 were mixed, adjusted to pH8.2 with 5M NaOH and made up to 1l with DDW.

B) Svendsen buffer:

1) 13.0g 5'5 diethyl barbituric acid, sodium salt,

2.02g 5'5 diethyl barbituric acid,

dissolved in boiling DDW and made up to 1l.

2) 45.2g TRIS, 56.2g glycine,

dissolved in DDW and made up to 1l.

Solutions 1 and 2 were mixed to provide a 5X concentrated stock.

C) 2X concentrated Svendsen gel buffer:

100ml 5 times concentrated Svendsen buffer were mixed with 150ml DDW and 1ml Triton X-100.

D) Destain: 45% (v/v) methanol, 10% (v/v) acetic acid.

E) Stain: 0.125% (w/v) Coomassie Brilliant Blue (Sigma) in destain.

2.8.2. Ouchterlony Immunodiffusion (Hudson and Hay, 1980)

Agarose (BDH) was dissolved in barbitone buffer at 1% (w/v) by boiling, cooled to 65°C, poured onto ethanol-acetone washed slides (3.5ml per slide) and allowed to set. Wells were formed by removing agar with a pasteur pipette connected to a water pump, giving one central, surrounded by six peripheral, wells. The wells were filled with 7μl of test solution, and plates developed in sealed plastic boxes with moist tissue paper at 6°C for 24 to 48 hours. Plates were washed in a large volume of 150mM NaCl overnight at room temperature to remove unreacted protein, dried with a hot air-drier, stained for 5-10 minutes and destained in several changes of destain.

2.8.3. Rocket Immunoelectrophoresis (Hames and Rickwood, 1981)

Agarose was melted in 1X Svendsen buffer at 1% (w/v) cooled to 56°C and antibody added to a final concentration of 0.1mg/ml. Plates were poured and wells punched out as described above. Wells were filled with 7μl of antigen solution, and gels electrophoresed in 1X Svendsen buffer towards the anode at 5-10 volts/cm for 3 hours. Gels were washed, dried and

stained as described for immunodiffusion.

2.8.4. Two-dimensional immunoelectrophoresis (Hames and Rickwood, 1981)

For the first dimension 20μl samples of antigen solution were electrophoresed in 1% (w/v) agarose gels in Svendsen buffer for 2.5 hours at 100 volts. The gel was transferred to one edge of a larger plate, trimmed, and agarose containing 0.1 mg/ml antiserum in Svendsen buffer was poured next to it to cover the plate. Electrophoresis in the second dimension was for 3 hours at 100 volts. Precipitation lines were visualised as described above.

2.8.5. Denaturation of actinidin

Actinidin was denatured with SDS both for use in immunisation and for subsequent testing of the antiserum: Actinidin was diluted to lmg/ml in distilled water with 1% (w/v) SDS, boiled for 10 minutes and dialysed against 0.1% (w/v) SDS at room temperature overnight.

2.9. Enzyme-linked immunosorbent assay (ELISA, Voller *et al.*, 1979)

2.9.1. Materials:

Polystyrene microtiter plates were purchased from Dynatech Laboratories Ltd. Anti-rabbit immunoglobulin (IgG) conjugated to horse radish peroxidase (HRP), developed in goats, was from Sigma.

2.9.2. Solutions:

A) Coating buffer (carbonate-bicarbonate pH 9.6): 1.59g Na_2CO_3, 2.93g $NaHCO_3$, dissolved in DDW to 1l and stored at 4°C.

B) PBS pH 7.4: 8.0g NaCl, 0.2g KH_2PO_4, 2.9g $Na_2HPO_4.12\ H_2O$, 0.2g KCl, made up to 1l with DDW.

C) PBS-Tween: as PBS with 0.5ml Tween 20 per 1.

D) Substrate solution: 0.1M sodium acetate was adjusted to pH 6.0 with 0.1M citric acid. To 20ml were added 150µl tetramethyl benzidine (10mg/ml in dimethyl sulphoxide) and 25 µl hydrogen peroxide prior to use.

2.9.3. Method:

All solutions were applied at 200µl per well. Wells were washed between applications by filling with PBS-Tween, leaving for 1 minute, and emptied by shaking of the plates.

2.9.3.1. Direct method:

Serial dilutions, usually two-fold, of actinidin or plant extracts were prepared using coating buffer, applied to wells and left at 6°C overnight. The wells were washed three times and blocked with 1% (w/v) BSA in coating buffer for 20 minutes at room temperature. After a further three washes, 100-fold diluted antibody in PBS-Tween was added and left for two hours at 37°C. Wells were washed three times and 500-fold diluted goat anti-rabbit IgG-HRP conjugate added. Incubation at 37°C for two hours was again followed by three washes. Substrate was added and left for approximately 10 minutes or until the blue colour was sufficiently developed. The reaction was stopped by addition of 50µl 2.5M sulphuric acid. For measurements of OD_{450}, 200µl were removed from each well and diluted to 600µl with DDW. Alternatively, optical densities were determined using a microelisa plate reader (Dynatech Laboratories, Ltd.).

2.9.3.2. Indirect (double antibody sandwich) method:

Microtiter wells were coated with actinidin antibody diluted to 0.02mg/ml with coating buffer at 6°C overnight. Plates were washed three times and blocked with 1% (w/v) BSA in coating buffer for 20 minutes at room temperature. After three washes, two-fold serial dilutions of antigen were added to the wells and incubated at 37°C for 2 hours. Wells were washed and 500-fold diluted actinidin antibody-HRP conjugate (see below)

was added, followed by incubation at 37°C for two hours. After three washes, substrate was added and activities determined as described in "direct method".

2.10. Conjugation of horseradish peroxidase to actinidin antibody by the periodate method (Voller *et al.*, 1979)

4mg horseradish peroxidase were dissolved in 1ml of DDW. 0.2ml freshly made 100mM sodium periodate was added and stirred for 20 minutes at room temperature. The solution was dialysed against 1mM sodium acetate buffer pH 4.4 at 6°C overnight. 20µl of 200mM sodium carbonate buffer pH 9.5 were added followed by 8mg immunoglobulin fraction (prepared as described above) dissolved in 1ml 10mM carbonate buffer pH 9.5. This was stored at room temperature for two hours.

100µl of sodium borohydride solution (4mg/ml, made up freshly in water) was added, and the solution left at 6°C for 2 hours. An equal amount of saturated ammonium sulphate solution was added, the precipitate centrifuged and washed twice with 50% saturated ammonium sulphate. The pellet was resuspended in 1ml PBS (see above) and dialyzed against PBS at 6°C overnight. BSA was added to a final concentration of 1% (w/v) and the conjugate filtered through a millipore filter (0.22µm). The filtrate was stored at -20°C in 40µl aliquots.

2.11. Extracts of Chinese gooseberry plant for use in immunoelectrophoresis and ELISA

A) Fresh tissue was homogenized with a polytron in protein extraction buffer containing: 50mM TRIS (adjusted to pH 8.5 with glacial acetic acid), 50mM potassium acetate, 5mM magnesium acetate, 2mM dithiothreitol (DTT). The extract was spun at 10,000 r.p.m. in an SS34 Sorvall rotor for

15 minutes at 4°C. The supernatant was precipitated either by adding an equal volume of 8% (w/v) cold TCA and leaving on ice for 30 minutes, or by adding acetone to a final concentration of 80% (v/v). After centrifugation as above, the pellet was resuspended in a small volume of 15mM sodium phosphate buffer pH 7.5.

B) Tissue was frozen in liquid nitrogen and ground to a fine powder in a pestle and mortar. The powder was homogenised in protein extraction buffer (as above) on ice, spun at 10,000 r.p.m. in a Sorvall HB-4 rotor, and the supernatant was precipitated with an equal volume of 8% (w/v) cold TCA, as in A. The final precipitate was resuspended in a small volume of DDW and boiled with an equal volume of cracking buffer to facilitate dissolution.

C) Fresh material was ground in 125mM TRIS pH 6.8 (1 ml per g of tissue), using a polytron, and the homogenate used directly or after removal of debris by centrifugation at 10,000 r.p.m. in an HB-4 Sorvall rotor for 15 minutes.

D) For ELISA tests, 50mg of freeze-dried plant material was ground with a pestle and mortar to a fine powder, homogenised in 3ml PBS-Tween, and centrifuged in Eppendorf tubes to remove debris. The supernatant was diluted with PBS-Tween in serial dilutions.

2.12. RNA extraction and purification

2.12.1. General precautions:

All equipment used in RNA purification was washed with Pyroneg, rinsed thoroughly with running tap water, then twice with distilled water, and sterilised at 180°C for 5 hours. Centrifuge tubes and pasteur pipettes

were siliconised with Repelcote (Hopkin and Williams, Chadwell Heath, Essex), rinsed withDDW, and sterilized as above. Polycarbonate centrifuge tubes and Eppendorf tubes were autoclaved for 20 minutes. Double distilled water (DDW) was used for all solutions, and was sterilized by autoclaving for 20 minutes. Disposable gloves were were worn during all manipulations involving RNA.

2.12.2. Phenol method (Kirby, 1964)

2.12.2.1. Solutions:

A) Extraction buffer: 200mM TRIS-HCl pH 9.0, 100mM LiCl, 25mM EDTA, 1% (w/v) SDS (autoclaved).

B) Phenol mix: 100g phenol, 100mg 8-hydroxyquinoline, 100ml chloroform, 4ml isoamyl alcohol, 200ml 10mM TRIS-HCl pH 7.5.

C) Resuspension buffer: 100mM Tris-HCl pH 9.0, 100mM LiCl, 10mM EDTA (autoclaved).

D) Caesium chloride solution: 5.7M CsCl in sterile: 10mM TRIS-HCl pH 7.5, 10mM EDTA.

E) 2M sodium acetate pH 5.5 (with glacial acetic acid, autoclaved).

2.12.2.2. Method

Fresh tissue was weighed and chopped roughly with a razor blade. To each g of tissue 1ml extraction buffer and 1ml phenol mix were added and homogenised using a Polytron. The suspension was centrifuged in an MSE bench centrifuge at 3,000 r.p.m. for 10 minutes, and the aqueous layer re-extracted twice with an equal volume of phenol mix. The volume of the clear aqueous phase was measured, nucleic acids were precipitated by addition of 2 volumes of ethanol and kept at -20°C for at least 4 hours. The precipitate was pelleted by centrifugation at 10,000 r.p.m. in a Sorvall SS-34 rotor for 10 minutes, drained and resuspended in x/2 ml (x=g

of tissue) of buffer C.

For caesium chloride centrifugation, 5ml polyallomer centrifuge tubes were soaked in 0.1% DEP, rinsed in sterile DDW followed by ethanol and allowed to dry. 1ml caesium chloride solution was added to each tube and filled to within 2mm of the top with resuspended nucleic acid solution. The tubes were centrifuged in an SW 50.1 rotor (Beckman ultracentrifuge) for 19 hours at 25,000 r.p.m. and 20°C. The supernatant was discarded and the pellet resuspended in a small volume (3-5ml) of sterile DDW (SDW). The RNA was precipitated by addition of 0.1 volume of 2M sodium acetate and 2 volumes of ethanol and kept at -20°C overnight or -70°C for 1 hour. The precipitate was recovered by centrifugation in an SS-34 or HB-4 rotor (Sorvall) at 10,000 r.p.m. for 10 minutes, washed twice with 80% ethanol, dried briefly under vacuum and resuspended in SDW, and stored in Eppendorf tubes under liquid nitrogen.

2.12.3. Guanidine method (Chirgwin *et al.*,1979)

2.12.3.1.Solutions:

A) Guanidinium thiocyanate stock: 50g guanidinium thiocyanate (Fluka purum grade), 0.5g sodium N-lauroyl sarcosine, 2.5ml of 1M sodium citrate pH 7.0), 0.7ml β-mercaptoethanol, 0.33ml Antifoam (30%, Sigma), DDW to 100ml, adjusted to 7.0 with 1N NAOH. The solution was filtered through Whatman No.1 paper and stored at room temperature for up to 1 month.

B) Guanidine hydrochloride solution: 7.5M guanidine-HCl (Fluka puriss), 0.025 volumes 1M sodium citrate pH 7.0, 5mM DTT, adjusted to pH 7.0 with 1N NaOH.

2.12.3.2. Method: Plant material was weighed and 3ml guanidinium thiocyanate solution added per g of tissue, and homogenised using a Polytron at room temperature. The extract was squeezed through two layers

of sterile Muslin and centrifuged at 8,500 r.p.m. in a Sorval SS-34 or HB-4 rotor for 10 minutes at 10°C. To the supernatant 0.025 volumes of 1M acetic acid and 0.75 volumes of ethanol were added, and kept at -20°C overnight. The precipitate was collected by centrifugation at 8,000 r.p.m. for 10 minutes at -10°C, resuspended in guanidine-HCl solution with a halving of the original extraction volume. RNA was precipitated by addition of 0.025 volumes 1M acetic acid and 0.5 volumes of ethanol at -20°C for at least 3 hours. The precipitate was recovered by centrifugation as before. RNA was reprecipitated from guanidine-HCl twice more with further halving of the volumes. The final pellet was dispersed in ethanol, centrifuged in an Eppendorf tube and washed repeatedly until it was white. The pellet was dried briefly under vacuum, resuspended in SDW, and centrifuged in a Minifuge for 5 minutes. The supernatant was transferred to a new tube and RNA precipitated with 0.1 volume 2M potassium acetate pH 5.5 and 2 volumes ethanol. The RNA precipitate was collected by centrifugation, washed several times with 95% ethanol, dried and resuspended in SDW.

2.12.4. CTAB Method (Murray and Thompson, 1980)

2.12.4.1.Solutions:

A) 2x Extraction buffer: 2% (w/v) CTAB, 100mM TRIS-HCl pH 8.0, 20mM EDTA, 1.4M NaCl (autoclaved).

B) Precipitation buffer: 1% (w/v) CTAB, 50mM TRIS-HCl pH 8.0, 10mM EDTA, 1% (v/v) β-mercaptoethanol (autoclaved before addition of β-mercaptoethanol).

C) 10% CTAB solution: 10% (w/v) CTAB, 0.7M NaCl (autoclaved).

D) 1M caesium chloride solution: 50mM TRIS-HCl pH 8.0, 5mM EDTA, 50mM NaCl 1M CsCl (autoclaved).

E) 5.7M caesium chloride solution: as 4, but containing 5.7M CsCl.

2.12.4.2. Method: Tissue was frozen in liquid nitrogen and ground to a fine powder with a mortar and pestle. The powder was transferred to a beaker, β-mercaptoethanol was added to a final concentration of 1% (v/v) and for each g of tissue 1ml of boiling 2x extraction buffer was added. The homogenate was transferred to a 55°C waterbath for 5 minutes, shaken with an equal volume of chloroform:octanol (24:1; v/v), and centrifuged in an HB-4 rotor at 8,500 r.p.m. for 10 minutes at 20°C. To the aqueous phase 0.1 volume of 10% (w/v) CTAB solution was added and the chloroform extraction repeated. Nucleic acids were precipitated from the aqueous phase by mixing with an equal volume of precipitation buffer and leaving at room temperature for 30 minutes. The precipitate was collected by centrifugation in an HB-4 rotor at 3,500 r.p.m for 10 minutes. The pellet was resuspended in 1M CsCl solution over approximately 45 minutes with occasional shaking, layered over a 1ml cushion of 5.7M CsCl solution and centrifuged in a Beckman SW50.1 rotor at 25,000 r.p.m. for 19 hours at 20°C. The pellet was resuspended in SDW on ice and precipitated with 0.1 volume of 2M sodium acetate pH5.5 and 2 volumes ethanol at -80°C for 30 minutes. The RNA precipitate was washed 3 times in 80% (v/v) ethanol, dried briefly under vacuum and dissolved in SDW.

2.12.5. Modified Kirby Method

This method was originally described by Covey and Grierson (1976) and obtained in an adapted version from S.Curzon, John Innes Institute, Norwich.

2.12.5.1. Solutions:

A) Phenol mixture: 500g phenol, 0.5g 8-hydroxyquinoline, 50mM TRIS-HCl pH8.3. Phenol was dissolved in 150ml buffer by warming to 37°C. The

organic phase was shaken three times with 200ml buffer, and the 8-hydroxyquinoline added. For use with modified Kirby solution, 70ml m-cresol was added per 500g phenol.

B) Modified Kirby: 1% (w/v) triisopropyl naphthalene sulphonate (TNS), 6% (w/v) 4-aminosalicylate, 50mM TRIS-HCl pH 8.3, 6% (v/v) phenol mixture (with cresol).

C) Phenol/chloroform: 50 parts phenol mixture (without cresol), 50 parts chloroform, 1 part isoamyl alcohol.

2.12.5.2. Method: Tissue was frozen in liquid nitrogen and ground to a powder with a mortar and pestle. To each g of tissue 1ml of modified Kirby solution and 1/100 x 2.5 g of sodium diethyldithiocarbamate were added with liquid nitrogen and ground. After rapid thawing a further 1ml (per g tissue) modified Kirby solution was added, mixed and shaken with 2ml (per g tissue) phenol/chloroform. The phases were separated by centrifugation in an HB-4 rotor at 10,000 r.p.m. for 10 minutes at 10°C, and the aqueous phase extracted twice more with phenol/chloroform. Nucleic acids were precipitated by addition of 4M lithium acetate pH 6.0 to a final of 0.2M and 2.5 volumes of ethanol at -20°C for at least 2 hours. The precipitate was collected by centrifugation at 10,000 r.p.m. for 5 minutes at 4°C, and resuspended in 0.5ml SDW (per g tissue). RNA was precipitated by adding an equal volume of 6M lithium acetate pH 6.0 and leaving on ice for 1-2 hours. After centrifugation at 10,000 r.p.m. in an HB-4 rotor for 10 minutes at 4°C, the pellet was washed twice with 3M lithium acetate pH 6.0, resuspended in 0.5ml SDW (per g tissue), and extracted once more with phenol/chloroform. RNA was precipitated from the aqueous phase with 0.1 volumeof 2M sodium acetate pH 5.5 and 2.5 volumes ethanol, left at -20°C for 2 hours, and centrifuged as above. The pellet was washed twice with 80% (v/v) ethanol, dried and resuspended in SDW.

2.13. Quantification of RNA and assessment of purity and intactness

2.13.1. Absorbance profile

RNA was diluted 100- to 200-fold with DDW and scanned in Quartz cuvettes in a Perkin-Elmer lambda-5 spectrophotometer in UV light between 200 and 300nm. The quantity of RNA was estimated on the basis of its OD_{260}, assuming that $1OD_{260}$ = 37μg RNA/ml.

Purity was assessed by calculating ratios of OD_{260}/OD_{280} and OD_{260}/OD_{230}. (The optical density scans in "Results" serve to illustrate differences in purity between RNA preparations, and dilution factors were therefore not quoted.)

2.13.2. Agarose gel electrophoresis of RNA (Maniatis *et al.*, 1982)

2.13.2.1. Solutions:

A) Running buffer: 10mM sodium phosphate buffer pH 7.0.

B) Sample buffer: 100mM " " " "

C) Glyoxal mixture (prepared just prior to use):

200μl glyoxal (deionized with Biorad Mixed Bed Resin)

250μl dimethyl sulphoxide, 50μl sample buffer

D) Loading buffer: 50% (v/v) glycerol, 1mM EDTA, 0.4% (w/v) bromophenol blue, 0.4% (w/v) xylene cyanol.

2.13.2.2. Method: 10x 10cm gels were prepared with 1% (w/v) agarose (Seakem) in running buffer. RNA samples (5-10μg) were dried under vacuum, resuspended in 20μl of glyoxal mixture and incubated at 50°C for 1 hour. To each sample 2.5μl loading buffer was added, loaded into wells, and electrophoresed in running buffer at 60volts for about 2 hours, or until the dye had travelled two thirds of the length of the gel. The gel was stained for 30 minutes in 2μg/ml ethidium bromide, destained in DDW for 1 hour and examined on a UV transilluminator.

2.14. Fractionation of RNA by oligo(dT) cellulose column chromatography

This was carried out as described in Maniatis *et al.* (1982).

2.14.1. Solutions:

A) Binding buffer: 10mM TRIS-HCl pH 7.4, 1mM EDTA, 0.1% (w/v) SDS, 500mM LiCl.

B) 2x Binding buffer: double strength binding buffer.

C) Elution buffer: 10mM TRIS-HCl pH 7.4, 1mM EDTA, 0.1% (w/v) SDS.

2.14.2. Method: Oligo(dT) cellulose (Pharmacia) was equilibrated in elution buffer and packed into a siliconised pasteur pipette, plugged with sterile polyallomer wool, to a height of 1-2cm. The column was washed with elution buffer, and equilibrated with several volumes of binding buffer. 1ml of RNA (1-2mg in SDW) was mixed with 1ml of 2x binding buffer and applied to the column. The eluate was collected and recycled through the column three times. The poly(A-) fraction was eluted with several column volumes of binding buffer (about 10ml total volume) and precipitated with 2 volumes of ethanol. The poly(A+) fraction was eluted with a small volume of elution buffer, collected in aliquots of 400µl and the OD_{260} determined for each fraction. Those with an OD_{260} greater than 0.1 were pooled and precipitated with 0.1 volume of 3M potassium acetate pH 5.5 and 2 volumes of ethanol at -20°C for 2 hours. RNA was recovered by centrifugation in a Minifuge for 10 minutes and washed twice with 80% ethanol, dried and resuspended in SDW. For poly(A++) RNA, the poly(A+) fraction, after elution from the column, was mixed with an equal volume of 2x binding buffer and passed several times over a second, smaller column of oligo(dT) cellulose before ethanol precipitation.

2.15. *In vitro* translation of RNA

2.15.1. Preparation of wheatgerm extract

This was carried out as described by Roberts and Paterson (1973).

2.15.1.1.Solutions:

A) Extraction buffer: 20mM HEPES pH 7.6, 100mM potassium acetate, 1mM magnesium acetate, 2mM $CaCl_2$, 6mM DTT (added after autoclaving).

B) Chromatography buffer: 10mM HEPES pH 7.6, 100mM potassium acetate, 4mM magnesium acetate, 1mM DTT (added after autoclaving).

2.15.1.2.Method: 5-6g wheatgerm were ground in a chilled mortar and pestle, with a broken pasteur pipette to increase efficiency of extraction. 10-12ml of extraction buffer were added and ground to a thick paste. This was centrifuged at 13,000 r.p.m. in an SS-34 rotor at 4°C for 20 minutes. The supernatant was removed and applied to a Sephadex G-5o column (20xlcm) equilibrated in chromatography buffer. The extract was eluted with chromatography buffer and, after the appearance of the first cloudy drop, a volume equivalent to that applied to the column was collected. The eluate was centrifuged in an SS-34 rototr at 10,000 r.p.m. for 10 minutes at 4°C, the supernatant collected and frozen as small beads (about 10µl) by dripping through a syringe with 21 gauge needle into liquid nitrogen, and stored in liquid nitrogen or at -80°C.

2.15.2. *In vitro* translation with wheatgerm

2.15.2.1. Solutions:

A) 40mM ATP adjusted to pH 7.0 with 1N KOH, followed by addition of creatine phosphate to 320mM.

B) 10mM GTP.

C) 12mM spermidine.

D) 200mM DTT.

E) Amino acid mixture containing 1mM of each of the 20 amino acids, except those present in the tritiated amino acid mix (see below).

F) 5mg/ml creatine phosphokinase dissolved in 50% glycerol.

(Solutions 1-6 were made up in SDW.)

G) Magnesium stock: 160mM HEPES pH 7.6 (with KOH), 720mM potassium acetate 1.5mM magnesium acetate.

H) High specific activity mixture of amino acids labelled with tritium containing leucine, lysine, phenylalanine, proline, and tyrosine (Amersham. TRK.550).

2.15.2.2. Method: Just prior to use, the "energy mixture" was prepared by mixing on ice the following: 10μl ATP with creatine phosphate (A), 10μl amino acid mixture (E), 10μl spermidine (C), 4μl GTP (B), 4μl DTT (D), 4μl creatine phosphokinase (F), 8μl SDW, 100μl tritiated amino acid mixture (8).

Translation assays of 20μl total volume contained:

1. energy mixture	7.5μl
2. magnesium stock	2.5μl
3. RNA (1mg/ml)	1-5μl
4. SDW	to 7.5μl
5. wheatgerm extract (just thawed)	2.5μl

After brief centrifugation in a Minifuge the tubes were incubated at 28°C for 60 minutes.

Aliquots were removed for estimation of incorporated radioactivity, blotted onto small pieces of Whatman No.1 filter paper soaked with a few drops of 20% TCA, and dried at room temperature for about 1 hour. The filters were immersed in boiling 5% TCA containing 0.1% (w/v) casein hydrolysate and allowed to stand for 5-10 minutes, washed with 2 to 3 changes of cold 5% TCA, followed by washes in ether:IMS (1:1) and ether

alone, and allowed to dry. The filters were placed into scintillation vials containing 6ml scintillant (1l toluene, 4g PPO (2,5 diphenyloxazole) and 0.8g POPOP (1,4-bis[5-phenyl-2-oxazolyl]-benzene; 2,2'-p-phenylene-bis[5-phenyloxazole]; Sigma), and counted for 5 minutes.

2.15.3. *In vitro* translation with rabbit reticulocyte lysate

This method was originally described by Pelham and Jackson (1976), and assays were set up as described by the manufacturers of the rabbit reticulocyte lysate (Amersham).

Lysate was stored under liquid nitrogen in 50µl aliquots which were thawed immediately prior to use.

Assays contained: 7µl lysate
1µl [^{35}S]methionine (Amersham SJ.204)
2µl RNA solution (or SDW)

and were incubated at 30°C for 60 minutes.

1µl aliquots were removed for estimation of incorporated radioactivity and treated as described above.

2.16. Immunoprecipitation of *in vitro* translation products

(Lomedico *et al.*, 1984)

2.16.1. Solutions:

A) 10mM TRIS pH 7.5, 20mM EDTA, 15mM NaCl, 0.5mg/ml bovine pancreatic RNase (Boehringer).

B) 25mM TRIS-HCl pH 7.4, 10mM EDTA, 350mM NaCl, 0.1% (w/v) SDS, 1% (w/v) sodium deoxycholate, 1% (v/v) Triton X-100.

2.16.2. Method: *In vitro* translation assays were combined in an Eppendorf tube to give a total of 10^6 c.p.m. for each treatment, mixed with 100-200µl solution A and allowed to stand at room temperature for 30

minutes. The suspension was centrifuged in a Minifuge for 1 minute, and to the supernatant 200μl solution B containing antibodies at a concentration of 1mg/ml were added, and rotated for 2-3 hours at room temperature. Anti-rabbit IgG-agarose beads (Sigma A-1027) were washed with several changes of solution 2, and 200μl of the suspension was added to each assay tube and rotated for a further 2-3 hours at room temperature. The beads were briefly centrifuged and washed with many changes of solution 2 at room temperature over several hours. The final pellet was resuspended in 40μl cracking buffer (same as for gel electrophoresis), boiled for 3 minute, centrifuged and the supernatant electrophoresed on SDS-PAGE.

2.17. Peptide mapping (after Hames and Rickwood, 1981)

Actinidin was adjusted to 4mg/ml with distilled water and diluted with an equal volume of cracking buffer (see above) to 2mg/ml and boiled for 3 minutes.

Trypsin was made up to 2mg/ml with cracking buffer (but not boiled) and diluted in a 1:1 serial dilution with half-strength cracking buffer.

Equal volumes of actinidin and trypsin solutions were mixed and incubated at room temperature or at 37°C for 0, 5 and 10 minutes. The reaction was stopped by boiling the samples for 3 minutes. 20μl aliquots were electrophoresed on SDS-urea gels.

2.18. cDNA synthesis (Gubler and Hoffman, 1983)

2.18.1. First strand synthesis

2.18.1.1. Solutions:

A) RVT1: (100μl)

	volume (μl)	final conc.
1M KCl	15	150mM
1M TRIS-HCl pH 8.3	15	150mM
1M $MgCl_2$	3	30mM
1M DTT	3	30mM
100mM spermidine	12	12mM
BSA (1mg/ml)	30	0.3mg/ml
SDW	22	

B) RVT2: 5mM each of dATP, dCTP, dGTP and dTTP.

C) 5X second strand mix: 100mM TRIS pH 7.5, 25mM $MgCl_2$, 50mM $(NH_4)_2SO_4$, 500mM KCl.

D) 200μM each of dATP, dCTP, dGTP and dTTP.

2.18.1.2. Method: Assays were set up by mixing 0.5 to 1 μg mRNA, 3μl oligo$(dT)_{12-18}$ (1mg/ml), 10μl RVT1, 8μl RVT2, 1μl $\alpha[^{32}P]$dCTP (>3000Ci/mmol, 10mCi/ml),1μl reverse transcriptase (Pharmacia or nbl), and adjusting the volume to 30μl. Incubation was for 1 hour at 42°C, and the reaction was stopped by adding 3μl of 250mM EDTA. Unincorporated nucleotides were removed by Sephadex G-50 column chromatography (see 2.20.) and the efficiency of first strand synthesis calculated from the fraction of incorporated label.

For the synthesis of radioactively labelled first strand cDNA as hybridisation probe, the concentration of dCTP was reduced to 1/8, and 5μl $\alpha[^{32}P]$-dCTP used instead of 1μl.

2.18.2. Second strand synthesis

Assays were set up by mixing 50-200ng DNA:RNA hybrid, 10μl solution D, 2.5μl BSA (1mg/ml, nuclease-free), 1μl RNase H (0.4 units/μl), 1μl DNA

Polymerase I (Boehringer, 5 units/μl), 10μl 5X second strand mix, 1μl $\alpha[^{32}]$dCTP (>3000Ci/mmol, 10mCi/ml), and made up to 50μl with SDW. Incubation was for 1 hour at 12°C followed by 1 hour at 22°C. Unincorporated nucleotides were removed by column chromatography as above, and double-stranded cDNA was phenol extracted and precipitated as above.

2.19. C-tailing of double-stranded cDNA

2.19.1. Solution:

5X tailing buffer: 500mM potassium cacodylate pH 7.2, 10mM $CoCl_2$, 1mM DTT.

2.19.2. Method: Assays were set up by mixing 25ng double-stranded cDNA, 6μl tailing buffer, 2μl of 1mM dCTP, 1μl of 0.05mM dideoxy CTP, 0.4μl α[32P]dCTP (>3000 Ci/mmol, 10mCi/ml), 1μl terminal deoxynucleotidyl transferase (BRL, 17u/μl), and adjusted to 30μl with SDW. Incubation was for 1 hour at 37°C. Unincorporated nucleotides were removed by column chromatography and levels of incorporation estimated from the fraction of incorporated label. The average length of C-tails was calculated on the basis of cDNA size estimated by denaturing polyacrylamide gel electrophoresis.

2.20. G-50 Sephadex column chromatography (Maniatis *et al.*, 1982)

A pasteur pipette was plugged with polyallomer wool and filled to within 0.5cm of the top with Sephadex G-50 (medium), previously equilibrated with TE. The sample containing radioactively labelled DNA was applied in a volume of approximately 100μl, eluted with TE, and fractions of 100μl were collected. After Cherenkov counting of individual fractions, those containing the first peak of radioactivity were pooled, extracted with phenol/chloroform (1:1) and chloroform/isoamylalcohol (24:1), and the DNA was precipitated by addition of 0.1 volume of 2M sodium acetate pH 5.6 and two volumes of ethanol at -70°C for 30 minutes. The precipitate was

collected by centrifugation in a Minifuge for 20 minutes, washed once with 80% (v/v) ethanol and resuspended in SDW.

2.21. Denaturing polyacrylamide gel electrophoresis of DNA

2.21.1 Solutions:

A) 10X TBE buffer: 0.9M TRIS, 0.9M boric acid, 25mM EDTA.

B) 30% (w/v) acrylamide, 1.2% (w/v) bisacrylamide.

C) 10% (w/v) ammonium persulphate (freshly prepared).

D) formamide-dye mix: 90% (v/v) formamide A.R., 25mM EDTA, 0.5% (w/v) bromophenol blue, 0.5% (w/v) xylene cyanol.

2.21.2. Method: 5% (w/v) acrylamide gels were prepared by mixing 2.5ml solution A, 21g urea, 8.22ml solution B, and DDW to 50ml with warming at 37°C, followed by 100µl of solution C and 75µl TEMED. DNA samples were mixed with an equal volume of formamide-dye solution and boiled for 3 minutes. Electrophoresis was carried out at 45mA in 0.5X TBE until the tracking dye was near the bottom of the gel. The gel was wrapped in cling film and autoradiographed overnight at -80°C.

2.22. Annealing of homopolymer-tailed cDNA with G-tailed pUC9

2.22.1. Solutions:

A) 5X annealing buffer: 50mM TRIS-HCl pH 7.6, 5mM EDTA, 750mM NaCl.

2.22.2. Method: Oligo(dG)-tailed pUC9 (Pharmacia) was diluted to 6.6µg/ml, and cDNA was adjusted to 1µg/ml with SDW. Annealing reactions contained 20µl of 5X annealing buffer in 100µl of total final volume, and varying ratios of plasmid to cDNA as shown in Results. Incubation for 5 minutes at 65°C was followed by 60 minutes at 44°C and cooling to 28°C.

2.23. Preparation of competent *E.coli* cells and transformation with recombinant plasmid or phage (Maniatis *et al.*, 1982)

2.23.1. Bacterial strains:

JM 83 *rk⁻, mk⁻, sup', ara, lac-pro, strA, thi-1, ΔlacZM15.*

JM 103 Δ(*lac pro*), *thi, strA, supE, endA, sbcB, hsdR⁻*, F'*traD36, proAB, lacI^q, ZΔM15.*

2.23.2. Solutions:

A) L-broth: 10g/l bactotryptone, 5g/l yeast extract, 5g/l NaCl, 1g/l glucose, adjusted to pH 7.5 with 1N NaOH.

B) L-plates: as L-broth, but containing 15g/l Bacto-agar.

C) 20X SSC: 3M NaCl, 300mM sodium citrate, adjusted to pH 7 with 10N NaOH.

D) top-agar: L-agar containing 0.7% (w/v) agar, containing 2mg/ml IPTG and 1.8mg/ml X-Gal.

2.23.3. Method: For transformation with pUC plasmid, *E.coli* JM83 were grown in 100ml L-broth (containing 100μg/ml streptomycin) from an overnight culture to OD_{600} 0.5 to 0.6, centrifuged in a pre-cooled Sorvall SS-34 rotor at 10000 r.p.m. for 5 minutes, and the pellet was resuspended in half of the original volume of ice-cold 100mM $MgCl_2$. After 5 minutes at 0°C, the cells were pelleted as above, resuspended in ice-cold 100mM $CaCl_2$, and left at 0°C for 20 minutes. Cells were pelleted, resuspended in one twentieth of the original volume of $CaCl_2$, and kept at 0°C until required.

Transformation mixtures contained 200μl cell suspension, 100μl of 1X SSC (ice-cold), and 10μl DNA solution, and were held at 0°C for 30 minutes. The cells were heat-shocked at 42°C for 2 minutes and transferred back to 0°C for 20 minutes. The suspensions were mixed with 4 volumes of L-broth and incubated at 37°C with shaking for 70 minutes. Cells were pelleted in Eppendorf tubes for 1 minute, resuspended in 100μl L-broth, plated onto L-agar containing 100μg/ml streptomycin, 50μg/ml ampicillin and 40μg/ml

X-gal, and incubated at 37°C overnight.

For transformation of *E.coli* JM 103 with M13 phage, competent cells were prepared as described above, but grown in the absence of antibiotics. 10μl of transforming DNA were mixed with 200μl cell suspension, and kept at 0°C for 40 minutes. The suspension was heat-shocked at 42°C for 2 minutes, mixed with 3ml top agar at 42°C, and poured over an L-agar plate.

2.24. Colony hybridisation

2.24.1. Solutions:

A) 5X Denhard's: 0.1% (w/v) PVP, 0.1% (w/v) Ficoll 400, 0.1% (w/v) BSA (Sigma, fraction V).

B) Filter hybridisation mix: 20ml of 5X Denhard's, 30ml of 20X SSC, 50ml DDW.

C) denatured salmon sperm DNA: DNA was dissolved in DDW at 5mg/ml and sonicated for 10 minutes, boiled for 5 minutes and stored frozen. Before addition to hybridisation solution, it was boiled briefly and cooled on ice.

D) Complete filter hybridisation mix (CFHM): Solution B containing 0.05mg/ml denatured salmon sperm DNA

2.24.2. Method: Recombinant transformant colonies were transferred to replica nitrocellulose filters (Schleicher and Schuell) and grown up overnight on L-agar. DNA was immobilised by placing the filters on Whatman 3mm paper saturated with the solutions and for times as follows: 500mM NaOH, 7 minutes; 1.5M NaCl, 100mM NaOH, 10 minutes; 1M TRIS pH 7.4, twice 2 minutes; 1.5M NaCl, 500mM TRIS pH 7.4, 4 minutes. The filters were blotted dry and baked at 80°C for 3 hours.

Filters were soaked in 6X SSC for 30 minutes at room temperature and pre-hybridised in CFHM for 1 hour. The radioactive probe was added and hybridisation carried out overnight, at the temperatures and using the

probes indicated in Results. Filters were washed over about 3 hours at the temperatures and with concentrations of SSC as indicated in Results. For autoradiography the filters were air-dried and wrapped in cling film, and X-ray film was exposed at -80°C with an intensifying screen.

2.25. Plasmid DNA "Miniprep" (Maniatis *et al.*, 1982)

1.5ml of an overnight culture were pelleted and the cells resuspended in 100µl lysis buffer containing 25mM TRIS-HCl pH 8.0, 10mM EDTA, 500mM glucose, and 1mg/ml lysozyme. After holding on ice for 10 minutes, 200µl of alkaline SDS (200mM NaOH, 1% (w/v) SDS) were added, followed after 5 minutes by 150µl of 3M potassium acetate pH 5.2, and kept on ice for 10 minutes. The precipitate was removed by centrifugation, the supernatant filtered through polyallomer wool, and DNA precipitated by addition of two volumes ethanol at -70°C for 15 minutes. The DNA was pelleted and resuspended in 100µl TE, re-precipitated and washed with 70% (v/v) ethanol. The final pellet was resuspended in 40 µl TE.

2.26. Large scale plasmid isolation (Maniatis et al., 1982)

500 ml cultures were grown in L-broth overnight, cells pelleted in a Sorvall GSA rotor at 8000 r.p.m. for 5 minutes at 4°C, resuspended in 20ml lysis solution (as above) and left at 0°C for 5 minutes. 40ml alkaline SDS (as above) were added at 0°C, and left for 10 minutes, followed by 30ml of 3M potassium acetate pH 5.2. After removal of the precipitate by centrifugation as before, the supernatant was filtered through polyallomer wool, and 60ml isopropanol were added to it. The DNA was pelleted as before, rinsed with 70% (v/v) ethanol, dried under a stream of nitrogen gas, and dissolved in 5ml TE. The volume was gravimetrically adjusted to 8.75g with TE, and 1.75ml ethidium bromide (5mg/ml) were added, followed by 10.4g CsCl, and mixed to dissolve. A gradient was formed by

centrifugation in a Ti rotor at 44,000 r.p.m. for 22 hours at 15°C, the plasmid DNA recovered with the help of a syringe and EtBr removed by repeated extraction with isopropanol saturated with 20X SSC.

2.27. Agarose gel electrophoresis of DNA and Southern blotting

Agarose gels of approximately 0.5cm thickness were prepared with 1% (w/v) agarose (Seakem) in ELFO buffer containing 40mM TRIS and 1mM EDTA. DNA was digested with restriction endonucleases using the conditions specified by the manufacturers (Pharmacia). DNA samples were mixed with a sixth volume of loading buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 30% (v/v) glycerol), and electrophoresed at 80V. The gels were depurinated in 250mM HCl for 7 minutes, rinsed in DDW, denatured in 500mM NaOH, 1.5M NaCl for 30 minutes, rinsed, and neutralised in 3M NaCl, 500mM TRIS-HCl pH 7.4 for 30 minutes. The gel was placed on a blotting tray, overlayed with nitrocellulose pre-soaked in 3X SSC, followed by Whatman 3mm paper and a stack of paper towels weighed down with a glass plate. After blotting overnight, the filter was rinsed briefly in 3X SSC, air-dried and baked at 80°C for 2 hours.

2.28. Recovery of DNA fragments from agarose gels

2.28.1. Electrophoresis onto DE81 paper: After electrophoresis, fragments were located under UV light and the gel was cut below the fragment of interest. A piece of DE81 paper, previously washed in 2.5M NaCl and stored in 1mM EDTA, was inserted into the cut and electrophoresis continued until the fragment reached the paper. This was removed, washed in several changes of DDW, and the DNA was eluted by vortexing in 450µl of high salt buffer (1.5M NaCl, 50mM TRIS-HCl pH 8.0, 1mM EDTA), incubating at 37°C for 1 hour, followed by 15 minutes at 65°C. Debris was removed by filtration through polyallomer wool, and EtBr removed by two extractions

with butanol. DNA was precipitated with 2 volumes ethanol and resuspended in 200µl SDW.

2.28.2. Low melting temperature agarose gels: Gels were prepared with 1% (w/v) low melting temperature agarose, and DNA electrophoresed as usual. The DNA fragments were visualised under long wave UV light, and cut out of the gel, removing excess gel. To each g of gel, 1.5ml SDW were added, boiled for 7 minutes, and stored frozen. Prior to use, the DNA was boiled for 5 minutes and held at 37°C until required.

2.29. Northern Blot

2.29.1. Solutions:

A) 5X gel running buffer: 200mM MOPS pH 7.0, 50mM sodium acetate, 5mM EDTA pH 8.0.

B) Loading buffer: 50% (v/v) glycerol, 1mM EDTA, 0.4% (w/v) bromophenol blue, 0.4% (w/v) xylene cyanol.

2.29.2. Method: The gel was prepared by melting 2g agarose in 50ml SDW and cooling to 60°C, mixing with 20ml of 5X gel buffer and formaldehyde to a final concentration of 2.2M, and adjusting the volume to 100ml with SDW. RNA was denatured by mixing 4.5µl RNA with 2µl 5X gel rinning buffer, 3.5µl formaldehyde, and 10µl formamide. Samples were incubated at 55°C for 15 minutes and electrophoresed in 1X gel running buffer at 40V. The gel was washed in several changes of DDW for 5 minutes, soaked in 50mM NaOH, 10mM NaCl for 45 minutes at room temperature, and neutralised in 100mM TRIS-HCl pH 7.5 for 45 minutes. After soaking in 20X SSC for 1 hour, transfer to nitrocellulose was carried out exactly as for Southern blotting.

2.30. Dot blot of DNA and RNA

DNA and RNA were immobilised on nitrocellulose by methods described by

Thomas (1980). DNA samples were dessicated and dissolved in 1M NaCl, 100mM NaOH, 10mM EDTA, and heated at 100°C for 5 minutes. The samples were spotted onto nitrocellulose filters, left to dry at room temperature for 30 minutes, rinsed in 3X SSC, blotted dry and baked at 80°C for 4 hours.

RNA samples were dissolved in 6X SSC and applied to nitrocellulose filters saturated with 20X SSC. The filters were air-dried and baked as above.

2.31. DNA sequencing

This was carried out by the method of Sanger *et al.* (1977)

2.31.1. Template preparation

Overnight cultures of *E.coli* JM103 were diluted 200-fold into 2ml of L-broth, grown for 1 hour at 37°C, and inoculated with single plaques. The cultures were incubated at 37°C with very fast shaking (300 r.p.m.) for 7 hours, decanted into Eppendorf tubes and centrifuged for 5 minutes. To 1 ml supernatant, 200 µl of 20% (w/v) PEG 6000, 2.5M NaCl were added, mixed and left at room temperature for 15 minutes. The phage was pelleted for 10 minutes, the supernatant removed by aspiration and repelleting using a drawn-out pasteur pipette, and resuspended in 100µl TE. The suspension was extracted by vortexing with 50µl phenol (equilibrated in TE), allowed to stand for 5 minutes and vortexed again. After centrifugation for 2 minutes, the phenol extraction was repeated, followed by extraction with chloroform:isoamylalcohol (50:1). DNA was precipitated with 0.1 volume of 2M sodium acetate pH 5.5, and 2.5 volumes ethanol at -70°C for 1 hour, centrifuged for 10 minutes, washed with 70% (v/v) ethanol, vacuum dried, and resuspended in 50 µl TE.

2.31.2. Sequencing gel

Sequencing gels were 39cm long and were prepared as wedge gels, with a single thickness of spacer (0.4mm) at the top and double thickness (0.8mm)

at the bottom of the gel. Acrylamide solution, sufficient for two gels, was prepared by mixing 17.1g acrylamide, 0.9g bisacrylamide, 150g urea, 30ml of 10X TBE, and made up to 300 ml with DDW. To 150ml of acrylamide solution 1.05ml of 10% (w/v) ammonium persulphate and 72µl TEMED were added and the gel poured.

2.31.3. Sequencing reactions

2.31.3.1. Solutions: A) dideoxy CTP 250mM in TE
dideoxy ATP 200mM "
dideoxy GTP 160mM "
dideoxy TTP 500mM "

B)	500mM dCTP	500mM dGTP	500mM dTTP	
C Mix:	10	200	200	µl
A Mix:	200	200	200	µl
G Mix:	200	10	200	µl
T Mix:	200	200	10	µl

Complete C, A, G and T Mix was prepared by mixing equal volumes of C Mix and dideoxy CTP, A Mix and dideoxy ATP, G Mix and dideoxy GTP, and T Mix and dideoxy TTP.

C) 10 seq: 100mM TRIS-HCl pH 7.5, 50mM $MgCl_2$.

D) Sequence chase mix: 0.5mM each of dCTP, dATP, dGTP and dTTP.

E) formamide dye mix: 90% (v/v) formamide, 10% (w/v) 500mM EDTA, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol.

F) 2µg/ml single stranded universal primer (17mer, Pharmacia).

G) Klenow mix (for 10 templates): 78µl TE, 2.2µl Klenow fragment, 10µl $\alpha[^{32}P]$dATP (>1000Ci/mmol; Amersham).

2.31.3.2. Protocol:

For 10 templates, primer mix was prepared with 5µl primer (F), 33µl 10seq (C), and 42µl SDW.

For each template 4 tubes were prepared and 2µl primer mix added, followed

by 2μl template. Annealing was carried out at 65°C for 5 minutes, followed by 15 minutes (or longer) at room temperature. 2μl of the appropriate complete mix was added to each tube, followed by 2μl Klenow mix. The assays were incubated at room temperature for 35 minutes, and 2μl chase mix were added and left for a further 15 minutes. 2μl formamide-dye mix were added, and the tubes heated to 80°C with open tops for 7 minutes to reduce the volume of liquid prior to electrophoresis. The gel was prefocused at 900V for 15 minutes, and samples were loaded into wells and electrophoresed at 1200V/60W until the lower tracking dye reached the bottom of the gel. The gel was fixed in 10% (v/v) acetic acid, 10% (v/v) ethanol for 15 minutes and dried down onto Whatman 3mm filter paper on a heated gel drier. Autoradiography was at -80°C overnight using intensifying screens.

2.32. Oligo-labelling of DNA fragments

This method was originally described by Feinberg and Vogelstein, 1983.

2.32.1. Solutions:

A) 2M TRIS-HCl pH 8.0 625μl
   5M $MgCl_2$ 25μl
   SDW 350μl
   β-mercaptoethanol 18μl
   100mM dATP (in TE) 5μl
   100mM dTTP " 5μl
   100mM dGTP " 5μl

B) 2M HEPES-NaOH pH 6.0.

C) 90 OD_{260} units/ml hexadeoxy ribonucleotides (Pharmacia), suspended in 3mM TRIS-HCL pH 7.0, 0.2mM EDTA.

D) Oligo-labelling buffer (OLB) was prepared by mixing solutions A, B and C in the ratios 2:5:3.

E) Stop solution: 20mM NaCl, 20mM TRIS-HCl pH 7.5, 2mM EDTA, 0.25% (w/v) SDS.

2.32.2. Method: Approximately 10ng of the DNA fragment to be labelled was boiled for 5 minutes, and transferred to 37°C prior to use.
Assays were assembled by mixing 3μl OLB, 0.6μl BSA, 10ng DNA, 1.5μl α[^{32}P]-dCTP, 0.6μl Klenow fragment, and made up to 15μl with SDW.
Incubation was at room temperature for 5 hours or overnight, and the reaction was stopped by addition of 85μl stop solution.
Unincorporated label was removed by Sephadex G-50 column chromatography.

2.33. End-labelling of synthetic oligonucleotide

(Maniatis *et al.*, 1982)

2.33.1. Solution:

10X Kinase buffer: 500mM TRIS-HCl pH 7.6, 100mM $MgCl_2$, 50mM DTT, 1mM spermidine, 1mM EDTA.

2.33.2. Method: 0.5 μg synthetic oligonucleotide was mixed with 1.5μl SDW, 2.5μl of 10X kinase buffer, 10μl γ[^{32}P]ATP (>3000Ci/mmol), and 2μl T_4 polynucleotide kinase (Amersham, 3.6 u/μl), and incubated at 37°C for 1 hour. Unincorporated label was removed by Sephadex G-50 column chromatography.

2.34. Isolation of plant DNA

The method was described by Murray and Thompson (1980), and followed the protocol decribed by Lichtenstein and Draper (1985).

2.34.1 Solutions:

A) extraction buffer: 1% (w/v) CTAB, 50mM TRIS-HCl pH 8.0, 10mM EDTA, 0.7M NaCl.

B) precipitation buffer: 1% (w/v) CTAB, 50mM TRIS-HCl, pH 8.0, 10mM EDTA, 1% β-mercaptoethanol, added just prior to use.

C) 10% CTAB solution: 10% (w/v) CTAB, 0.7M NaCl.

2.34.2. Method: 0.1g of freeze-dried plant material was ground to a fine powder with alumina in a mortar and pestle, immersed in 600µl extraction buffer, and incubated at 56°C for 10 minutes. The emulsion was extracted with 600µl chloroform:octanol (24:1 v/v), and centrifuged in a microcentrifuge for 10 minutes. The interface was re-extracted once with 100µl extraction buffer and aqueous phases were pooled. After addition of 1/10 volume of 10%CTAB solution, the aqueous phase was re-extracted with chloroform: octanol. Nucleic acid was precipitated by addition of an equal volume of precipitation buffer and standing at room temperature for 20 minutes. The precipitate was pelleted by centrifugation for 15 minutes and thoroughly drained. The pellet was resuspended in 400µl of 1M NaCl, and precipitated with two volumes of ethanol at -20°C overnight. After centrifugation in a microcentrifuge the pellet was washed twice in 70% ethanol, and resuspended in 100µl of SDW.

2.35. Synthetic oligonucleotide

This was synthesised by J.Keyte, Department of Biochemistry, University of Leicester.

3. Identification of actinidin in various parts of the Chinese gooseberry plant

3.1. Introduction

A study investigating the levels of actinidin protein in different organs of the Chinese gooseberry plant was considered important for two reasons:

1) It was expected to reveal whether actinidin is a fruit-specific protein or whether it is also present in other parts of the plant.

2) In order to establish a cDNA library, it would be preferable to chose, as a source of RNA, a tissue containing large amounts of the message of interest, and protein levels may provide an initial indication as to which tissue may be most suitable. Although fruit tissue may have been expected to be the most likely material to fulfill these requirements, it was found to contain only small amounts of RNA, and was not amenable to a variety of RNA extraction methods (see Chapter 4). Although ultimately proved unnecessary, the possibility of using tissue other than fruit as a source of RNA for cDNA cloning was therefore considered.

The most specific and sensitive methods for detecting protein in tissue extracts involve the use of antibodies. This Chapter describes the purification of actinidin from Chinese gooseberry fruit, the raising and testing of antibodies to the protein, and the use of the antiserum in estimating actinidin levels in various organs of the Chinese gooseberry plant.

3.2. Results

3.2.1. Growth of seedlings

Initially seeds were removed from ripe fruit manually, imbibed in running tap water for two days and germinated on filter paper. After 6 days radicles appeared on 6 out of 200 seeds. The remaining seeds were given various treatments to try to improve germination: Seed coats were completely (a), or partially (b), removed, and seeds placed on fresh filter paper; some seeds were only placed on new filter paper (c), others were left unchanged (d). Each treatment was also divided into light and dark incubation. After a further 7 days all seeds of treatment (a) and most of treatment (b) had germinated, but (c) and (d) showed only about 5% germination. The results clearly suggest an inhibitory function of the seed coat. However, seeds which were subsequently recovered from homogenised fruit extracts (using a blender) germinated more readily (30% of seeds within two weeks) without any treatment, presumably because their seed coats had been damaged slightly during homogenisation. The number of seedlings established with this method was sufficient and for most purposes 6 months old plants were used.

3.2.2. Extraction and purification of actinidin

First attempts at preparing actinidin, following the original method by Carne and Moore (1978), encountered problems at the column fractionation stage due to the accumulation of a gelatinous mass, causing complete blockage, at the top of the column. Addition of polyethylene imine, which precipitates nucleic acids and carbohydrates, was suggested by G.Whitelam (this laboratory) and prevented gelation of the extract in further extractions.

Figure 3.1 shows the elution profile of protein and protease activity on ion exchange column, from an extraction using 250g of fruit (OD_{280} values for fractions 27 to 31 were not determined). Although the position of the protein peak is not known, it appears that some protein without protease activity eluted before the main protease peak, which is in agreement with published results (Boland and Hardman, 1972; Carne and Moore,1978). The majority of activity eluted in five fractions (22.5ml total). Attempts at crystallising the enzyme by dialysis against a 20% saturated solution of ammonium sulphate in potassium phosphate buffer (Carne and Moore, 1978) were unsuccessful. Therefore the pooled fractions were first precipitated with 33% saturated ammonium sulphate, and the precipitate pelleted. Upon further addition of ammonium sulphate to 95% saturation more protein precipitated. The "33%" and "95% precipitates" were dialysed separately, after which the 33% fraction contained insoluble precipitate which was removed by centrifugation. The yield of total protein was 16mg in the 33% fraction and 34mg in the 95% fraction.

Samples from several stages of actinidin extraction were electrophoresed on a 10% acrylamide gel (Figure 3.2a) in order to follow the progress of purification. The major protein in the original extract (lane 2) appears to be the same as the final product, although concentrating the extract for application of a larger protein sample may have allowed detection of further protein bands. A protein of lower molecular weight, present both in the original extract and the calcium phosphate supernatant, was apparently removed during column chromatography (compare lanes 10 and 11). All final actinidin preparations were grossly overloaded and contained some small protein impurities which may not have been detected with "normal loading". It is interesting that the original extract contained such few additional proteins and apparently none of higher

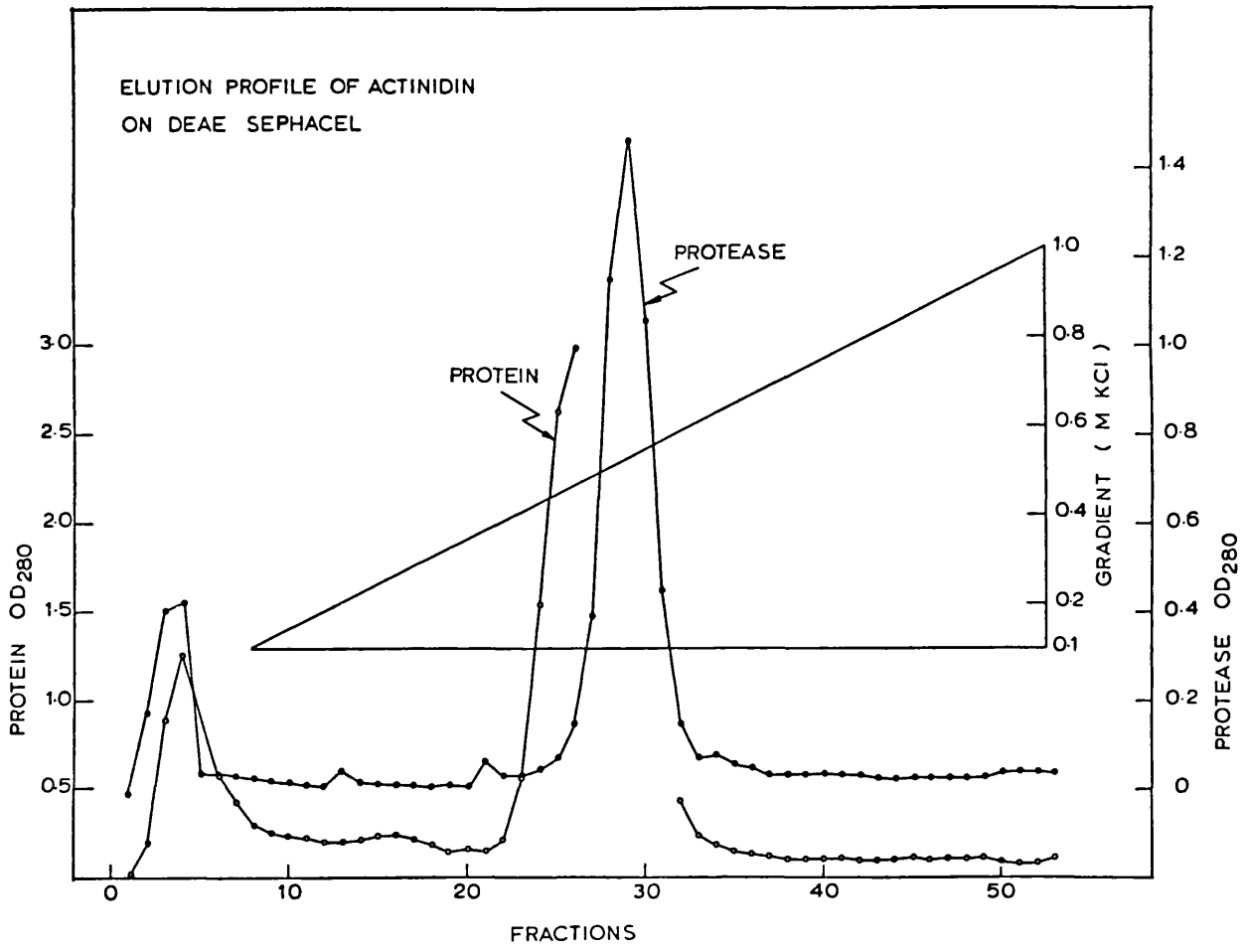


Figure 3.1. Purification of actinidin by column chromatography on DEAE Sephacel (first batch)

Protein was eluted with 2l of potassium chloride gradient (0.1M to 1.0M) and collected in 5ml fractions.
Protein concentration was estimated by OD_{280} determination of fractions. Protease activity was measured by casein hydrolysis as described in Methods.

molecular weight. Although McDowall (1969) found that actinidin accounted for 50% of the protein in centrifuged extract, the results from SDS-PAGE suggest that this value may be as much as 75%. It is possible that very ripe fruit such as that used for actinidin preparation contains very little other protein and this may or may not be a result of hydrolysis by the protease following cellular breakdown during ripening. Alternatively, most other proteins may have been digested by actinidin during the extraction procedure.

Since the actinidin purified in this way was to be used for the production of antibodies it would be desirable to purify the protein to give only a single band on gel electrophoresis. Although this may have been achieved by eluting a single band from a preparative gel, an alternative approach was tested: actinidin was prepared from 420g fruit as before but the KCl gradient for elution was more shallow (0.3M to 0.7M) in the hope that some of the better separated fractions contained less impurities. Figure 3.3 shows that protease activity eluted in a larger volume (60ml) and that the elution of protein was more closely followed by protease elution. The fractions were separately precipitated with ammonium sulphate to 75% saturation. The change in composition of the small molecular weight contaminants can be seen in figure 3.2b which shows gel electrophoresis of freeze-dried samples from the different stages and column fractions. The original extract contained a trace of a higher molecular weight protein and a number of smaller proteins whose composition was the same in samples from all purification stages up to column chromatography. The column fractions also contained low molecular weight contaminants which comigrated with the proteins of the unfractionated extract. Very little difference in the composition of low molecular weight impurities can be detected between early and late fractions, with the exception of one band

Figure 3.2.A. Actinidin purification (first batch) monitored by SDS-PAGE

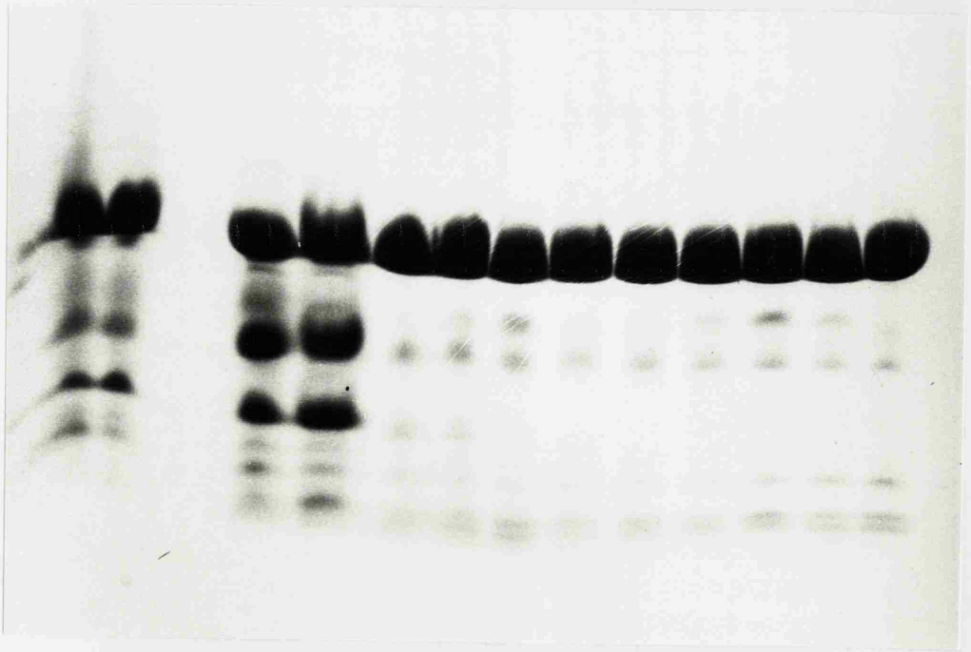
(10% acrylamide gel)

1 papain, 10μg
2 original extract after centrifugation, 20μl undiluted
3 " " " " 2oμl of 20-fold dilution
4 extract after PEI treatment, 20μl undiluted
5 " " " 20μl of 10-fold dilution
6 supernatant of first ammonium sulphate precipitation, 20μl
7 " " " " " "
8 papain, 10μg
9 final actinidin preparation after dialysis, "33% precipitate" , 80μg
10 " " " " " " , 160μg
11 supernatant of calcium phosphate treatment, 20 μl
12 insoluble pellet of 33% "precipitate"
13 " " " "
14 final actinidin preparation after dialysis, "99% precipitate", 228μg
15 " " " " " " , 114μg

Figure 3.2.B. Actinidin purification (second batch) monitored by SDS-PAGE
(12-20% gradient gel)
1 250μl centrifuged extract
2 400μl PEI-treated extract
3 300μl dialyzed ammonium sulphate precipitate
4 300μl calcium phosphate supernatant
5-13 fractions 25 to 33 from DEAE-Sephacel column, 200μg protein each

A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



B

1 2 3 4 5 6 7 8 9 10 11 12 13

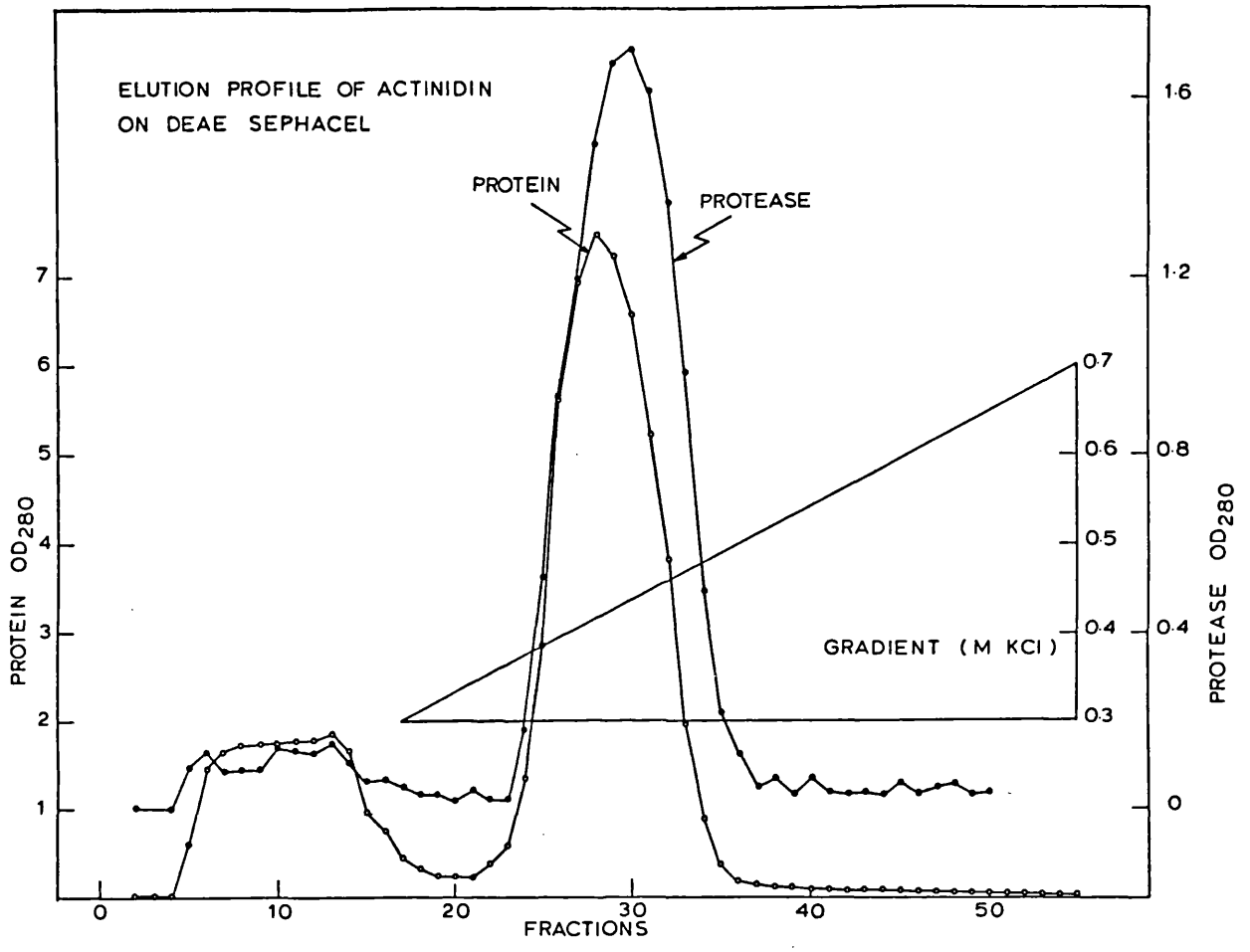


Figure 3.3. Purification of actinidin by column chromatography on DEAE Sephacel (second batch)

Protein was eluted with 2l of a shallower gradient (0.3M to 0.7M) of potassium chloride than that used for the first batch (compare Figure 3.1). Fractions of 5ml were collected, and protein concentration was estimated by OD_{280} determination of fractions. Protease activity was measured by casein hydrolysis as described in Methods.

present in early and absent in late fractions (see arrow). A slight increase in the concentration of the smallest proteins during the purification procedure and remaining in the column fractions may indicate hydrolysis of proteins by the enzyme.

Protein and protease activity were estimated for each fraction (Table A.1, Appendix). Although the protease assay was only comparative rather than quantitative, it is possible to express relative protease activity in terms of total protein per fraction. The plot of specific activity in Figure 3.4 suggests that the later column fractions were much more active against casein than the early ones. Clearly, comparing these results with those from SDS-PAGE, this difference in activity cannot be explained by a difference in the amount of actinidin present, but suggests that most of the protein in early fractions consisted of inactive enzyme. Enzyme activity of individual fractions was compared to activity of known amounts of papain for which a standard curve was constructed (see Figure A.1., Appendix). From the results in Table A.1 (Appendix) it appears that a total of 135 mg of actinidin from all fractions was equivalent to on 2.6mg of papain, or expressed differently, actinidin had only 1.9% of the activity of papain. This difference is unlikely to be a result of differences in optimal reaction conditions for the two enzymes, since both are active over a wide range of pH values (McDowall, 1970). Taken together the results suggest that actinidin was purified mainly in its inactive form. This is consistent with the results obtained by McDowall (1970) who obtained twice crystallised actinidin containing 70% inert protein, which was partially resolved on Sephadex G-50 gel filtration, inactive protease eluting ahead of the active form. Boland and Hardman (1972) also identified two forms of the enzyme, with identical molecular weight and charge, but resolvable on gel filtration, and concluded that

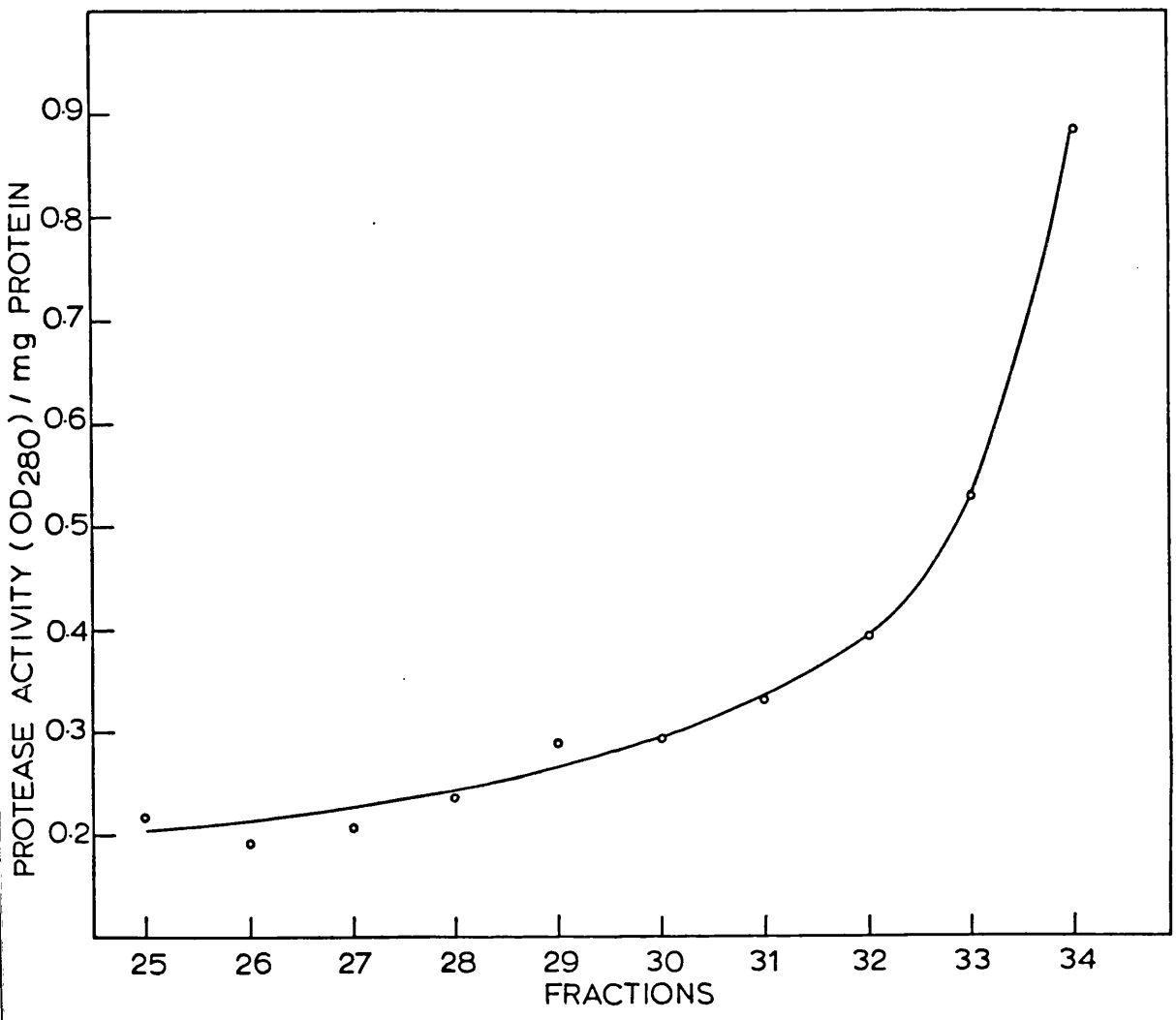


Figure 3.4. Fractionation of fruit extract DEAE-Sephacel

Specific activity of actinidin in column fractions:
Protease activity was determined for each fraction by casein hydrolysis, and specific activity expressed as OD_{280} of the supernatant per g protein in the fraction.

the active site thiol group may have been oxidised to a sulphonate giving rise to inactive enzyme.

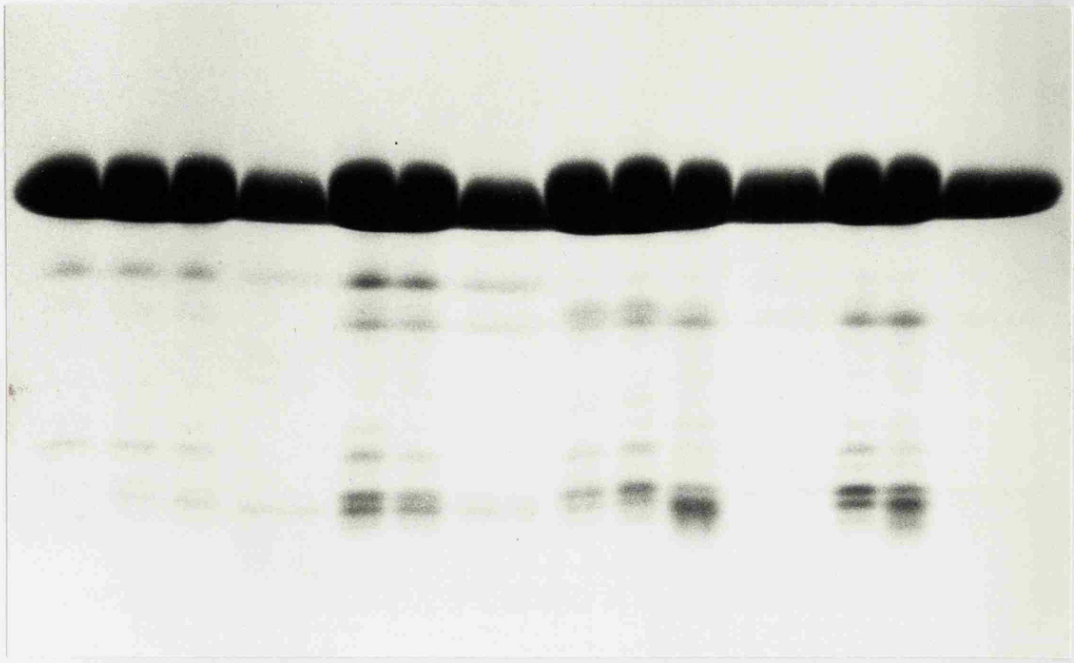
The purpose for which actinidin was purified does not require the enzyme to be active and indeed an inert protease preparation may be a preferable antigen compared with a highly active enzyme.

3.2.3. Stability of actinidin

The possibility that auto-digestion of actinidin could present difficulties in the production of antibodies to the enzyme was considered. A sample of crystalline actinidin (a gift from C.H.Moore in New Zealand), was included in a study of autolysis, using actinidin from the first batch and from two fractions of the second batch. Aliquots of actinidin from the four sources were incubated at 40°C for 24 hours in the presence or absence of cysteine, using the assay protocol for protease activity. All samples were freeze-dried, together with control samples which had not been incubated, and 200μg of protein were electrophoresed on a gradient SDS-PAGE gel shown in Figure 3.5. A comparison of untreated samples (lanes 1,5,8 and 12) suggests that the enzyme from New Zealand (lane 1) contained less low molecular weight contaminants, although those bands present had the same electrophoretic mobility as the contaminants of the first batch (lane 5). Actinidin extracted in the second batch showed slightly more impurities, the largest contaminating band from the first batch being absent, but with more of the fast migrating bands.

Freeze-drying appears not to have affected the composition of proteins as judged from SDS-PAGE (compare lanes 2,8 and 9).

Incubation of actinidin in the absence of cysteine resulted in a slight increase in low molecular weight proteins only in the two samples from the second batch (lanes 10 and 13). After incubation in the presence of



1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 3.5. Stability of actinidin: SDS-PAGE

1- 4	200µg actinidin from New Zealand
5- 7	200µg actinidin from first batch
8-11	200µg actinidin, fraction 27 of second batch
12-14	200µg actinidin, fraction 30 of second batch

1,8	untreated control
2,5,9,12	freeze-dried control
3,6,10,13	incubated without cysteine and freeze-dried
4,7,11,14	incubated in presence of cysteine and freeze-dried

cysteine (lanes 4, 7, 11 and 14) virtually all contaminants from fractions 27 and 30 of the second batch were absent, but only slightly reduced in the samples from New Zealand and the first batch. Although difficult to assess, the amounts of intact enzyme seemed not to be reduced by any of the treatments and thus it appears that actinidin is not a good substrate for itself. Alternatively, if actinidin is present mostly in the inactive form it is possible that this would prevent auto-digestion. It is worth noting that the actinidin supplied by Moore had been prepared in crystalline form, migrating as a single band on denaturing PAGE. The presence of low molecular weight impurities would not have been detected with "normal " loading of the gel (e.g. 10μg protein), and it is possible that such small amounts of impurities do not interfere with crystallisation. However the fact that even a crystalline preparation of actinidin contained impurities could indicate that these are products of auto-digestion.

In conclusion, it is unlikely that actinidin will be unsuitable for raising antibodies due to auto-digestion before recognition as an immunogen, particularly if the low molecular weight contaminants are indeed breakdown products of actinidin.

3.2.4. Raising and testing antibodies to actinidin

Antiserum was raised in rabbits to native and SDS-denatured actinidin which had been obtained from C.H.Moore in New Zealand. Antiserum raised to denatured actinidin did not precipitate native or denatured protein. It is unlikely that denaturation of actinidin had destroyed its antigenicity since SDS-treated proteins can elicit strong antibody responses (Goding, 1983). Since rabbits are highly variable in their immune responsiveness (Goding, 1983) it seems more likely that this was

Figure 3.6. Ouchterlony immunodiffusion of partially purified antibodies to native actinidin

All serial dilutions were two-fold.

1. centre: antibody, 10mg/ml
   peripheral: serial dilution of apex extract

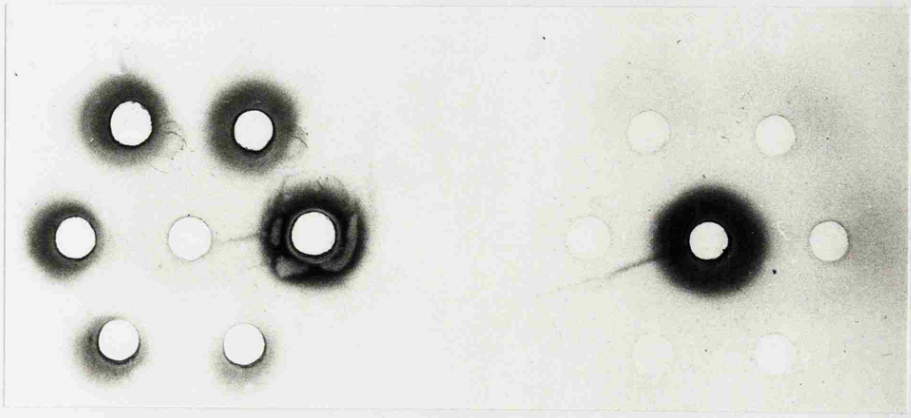
2. centre: apex extract, undiluted
   peripheral: serial dilution of antibody, from 10mg/ml

3. centre: antibody, 10mg/ml
   peripheral: serial dilution of actinidin, from 10mg/ml

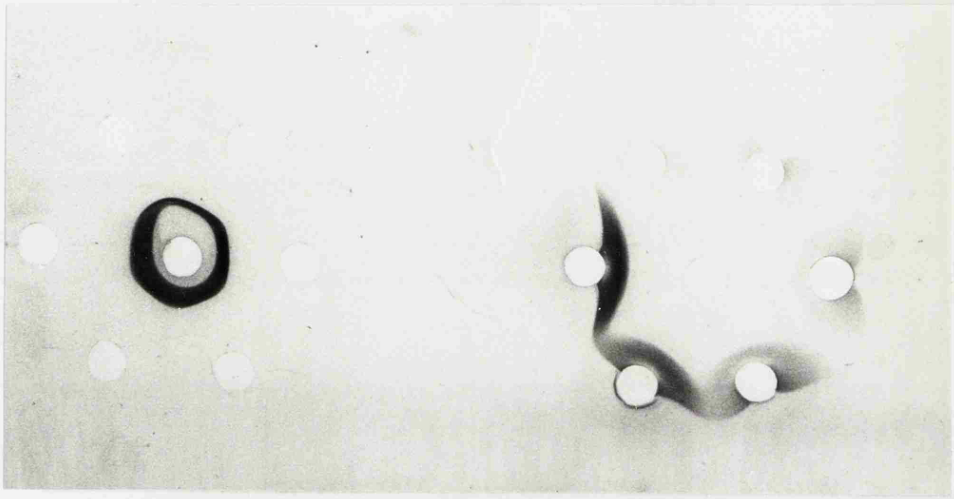
4. centre: actinidin, 10mg/ml
   peripheral: serial dilution of antibody, from 10mg/ml

5. centre: antibody, 20mg/ml
   peripheral: serial dilution of actinidin, from 10mg/ml

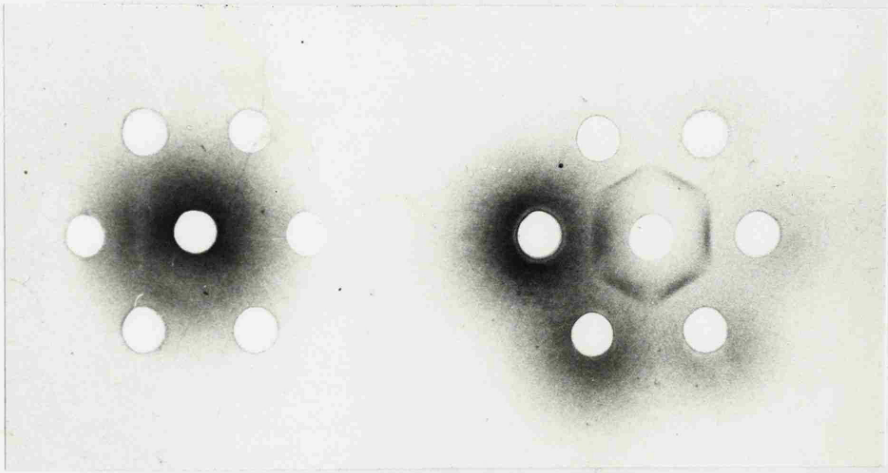
6. centre: actinidin, 10mg/ml
   peripheral: serial dilution of antibody, from 20mg/ml



1 2



3 4



5 6

the cause of failure.

Figure 3.6 shows an immunodiffusion test of the final antiserum preparation to native antigen and illustrates different precipitation patterns obtained with different amounts of antiserum (figures 3.6.1 and 3.6.2). For comparison Figure 3.6.3 shows a negative reaction with plant extract containing protein which could not be removed from the gel by washing.

Since both the actinidin preparation used for raising antibodies, and the actinidin used in testing antiserum, contained unidentified impurities, the possibility had to be considered that the observed reaction was due to antibodies raised to the impurities rather than actinidin.

The size of arcs obtained in rocket immunoelectrophoresis is proportional to the amount of antigen applied, and the results in Figure 3.7.A suggest that the antiserum was most likely to be specific for the major protein in the preparation, as the size of rockets from small amounts of contaminating protein would have been much smaller if detectable at all.

3.2.5. Estimation of actinidin in plant extracts

Young plants were divided into apex (which included very small leaves), fully expanded leaves, and roots, and extracts were prepared as described in Methods. The fruit used in immunoelectrophoresis was ripe fruit from the local grocer, but for ELISA, fully-grown, unripe fruit obtained from Wisley Gardens, was used.

3.2.5.1. Immunoelectrophoresis

The standard curve obtained for rocket immunoelectrophoresis of actinidin (Figure 3.8) is not a straight line as expected (Hames and Rickwood,

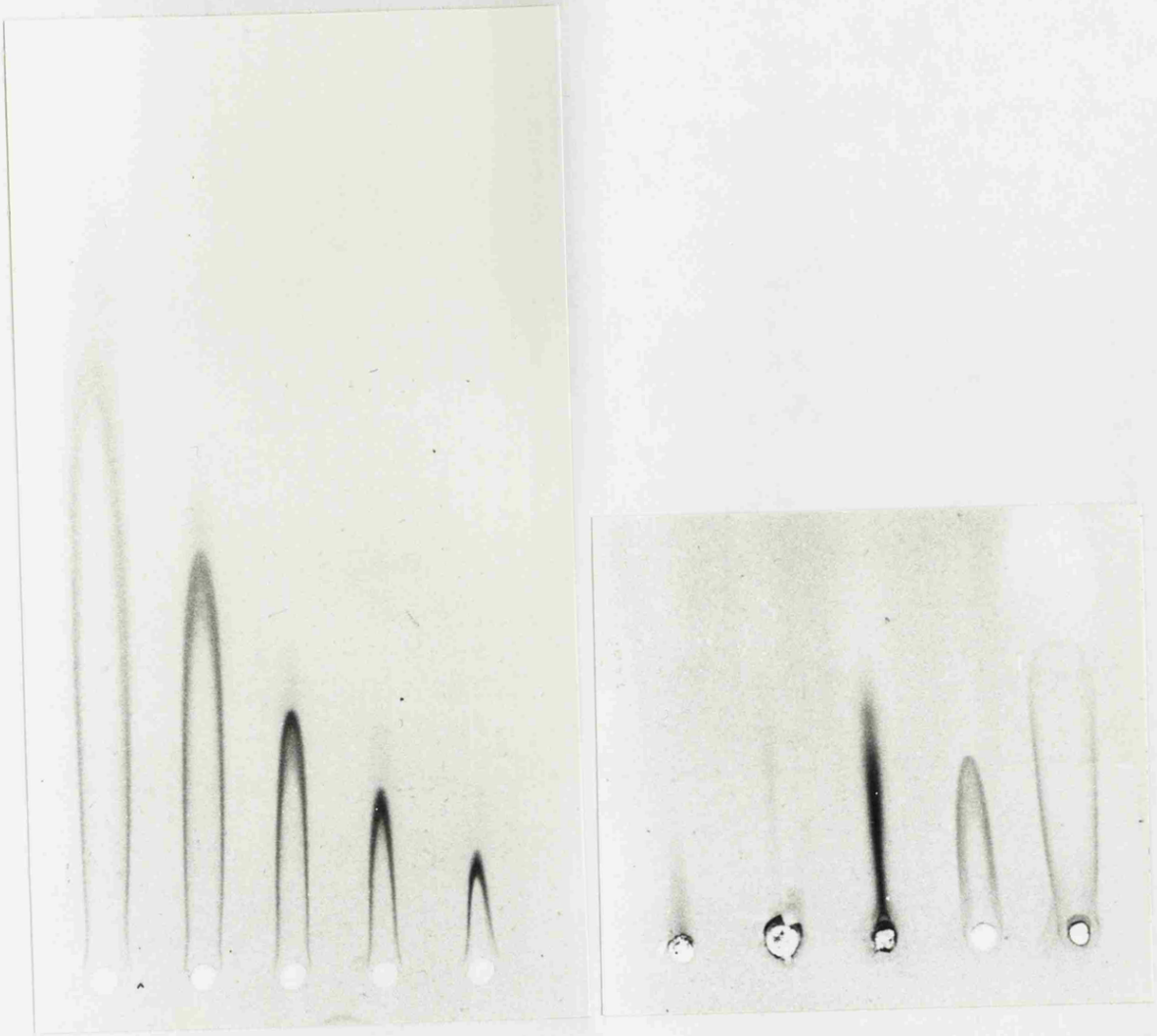


Figure 3.7.A. Rocket immunoelectrophoresis of actinidin

A.1. 0.5mg/ml
A.2. 0.25mg/ml
A.3. 0.10mg/ml
A.4. 0.05mg/ml
A.5. 0.025mg/ml

Figure 3.7.B. Rocket immunoelectrophoresis of plant extracts

B. 6. root extract
B. 7. leaf extract
B. 8. apex extract
B. 9. fruit extract, ten-fold diluted
B.10. fruit extract, undiluted

Root and leaf extracts were prepared by method B) as described in "methods". Apex and fruit were extracted by grinding in TRIS pH 6.8, at 1ml/g of tissue, centrifuged and supernatant used directly. Both gels contained 0.1mg/ml antibody.

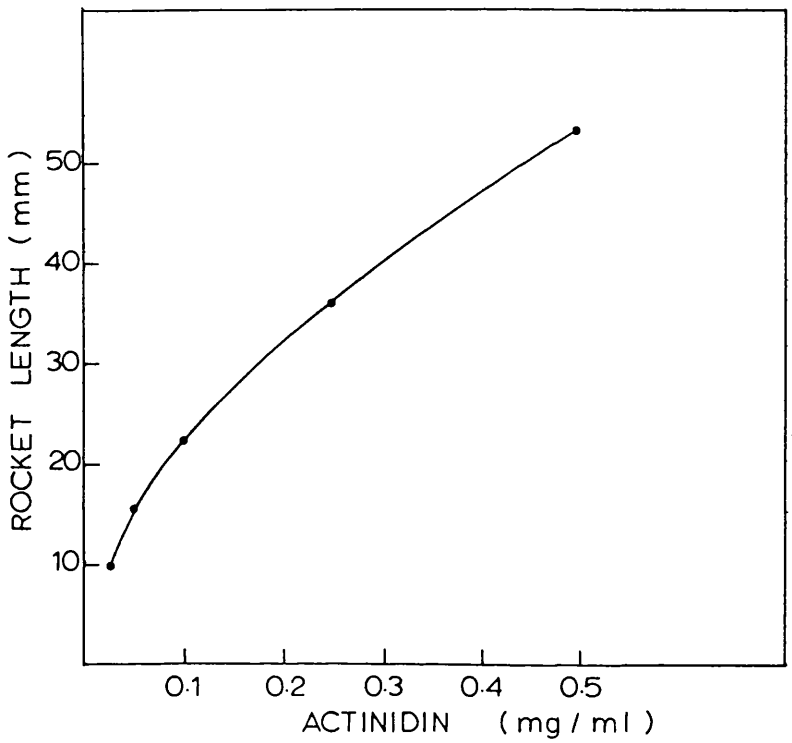


Figure 3.8. Rocket immunoelectrophoresis of actinidin

Standard curve of actinidin.
Plot of rocket length versus concentration of actinidin in the sample.

1981), although it approaches linearity for the higher actinidin concentrations. According to Hames and Rickwood the standard curve does not pass through the origin, allowing accurate quantitation only for rockets of at least 10mm length. The curve obtained in Figure 3.8 suggests that it may in fact pass through the origin. Although reasons for this discrepancy are not known, it would seem reasonable to reject rocket arcs less than 10mm long.

Figure 3.7.B shows results of immunoelectrophoresis of plant extracts under conditions comparable to those for the actinidin standards (lanes 1 to 5). Root and leaf extracts show only faint trails of protein. Apex extract gave rise to a line resembling a small precipitin arc under a thick trail of protein, but this arc was only 1mm in length, and it is possible that it was an artefact of protein migration rather than a precipitin arc. Undiluted extract from ripe fruit resulted in an unresolved rocket, probably due to insufficient electrophoresis time, but a ten-fold dilution of fruit extract produced a precipitin arc of 15mm length. Considering the overall dilution factor of 20, the content of actinidin can be calculated at 0.9mg per g fruit (fresh weight).

Rocket immunoelectrophoresis was repeated with extracts prepared by a different method (see legend) and using a lower concentration of antibody in the agar to increase rocket size (Figure 3.9.A). Again apex extract may have given rise to a very faint pricipitin line on the outside of the central protein trail but this cannot be identified unambiguously as an arc. Leaf and root extract showed negative results.

In order to investigate whether other proteins in apex extract interfered with detection of actinidin by this method, the extract was subjected to two-dimensional (2-D) immunoelectrophoresis: Proteins were

Figure 3.9.A. Rocket immunoelectrophoresis of plant extracts

Extracts "A" were centrifuged homogenates in protein extraction buffer, extracts "B" were acetone precipitated proteins resuspended in phosphate buffer (see Methods).
The gel contained 50μg/ml antibody.

A.1 apex extract A
A.2 " " B
A.3 actinidin, 10μg/nml

A.4 mature leaf extract A
A.5 " " " B
A.6 actinidin, 10μg/ml

A.7 root extract A
A.8 " " B
A.9 actinidin, 10μg/ml

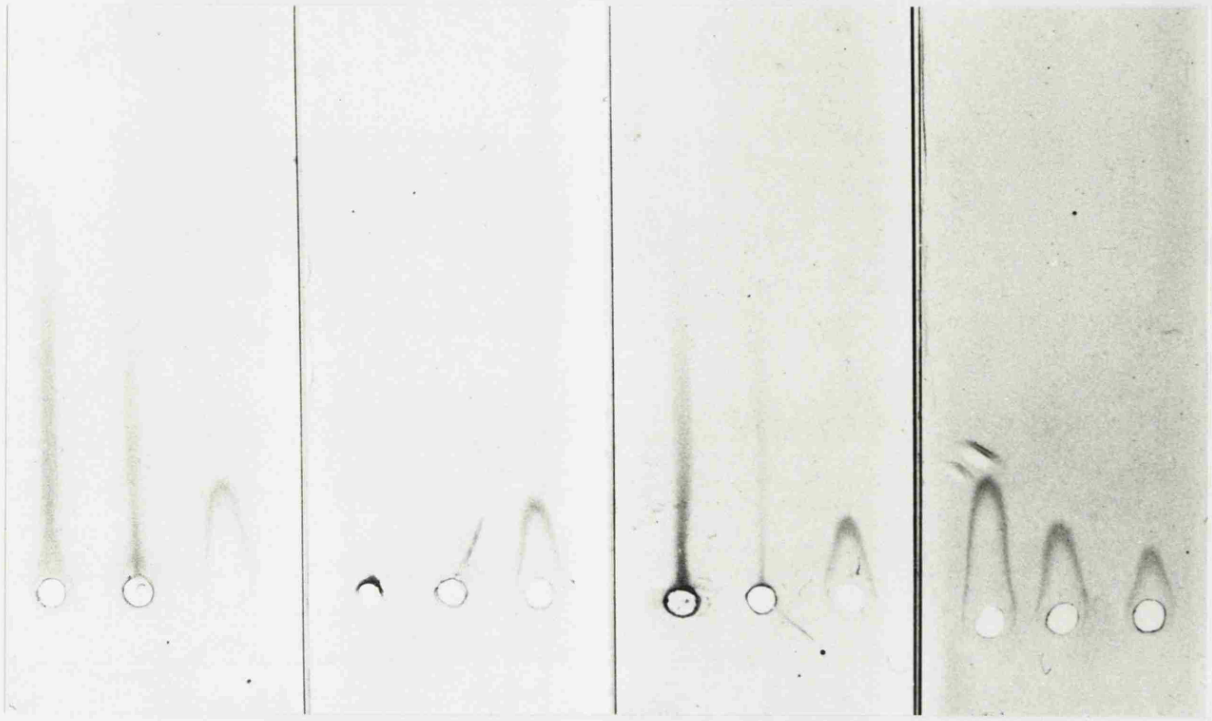
10 actinidin, 20μg/ml
11 " 10μg/ml
12 " 5μg/ml

Figure 3.9.B. Two-dimensional immunoelectrophoresis)

B.1. apex extract equivalent to 1.24mg dry weight of apex with 50μg/ml actinidin

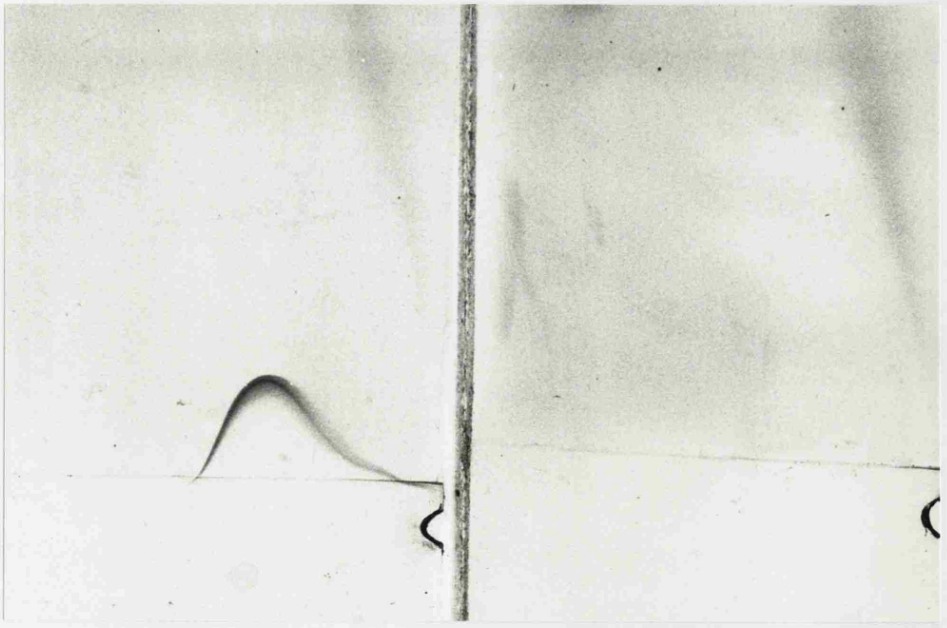
B.2. apex extract equivalent to 2.24mg dry weight of apex without added actinidin

A



1 2 3 4 5 6 7 8 9 10 11 12

B



1 2

electrophoresed in the first dimension according to charge, and in the second dimension into agar containing antibody. Figure 3.9.B shows the result of 2-D immunoelectrophoresis of a concentrated apex extract (3.9.B.2) and the same extract mixed with an equal volume of 100μg/ml actinidin (3.9.B.1). The final concentration of actinidin was 50μg/ml and a total of 1μg was applied to the well. The precipitin arc was slightly wider than expected from published examples (Hames and Rickwood, 1981) and it is possible that the viscosity of the extract interfered with migration of proteins as narrow bands in the first dimension. However, it can be concluded that the method would be suitable for the detection if not accurate measurement of actinidin in plant extracts. No precipitin arc was found when apex extract was electrophoresed alone, suggesting that actinidin, if present in the apex, is at a concentration not detectable by this method.

In conclusion, the methods involving immunoelectrophoresis were suitable for the identification of actinidin in fruit, but not sufficiently sensitive to detect the protein in other parts of the plant. The range of sensitivity of rocket immunoelectrophoresis was from 25 to 250 μg/ml of actinidin, a level not found in the most concentrated plant extracts other than fruit. Therefore a more sensitive test, also using antibodies, the enzyme-linked immunosorbent assay (ELISA) was employed to determine whether actinidin was present at all in non-fruit tissue.

3.2.5.2. ELISA

Initially this test was carried out using the direct method in which microtiter wells were coated with antigen and then reacted with antibody. However, with this method root extracts gave consistently high values

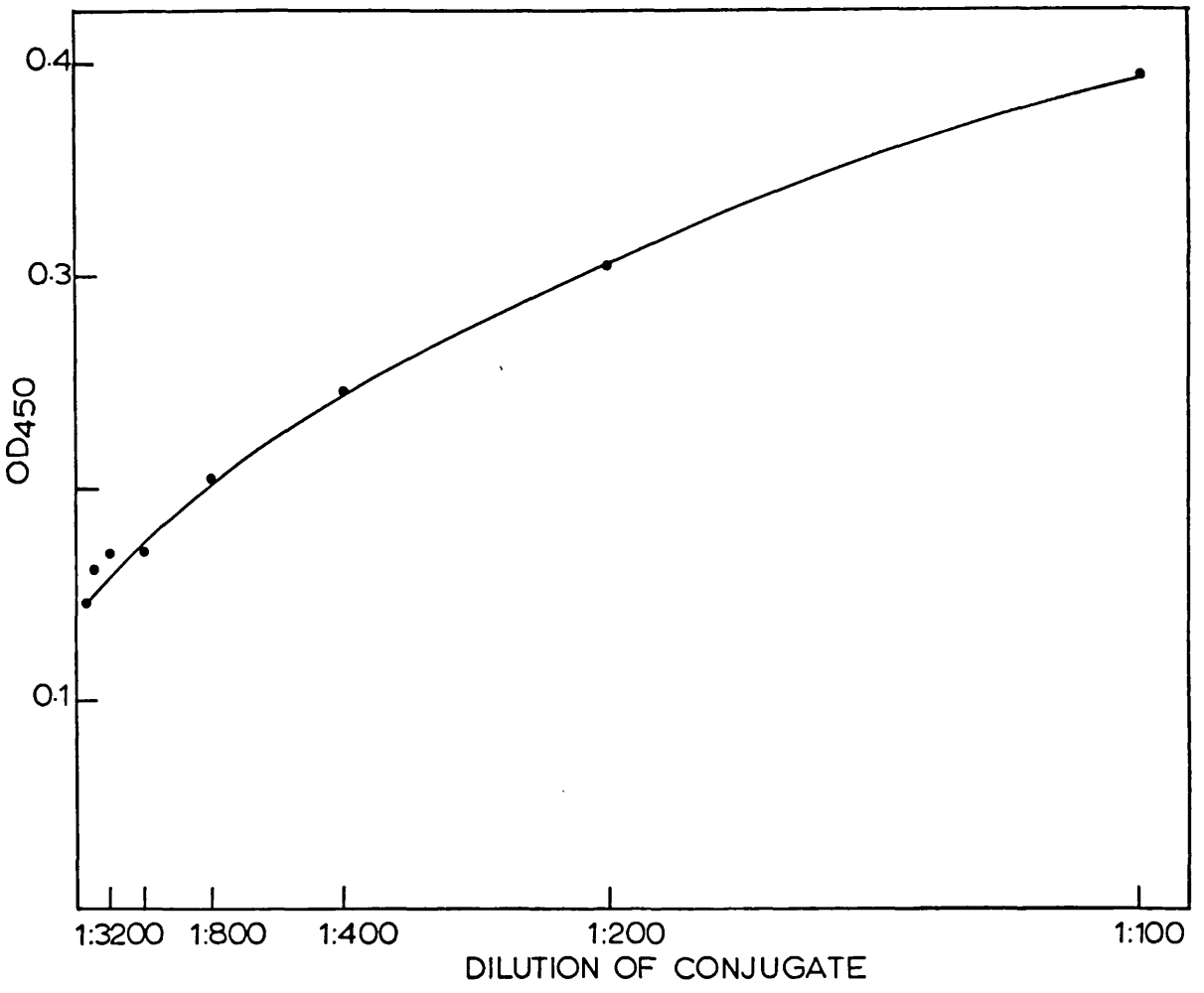


Figure 3.10. ELISA of actinidin

Testing of antibody-HRP conjugate against a constant concentration of actinidin.
The background level of OD_{450}, in the absence of actinidin, was identical to the OD_{450} of the lowest actinidin concentration tested.

which were subsequently found to be due to interference with the peroxidase assay by peroxidase present in the root tissue: A small amount of root extract added to substrate solution resulted in rapid colour change. This problem of endogenous peroxidase enzymes should be overcome by using the indirect double antibody sandwich method in which only antigen recognised by the specific antibody immobilised on the well surface is retained. For this purpose actinidin antibody was conjugated to horseradish peroxidase (HRP).

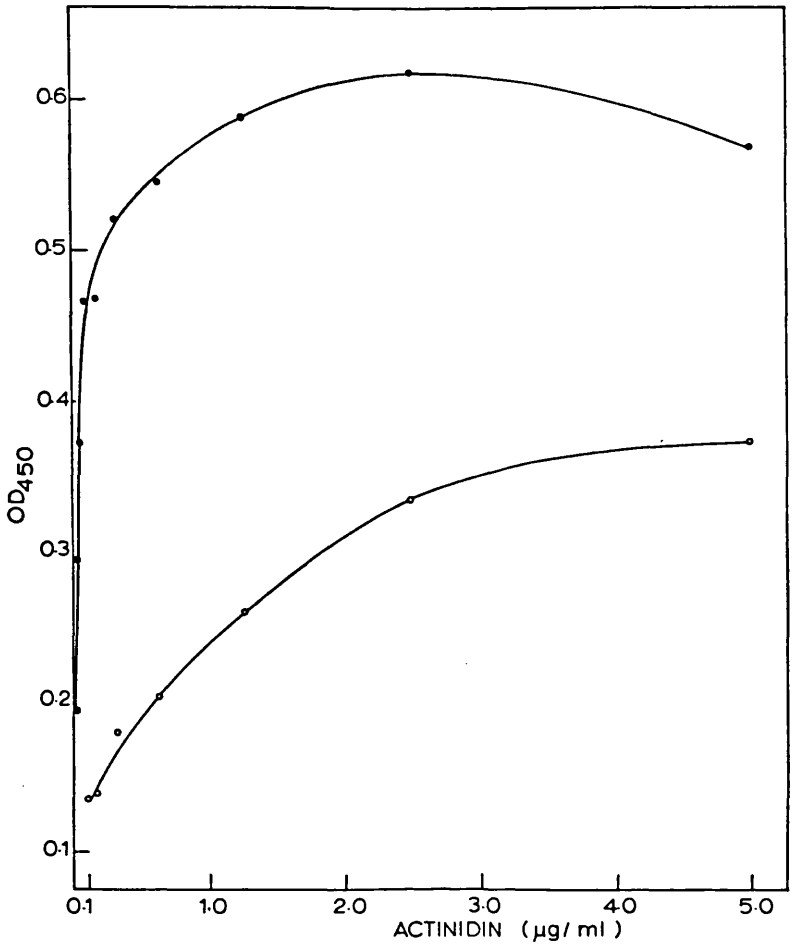
Although generally a 500-fold dilution of antibody-HRP conjugate was used, the dependence of the assay on conjugate concentration was first tested against a constant concentration of actinidin. The microtiter wells were coated with 1μg/ml actinidin and reacted with a two-fold serial dilution of conjugate from 1:100. The plot of optical density obtained with the peroxidase assay versus the dilution of the conjugate (Figure 3.10) suggests that the sensitivity of the assay may be limited by conjugate at high dilutions. However, subsequent tests using 500-fold dilution gave reproducible results with actinidin standards and were adequate in detecting low levels of actinidin in plant extracts.

Using the same serial dilution of actinidin in a direct and double antibody sandwich test, and both using the antibody conjugate prepared as described, the two methods were compared. The results in Figure 3.11 were plotted on two different scales to highlight differences for higher and lower concentrations of actinidin. Figure 3.11.A illustrates clearly that the sandwich method was much more sensitive, as actinidin was detected at a level of 0.005μg/ml. Although the background level using pre-immune serum was not determined in this test, this was likely to be close to the values obtained for the lowest actinidin concentrations in the direct

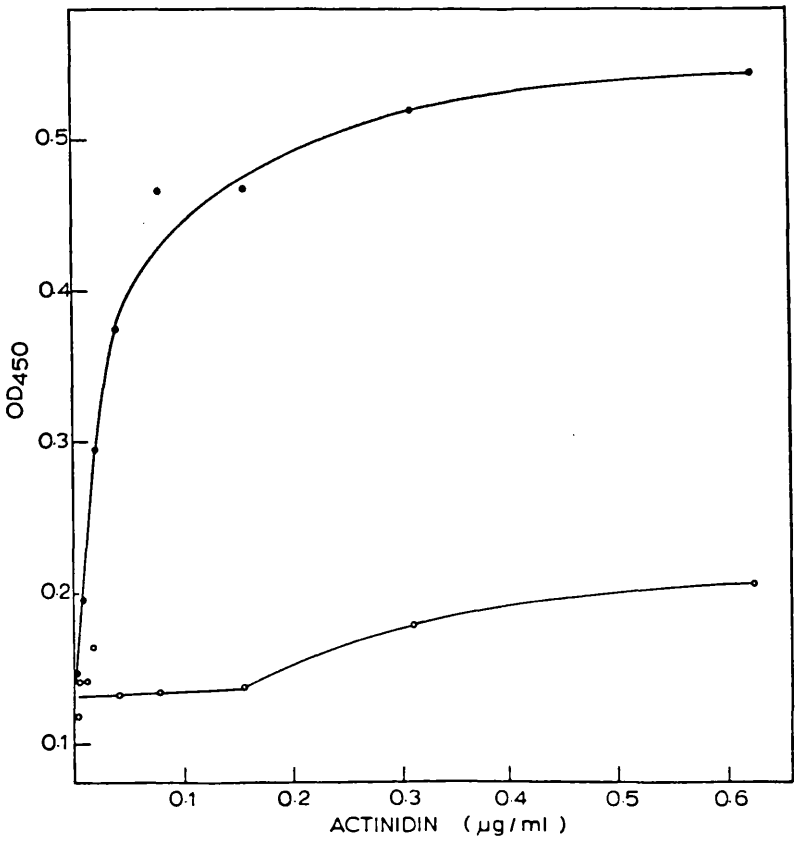
Figure 3.11. Comparison of direct and double antibody sandwich methods of ELISA using actinidin antibody-HRP conjugate

open circles: direct method
filled circles: double antibody sandwich method

A) results plotted to highlight low actinidin concentrations
B) same results to include high actinidin concentration


0·6
0·5
0·4
0·3
0·2
0·1
OD450
0·1
1·0
2·0
3·0
4·0
5·0
ACTINIDIN (µg/ml)

A


0·5
0·4
0·3
0·2
0·1
OD450
0·1
0·2
0·3
0·4
0·5
0·6
ACTINIDIN (µg/ml)

B

method. The lowest level of actinidin detectable with the direct method was 0.3μg/ml and this concentration was near the upper limit of sensitivity of the sandwich method. The double antibody sandwich ELISA was thus found to be approximately a thousand times more sensitive than rocket immunoelectrophoresis.

For the measurement of actinidin in different parts of the plant, extracts were prepared as described in Methods and were at a concentration of 0.017g dry weight per ml, from which two-fold serial dilutions were made. Fruit extract was diluted one thousand-fold and then diluted in two-fold steps. For each serial dilution a control assay with pre-immune antibody was included and values were subtracted from the test values. The original optical density (OD) readings (Table A.2), the actinidin standard curve (Figure A.2.), and the results from two independent assays (Table A.3), are presented in the Appendix.

The optical densities obtained for the actinidin standard background with pre-immune antibody varied between 0.26 and 0.32 and indicate a certain degree of variability of the assay. This was not unexpected as the test involves many steps, including washing, from which inaccuracies may result. In addition it is likely that the microtiter wells are not uniform in their capacity to adsorb proteins. As the serial dilution of actinidin was of uniform consistency (not viscous) throughout the dilutions, the test readings were corrected using the average from all backgrounds. However, plant extracts other than fruit occasionally showed high backgrounds with pre-immune serum in concentrated dilutions, and these were therefore treated separately. The readings obtained for root extracts were so similar for test and control assays that root was concluded to contain levels of actinidin that are too low to be detected

by this method, or less than 0.3μg/g dry weight of tissue. The high values found in the concentrated controls suggest that the interference by non-actinidin proteins, possibly peroxidase, could not be eliminated with the double antibody sandwich method. Occasionally plant extracts other than fruit appeared to contain higher concentrations of actinidin in the more dilute samples. This was likely to be caused by the viscosity of the concentrated extracts, particularly of leaf and apex, which may have interfered with the free interaction of antigen and antibody by reducing fluidity. Therefore the less viscous dilute samples are likely to reflect more accurately levels of actinidin measured by this assay. Fruit extract was tested in two separate serial dilutions. Only greater than 8000-fold dilutions were within the sensitivity range of the assay. Table 3.1 summarises the results.

Table 3.1. Average levels of actinidin

	dry weight	fresh weight
leaf	6.4 μg/g	0.65 μg/g
apex	2.05 μg/g	0.23 μg/g
fruit	24.1 mg/g	4.1 mg/g
root	less than 0.3 μg/g	

Leaf tissue was found to contain three times as much actinidin as apex tissue. The level of actinidin in fruit was 3,800 times that in leaf and 11,800 times greater than in apex tissue, on a dry weight basis. This difference was even greater when calculated on the basis of fresh weight. Comparing the results for fruit with those from rocket immunoelectrophoresis, it appears that ripe fruit obtained locally had only one fifth of the amount of actinidin present in unripe mature fruit

from Wisley Gardens.

Although the two separate ELISA's using different plant extracts gave very similar results, the accuracy of actinidin measurement in tissues is likely to be limited by the efficiency of actinidin extraction, which will vary between methods of homogenisation, and probably between tissues as these have widely differing consistencies. However, the purpose for which actinidin levels in different parts of the plant were determined does not require greater accuracy of the results, as these showed differences of several orders of magnitude between fruit and other parts of the plant.

3.3. Conclusions

Summarising the results of this chapter, actinidin purified from locally purchased Chinese gooseberry fruit reacted with antibodies raised to native actinidin from New Zealand. Thus, although it is possible that the two preparations of actinidin differ since they were obtained from different sources, their antigenic properties were sufficiently similar to allow precipitation by the same antibodies.

Using these antibodies in immunodetection studies, levels of actinidin in fruit were found to be approximately four orders of magnitude greater than those in apex or leaf tissue. This difference in actinidin content between fruit and other parts of the plant is clearly striking, and suggests that actinidin gene expression, at least at the level of the protein, must be regulated in a highly tissue-specific manner.

In the majority of cases studied, gene expression is regulated at the level of transcription (Darnell Jr., 1978). If this applies to actinidin, it may be expected that fruit tissue contains a large amount of actinidin-specific mRNA, and thus would seem most suitable as a source of RNA for use in the construction of a cDNA library.

4. RNA isolation and *in vitro* translation

4.1. Introduction

In the preceding chapter actinidin was demonstrated to be predominantly synthesised in the fruit of the Chinese gooseberry, where it represents the major soluble protein as judged from SDS-PAGE of fruit extracts. It was necessary to examine whether this difference in the level of actinidin between fruit and other tissues is reflected by differences in the levels of actinidin-specific RNA. If actinidin-encoding RNA is present in significant amounts in fruit tissue, this would clearly be the most suitable source of RNA for cDNA cloning.

4.1.1. Purification of RNA

The first part of this chapter presents results obtained from RNA preparations derived from fruit and apex or young leaf tissue.
A number of methods for the isolation of intact mRNA have been developed to suit different systems (Taylor, 1979). It was expected that established methods would have to be modified in order to be applicable to Chinese gooseberry tissue, a system which had not been previously studied.
The most important objective in the initial stages of tissue homogenisation is the inactivation of ribonuclease, a class of RNA-hydrolysing enzymes present in subcellular components of all tissues, which often occur in chemical reagents derived from biological sources and on human hands (Taylor, 1979). Therefore most methods are designed to denature proteins and inactivate enzymes as rapidly as possible.
Two of the methods described in this section (both modified from Kirby, 1964) employ phenol as a denaturant, and involve the partitioning of nucleic acids into the aqueous phase from an emulsion containing liquefied phenol and chloroform, with denatured protein compacting at the

interface. Extraction buffers are optimised to facilitate protein denaturation and dissolution of nucleoprotein complexes and to inhibit ribonuclease by addition of SDS, triisopropylnaphthalene sulphonate (TNS), EDTA, and β-mercaptoethanol. An alkaline pH , the use of temperatures around 20°C, and the addition of chloroform, are designed to prevent poly-adenylated RNA from binding to denatured proteins at the interface. Isoamyl alcohol and 8-hydroxy quinoline are used as antifoaming agents, and m-cresol as antifreeze and protein denaturant. Partially purified nucleic acid is precipitated by addition of ethanol to the acidified aqueous phase, and RNA is separated from DNA either by centrifugation through a caesium chloride cushion, or by selective precipitation in 3M sodium acetate (or 3M lithium acetate) or 2M lithium chloride. Both DNA and most carbohydrates are soluble in these salt solutions (Palmiter, 1974).

The third method described involves the rapid dissociation of ribonucleoprotein complexes and denaturation of proteins by a concentrated solution of the salt guanidinium thiocyanate, in the presence of the reducing agent β-mercaptoethanol and detergent sodium N-lauroyl sarcosine (Chirgwin *et al.*, 1979). RNA is selectively precipitated from a clear aqueous phase by the addition of ethanol and re-precipitated from guanidine-hydrochloride solution several times.

The fourth method was initially developed for the extraction of plant DNA (Murray and Thompson, 1980) and claimed to be applicable also to the isolation of intact RNA free from carbohydrates. Lyophilised ground tissue is homogenised in buffer containing cetyltrimethylammonium bromide (CTAB), a high salt concentration, and the detergent β-mercaptoethanol. Nucleic acid is soluble in CTAB at high salt concentrations, whereas proteins and most carbohydrates are insoluble and these are removed during chloroform extraction of the suspension. Nucleic acid is precipitated in

the form of a CTAB-nucleic acid complex by lowering of the salt concentration, and RNA is separated from DNA either by caesium chloride centrifugation or salt precipitation.

The purification of mRNA by affinity chromatography on oligo(dT) cellulose columns capitalises on the presence of a poly(A) tail on all mRNA's (with the exception of histone mRNA's), which can hybridise with a synthetic oligo(dT) molecule (12 -18 nucleotides long) under high salt conditions. The poly(A+) fraction can be eluted from the column with a buffer of low salt concentration. The considerable amount of non-poly(A+) RNA present in this fraction after one fractionation cycle can be virtually eliminated by a second cycle.

The main problem encountered during the isolation of RNA from fruit was the co-purification of contaminating substances, probably carbohydrate, with poly(A+) RNA during fractionation on oligo(dT) cellulose, also reported by Mozer (1980). Since carbohydrate is known to interfere with template activity of mRNA (Gordon and Payne,1976) considerable effort was directed towards obtaining highly purified total RNA.

4.1.2. Assessment of purity of RNA

The quality of RNA isolated from Chinese gooseberry plant material was assessed on the basis of one or more of the criteria listed below, and compared to that of a control RNA isolated from pea , a tissue which was amenable to a variety of purification methods.

a) Absorption profile between 230 and 280 nm.

Nucleic acids absorb UV light with a maximum at 260 nm, and in the absence of contaminants the ratios of OD_{260}/OD_{280} and OD_{260}/OD_{230} are equal to or better than 2, indicating minimal contamination by protein and carbohydrate, respectively (Taylor and Powell, 1982).

b) Agarose gel electrophoresis.

The intactness of ribosomal RNA, migrating as discrete bands in agarose gels, can be regarded as an indication that digestion of RNA by contaminating ribonuclease was minimal.

c) *In vitro* translation.

The quality of mRNA was assessed either before or after purification on oligo(dT) cellulose, by its translatability in cell-free protein synthesising systems. Both the wheatgerm system (Roberts and Paterson, 1973) and the rabbit reticulocyte lysate system (Schimke *et al.*, 1974) were used. The incorporation of radio-labelled amino acids into TCA-precipitable label was compared with the level of a control containing no added RNA. Products of *in vitro* protein synthesis were also separated on acrylamide gels and subjected to autoradiography. The presence of discrete high molecular weight bands was taken as an indication of the intactness of mRNA although it also reflects the efficiency of the *in vitro* translation system itself.

4.1.3. Immunoprecipitation of *in vitro* translation products

This chapter also describes results obtained from immunoprecipitation of *in vitro* translation products with anti-actinidin antibodies, using apex and fruit RNA to direct protein synthesis.

The aim of the immunoprecipitation experiments was to demonstrate the presence of actinidin-specific message in mRNA from fruit before this was used as a template in cDNA synthesis.

The ability of antibodies, raised to mature proteins, to immunoprecipitate products of *in vitro* translation depends on the presence of antigenic sites on the *in vitro* product similar to those of the mature protein. Since many proteins are synthesised as larger precursors which are not processed to the mature forms in the wheatgerm and rabbit reticulocyte systems, actinidin antibodies may not react with the *in vitro* translation

product.

Antibodies which were raised to mature proteins have been used successfully to immunoprecipitate precursor forms of the protein, for example the aspartate proteinase chymosin (from calf; Harris *et al.*, 1982), but in the cases of gastrin (Noyes, 1979) and papain (McKee *et al.*, 1986) specific antibodies did not immunoprecipitate translation products. It is not clear whether the absence of immunoprecipitates was due to non-recognition by the antibodies, to the low levels of *in vitro* synthesised proteins, or even to the unavailability for translation in the *in vitro* systems of the mRNAs in question. Therefore, although a positive result would strongly suggest the presence of actinidin-encoding mRNA, a negative result would not be proof of its absence.

The immunoprecipitation of an *in vitro* translation product indicates that the protein to which the antibodies were raised and the *in vitro* product are antigenically related, but, although it strongly suggests identity of the two, it does not prove it, since otherwise unrelated proteins may have similar antigenic properties. Peptide mapping of the purified protein and the immunoprecipitated *in vitro* product has been successfully employed in the identification of *in vitro* synthesised storage protein Gl from French bean (Hall, 1978) and rat malic enzyme (Magnuson, 1983). The proteins were digested with several different proteases, for example trypsin and staphylococcal protease V8, and separated on SDS-gels. Identical peptide patterns obtained from digests of the mature and the *in vitro* synthesised proteins were evidence for the identity of the two proteins.

This section includes results from peptide mapping of actinidin with trypsin and an attempt at mapping an immunoprecipitate of *in vitro* translation products from fruit RNA.

4.2. Results

4.2.1. RNA purification

4.2.1.1. "Phenol method" (Kirby, 1964)

The "phenol method" involving the purification of RNA by caesium chloride centrifugation was tested with dark grown pea apices, and a yield of 0.5mg total RNA/g tissue was obtained. Although the absorbance profile of the final preparation was not determined, mRNA was successfully purified by oligo(dT) cellulose column chromatography. After one column cycle, between 1.2% and 1.9% of RNA eluted in the poly(A+) fraction. When this poly(A+) RNA was used to direct protein synthesis in a wheatgerm translation system, it was up to 20 times more active than a control containing no added RNA. The incorporation of TCA-precipitable radioactivity was dependent on added RNA since doubling of the amount of RNA added doubled the amount of label incorporated. The results suggest that this method was suitable for the production of active mRNA from peas. When the same method was applied to ripe Chinese gooseberry fruit, the only fruit material available initially, the ethanol precipitate containing total nucleic acid was jelly-like in consistency. The pellet after caesium chloride centrifugation consisted of a thick gelatinous mass, purple in colour, far exceeding the expected volume of RNA. Attempts at removing the suspected carbohydrates from the pellet by washing with 2M lithium chloride were unsuccessful. The absorbance scan (Figure 4.1.a) of a ten-fold dilution of the pellet suggests that a trace of RNA, represented by the shoulder between 280 and 260nm, was contained in a vast excess of carbohydrate absorbing in the region of 230nm, where the scan was off-scale. In a modification of the method, not including the caesium chloride gradient, the ethanol precipitate was washed with 20ml cold 2M lithium chloride, which should dissolve both DNA and

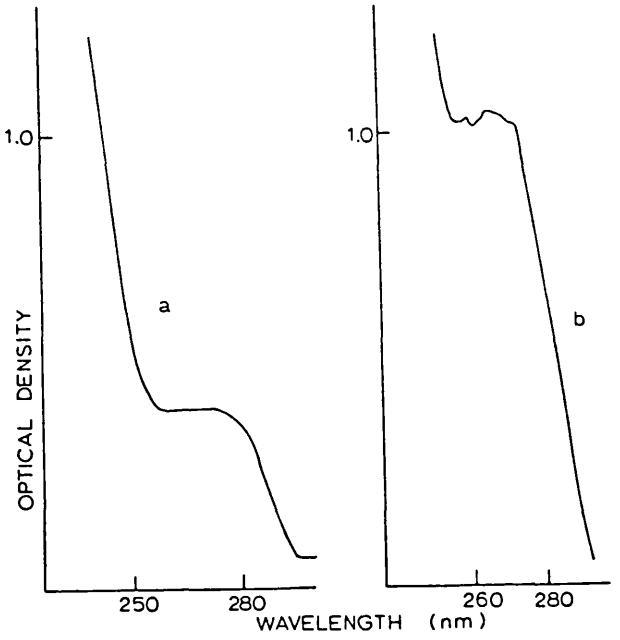


Figure 4.1. Absorbance profile of RNA from ripe Kiwi fruit
RNA was extracted by the phenol method.

a) pellet from caesium chloride gradient washed with 2M lithium chloride
b) ethanol precipitate washed with 3M lithium chloride

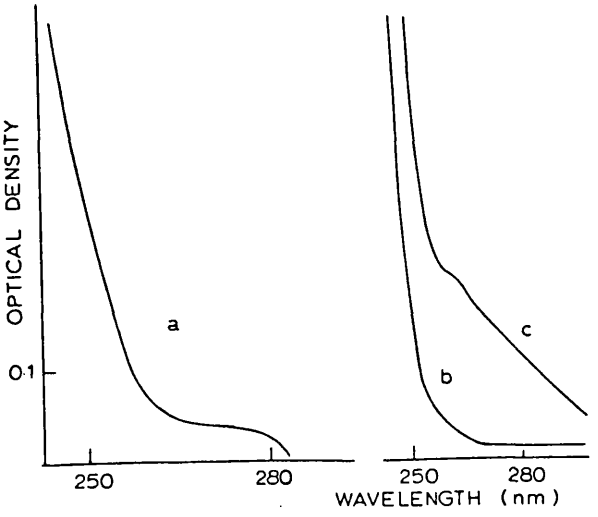


Figure 4.2. Absorbance profiles of RNA
RNA was extracted by the guanidine method.

a) unripe Kiwifruit
b) 100-fold dilution of guanidine hydrochloride solution
c) dark-grown peas

carbohydrates. A small peak is differentiated at 260nm within a wide shoulder presumably representing contaminating protein and some nucleic acid, but the rapid increase in absorbance at wavelengths below 260nm again suggests that the pellet consisted mainly of carbohydrate (Figure 4.1.b). This method of RNA purification was therefore rejected as unsuitable for Kiwi fruit.

4.2.2.2. Guanidine method

Initial extractions with this method using Chinese gooseberry fruit and apex material yielded preparations which did not resemble RNA in their absorbance profiles <Figure 4.2.a) When tested with the "control" tissue pea, the same results were obtained (Figure 4.2.c). Since pea tissue gave excellent results with this method for other workers, consultation with A.Greenland (Biocentre, Leicester University) showed that the quality of guanidine hydrochloride, recommended by the authors of the method (Sigma, practical grade), was inadequate. A scan of dilute guanidine hydrochloride solution is shown in Figure 4.2.b and closely resembles that of the final preparations obtained, suggesting that impurities in the chemical co-purified with RNA. Subsequently guanidine hydrochloride "puriss grade" from Fluka was used and gave good yields of apparently clean RNA from light grown pea tissue (Figure 4.3.a), which was effective in *in vitro* translation (see Table 4.1).

Kiwi apex RNA preparations were found to be more variable with regard to their absorbance profiles (Figure 4.3.b and c). Table 4.1 summarises the results from three independent extractions (E7, E8 and E10). The ratio of OD_{260}/OD_{230} was found to be particularly variable. A low ratio correlated with a lower efficiency in *in vitro* translation, suggesting the inhibition of translation by contaminants absorbing near 230nm. Extract E7 was fractionated into poly(A+) and poly(A-) fractions, and although the

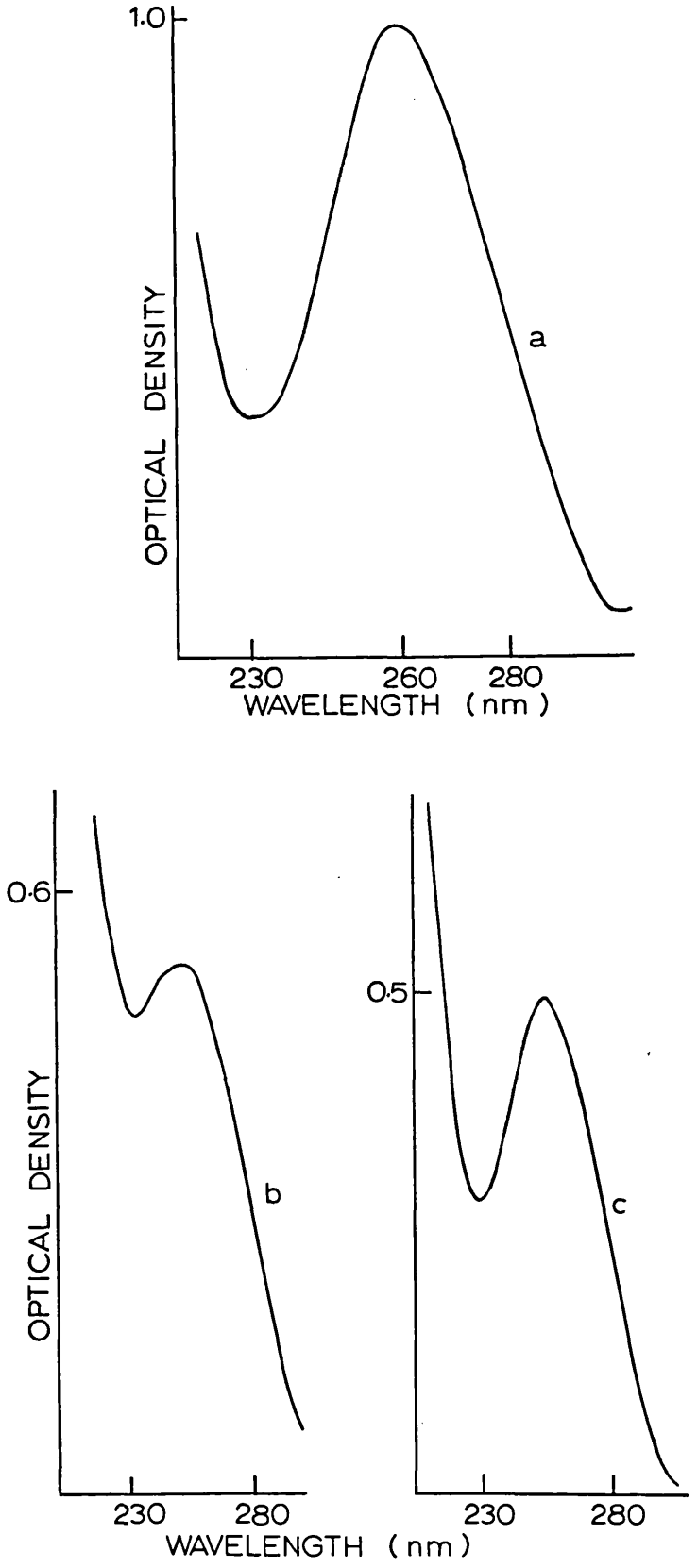


Figure 4.3. Absorbance profiles of RNA
RNA was extracted by the guanidine method.

a) light-grown peas (E6)
b) Kiwi apex (E7)
c) Kiwi apex (E8)

Table 4.1. Assessment of RNA quality

Tissue	Yield (mg/ g tissue)	OD_{260} / OD_{280}	OD_{260} / OD_{230}	Translatability[1] a	b	c	Intactness[2] of rRNA
1) Phenol method							
pea	o.5	n.d.	n.d.	A+	WG -	20x	n.d.
Kiwi fruit	n.d.	1	1	n.d.			n.d.
2) Guanidine method							
E6 pea	1.4	1.95	2.67	total	WG +	40x	intact
				A+	WG +	20x	
E7 Kiwi apex	0.76	1.78	1.1	A+	WG +	4x	n.d.
				A-	WG +	9x	intact
E8 Kiwi apex	0.46	1.68	2.06	total	WG -	10x	intact
					WG +	17x	
E10 Kiwi apex	0.46	1.67	1.37	total	WG +	14x	intact
3) CTAB method							
E30 Kiwi leaf	o.08	1.94	2.8	n.d.			degraded
E33 Kiwi apex	0.18	2.04	2.39	n.d.			degraded
E34 Kiwi fruit	0.009	2.01	1.85	n.d.			degraded

1) Translatability: a) template RNA
b) WG+ = high efficiency wheatgerm
WG- = low " "
c) incorporation of radioactivity compared with a "no RNA" control

2) Intactness judged by agarose gel electrophoresis.

individual fractions were not scanned it appeared that the inhibitory substances present in total RNA co-purified with poly(A+) RNA, resulting in a lower efficiency of translation of this fraction compared to that of the poly(A-) fraction, as shown in Table 4.1.

Translation products from total apex RNA in a wheatgerm system were separated on SDS-PAGE and compared with those from pea RNA. The autoradiograph (Figure 4.4) shows distinct bands, although the majority of proteins synthesised appeared to belong to the low molecular weight class. The majority of the protein bands were common to both plant species. In this gel system, purified actinidin (M_r 23.6) co-migrated with the 30 kDa marker, as indicated in lane 7. This anomaly is discussed below.

The guanidine method was applied to unripe, mature fruit with several modifications, and the problems encountered are sumarised as follows. The initial extract in guanidine thiocyanate retained a slight cloudiness, even after prolonged centrifugation. The ethanol precipitate was difficult to dissolve in guanidinium hydrochloride and gave rise to a very viscous, although clear, solution. Precipitation of RNA with 0.5 volumes of ethanol resulted in the complete absence of pelletable material. When the volume of ethanol was increased to a final of 0.75 volumes (as in the first precipitation step) a thick, viscous precipitate formed again. These observations clearly suggest that a large amount of carbohydrate was present in fruit, which may have interfered with the selective precipitation of RNA with 0.5 volumes ethanol, and was insoluble in higher ethanol concentrations. Since subsequent successful RNA extractions yielded at best 0.05mg total RNA/g of fruit tissue, it is, however, possible that this low level of RNA was the cause of failure rather than the interference of carbohydrate. When optimising this method, Chirgwin *et al.* (1979) found that RNA concentrations of greater than 25μg/ml should

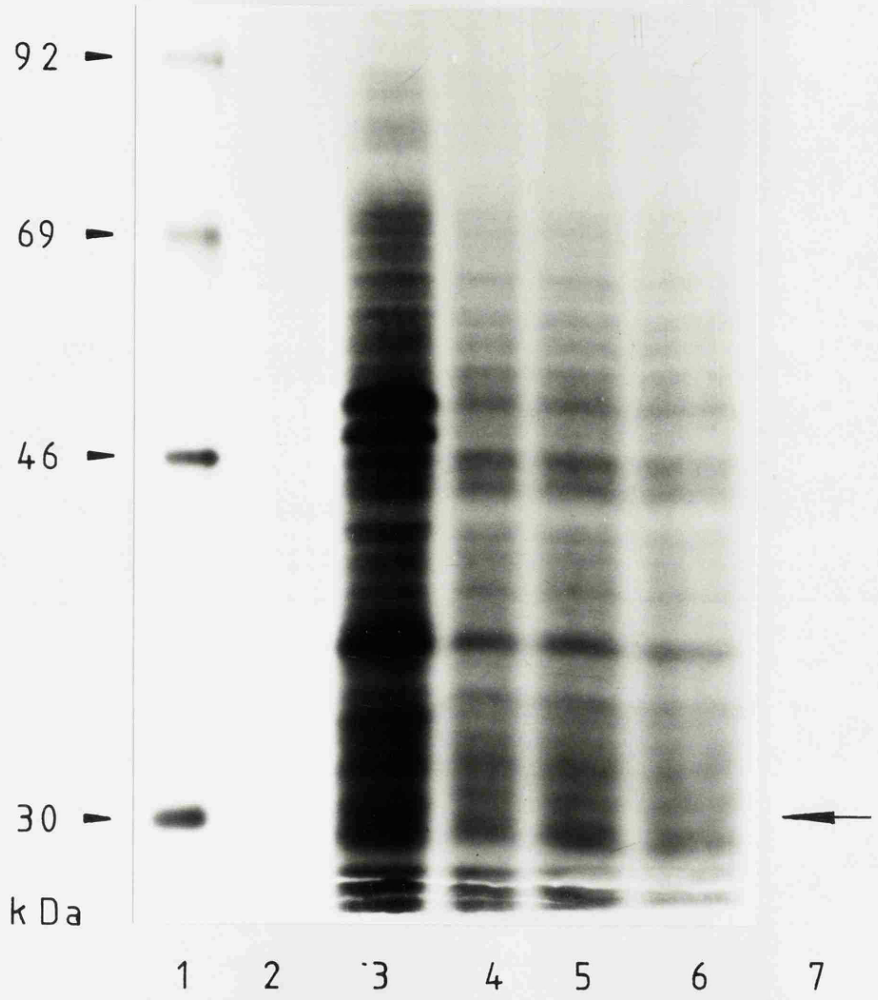


Figure 4.4. Autoradiograph of SDS-PAGE of *in vitro* translation

	c.p.m.
1. radioactive size markers	
2. control without added RNA	11,300
3. pea RNA	282,000
4. Kiwi apex RNA	154,800
5. " " "	172,000
6. " " "	167,600
7. position of actinidin on stained gel	

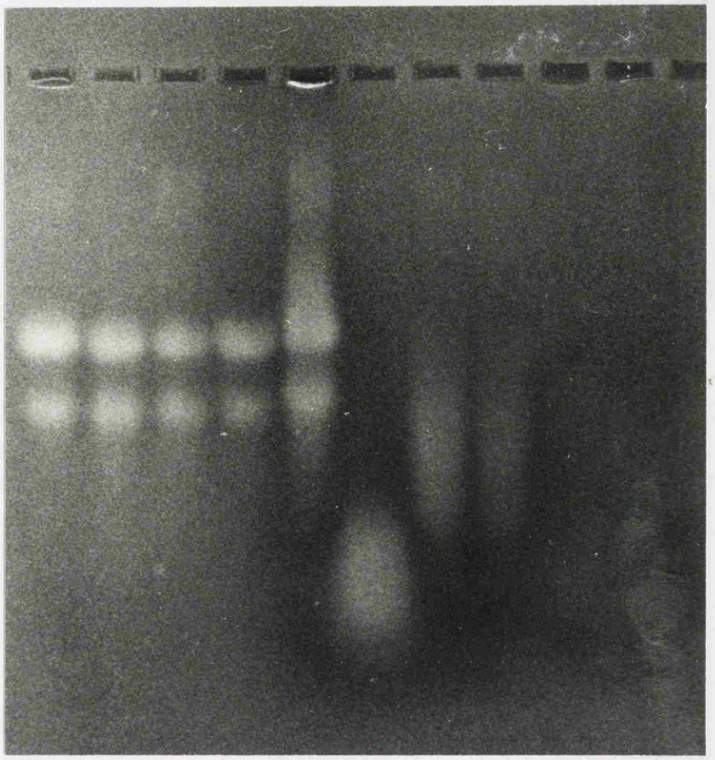
be maintained throughout to ensure precipitation with ethanol. Assuming that RNA was present at approximately 0.05mg/g fruit, the concentration of RNA would have been 20μg/ml in the first guanidinium hydrochloride wash, and this may well have been too low to precipitate.

Again it was concluded that this method was not suitable for the purification of RNA from Kiwi fruit. It was successful, if somewhat variable, with apex tissue, and resulted in the co-purification of contaminants with poly(A+) RNA on oligo(dT) cellulose.

4.2.1.3. CTAB method

Total nucleic acid was extracted from Chinese gooseberry tissue by the CTAB method as described, and RNA separated from DNA by caesium chloride centrifugation. The results are summarised in Table 4.1. Very little RNA was recovered compared with other methods. Although the absorbance ratios indicate high purity of both leaf and apex RNA from Chinese gooseberry, the results from gel electrophoresis (Figure 4.5) show that the ribosomal RNA was degraded, the degree of degradation being greatest for leaf and fruit RNA. It seems that the purity of RNA extracted with this method was achieved at the expense of both yield and intactness.

Although other workers apparently succeeded in preparing intact RNA with this method (Taylor and Powell, 1982), it clearly must involve steps in which RNA is insufficiently protected from degradation. The authors point out that "heating of pellets to facilitate resuspension may result in RNA breakdown", although this is recommended for the preparation of DNA. It is likely that tissues differ in their contents of ribonuclease and that some may be more amenable than others to RNA purification by this method. The CTAB method was rejected without further attempts at improvement.



1 2 3 4 5 6 7 8 9 10

Figure 4.5. Glyoxal gel of RNA

1. 10µg poly (A-) RNA from Kiwi apex (E7)
2. 8µg poly (A-) RNA from peas (E6)
3. 10µg total RNA from Kiwi apex (E9)
4. " " " " (E10)
5. " " " " (E29)
6. 39µg total RNA from young Kiwi leaves (E30)
7. 22µg total RNA from Kiwi apex (E33)
8. 11µg " " " " "
9. 5µg total RNA from Kiwi fruit (E34)
10. 10µg " " " " "

1-5 RNA extracted with the guanidine method
6-10 " " " " CTAB method

4.2.1.4. Modified Kirby method

This method was first tested on pea apices and the results are shown in Table 4.2 (E50). After the initial ethanol precipitation of total nucleic acids, RNA was selectively precipitated with 2M lithium chloride (A) or 3M lithium acetate (B). Since the lithium acetate precipitation resulted in a higher yield and higher OD_{260}/OD_{230} ratio this salt was used in all further experiments.

Table 4.2 summarises some of the results obtained with Chinese gooseberry fruit and apex or young leaves. The yield of RNA from leaf or apex tissue was much lower than the yield of RNA from pea tissue with the same method. It was necessary also to repeat precipitation of RNA from 3M lithium acetate and phenol extractions in order to improve purity (compare E62.A and B, Table 4.2). Generally about 30% of RNA was lost with each additional precipitation and phenol extraction.

Extractions with the modified Kirby method were necessarily subject to variations between experiments. Variables include the amount of starting material, the extent of pulverisation in liquid nitrogen, recovery of aqueous phases after several phenol extractions, and efficiency of salt precipitation. The highest yield/g tissue was obtained when a small amount of leaf tissue (5g) was extracted with a larger than recommended volume of buffer (3.3ml/g), and including the interface in each subsequent phenol extraction. The final aqueous phase was extracted once with chloroform only. After ethanol precipitation the pellet was resuspended in a larger than recommended volume of distilled water again, and precipitated with 3M lithium acetate on ice for two hours. The precipitate was pelleted at 13,000 r.p.m. for 20 minutes (instead of 10,000 r.p.m. for 10 minutes), and the pellet washed twice with 3M lithium acetate and dissolved in distilled water. Further steps were as described

Table 4.2. Assessment of RNA purified by the "Modified Kirby" method

Tissue	yield	OD_{260} / OD_{280}	OD_{260} / OD_{230}	translatability[1]
E50 pea apex A	0.96	1.89	2.08	n.d.
" " " B	1.52	1.85	2.28	15x
E62 Kiwi apex A	0.11	2.10	1.86	n.d.
" " " B	0.07	1.83	2.30	9x
E71 Kiwi leaf	0.19	1.82	2.60	n.d.
E54 Kiwi fruit	0.04	1.75	1.85	n.d.
E56 " "	0.05	1.79	1.51	n.d.
E57 " "	0.04	1.87	1.50	n.d.
E60 " "	0.02	2.04	1.92	5x
E63 " "	0.04	1.84	2.06	2x
E61 " " 10d.[2)]	0.01	1.85	1.29	5x
E59 " " 21d.[3)]	0.023	2.07	1.60	4x
E80 " " S+S[4)]	0.06	1.75	2.43	n.d.
E64 " " ripe[5)]	0.001	1.80	1.62	n.d.

1) Translatability of total RNA in a rabbit reticulocyte lysate compared with a "no RNA" control
2) unripe fruit kept at 10°C for 10 days
3) " " " " " " 21 "
4) unripe fruit with skin and seeds removed prior to RNA extraction
5) ripe soft fruit

in methods. The final RNA preparation had a very high OD_{260}/OD_{230} ratio (see Table 4.2, E71).

Unlike pea tissue, Chinese gooseberry leaf and apex extract was very viscous, probably due to the presence of carbohydrate as reflected in absorbance ratios. This apparently reduced the efficiency of deproteinisation during initial phenol extractions, since an additional phenol step was always required after salt precipitation of RNA, with a considerable amount of precipitate accumulating at the interface. However, with these slight modifications of the original method, good quality RNA at high yields was obtained from Chinese gooseberry leaf and apex tissue.

When applied to fruit tissue, the yield of RNA was very low at approximately 0.04mg/g tissue. The absorbance profiles in Figure 4.6 show how the purity of RNA was improved if an additional salt precipitation, followed by phenol extraction, was carried out beyond the standard protocol. The purity of RNA as reflected in its absorbance in the region of 230nm was generally very variable between experiments, as can be seen from the examples in Table 4.2 (E54, E56, E57, E60, E63). The absorbance ratio of OD_{260}/OD_{230} of this RNA was never better than about 2 even after repeated salt and phenol washes.

The contaminant absorbing at this wavelength was to cause problems during the fractionation of RNA on oligo(dT) cellulose. Figure 4.7.B shows an example of the absorbance profile of the poly(A+) fraction of fruit RNA. Whatever RNA it contained was clearly contaminated by matter absorbing around 230nm. The original RNA applied to the column (Figure 4.7.a) had an OD_{260}/OD_{230} ratio of 1.83, which was improved in the poly(A-) fraction to 2.03 after column chromatography. Clearly the results suggest that the contaminant present in fruit RNA co-purified with poly(A+) RNA by

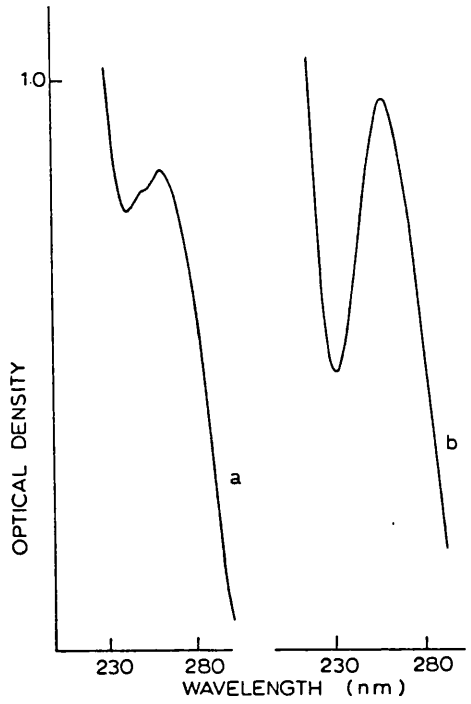


Figure 4.6. Absorbance profiles of RNA from Kiwi fruit
RNA was extracted by the modified Kirby method.
a) RNA precipitated from aqueous phase with 3M lithium acetate and pellet washed twice with 3M lithium acetate.
b) the same RNA after an additional precipitation with 3M lithium acetate and phenol extraction.

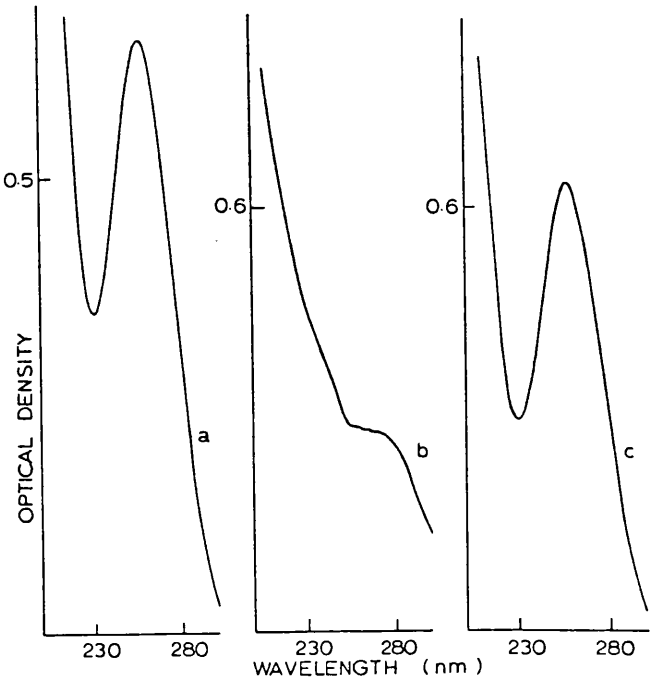


Figure 4.7. Oligo (dT cellulose column chromatography of Kiwi fruit RNA: Absorbance profiles
a) total RNA applied to oligo(dT) cellulose column
b) poly (A+) fraction
c) poly (A-) fraction

non-specific adsorption to the column under conditions required for specific binding of poly(A+)RNA, eluting from the column at lower salt concentrations.

It was therefore necessary to improve the purity of total RNA used for fractionation. Two different approaches resulted in "cleaner" RNA. Results from the first are shown in Figure 4.8. Fruit RNA was extracted by the standard modified Kirby method and showed a very poor absorbance profile (Figure 4.8.a). The RNA, dissolved in 1ml of distilled water, was centrifuged in an HB-4 rotor for 10 minutes at 10,000 r.p.m.. The profile of the supernatant was considerably improved (Figure 4.8.b), and it appeared that this was the result of the contaminant sedimenting during centrifugation (Figure 4.8.c). After a further spin of the supernatant for 1 hour at 13,000 r.p.m. at 4°C, its purity was further improved (d), and although the pellet (e) apparently contained significant amounts of RNA, it also contained a large amount of impurities absorbing near 230nm. The pellet was visible in the form of a dark yellow film. After precipitation with ethanol the supernatant showed an excellent absorbance profile (f) and a ratio OD_{260}/OD_{230} of 2.65.

The conclusions are that the contaminant was likely to be a macromolecule, possibly polysaccharide, which co-purified with RNA. In addition, a fraction of the contaminant may have been ethanol soluble, although it neverless co-precipitated with RNA in the presence of the macromolecular contaminant. During the process of centrifugation about half of the RNA was lost, a consequence which must be acceptable in view of the great improvement of purity.

In a different approach, fruit was freed from seeds and skin (in a very laborious exercise) while still frozen, and extracted by the standard protocol. The reasoning was that seeds and skin of the Chinese gooseberry

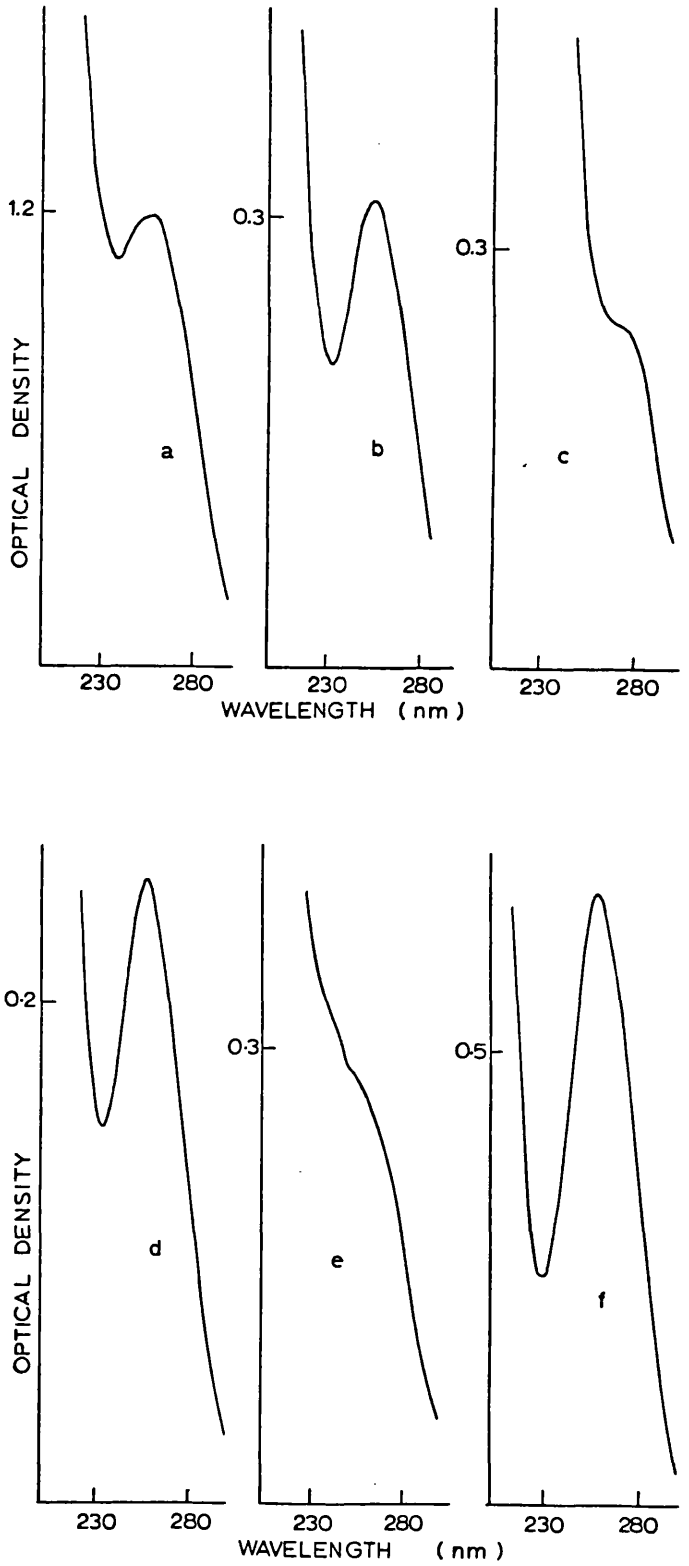


Figure 4.8. Absorbance profiles of successive stages during purification of RNA by centrifugation

a) RNA after extraction by the standard modified Kirby method
b) supernatant after 10 minute centrifugation of RNA
c) pellet of b) dissolved in distilled water
d) supernatant of 1hour centrifugation of supernatant b)
e) pellet of d)
f) supernatant of second centrifugation (d) after ethanol precipitation

fruit are known to contain large amounts of tannin (Schmid , 1978), which may have been a substantial constituent of the contaminating fraction. The results are shown in Table 4.2 (E80), and indeed suggest that a significant fraction of co-purifying contaminant may be present in the seeds and skin of the fruit. The yield of total RNA/g fruit tissue was greater than previously obtained, presumably because fewer precipitations and phenol extractions were required. Therefore, in all subsequent extractions, the seeds and skin were removed.
Table 4.2 also shows the results from an RNA extraction using soft, ripe fruit (E64). The yield of total RNA was only 37μg from 38g fruit, and it is doubtful whether any of this was intact mRNA.

4.2.2. *In vitro* translation of Kiwi fruit and apex RNA

Total RNA prepared from fruit and apex tissue by the modified Kirby method was translated *in vitro*, and the products were separated on SDS-PAGE and autoradiographed. Figure 4.9 shows a comparison of translation products from apex and fruit total RNA. Although the efficiency of incorporation of labelled amino acids by fruit RNA was only 4 times that of the no RNA control, high molecular weight proteins were synthesised, indicating that a reasonable percentage of this RNA was intact. Comparing the distribution of products from fruit and leaf RNA, the majority of protein bands were common to both, but two major bands at M_r 33 and 50 kDa were characteristic of leaf proteins, the smaller of which may represent the precursor of the chlorophyll a/b binding protein. Fruit RNA directed the synthesis of a major double band at approximately 39 and 40 kDa of which the smaller was more abundant. Only a trace of the lower band was synthesised by leaf RNA, suggesting that these two abundant proteins were products from fruit-specific RNA which, assuming all mRNAs are translated

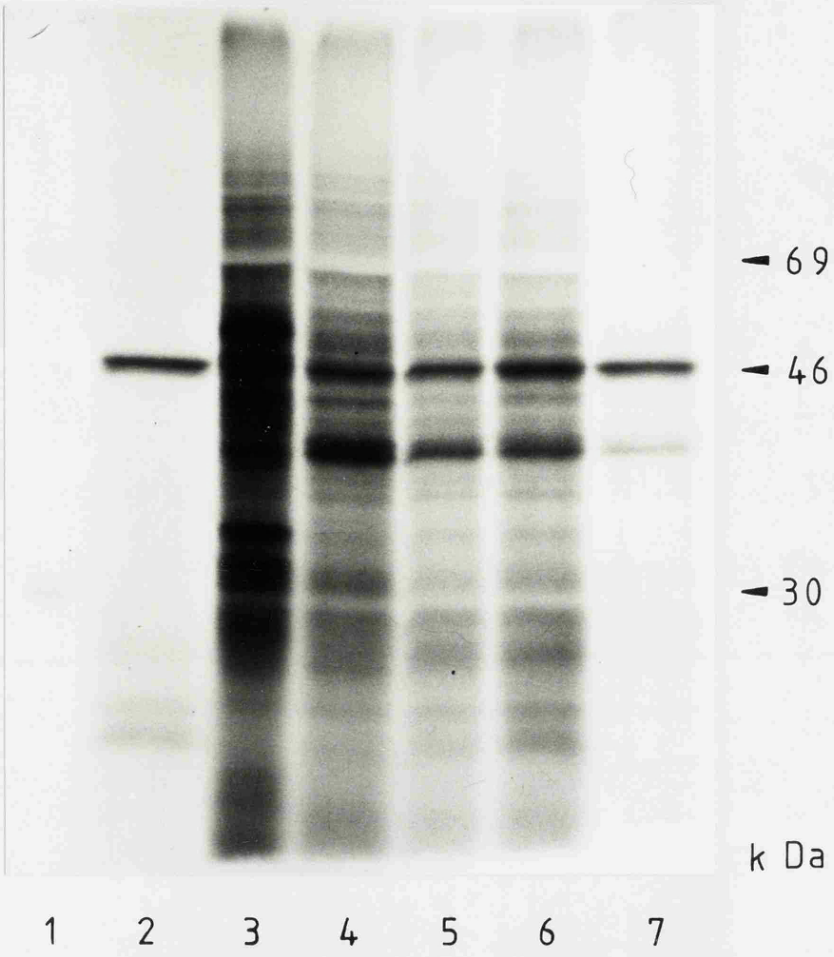


Figure 4.9. SDS-PAGE of *in vitro* translation
12% acrylamide gel

1 molecular weight markers
2 control without added RNA
3 RNA from Kiwi apex (E62)
4 " " unripe Kiwi fruit (E60)
5 " " " " " (E61), stored at 10°C for 10 days
6 " " " " " (E59), " " " " 21 "
7 " " " " " (E63)

lanes 3-7: about 100,000 c.p.m. were loaded per track.
Autoradiograph exposed for 2 days.

at equal rates, appeared to constitute a large fraction of the mRNA population. No apparent difference was observed between the distribution of *in vitro* products of RNA from fruit frozen immediately after harvest and that stored for 10 and 21 days at 10°C.

4.2.3. Immunoprecipitation of *in vitro* translation products

In vitro translation products from fruit RNA were immunoprecipitated with anti-actinidin antibodies (Ab) and pre-immune control Ab, and separated on SDS-PAGE. The autoradiograph (Figure 4.10) shows that the lower band of the doublet, migrating almost as a single band of 39 kDa on this gel, was immunoprecipitated with immune Ab but not pre-immune Ab (lane 5 and 7, respectively). Prolonged exposure of X-ray film revealed two additional bands of approximately M_r 46 and 34.5 kDa (lane 6). The smaller of these was not particularly abundant in the total translation products (lane 4) but its absence from the supernatant of the immunoprecipitate with anti-actinidin Ab (lane 8) and presence in that of the pre-immune Ab (lane 9) suggests that it was selectively immunoprecipitated rather than non-specifically "trapped" in the precipitate.

In the second experiment both apex and fruit *in vitro* translations were immunoprecipitated with anti-actinidin Ab, and fruit *in vitro* products were also immunoprecipitated with anti-papain Ab. The results are shown in Figure 4.11. Apex translation products were not specifically precipitated as both immune and pre-immune Ab gave rise to the same faint background pattern (lanes 4 and 5). Anti-papain Ab did not react with *in vitro* translation products from fruit RNA (lane 9). The immunoprecipitate in lane 7 again corresponded to the lower of the double band in total translation products (lane 6), however, in this experiment the lower band was much less represented than previously, although the same template RNA

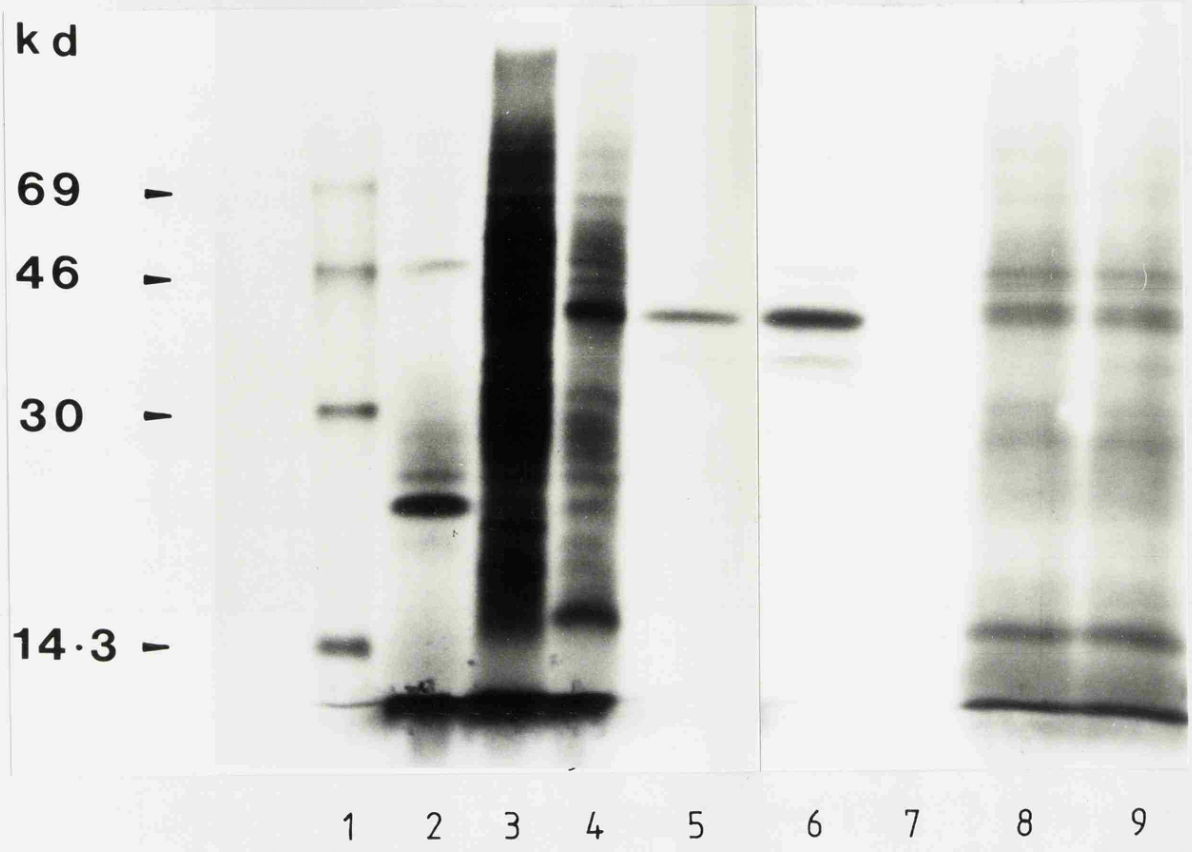


Figure 4.10. Immunoprecipitation of *in vitro* translation products
autoradiograph of SDS-PAGE

	c.p.m.
1 radioactive size markers	
2 control without added RNA	180,000
3 total Kiwi apex RNA	190,000
4 total fruit RNA	72,000
5 immunoprecipitate with anti-actinidin Ab	
6 same as 5	
7 " " pre-immune Ab	
8 supernatant of immunoprecipitate 5 (anti-actinidin Ab)	
9 " " 6 (pre-immune Ab)	

lanes 1-5: autoradiograph exposed for 2 days
lanes 6-9: " " " 14 "

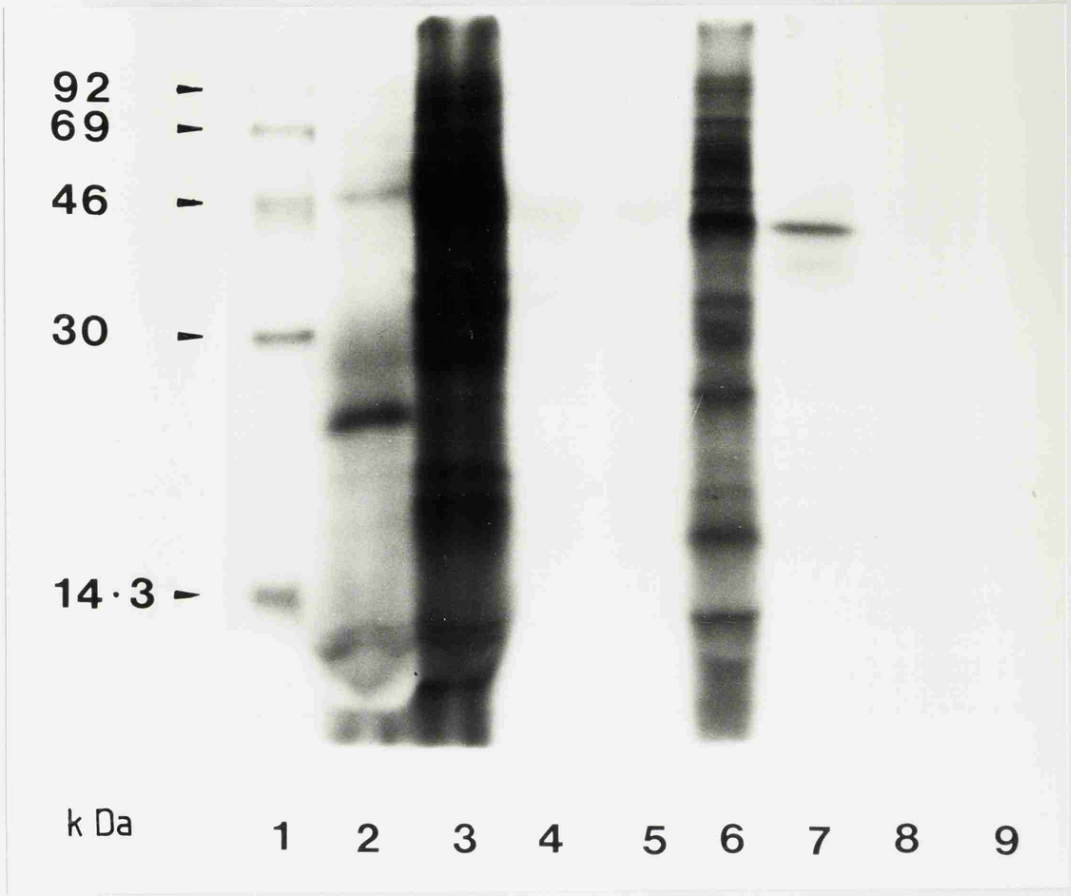


Figure 4.11. Immunoprecipitation of *in vitro* translation products
Autoradiograph of SDS-PAGE

1 radioactive size markers
2 control without added RNA
3 leaf total RNA
4 leaf translation products immunoprecipitated with anti-actinidin Ab
5 " " " " pre-immune Ab
6 fruit total RNA
7 fruit translation products immunoprecipitated with anti-actinidin Ab
8 " " " " pre-immune Ab
9 " " " " anti-papain Ab

Autoradiograph exposed for 4 days.

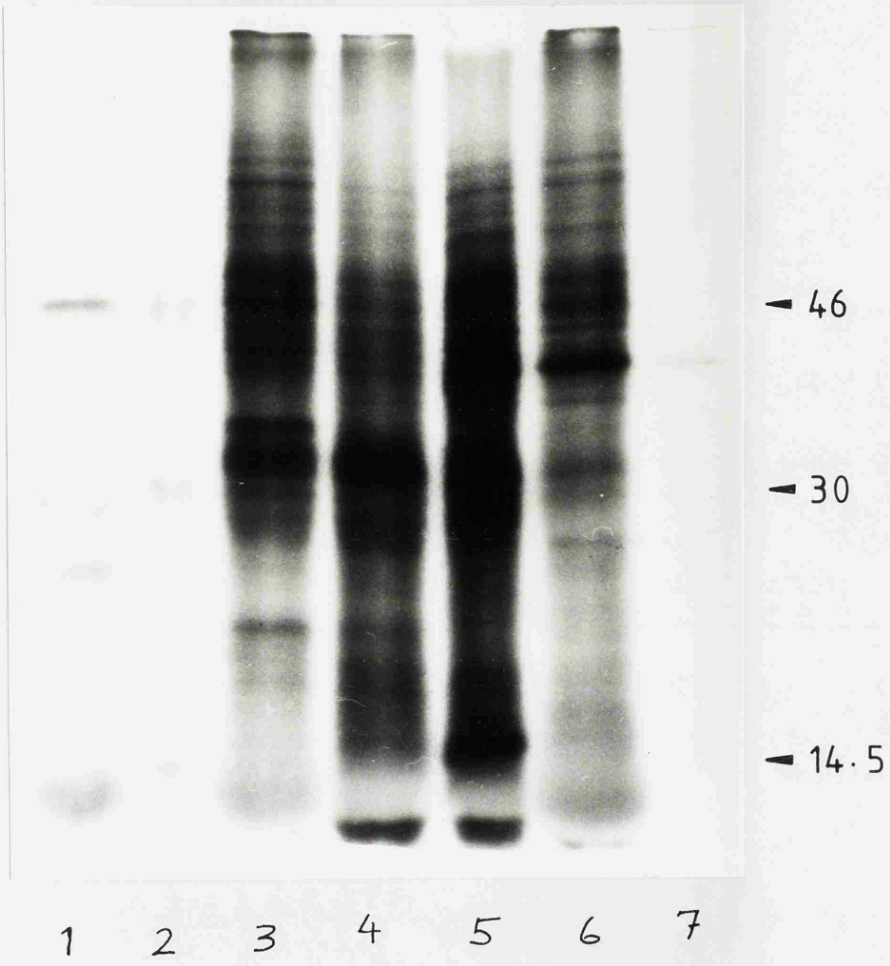


Figure 4.12. Immunoprecipitation of *in vitro* translation products
(autoradiograph of 12% SDS gel)

1. control without RNA
2. molecular weight markers
3. total Kiwi apex RNA
4. poly (A+) Kiwi apex RNA
5. poly (A+) Kiwi fruit RNA
6. total Kiwi fruit RNA
7. Immunoprecipitation of poly A(+) Kiwi fruit RNA

had been used for the *in vitro* translations.

The immunoprecipitation experiments were carried out with translation products from total RNA. However, mRNA does not require a poly(A) tail in order to be active as a template for *in vitro* protein synthesis (Darnell, 1982). Although mRNA coding for a protease is expected to possess a poly(A) tail, as do other non-histone mRNAs, it would be important to demonstrate the synthesis of immunoprecipitable protein from the poly(A+) fraction since this will be used as a template for cDNA synthesis. Figure 4.12 shows the results from an immunoprecipitation of *in vitro* translation products from fruit poly(A+) RNA (lane7). The precipitated protein co-migrated with the abundant species from total fruit RNA (lane 6), suggesting that mRNA coding for this protein was indeed present in the poly(A+) fraction. However, comparison of translation products from total and poly(A+) RNA, both from apex and fruit, clearly shows differences in the relative abundance of proteins synthesised. In particular, the immunoprecipitated protein was by no means the most abundant protein synthesised from poly(A+) RNA as was the case with total RNA.

4.2.4. Peptide mapping

Although the immunoprecipitated *in vitro* product was clearly antigenically related to actinidin, and may well represent an unprocessed precursor of the protease, which would account for the difference in size of up to 15 kDa, an attempt was made to further identify it by peptide mapping. Figure 4.13 shows the peptide map obtained from a digest of actinidin, using decreasing amounts of trypsin. The highest concentration of trypsin resulted in the complete disappearance of intact actinidin (lane 3) and the production of a number of smaller peptides. Although most of the peptides were of higher molecular weight than expected since trypsin

digestion should yield 12 fragments (Carne and Moore, 1978), and most bands are therefore likely to be products of limited proteolysis, the resulting pattern was reproducibly obtained in different experiments.

An immunoprecipitate of fruit *in vitro* proteins was treated in the same way as actinidin: after denaturation by boiling in cracking buffer, half of the immunoprecipitate was incubated with an equal volume of trypsin (2mg/ml) for 10 minutes at 37°C and the reaction stopped by boiling. The untreated and treated immunoprecipitates were electrophoresed in parallel with size markers and tryptic digests of actinidin. The stained dried gel and autoradiograph are shown separately in Figure 4.14.a and b, and superimposed in Figure 4.14.c. The undigested immunoprecipitate can be seen in lane 3a, but no trace of radioactive peptides was recovered in lane 4a, even after prolonged exposure of the film. Lane 3b contains the immunoglobulins associated with the immunoprecipitate; after tryptic digestion these were completely removed (lane 8a) and only trypsin remained, co-migrating with actinidin. The complete absence of both labelled and unlabelled peptides from the digestion of immunoprecipitate and immunoglobulins was unexpected and seems to be the result of complete hydrolysis of the proteins.

Although actinidin appeared to migrate close to the 24kDa molecular weight marker on this gel system, it is unfortunate that the same gel was not suitable for the separation of the high molecular weight protein markers (lane 1a). However, even if these markers were more clearly separated it is doubtful whether the size of the immunoprecipitate could be estimated accurately, since it would be outside of the linear range of this gel system (Swank and Munkres, 1970).

Although the results of this experiments were inconclusive, further attempts at improving conditions for peptide mapping were decided against

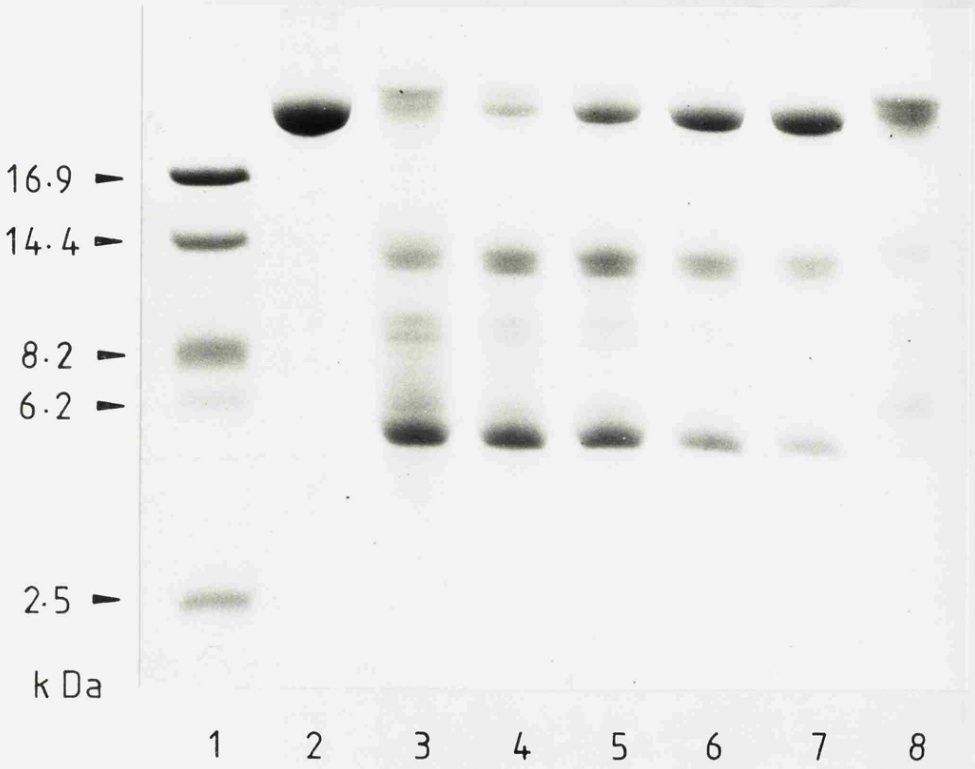


Figure 4.13. Peptide mapping: SDS-urea gel of actinidin tryptic digests

1 molecular weight markers
2 20μg actinidin
3 " " with 20μg trypsin
4 " " " 10 " "
5 " " " 5 " "
6 " " " 2.5 "
7 " " " 1.25 "
8 20μg trypsin

Figure 4.14. Peptide mapping of actinidin and immunoprecipitate of *in vitro* translation products

1 radioactive high molecular weight markers and 24 and 30 kDa non-radioactive markers

2 actinidin

3 immunoprecipitate of Kiwi fruit *in vitro* translation products

4 " digested with trypsin (20μg)

5 tryptic digest of actinidin (20μg with 20μg trypsin)

6 " " " (" " 10 " ")

7 " " " (" " 2.5 ")

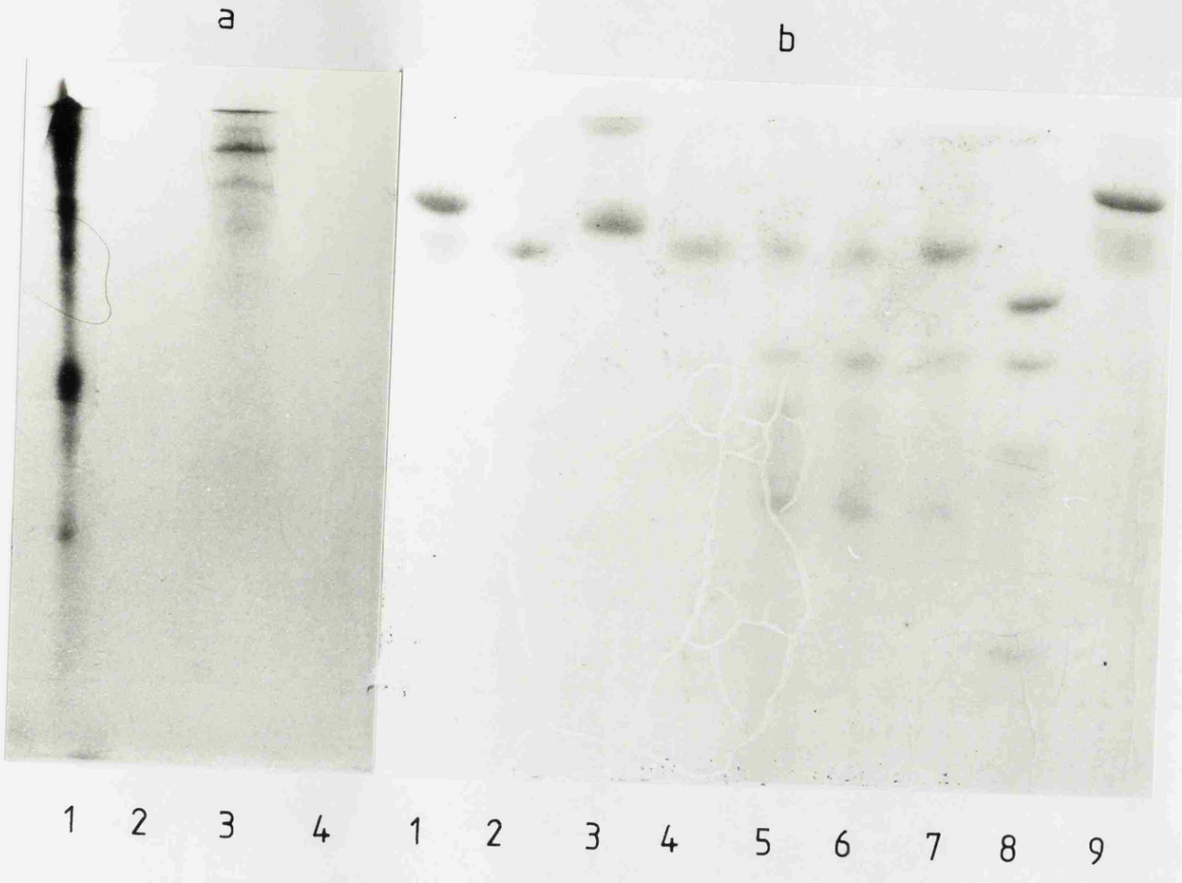
8 low molecular weight markers

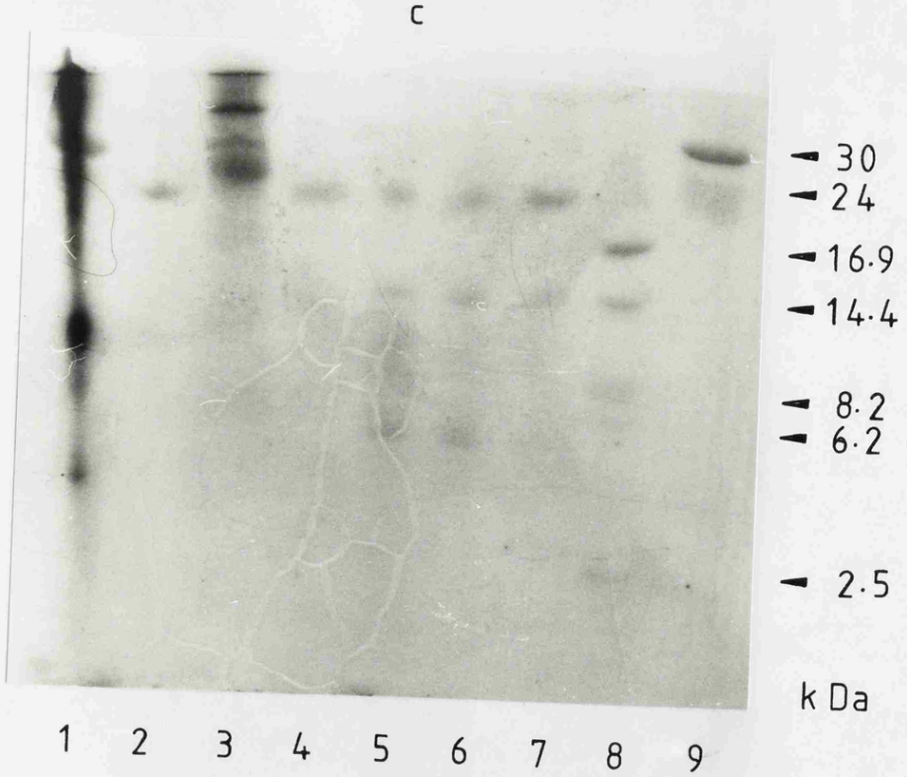
9 24 and 30 kDa markers

a) autoradiograph of gel

b) stained and dried gel

c) a) and b) superimposed


a
b
1 2 3 4
1 2 3 4 5 6 7 8 9


c
30
24
16.9
14.4
8.2
6.2
2.5
k Da
1 2 3 4 5 6 7 8 9

in view of the very limited availability of fruit mRNA and also of the expected effort and time involved.

4.3. Conclusions

4.3.1. RNA purification methods

The modified Kirby method, successful in the isolation of RNA from all the tissues used, had several features which were crucial to its success. The initial grinding of tissue at the temperature of liquid nitrogen suspended the action of ribonuclease. Since half of the final volume of extraction buffer, containing powerful denaturing agents and detergents, was frozen and ground with the tissue, the action of ribonuclease was likely to be minimised during thawing of the extract. Thus the initial stages of this method would ensure a good recovery of undegraded RNA. RNA was selectively precipitated from the aqueous nucleic acid solution with salt in which DNA and carbohydrate are soluble. This appeared to be an efficient method for the removal of most carbohydrate. Washing of ethanol precipitate containing nucleic acids and carbohydrate with salt, on the other hand, was ineffective in removing the carbohydrate. Even ripe, soft fruit containing only a trace of RNA in a vast excess of carbohydrate was amenable to this method, suggesting that the differential precipitation of RNA with salt was indeed very efficient. This is also reflected in the fact that the yield of RNA from the control tissue, pea, was higher than with any other method. Thus, the modified Kirby method appears to be suitable for the isolation and purification of RNA from a number of tissues which give poor results with other methods.

4.3.2. *In vitro* translation systems

Initially the wheatgerm translation system was used. However, the efficiency of incorporation of labelled amino acids varied greatly between different wheatgerm preparations and never reached the level of a highly efficient "control wheatgerm" prepared by J.Matthews (this department). This is reflected in the results shown in Table 4.1 for experiment E6, where total RNA from peas appeared to be twice as efficient as poly(A+) RNA at directing protein synthesis. An inefficient protein synthesising system would not be acceptable as this may be selective in synthesising only small proteins, with the result that the distribution of *in vitro* products does not reflect the distribution of mRNA's. Therefore commercially purchased rabbit reticulocyte lysate was used subsequently, and this gave more reproducible results. The disadvantage of rabbit reticulocyte lysate was the presence of translatable endogenous RNA which increased the background of incorporation of labelled amino acids and gave rise to a strong protein band on autoradiographs of acrylamide gels, although in the presence of good quality mRNA, translation of the endogenous RNA was significantly reduced.

4.3.3. *In vitro* translation and immunoprecipitation of Chinese gooseberry RNA

Comparison of *in vitro* translation products from total and poly(A+) RNA, both from apex and fruit, showed considerable differences in the distribution of proteins. In particular, the immunoprecipitated protein was much less represented in the products from poly(A+) RNA. It is possible that some mRNAs were preferentially translated in the total RNA, giving rise to the predominant immunoprecipitable protein. However, the fact that poly(A+) RNA was obtained by an additional fractionation step

would support the view that translation products from total RNA were more representative of the composition of mRNA *in vivo*. As a result of gradual shortening of the poly(A) tail of mRNA molecules during their life time (Taylor, 1979) each mRNA species would be expected to be composed of a mixture of molecules with different tail lengths. Since the ability of mRNA to bind to the affinity column depends on the presence of a poly(A) tail at least 20 nucleotides long (Taylor, 1979) differences in the affinity of mRNA molecules due to different tail lengths may have given rise to the selective purification of a number of RNA species. The observed reduction in the immunoprecipitable protein in translations from poly(A+) RNA may indicate that the majority of the mRNAs coding for this protein may have possessed short poly(A) tails.

However, the fact that immumoprecipitate was obtained with poly(A+)-directed translation products showed that the poly(A+) fraction contained the mRNA of interest and that this was poly-adenylated. Thus poly(A+) RNA from Kiwi fruit would be suitable for cDNA cloning, although the RNA of interest did not represent the major species of the poly(A+) fraction and would therefore not result in a cDNA library containing a large proportion of the desired clones.

The immunoprecipitated *in vitro* synthesised protein was considerably larger than actinidin. Assuming that the *in vitro* and *in vivo* proteins are products of the same gene, the results suggest that actinidin is synthesised as a larger precursor. The difference in size amounted to approximately 15 kDa, although it is possible that, like actinidin, the precursor migrates abnormally on SDS-PAGE. Actinidin has an unusually low isoelectric point, as a result of a large proportion of negatively charged side chains on its surface. It is possible that this characteristic may

have interfered with denaturation and affected its migration on SDS-PAGE, since in the presence of the strong protein denaturant, urea, it migrated at the expected position. Examples of aberrant migration on SDS-PAGE of proteins have been reported previously (De Vries *et al.*, 1985, Faye and Chrispeels, 1987).

Although immunoprecipitation of the *in vitro* translated protein does not prove its identity as a precursor of actinidin, the fact that this translation product was also the most abundant protein snthesised by fruit RNA strongly supports such a conclusion.

The presence of a close doublet constituting the major protein synthesised from total fruit RNA is interesting, particularly as the relative amounts of the two species appeared to vary between *in vitro* translation experiments, using the same batch of RNA as a template. The immunoprecipitate corresponded to the lower band in all cases. It seems tempting to speculate that some form of auto-catalytic processing may be responsible for the conversion of the larger into the smaller form of the protein precursor.

5. CDNA cloning and screening of recombinants

5.1. Introduction

RNA fractionated twice on oligo(dT) from unripe but fully grown Kiwi fruit was used as a template for cDNA synthesis. The method of Gubler and Hoffman (1983) was chosen as it avoids the use of S_1 nuclease treatment of cDNA, and higher yields of cDNA can be expected. Briefly, first strand cDNA synthesis is primed by annealing of oligo(dT) to the poly(A) tail, and cDNA synthesis catalysed by reverse transcriptase. The RNA strand of the RNA-cDNA hybrid is replaced by DNA in a nick translation reaction with DNA polymerase I in the presence of RNase H.

After C-tailing of the double stranded cDNA, this was cloned into a plasmid vector, pUC9. This has a selectable marker, ampicillin resistance, which only allows transformants to grow on agar containing the antibiotic. In addition, since the cloning site is within the β-galacosidase gene, the insertion of cDNA results in inactivation of the gene. Recombinant transformant produce white colonies on agar containing the lactose analog and indicator X-gal, while non-recombinants form blue colonies. The presence of a number of unique restriction sites within the multiple cloning site is useful for subcloning and orientation of the insert.

For the identification of recombinant clones a combination of methods was used. Screening of cDNA libraries with synthetic oligonucleotide probes has been successful in the identification of, for example, sequences coding for glycinin (Negoro, 1985), papain (McKee *et al.*, 1986), and human gastrin (Noyes *et al.*, 1979). Since the complete amino acid sequence of actinidin is known, an oligonucleotide was designed which would be complementary to a region coding for amino acids 190 to 194 (see below). This region was chosen because it contained the minimum number of codon

Synthetic oligonucleotide

```
                 190                 194
                 glu glu gly tyr met
mRNA        5'   GAA GAA GGX UAU AUG  3'
                   G   G       C

oligo       3'   CTT CTT CCX ATA TAC  5'
                   C   C       G
```

degeneracies, requiring the synthesis of a 32-fold degenerate oligonucleotide. However, since screening of recombinants by colony hybridisation with the synthetic oligonucleotide was unsuccessful, the library was pre-screened by differential hybridisation, using radioactively labelled first strand cDNA from fruit and leaf RNA. As this is a quantitative method only abundant mRNA species are expected to produce radioactive probes detectable upon hybridisation with their complementary cDNA sequences. Since actinidin mRNA was shown to be abundant in fruit but not leaf, those colonies hybridising with the leaf probe were eliminated from further screening and colonies hybridising only with the fruit probe were selected for further analysis. In the hope to synthesise a more specific cDNA probe from fruit RNA, reverse transcription was primed with the synthetic oligonucleotide instead of oligo(dT) as described by Chan *et al.* (1979). Putative positive clones were further identified by dot blot hybridisation of their isolated plasmid DNAs with the synthetic oligonucleotide. The cDNA inserts of positive clones were sized by gel electrophoresis and confirmed by Southern blotting.

5.2. Results

5.2.1. cDNA synthesis

Both first and second strand cDNA synthesis were carried out in several parallel reactions, with one reaction containing ^{32}P-labelled dCTP to allow estimations of yield and size of the cDNA. The efficiency of first strand synthesis, expressed as the amount of cDNA obtained as a percentage of mRNA added, varied between 10% and 90% for different experiments. It appeared to be largely dependent on the quality of reverse transcriptase, since a new batch of enzyme yielded 90% efficiency compared with 30% efficiency obtained with enzyme which had been stored for several months. In several experiments of second strand synthesis, no more than 13% of cDNA:RNA hybrid was converted into double-stranded cDNA. However, the yield of first strand cDNA was always estimated from the fraction of incorporated radioactivity eluting from the Sephadex column, without confirmation of the recovery after phenol extraction and precipitation. If recovery was low, the efficiency of second strand synthesis would have been underestimated.

Samples of cDNA labelled with ^{32}P-dCTP in the first or second strand were electrophoresed on denaturing polyacrylamide gel, and autoradiographed. Figure 5.1 shows long smears both for first and second strand cDNA, with greatest intensity around 600 bases, suggesting that the cDNA may be suitable for cloning. The presence of a distinct band in the track of first strand cDNA around 500 bases suggests the presence of a predominant size class, although its significance is not clear.

5.2.2. C-Tailing of double-stranded cDNA

C-tailing was carried out by the dideoxy chain termination method (communicated by A.Greenland) as this eliminated the necessity of pilot

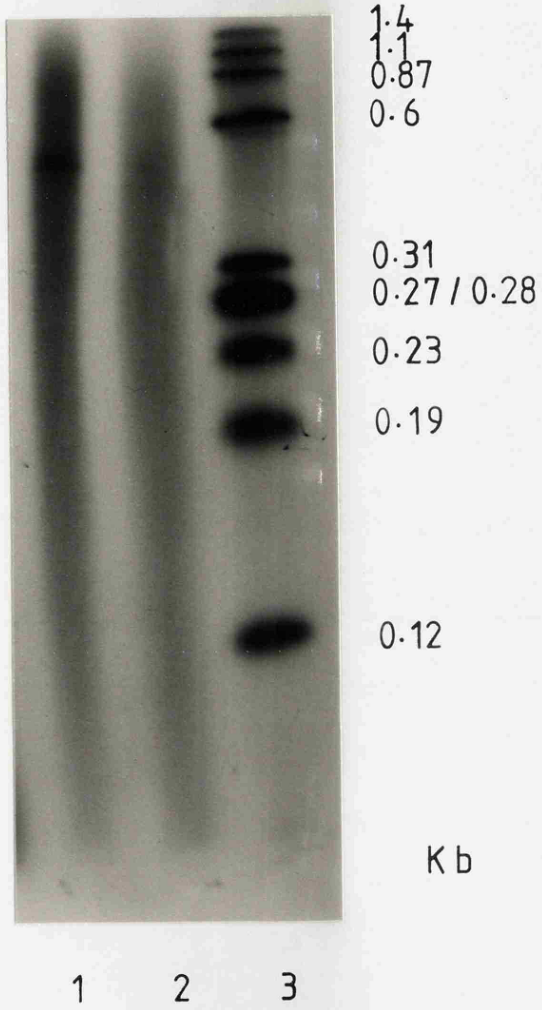


Figure 5.1. Size estimation of cDNA
cDNA was electrophoresed on denaturing polyacrylamide gel and autoradiographed for 4 days.
Lane 1: ^{32}P-labelled first strand cDNA.
Lane 2: double stranded cDNA labelled only in the second strand.
Lane 3: ^{32}P-labelled ΦX174 DNA digested with HaeIII.

assays to monitor the tailing reaction. It was expected to yield a proportion of cDNA with suitable tail length, independent of the number of 3' ends present. From the fraction of incorporated radioactivity it was estimated that on average 24 residues of dC were added to each 3' end.

5.2.3. Annealing and transformation

C-tailed cDNA was annealed to G-tailed pUC9 vector DNA, and transformed into competent *E.coli* JM83 host cells. The yields obtained with different ratios of plasmid to vector DNA are shown in Table 5.1.

Table 5.1 Yield of recombinant transformants

amount cDNA (ng)	ratio (weight) plasmid:cDNA	no.of transformants blue	white	percentage recombinants
1.0	3.3:1	66	114	63
0.5	6.6:1	41	74	64
0.5	13.2:1	82	57	41
0.5	26.4:1	121	76	39
0.5	52.8:1	276	48	15
0	plasmid only	190	2	

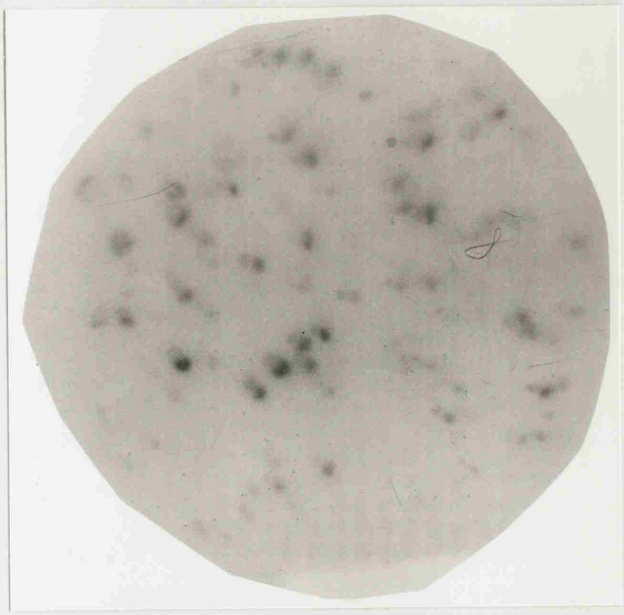
The highest yield was obtained with the lowest plasmid:cDNA ratio, and this ratio was used in further transformation experiments. From a control transformation with pUC18 DNA, the transformation efficiency was calculated as 4.3×10^4 transformants per μg of plasmid DNA, which was lower than expected. The highest yield of recombinants obtained per μg of cDNA was 1.14×10^5, or 2×10^3 transformants per μg RNA. The background transformation of non-recombinants was 3.8×10^4 blue colonies per μg Pst-cut, G-tailed plasmid pUC9.

From a total of 0.025 μg cDNA 1,765 recombinants were obtained. Approximately one thousand recombinants were transferred to replicate nitrocellulose filters, one of these retained as a master and the others used for colony hybridisation.

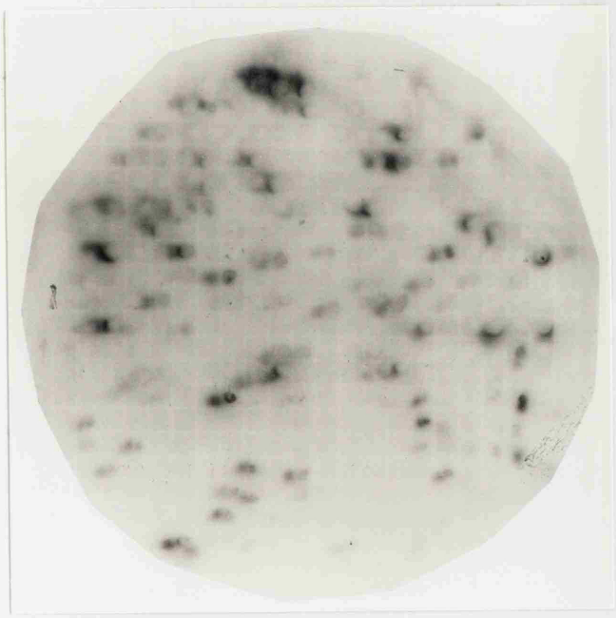
5.2.4. Screening of recombinant clones

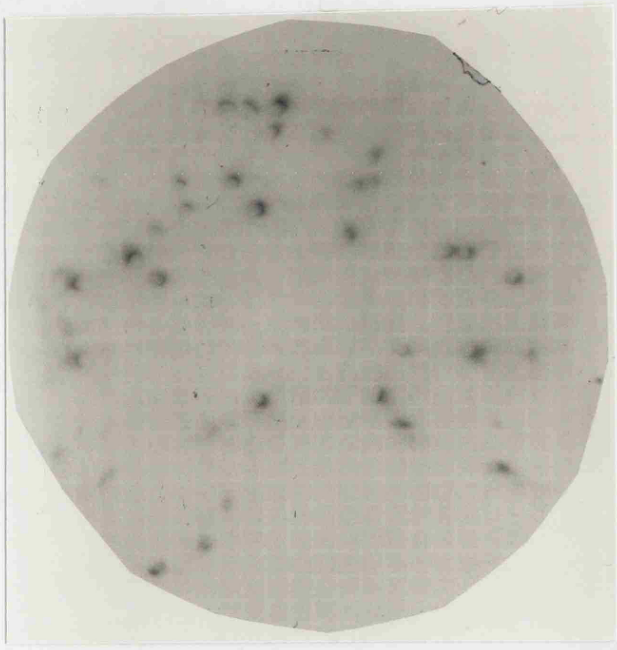
Initially one set of filters representing the whole library was hybridised with ^{32}P-labelled synthetic oligonucleotide. The hybridisation was carried out at 35°C, and filters were washed at the same temperature. The autoradiograph (not shown) showed only faint, uniform background fogging, but none of the colonies gave a strong signal. In a repeat of the experiment with hybridisation and washing temperatures of 25°C some colonies gave very strong signals, but these appeared to be due to non-specific labelling as they were clustered non-randomly in several areas of the filters, and since some of the negative control colonies showed positive signals. Colony hybridisation using the synthetic oligonucleotide was subsequently confirmed to give false positive results, as illustrated in Figure 5.2.

The cDNA library was therefore screened by differential hybridisation with radioactively labelled cDNA from fruit and leaf RNA. Figure 5.2 shows an example of a set of replica filters hybridised with three different probes: a) leaf cDNA, b) fruit cDNA primed with oligo(dT), c) and fruit cDNA specifically primed with the synthetic oligonucleotide. The results from screening of approximately 1000 colonies are summarised as follows: out of 169 colonies showing positive hybridisation with the leaf cDNA probe, 164 also hybridised with the fruit probe b. These clones were likely to contain sequences expressed at high levels in both tissues and were therefore eliminated from further screening. A further 187 colonies hybridised to the fruit, but not the leaf probe, and these were likely to



a





b

c

Figure 5.2. Differential hybridisation of cDNA library

Replica nitrocellulose filters were hybridised with ^{32}P-labelled first strand cDNA, which was synthesised using the following templates:
a) leaf mRNA.
b) fruit mRNA primed with oligo(dT).
c) fruit mRNA primed with synthetic oligonucleotide.

represent abundant fruit-specific sequences. Specifically-primed fruit cDNA (probe c) hybridised with approximately 160 colonies, most of which did not also hybridise with the leaf probe. Some of the colonies hybridising strongly with the specific-primed probe showed only weak signals with the oligo(dT)-primed fruit probe. 48 clones were chosen for further analysis by dot blot hybridisation of their plasmid DNAs. Plasmid DNA was isolated, immoblised on nitrocellulose and screened with radioactively labelled synthetic oligonucleotide. Figure 5.3 shows autoradiographs of the filter after washing at low and high stringency, and clearly suggests that clones 1, 3 to 8, 13 and 41 contain sequences related to the oligonucleotide. It is interesting that the majority of positively hybridising DNAs are derived from colonies giving the strongest signals on colony hybridisation with the specifically primed fruit probe. Plasmids from an additional 41 clones were isolated and screened in the same way, and one more positive clone (no. 89) was found.
Re-screening of clones containing positively hybridising plasmid DNA by colony hybridisation again showed false positives (Figure 5.3.d.).

5.2.5. Characterisation of positive recombinants

In order to determine the size of cDNA inserts and to test whether the hybridisation to the oligonucleotide is specific for the cDNA insert rather than vector or chromosomal sequences, plasmid DNA from all positive clones and from a non-recombinant as well as a negative recombinant clone were PstI-digested and electrophoresed on agarose gel. Insert sizes were determined from the ethidium bromide-stained gel (Figure 5.4.a), and the gel was blotted and hybridised to synthetic oligonucleotide (Figure 5.4.b). The size distribution of inserts is shown in table 5.2. The autoradiograph of the Southern blot shows that the cDNA inserts rather than vector DNA hybridised with the probe. The additional faint bands detected by the probe are due to undigested or partially digested

Figure 5.3. Dot blot hybridisation of isolated plasmid DNA

a) and b) Initial screen of clones:

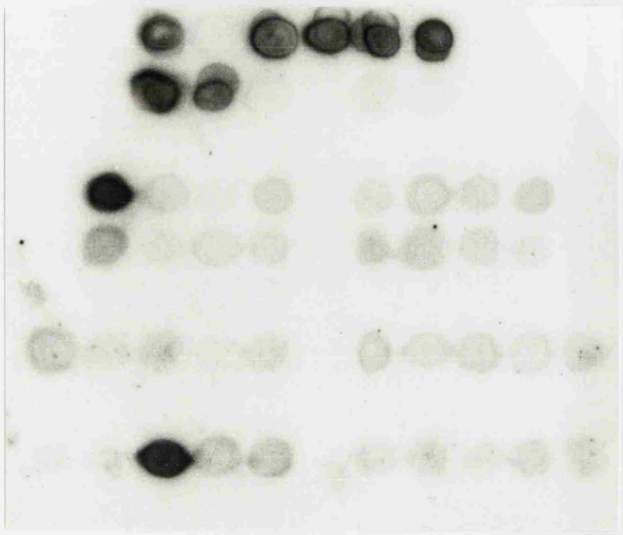
a) Plasmid DNA from 48 clones, immobilised on nitrocellulose, was hybridised with ^{32}P-labelled synthetic oligonucleotide at 20°C and washed at 24°C (6X SSC).

b) the same filter after an additional wash at 35°C (6X SSC).

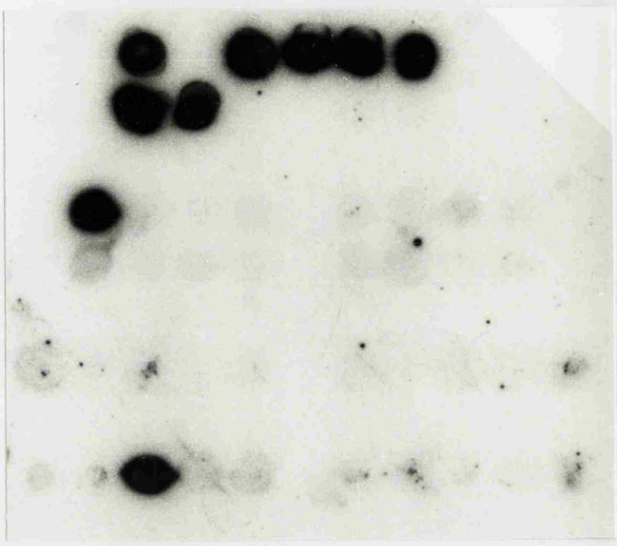
c) Plasmid DNA hybridising with probe in the initial screen, was screened again with a further 40 clones. Hybridisation was as above, and washing at 35°C.

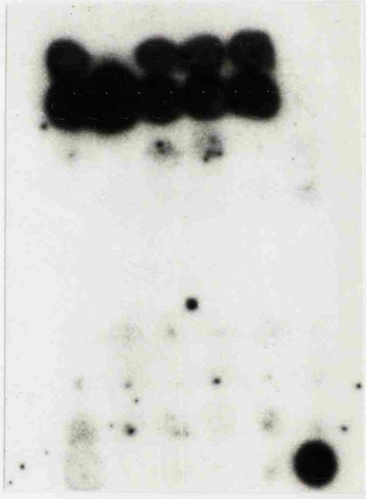
d) Colony hybridisation of clones as in c, arranged in the same order as the dot blot. Colonies showing positive signals do not correspond to positively hybridising plasmid DNA.

a

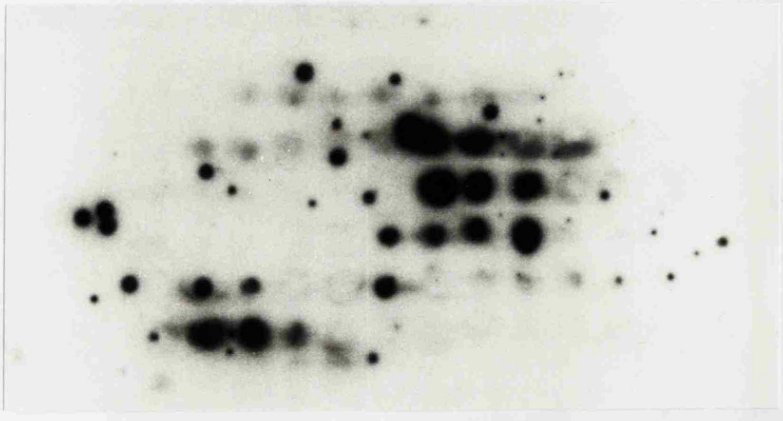


b





c



d

Figure 5.4. Size estimation of cDNA inserts and Southern blot

2μg of plasmid DNA was digested with PstI and electrophoresed on a 1% agarose gel.

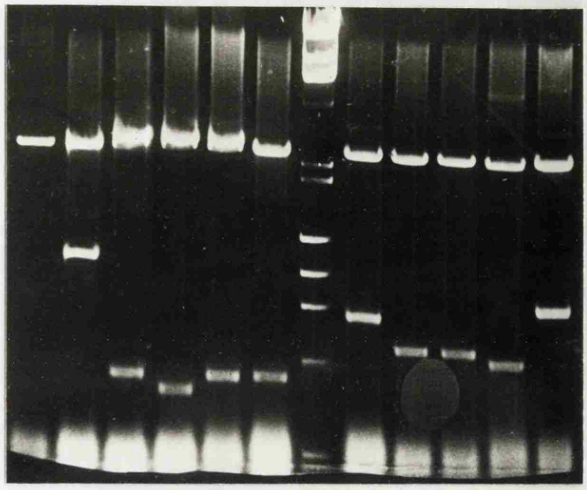
a) Ethidium bromide-stained DNA visualised under UV-light.

b) Southern blot of the same gel, hybridised with ^{32}P-labelled synthetic oligonucleotide at 20°C, washed at 25°C in 6X SSC, and autoradiographed for 3 days.

c) A similar Southern blot hybridised with ^{32}P-labelled, nick-translated insert from clone 1, at 65°C, washed at 65°C in 1X SSC, 0.1% (w/v) SDS, and autoradiographed for 2 days.

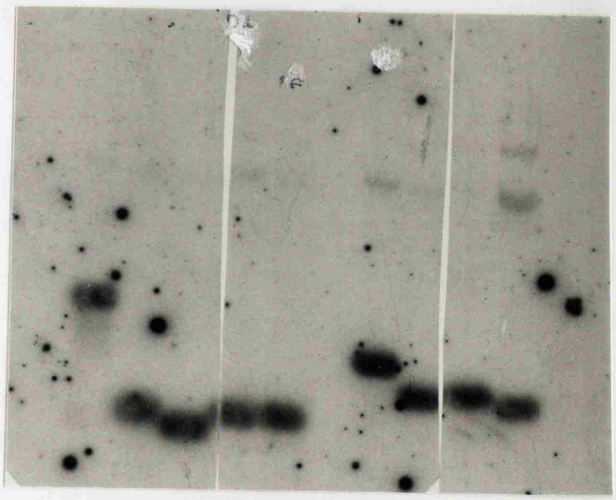
Lanes:

1	non-recombinant control	8	clone 7
2	clone 1	9	clone 8
3	clone 3	10	clone 13
4	clone 4	11	clone 41,
6	clone 6	12	clone 57, negative recombinant control
7	λHindIII markers		



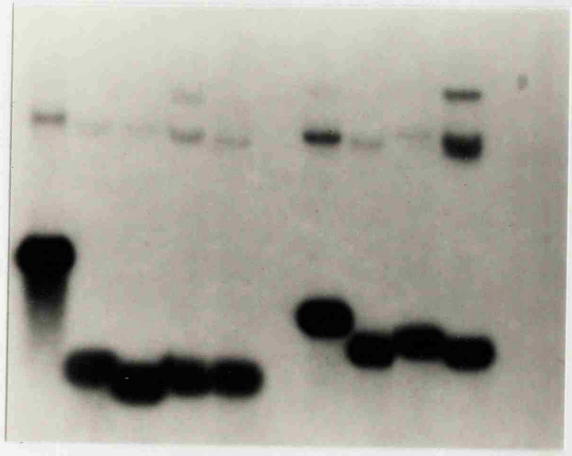
a

1 2 3 4 5 6 7 8 9 10 11 12



b

1 2 3 4 5 6 7 8 9 10 11 12



c

2 3 4 5 6 7 8 9 10 11 12

Table 5.2: Size distribution of cDNA inserts

Clone No.	size (bp)	Clone No.	size (bp)
1	1,250	7	810
3	570	8	630
4	500	13	630
5	540	41	575
6	540	89	570

(linearised) plasmid rather than vector DNA. Clone 89, which was not included in this experiment, was subsequently confirmed to contain an insert of 570 bp which specifically hybridised with the probe.
The Southern blot was washed and rehybridised with ^{32}P-labelled insert from clone 1, and the strong hybridisation indeed confirms that all cDNA inserts have highly homologous sequences (Figure 5.4.c).

5.3. Conclusions

The efficiency of cDNA cloning was not very high at $2x10^3$ colonies per μg RNA, compared to reported efficiencies of 10^6 recombinants per μg RNA (Gubler and Hoffman, 1983). This was apparently due to low yields in every step, including first and second strand cDNA synthesis as well as transformation itself. In addition the background level of transformants obtained with vector DNA only was very high and suggests that a considerable proportion of vector DNA was not digested before it was G-tailed. These non-recombinant molecules are likely to have competed with recombinant transformation.
Ten clones were identified whose inserts hybridised to the oligonucleotide probe and to each other. The majority of inserts were small, but one (clone 1) was sufficiently long to contain a considerable portion of the

coding sequence of the gene. Since the probe used to detect the recombinants was complementary to the 3' coding region of the RNA (carboxy terminus of polypeptide), the small inserts must have arisen as a result of incomplete first strand rather than second strand cDNA synthesis. In retrospect it would have been useful to rescreen the whole library by colony hybridisation with one of the positive cDNA inserts, as there may have been further related clones. However, the fact that clone 1, chosen first for its very intense hybridisation signal, also had the largest cDNA insert, suggests that even longer cDNA clones are unlikely to have been present in the library.

It is not clear why the synthetic oligonucleotide failed to detect complemetary sequences in colony hybridisation, while it hybridised specifically with these sequences in Southern blot or dot blot hybridisation. However, it is possible that the conditions of hybridisation and washing were too stringent in the first and not stringent enough in the second experiment. Although in the majority of reported cases, mixed synthetic oligonucleotides have been successfully used in colony hybridisations (see above), San Segundo *et al.* (1985) failed to identify cathepsin B encoding cDNA clones by this method, using two mixed synthetic oligonucleotides of 14 and 17 bases.

Although hybridisation to the synthetic oligonucleotide does not unequivocally prove that the cDNA inserts contain sequences coding for actinidin, the probability of a similar sequence occurring in an unrelated gene is very small indeed (Lathe, 1985). Therefore rather than confirming the identity of the clones by methods such as hybrid release translation (Paterson *et al.*, 1977), it was decided to sequence the two largest clones.

6. Sequence Analysis of cDNA clones pAC.1 and pAC.7

6.1. Introduction

Briefly, the method chosen for sequencing involves the use of the sequencing vectors M13 mp18 and M13 mp19 (Yanisch-Perron *et al.*, 1985), which are identical except for the orientation of their multiple cloning site, thus facilitating the sequencing of fragments on both strands. Fragments resulting from digestion of the DNA with two enzymes, creating two different sticky ends compatible with restriction sites in the M13 vectors, were subcloned into both vectors. Single stranded templates prepared from the phage were sequenced by the chain termination method of Sanger *et al.* (1977), using a synthetic primer which hybridises to a sequence in the vector DNA (Heidecker *et al.*, 1980).

6.2. Results

6.2.1. Sequencing strategy: Figure 6.1 shows the restriction maps obtained for clones pAC.1 and pAC.7. The sequencing strategy is indicated by lines below the maps.

Although the whole sequence of pAC.1 could be determined on one strand by forced cloning, i.e. subcloning of fragments with two different ends, using the enzymes BamHI, HaeIII, HincII, MspI and TaqI, this was not always possible for confirmation on the opposite strand. Therefore templates were prepared by shotgun cloning: the insert contained multiple sites for MspI and RsaI, and fragments produced by digestion with these enzymes were subcloned into M13 mp18. Templates were prepared from up to 16 clones from each ligation, and their identities determined by parallel electrophoresis of only one sequencing reaction. From each group of identical templates, the sequence of one was determined fully. In this

way the sequence of pAC.1 was determined completely on both strands. Most of the sequence of pAC.7 was determined on both strands by forced cloning, except for 180 bp of the 3'end, because of interference by a long homopolymer (G.C) tail, and a region of 60 bp around the stop codon, due to the lack of a suitable restriction site in this region.
However, the sequence of regions determined on only one strand can be accepted with confidence because 1) both regions were confirmed by sequencing of more than one template and 2) sequencing of opposite strands from the other regions always confirmed the sequence determined on the first strand without ambiguities.

6.3. Analysis of sequencing results

6.3.1. Nucleotide sequences: The nucleotide sequences determined for both pAC.1 and pAC.7 are shown in Figure 6.2, together with the deduced amino acid sequences. The large cDNA clone, pAC.1, has an open reading frame of 908 base pairs coding for 302 amino acids, and a 3' untranslated region of 229 bp excluding the poly(A) tail of 8 residues. The shorter clone, pAC.7, has an open reading frame of 552 base pairs coding for 184 amino acids, followed by a 3' untranslated region of 225 base pairs plus 32 residues of poly(A) tail. An alignment of restriction maps (Figure 6.1), constructed from the DNA sequence, shows that two of the restriction sites present in pAC.1 are lacking in equivalent positions of pAC.7.
Within the region of homology the two sequences differ by 7 nucleotides, four of these in the coding, and three in the 3' untranslated region (Figure 6.2). In addition, pAC.1 has an insertion of four nucleotides immediately upstream from the poly-adenylation site.
The 3' untranslated sequences of both clones contain three regions homologous to the consensus sequence for plant polyadenylation signals

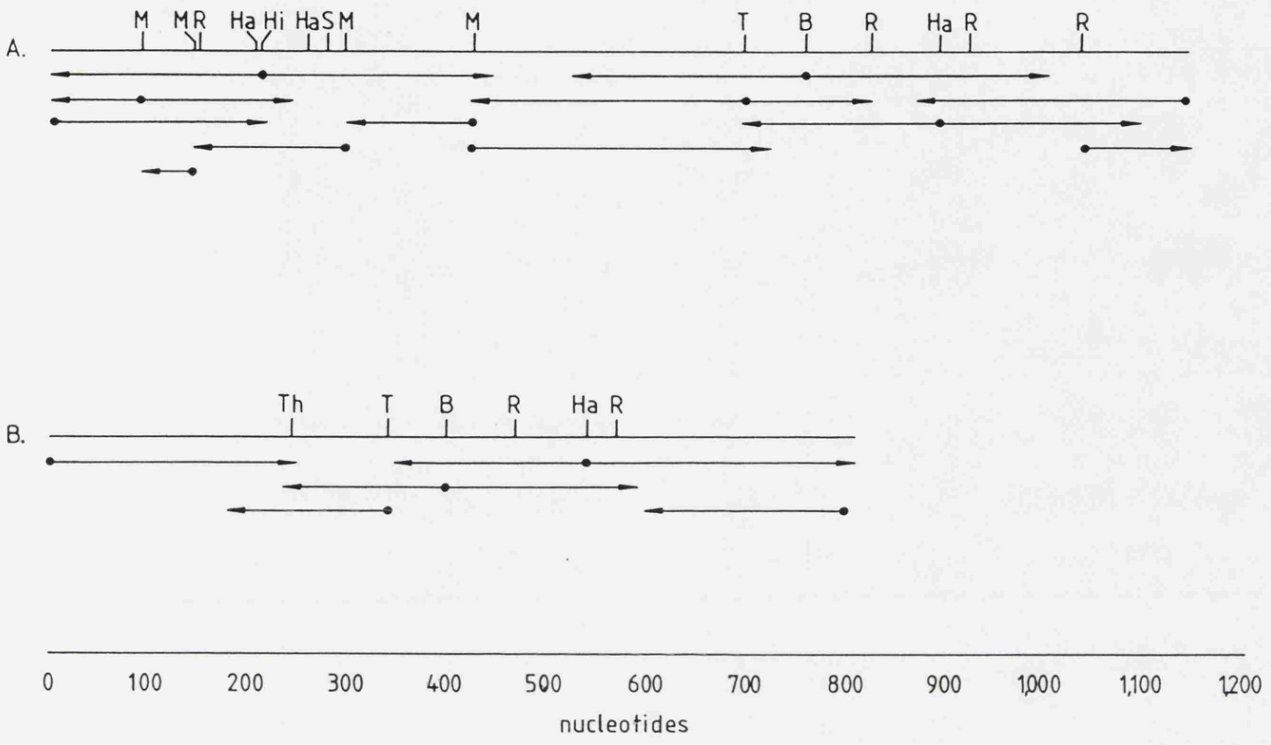
Figure 6.1. Restriction analysis

A and B: restriction maps, with sequencing strategy indicated by arrows below, of pAC.1 (A) and pAC.7 (B).

C: aligned restriction maps of pAC.1 and pAC.7, indicating the coding regions for the mature enzyme domain (black boxes) and for the additional polypeptide sequences (white boxes), as well as the non-coding regions (single line).

Restriction enzymes:

B = BamHI	R = RsaI
Ha = HaeIII	S = Sau3A
Hi = HincII	T = TaqI
M = MspI	Th = ThaI


A.
M
MR
Ha
Hi
HaS
M
M
T
B
R
Ha
R
R
B.
Th
T
B
R
Ha
R
0
100
200
300
400
500
600
700
800
900
1,000
1,100
1,200
nucleotides

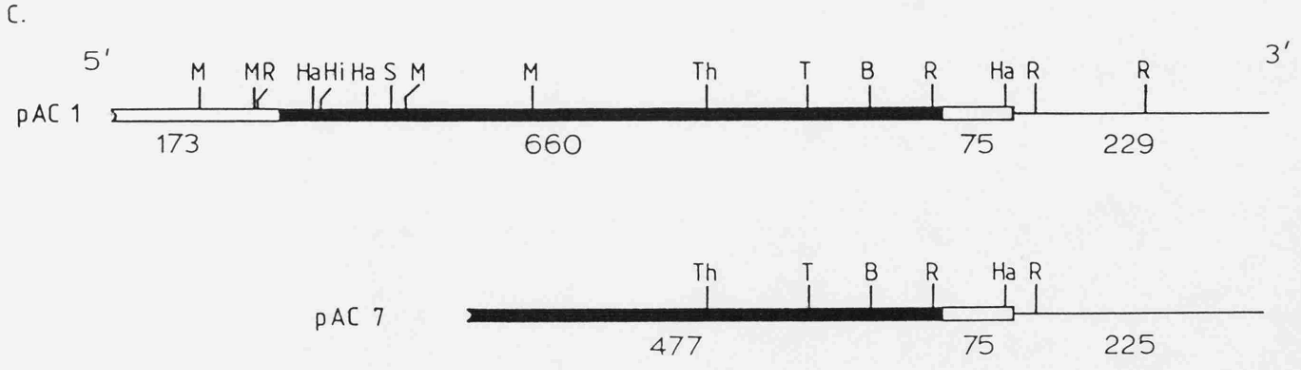

C.
5'
3'
pAC 1
M
MR
Ha
Hi
Ha
S
M
M
Th
T
B
R
Ha
R
R
173
660
75
229
pAC 7
Th
T
B
R
Ha
R
477
75
225

Figure 6.2. Nucleotide sequences of pAC.1 and pAC.7

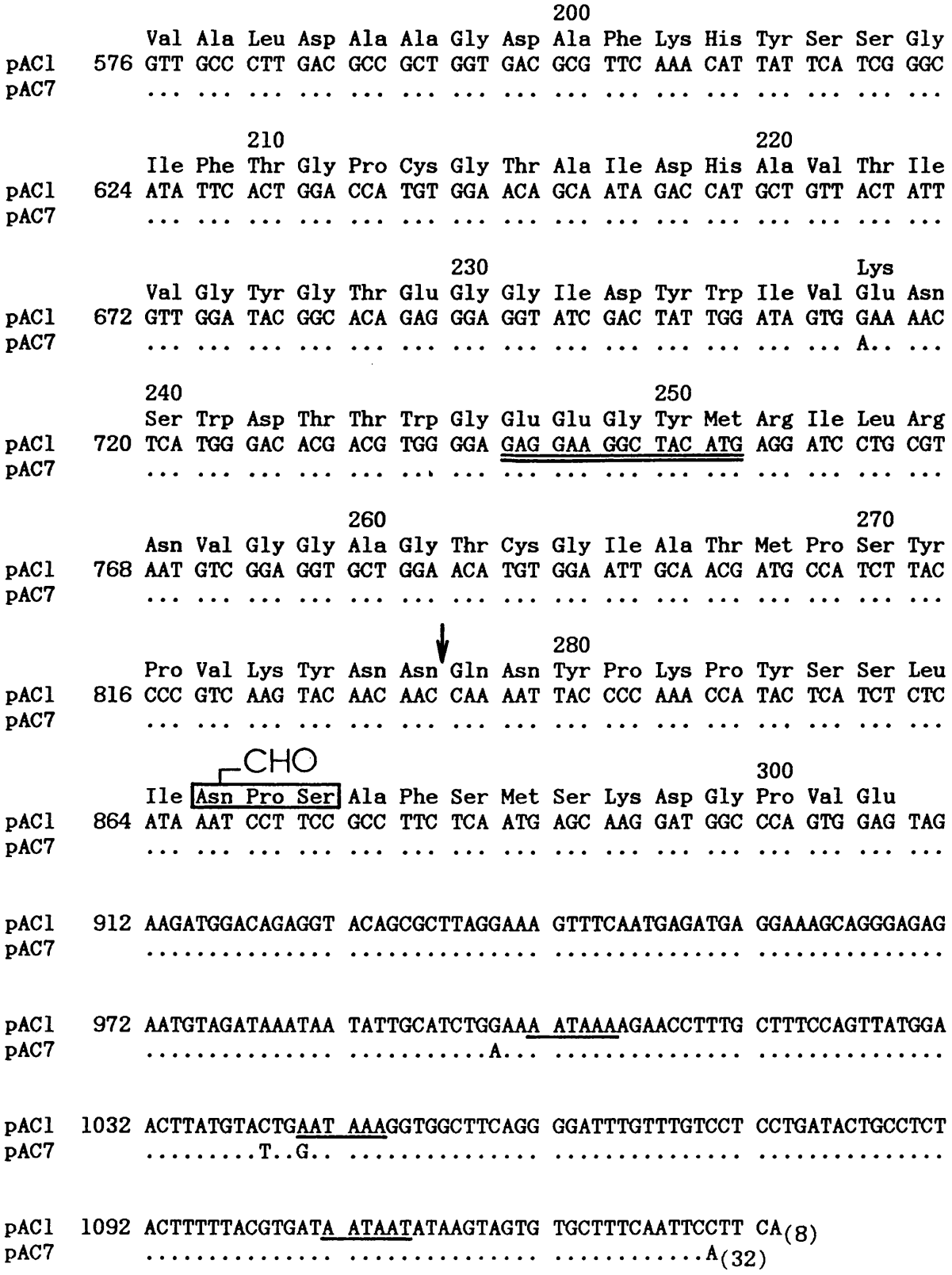
The complete nucleotide sequence of pAC.1 is shown, as well as its deduced amino acid sequence (above). Only differences of the nucleotide sequence of pAC.7 are shown, and similarities are indicated by dots. Differences in the deduced amino acid sequence of pAC.7 are shown above the amino acid sequence of pAC.1.

Polyadenylation signals are underlined.

The 15 nucleotide sequence corresponding to the synthetic oligonucleotide is double-underlined.

Potential glycosylation sites (CHO) are boxed.

```
                                                    10          ⌐CHO
            Leu Arg Phe Ile Asp Glu His Asn Ala Asp Thr Asn Arg Ser Tyr
pAC1     1  CT TTG AGG TTC ATT GAC GAG CAC AAT GCA GAC ACA AAC CGT AGT TAC

                          20                                          30
            Lys Val Gly Leu Asn Gln Phe Ala Asp Leu Thr Gly Glu Glu Phe Arg
pAC1    48 AAG GTG GGC TTG AAC CAG TTT GCT GAC CTG ACT GGT GAG GAG TTC CGG

                                        40          ⌐CHO
            Ser Thr Tyr Leu Gly Phe Thr Gly Gly Ser Asn Lys Thr Lys Val Ser
pAC1    96 TCC ACT TAC TTG GGA TTT ACA GGC GGT TCA AAT AAG ACG AAG GTG AGC

                    50                                  ↓       60
            Asn Arg Tyr Glu Pro Arg Val Ser Gln Val Leu Pro Ser Tyr Val Asp
pAC1   144 AAC CGG TAC GAG CCC CGA GTC AGC CAA GTA TTG CCG AGT TAT GTT GAT

                                70
            Trp Arg Ser Ala Gly Ala Val Val Asp Ile Lys Ser Gln Gly Glu Cys
pAC1   192 TGG AGG TCG GCT GGG GCC GTG GTT GAC ATC AAA TCC CAG GGT GAA TGT

            80                                          90
            Gly Gly Cys Trp Ala Phe Ser Ala Ile Ala Thr Val Glu Gly Ile Asn
pAC1   240 GGG GGT TGT TGG GCT TTT TCG GCC ATC GCC ACG GTG GAA GGG ATC AAC

                          100                                         110
            Lys Ile Val Thr Gly Val Leu Ile Ser Leu Ser Glu Gln Glu Leu Ile
pAC1   288 AAG ATA GTG ACC GGA GTC TTA ATT TCG CTG TCA GAA CAA GAG CTT ATA

                                        120
            Gly Cys Gly Gly Thr Gln Asn Thr Arg Gly Cys Asn Gly Gly Tyr Ile
pAC1   336 GGT TGC GGT GGG ACA CAA AAC ACC AGG GGC TGC AAT GGC GGT TAC ATA
pAC7                                  ... ... ... ... ... ... ... ... ...

                    130                                 140         Glu
            Thr Asp Gly Phe Gln Phe Ile Ile Asn Asn Gly Gly Ile Asn Thr Gly
pAC1   384 ACC GAC GGG TTT CAG TTC ATT ATC AAC AAC GGT GGG ATT AAC ACC GGG
pAC7       ... ... ... ... ... ... ..C ... ... ... ... ... ... ... ... .A.

                                150
            Glu Asn Tyr Pro Tyr Thr Ala Gln Asp Gly Glu Cys Asn Leu Asp Leu
pAC1   432 GAA AAT TAT CCC TAC ACG GCT CAA GAT GGT GAA TGC AAC TTG GAC TTA
pAC7       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

            160                                         170 Glu
            Gln Asn Glu Lys Tyr Val Thr Ile Asp Thr Tyr Gly Asn Val Pro Tyr
pAC1   480 CAA AAT GAA AAG TAT GTT ACA ATT GAT ACT TAT GGA AAT GTT CCA TAT
pAC7       ... ... ... ... ... ... ... ... ... ... ... .A. ... ... ... ...

                          180                                         190
            Asn Asn Glu Trp Ala Leu Gln Thr Ala Val Thr Tyr Gln Pro Val Ser
pAC1   528 AAC AAC GAG TGG GCA TTG CAA ACA GCA GTG ACA TAC CAA CCT GTG AGC
pAC7       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```


200
Val Ala Leu Asp Ala Ala Gly Asp Ala Phe Lys His Tyr Ser Ser Gly
pAC1 576 GTT GCC CTT GAC GCC GCT GGT GAC GCG TTC AAA CAT TAT TCA TCG GGC
pAC7 ... ...
210
220
Ile Phe Thr Gly Pro Cys Gly Thr Ala Ile Asp His Ala Val Thr Ile
pAC1 624 ATA TTC ACT GGA CCA TGT GGA ACA GCA ATA GAC CAT GCT GTT ACT ATT
pAC7 ... ...
230
Lys
Val Gly Tyr Gly Thr Glu Gly Gly Ile Asp Tyr Trp Ile Val Glu Asn
pAC1 672 GTT GGA TAC GGC ACA GAG GGA GGT ATC GAC TAT TGG ATA GTG GAA AAC
pAC7 ... ... A.. ...
240
250
Ser Trp Asp Thr Thr Trp Gly Glu Glu Gly Tyr Met Arg Ile Leu Arg
pAC1 720 TCA TGG GAC ACG ACG TGG GGA GAG GAA GGC TAC ATG AGG ATC CTG CGT
pAC7 ... ...
260
270
Asn Val Gly Gly Ala Gly Thr Cys Gly Ile Ala Thr Met Pro Ser Tyr
pAC1 768 AAT GTC GGA GGT GCT GGA ACA TGT GGA ATT GCA ACG ATG CCA TCT TAC
pAC7 ... ...
280
Pro Val Lys Tyr Asn Asn Gln Asn Tyr Pro Lys Pro Tyr Ser Ser Leu
pAC1 816 CCC GTC AAG TAC AAC AAC CAA AAT TAC CCC AAA CCA TAC TCA TCT CTC
pAC7 ... ...
CHO
300
Ile Asn Pro Ser Ala Phe Ser Met Ser Lys Asp Gly Pro Val Glu
pAC1 864 ATA AAT CCT TCC GCC TTC TCA ATG AGC AAG GAT GGC CCA GTG GAG TAG
pAC7 ... ...
pAC1 912 AAGATGGACAGAGGT ACAGCGCTTAGGAAA GTTTCAATGAGATGA GGAAAGCAGGGAGAG
pAC7 ....
pAC1 972 AATGTAGATAAATAA TATTGCATCTGGAAA ATAAAAGAACCTTTG CTTTCCAGTTATGGA
pAC7 ....A... ....
pAC1 1032 ACTTATGTACTGAAT AAAGGTGGCTTCAGG GGATTTGTTTGTCCT CCTGATACTGCCTCT
pAC7 .....T..G.. ....
pAC1 1092 ACTTTTTACGTGATA ATAATATAAGTAGTG TGCTTTCAATTCCTT CA(8)
pAC7 ....A(32)

(Lycett *et al.*, 1983), the most 3' of which may represent two overlapping signals, positioned 26 (pAC.1) or 22 bases (pAC.7) upstream from the polyadenylation site.

6.3.2. Deduced amino acid sequence: The deduced amino acid sequences of both pAC.1 and pAC.7, shown superimposed on the nucleotide sequence in Figure 6.2, differ by three residues. pAC.1 encodes all 220 amino acids of a protein 91% homologous to actinidin, with 20 amino acid substitutions compared to the published sequence (Figure 6.3). The pAC.7 encoded sequence corresponds to amino acid residues 62 to 220 of actinidin, and is slightly more similar to the published sequence as it shows all but three of the substitutions found in the pAC.1 encoded protein.

Both cDNA clones have coding potential for 25 amino acids beyond the mature carboxy terminus of actinidin. It has a molecular weight of 2,754 Kd and is characterised by a high content of polar residues as illustrated in the hydrophobicity profile (Figure 6.4), and an excess of both Pro and Ser residues (16% and 20%, respectively). The sequence Asn-Pro-Ser conforms to the canonical sequence Asn-X-Ser, indicating a potential glycosylation site.

pAC.1 encodes 57 residues of an amino-terminal extension of molecular weight 6,508 Kd, which is likely to be longer since it contains no Met residue. Thus pAC.1 does not represent a complete cDNA.

The amino-terminal extension, or propeptide, is very polar in nature as can be seen from the hydrophobicity profile (Figure 6.4). It contains 72% polar amino acids, many of which are charged, compared with 64% in the mature enzyme, and its content of positively charged amino acids (16%) is particularly striking since the mature enzyme contains only 5% positive residues. Two Asn residues of the propepide may be sites of glycosylation, as indicated in Figure 6.2.

The M_r of the whole polypeptide encoded by pAC.1 is 32,709 kDa.

Figure 6.3. Alignment of the deduced amino acid sequences of pAC.1 and pAC.7 with the published amino acid sequence of actinidin

a) published amino acid sequence (Carne and Moore, 1978; Baker, 1980)

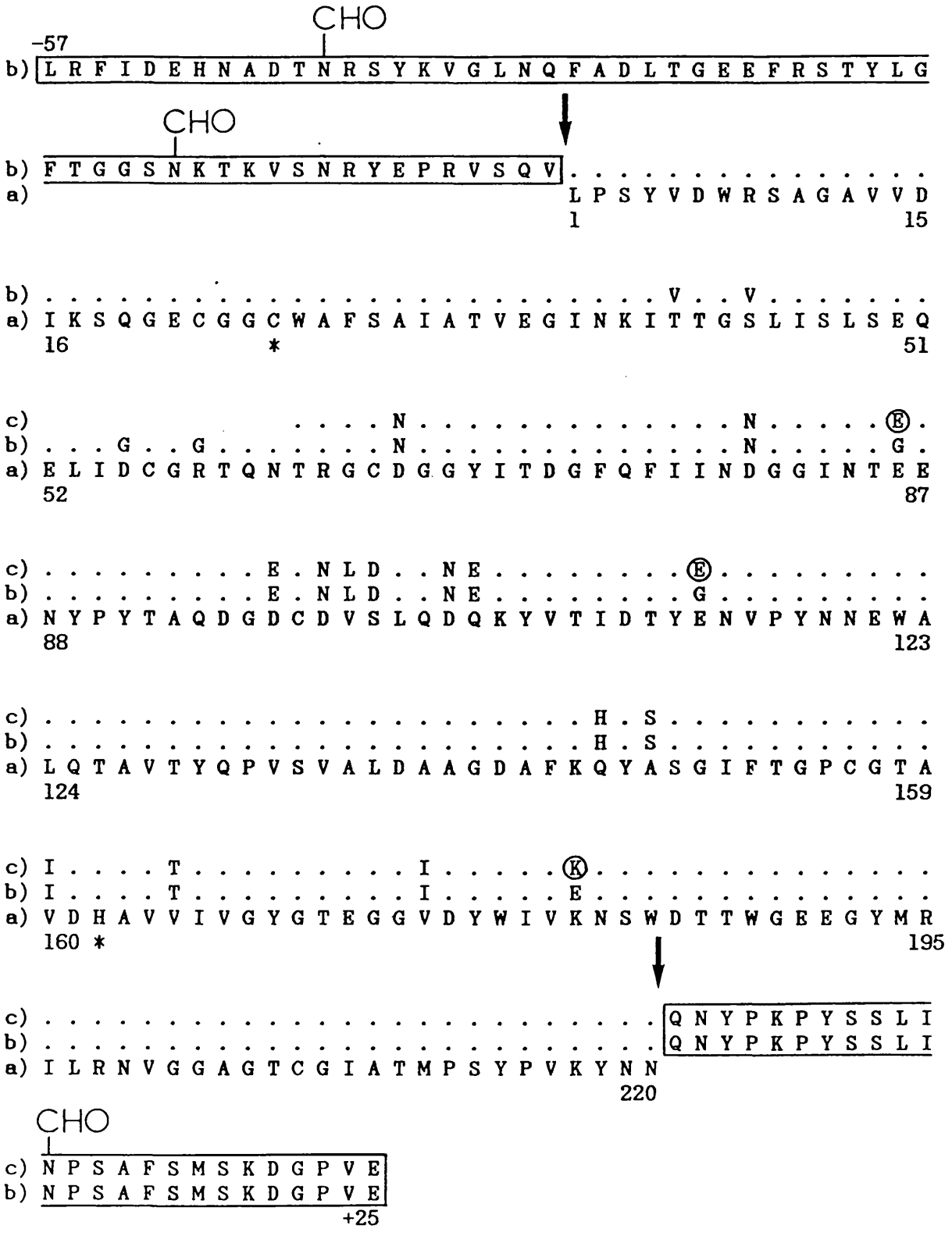
b) sequence deduced from pAC.1

c) sequence deduced from pAC.7

Residues not present in mature actinidin are boxed, and potential glycosylation sites indicated by "CHO".

Arrows indicate putative cleavage sites.

Differences between pAC.1 and pAC.7 are circled in the pAC.7-derived sequence, and correspond to the published sequence.


CHO
-57
b) L R F I D E H N A D T N R S Y K V G L N Q F A D L T G E E F R S T Y L G
CHO
b) F T G G S N K T K V S N R Y E P R V S Q V . . . . . .
a) L P S Y V D W R S A G A V V D
1 15
b) . . . . . V . . V . . . . . .
a) I K S Q G E C G G C W A F S A I A T V E G I N K I T T G S L I S L S E Q
16 * 51
c) . . . . N . . . . . N . . . . . E .
b) . . . G . . G . . . . . N . . . . . N . . . . . G .
a) E L I D C G R T Q N T R G C D G G Y I T D G F Q F I I N D G G I N T E E
52 87
c) . . . . . E . N L D . . N E . . . . . E . . . . . .
b) . . . . . E . N L D . . N E . . . . . G . . . . . .
a) N Y P Y T A Q D G D C D V S L Q D Q K Y V T I D T Y E N V P Y N N E W A
88 123
c) . . . . . H . S . . . . . .
b) . . . . . H . S . . . . . .
a) L Q T A V T Y Q P V S V A L D A A G D A F K Q Y A S G I F T G P C G T A
124 159
c) I . . . . T . . . . . I . . . . . K . . . . . .
b) I . . . . T . . . . . I . . . . . E . . . . . .
a) V D H A V V I V G Y G T E G G V D Y W I V K N S W D T T W G E E G Y M R
160 * 195
c) . . . . . Q N Y P K P Y S S L I
b) . . . . . Q N Y P K P Y S S L I
a) I L R N V G G A G T C G I A T M P S Y P V K Y N N
220
CHO
c) N P S A F S M S K D G P V E
b) N P S A F S M S K D G P V E
+25

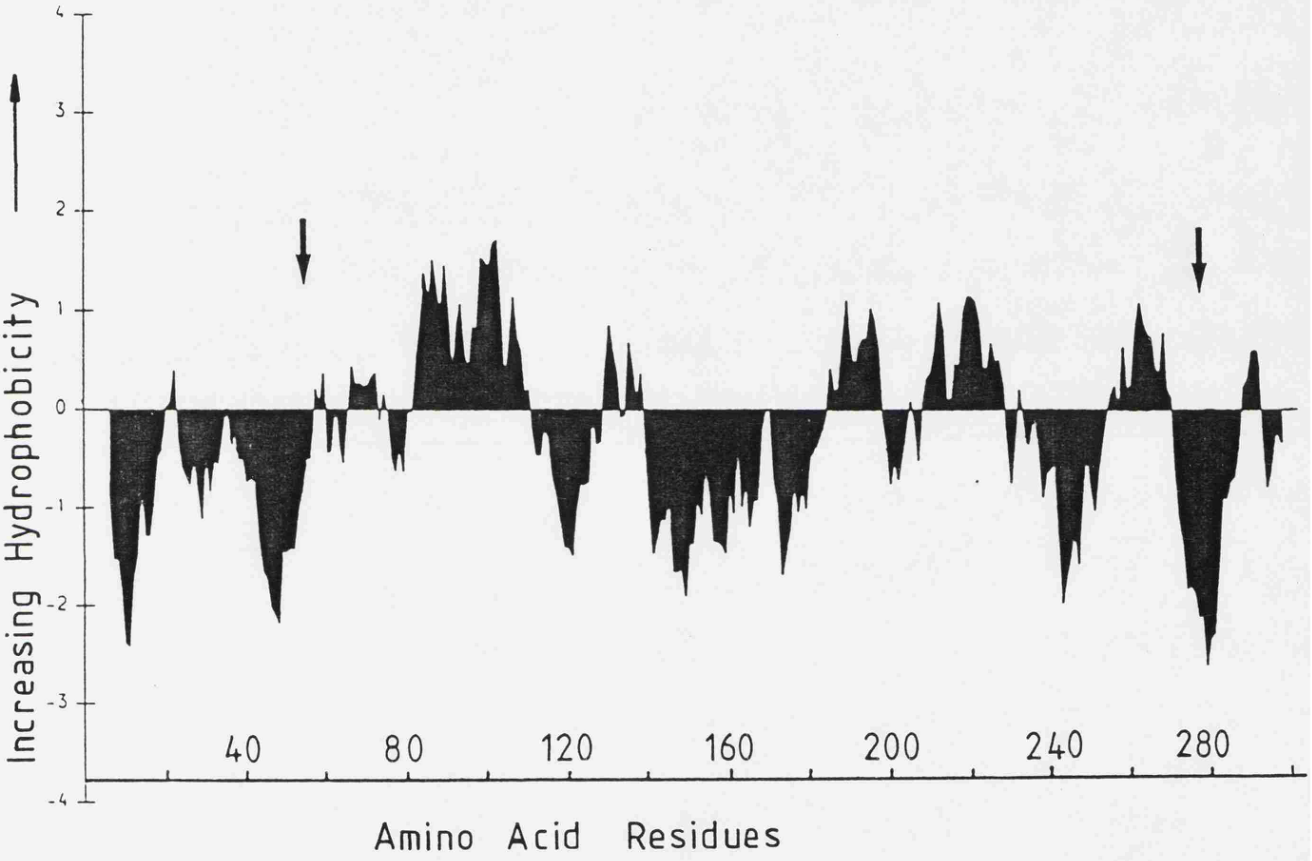


Figure 6.4. Hydrophobicity profile of the pAC.1-derived polypeptide

Relative hydrophobicities represent a moving window average of 11 amino acids.

Arrows indicate putative proteolytic cleavage sites.

6.3.3. Codon usage

An analysis of codon usage of clone pAC.1 is shown in Table 6.1. The nucleotide composition of the coding regions shows almost equal frequencies of all bases, but the non-coding region is 63% A/T-rich. Only three codons, CGC, AGA and CTA, were not used at all. A strong codon preference is not apparent, although some codons for Glu and Val, used preferentially in the region corresponding to the mature protein, are under-represented in the region coding for the amino-terminal extension. Unlike in a number of other plant genes (Lycett *et al.*, 1983), the third codon position in the region corresponding to mature actinidin has no preference for A/T residues. The region encoding the amino-terminal extension shows a clear preference (65%) for G and C residues in the third codon position, similar to the situation found in animal genes.
The dinucleotide CG and the trinucleotide CXG, which are sites of cytosine methylation in many plant genes (Gruenbaum *et al.*, 1981), are also under-represented in pAC.1 and pAC.7: CG occurs at 44% and CXG at 36% of the frequency expected from a random sequence.

6.4. Detailed comparison of the cDNA coding regions corresponding to mature actinidin with reference also to other cysteine proteinases.

The amino acid sequence of the pAC.1 encoded protein was aligned with the published sequence of actinidin and those of other related cysteine proteinases (Figure 6.5). The following analysis of amino acid substitutions in the cDNA encoded proteins compared to the published amino acid sequence of actinidin is based on the physico-chemical properties, size, hydrophobicity and charge characteristics of amino acids (Taylor, 1986) and applies to both clones unless otherwise stated.
The largest number of substitutions are concentrated in a region of lowest

Table 6.1. Codon Usage

		a	b	c			a	b	c			a	b	c
Ala	GCT	1	6	7	Gly	GGT	2	10	12	Pro	CCT	0	2	2
	GCC	0	6	6		GGC	2	6	8		CCC	1	3	4
	GCA	1	4	5		GGA	1	10	11		CCA	0	5	5
	GCG	0	1	1		GGG	0	7	7		CCG	0	1	1
Arg	CGT	1	1	2	His	CAT	0	2	2	Ser	TCT	0	2	2
	CGC	0	0	0		CAC	1	0	1		TCC	1	2	3
	CGA	1	0	1							TCA	1	5	6
	CGG	2	0	2	Ile	ATT	1	6	7		TCG	0	4	4
	AGA	0	0	0		ATC	0	6	6		AGT	1	1	2
	AGG	1	3	4		ATA	0	7	7		AGC	2	2	4
Asn	AAT	2	7	9	Leu	TTA	0	2	2	Thr	ACT	2	3	5
	AAC	3	11	14		TTG	3	3	6		ACC	0	4	4
						CTT	0	2	2		ACA	2	7	9
Asp	GAT	0	4	4		CTC	0	1	1		ACG	1	5	6
	GAC	3	8	11		CTA	0	0	0					
						CTG	1	2	3	Trp	TGG	0	6	6
Cys	TGT	0	4	4										
	TGC	0	3	3	Lys	AAA	0	3	3	Tyr	TAT	0	7	7
						AAG	3	4	7		TAC	3	9	12
Gln	CAA	1	7	8										
	CAG	1	2	3	Met	ATG	0	3	3	Val	GTT	0	7	7
											GTC	1	3	4
					Phe	TTT	2	2	4		GTA	1	0	1
						TTC	2	4	6		GTG	2	7	9

a) codon frequencies in prosegment of pAC.1.
b) codon frequencies in mature protein coding region plus carboxy-terminal extension of pAC.1.
c) codon frequencies of the whole coding region of pAC.1.

Figure 6.5. Alignment of cysteine proteinases:

Mature enzyme domains and carboxy-terminal extensions

PAC1 = amino acid sequence deduced from pAC.1

ACT = published amino acid sequence of actinidin

PAP = papain

SB = stem bromelain

ALEU = aleurain

AS-A = asclepain A

AS-B = asclepain B

CA-B = cathepsin B

CA-H = cathepsin H

CP-1 = cysteine proteinase 1 from slime mould

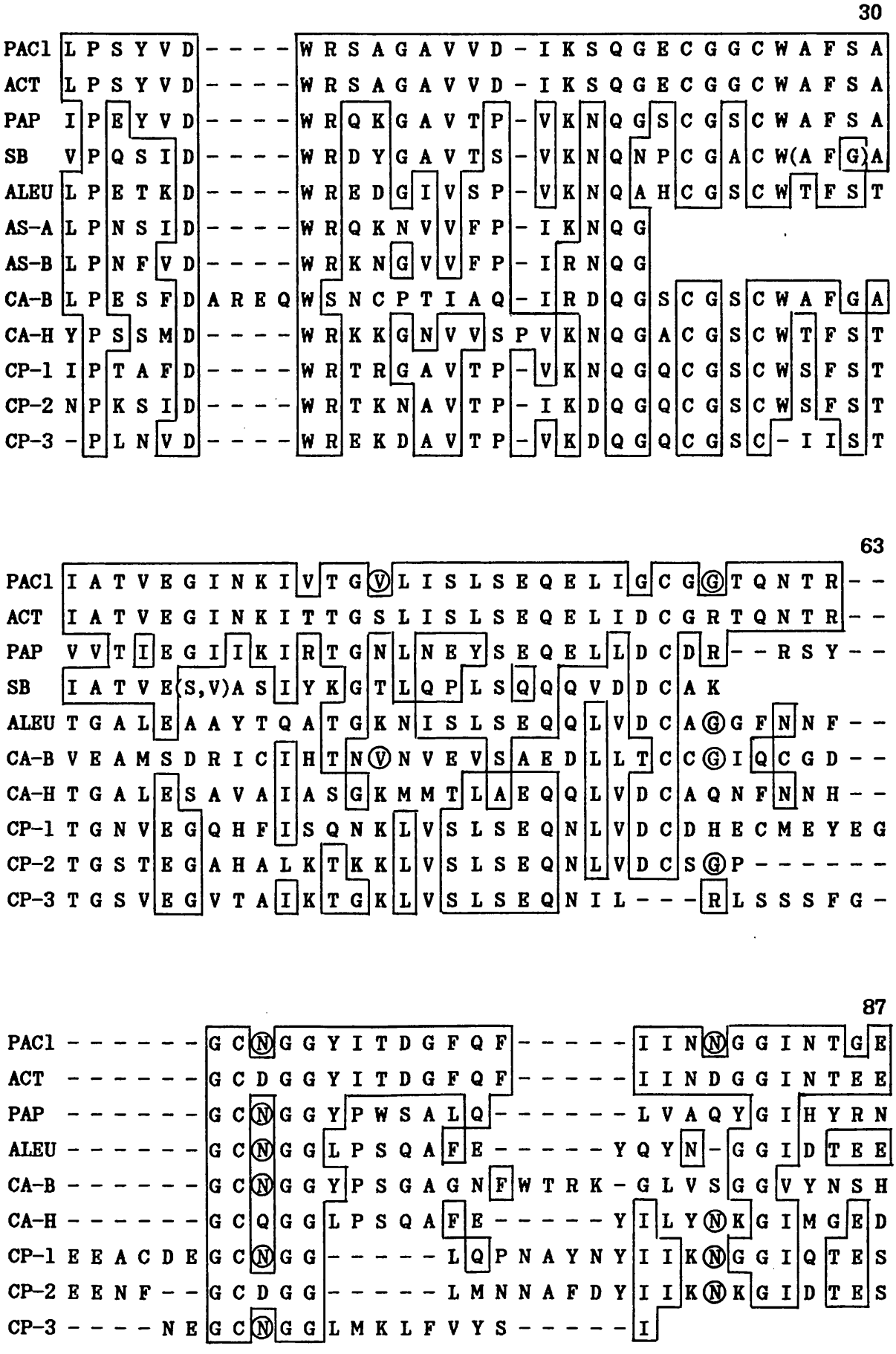
CP-2 = cysteine proteinase 2 " " "

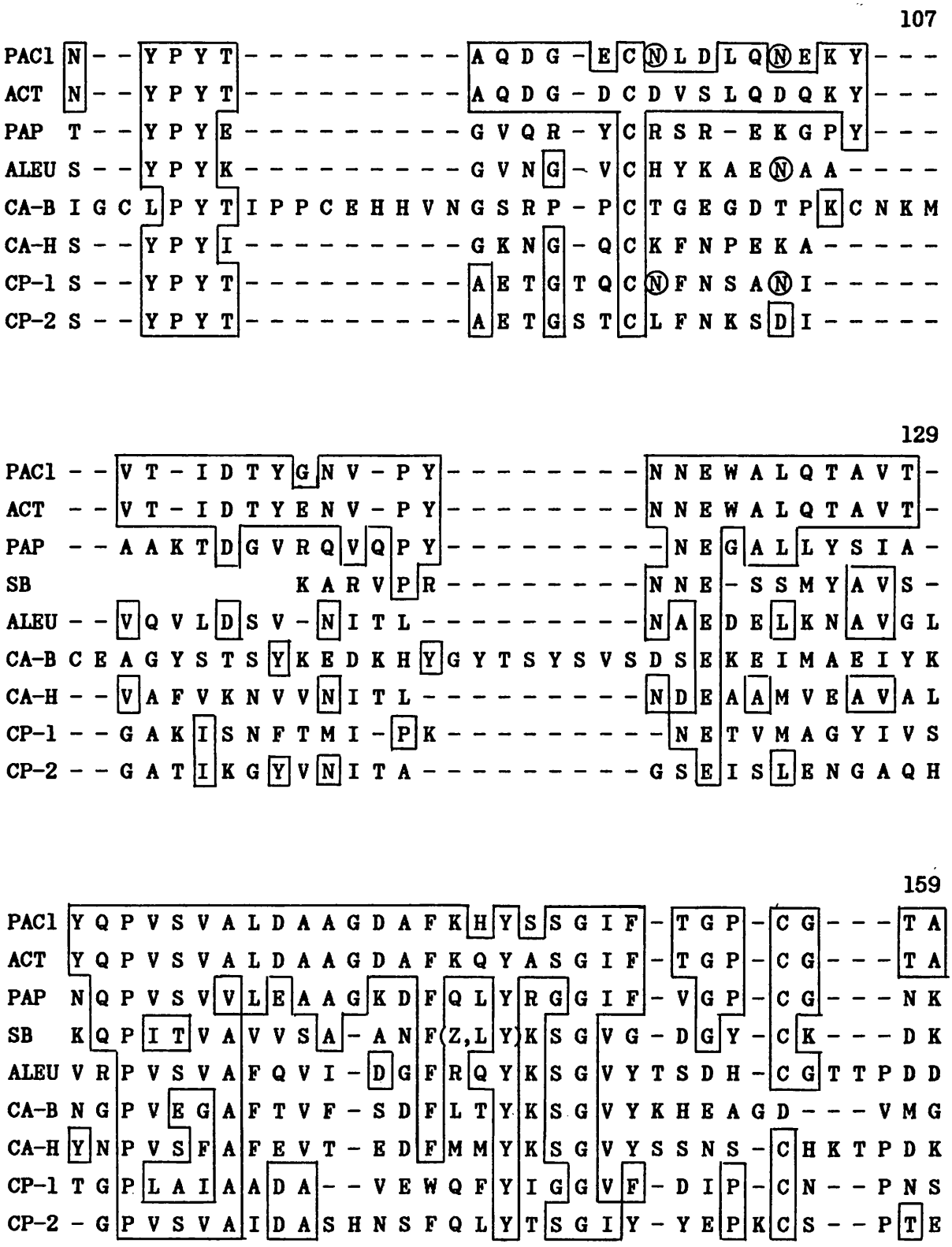
CP-3 = cysteine proteinase 3 " " "

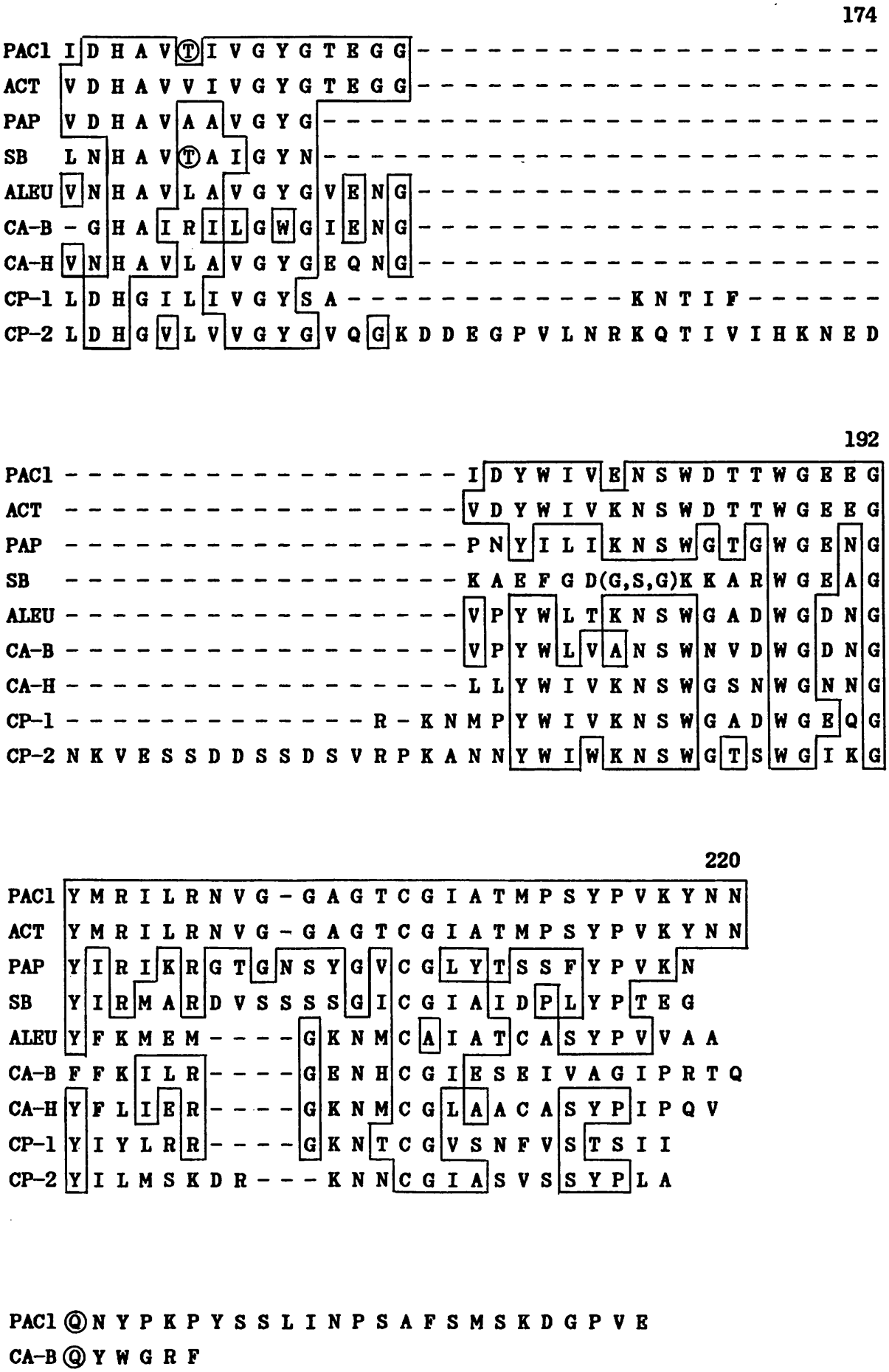
Numbering is based on the published amino acid sequence of actinidin.

Homologies with the published amino acid sequence are boxed.

Residues in PAC1 differing from the published sequence are circled if showing homologies with other sequences.


30
PACl L P S Y V D - - - - W R S A G A V V D - I K S Q G E C G G C W A F S A
ACT L P S Y V D - - - - W R S A G A V V D - I K S Q G E C G G C W A F S A
PAP I P E Y V D - - - - W R Q K G A V T P - V K N Q G S C G S C W A F S A
SB V P Q S I D - - - - W R D Y G A V T S - V K N Q N P C G A C W(A F G)A
ALEU L P E T K D - - - - W R E D G I V S P - V K N Q A H C G S C W T F S T
AS-A L P N S I D - - - - W R Q K N V V F P - I K N Q G
AS-B L P N F V D - - - - W R K N G V V F P - I R N Q G
CA-B L P E S F D A R E Q W S N C P T I A Q - I R D Q G S C G S C W A F G A
CA-H Y P S S M D - - - - W R K K G N V V S P V K N Q G A C G S C W T F S T
CP-1 I P T A F D - - - - W R T R G A V T P - V K N Q G Q C G S C W S F S T
CP-2 N P K S I D - - - - W R T K N A V T P - I K D Q G Q C G S C W S F S T
CP-3 - P L N V D - - - - W R E K D A V T P - V K D Q G Q C G S C - I I S T
63
PACl I A T V E G I N K I V T G Ⓥ L I S L S E Q E L I G C G Ⓖ T Q N T R - -
ACT I A T V E G I N K I T T G S L I S L S E Q E L I D C G R T Q N T R - -
PAP V V T I E G I I K I R T G N L N E Y S E Q E L L D C D R - - R S Y - -
SB I A T V E(S,V)A S I Y K G T L Q P L S Q Q Q V D D C A K
ALEU T G A L E A A Y T Q A T G K N I S L S E Q Q L V D C A Ⓖ G F N N F - -
CA-B V E A M S D R I C I H T N Ⓥ N V E V S A E D L L T C C Ⓖ I Q C G D - -
CA-H T G A L E S A V A I A S G K M M T L A E Q Q L V D C A Q N F N N H - -
CP-1 T G N V E G Q H F I S Q N K L V S L S E Q N L V D C D H E C M E Y E G
CP-2 T G S T E G A H A L K T K K L V S L S E Q N L V D C S Ⓖ P - - - - - -
CP-3 T G S V E G V T A I K T G K L V S L S E Q N I L - - - R L S S S F G -
87
PACl - - - - - - G C Ⓝ G G Y I T D G F Q F - - - - - I I N Ⓝ G G I N T G E
ACT - - - - - - G C D G G Y I T D G F Q F - - - - - I I N D G G I N T E E
PAP - - - - - - G C Ⓝ G G Y P W S A L Q - - - - - - L V A Q Y G I H Y R N
ALEU - - - - - - G C Ⓝ G G L P S Q A F E - - - - - Y Q Y N - G G I D T E E
CA-B - - - - - - G C Ⓝ G G Y P S G A G N F W T R K - G L V S G G V Y N S H
CA-H - - - - - - G C Q G G L P S Q A F E - - - - - Y I L Y Ⓝ K G I M G E D
CP-1 E E A C D E G C Ⓝ G G - - - - - L Q P N A Y N Y I I K Ⓝ G G I Q T E S
CP-2 E E N F - - G C D G G - - - - - L M N N A F D Y I I K Ⓝ K G I D T E S
CP-3 - - - - N E G C Ⓝ G G L M K L F V Y S - - - - - I


107
PAC1 N - - Y P Y T - - - - - - - - - - A Q D G - E C N L D L Q N E K Y - - -
ACT N - - Y P Y T - - - - - - - - - - A Q D G - D C D V S L Q D Q K Y - - -
PAP T - - Y P Y E - - - - - - - - - - G V Q R - Y C R S R - E K G P Y - - -
ALEU S - - Y P Y K - - - - - - - - - - G V N G - V C H Y K A E N A A - - - -
CA-B I G C L P Y T I P P C E H H V N G S R P - P C T G E G D T P K C N K M
CA-H S - - Y P Y I - - - - - - - - - - G K N G - Q C K F N P E K A - - - - -
CP-1 S - - Y P Y T - - - - - - - - - - A E T G T Q C N F N S A N I - - - - -
CP-2 S - - Y P Y T - - - - - - - - - - A E T G S T C L F N K S D I - - - - -
129
PAC1 - - V T - I D T Y G N V - P Y - - - - - - - - - N N E W A L Q T A V T -
ACT - - V T - I D T Y E N V - P Y - - - - - - - - - N N E W A L Q T A V T -
PAP - - A A K T D G V R Q V Q P Y - - - - - - - - - - N E G A L L Y S I A -
SB K A R V P R - - - - - - - - - N N E - S S M Y A V S -
ALEU - - V Q V L D S V - N I T L - - - - - - - - - - N A E D E L K N A V G L
CA-B C E A G Y S T S Y K E D K H Y G Y T S Y S V S D S E K E I M A E I Y K
CA-H - - V A F V K N V V N I T L - - - - - - - - - - N D E A A M V E A V A L
CP-1 - - G A K I S N F T M I - P K - - - - - - - - - - N E T V M A G Y I V S
CP-2 - - G A T I K G Y V N I T A - - - - - - - - - - G S E I S L E N G A Q H
159
PAC1 Y Q P V S V A L D A A G D A F K H Y S S G I F - T G P - C G - - - T A
ACT Y Q P V S V A L D A A G D A F K Q Y A S G I F - T G P - C G - - - T A
PAP N Q P V S V V L E A A G K D F Q L Y R G G I F - V G P - C G - - - N K
SB K Q P I T V A V V S A - A N F (Z,L Y) K S G V G - D G Y - C K - - - D K
ALEU V R P V S V A F Q V I - D G F R Q Y K S G V Y T S D H - C G T T P D D
CA-B N G P V E G A F T V F - S D F L T Y K S G V Y K H E A G D - - - V M G
CA-H Y N P V S F A F E V T - E D F M M Y K S G V Y S S N S - C H K T P D K
CP-1 T G P L A I A A D A - - V E W Q F Y I G G V F - D I P - C N - - P N S
CP-2 - G P V S V A I D A S H N S F Q L Y T S G I Y - Y E P K C S - - P T E


174
PACl I D H A V T I V G Y G T E G G -
ACT V D H A V V I V G Y G T E G G -
PAP V D H A V A A V G Y G -
SB L N H A V T A I G Y N -
ALEU V N H A V L A V G Y G V E N G -
CA-B - G H A I R I L G W G I E N G -
CA-H V N H A V L A V G Y G E Q N G -
CP-1 L D H G I L I V G Y S A - - - - - - - - - - - - - - - K N T I F - - - - - -
CP-2 L D H G V L V V G Y G V Q G K D D E G P V L N R K Q T I V I H K N E D
192
PACl - I D Y W I V E N S W D T T W G E E G
ACT - V D Y W I V K N S W D T T W G E E G
PAP - P N Y I L I K N S W G T G W G E N G
SB - K A E F G D(G,S,G)K K A R W G E A G
ALEU - V P Y W L T K N S W G A D W G D N G
CA-B - V P Y W L V A N S W N V D W G D N G
CA-H - L L Y W I V K N S W G S N W G N N G
CP-1 - - - - - - - - - - - - - - - - R - K N M P Y W I V K N S W G A D W G E Q G
CP-2 N K V E S S D D S S D S V R P K A N N Y W I W K N S W G T S W G I K G
220
PACl Y M R I L R N V G - G A G T C G I A T M P S Y P V K Y N N
ACT Y M R I L R N V G - G A G T C G I A T M P S Y P V K Y N N
PAP Y I R I K R G T G N S Y G V C G L Y T S S F Y P V K N
SB Y I R M A R D V S S S S G I C G I A I D P L Y P T E G
ALEU Y F K M E M - - - - G K N M C A I A T C A S Y P V V A A
CA-B F F K I L R - - - - G E N H C G I E S E I V A G I P R T Q
CA-H Y F L I E R - - - - G K N M C G L A A C A S Y P I P Q V
CP-1 Y I Y L R R - - - - G K N T C G V S N F V S T S I I
CP-2 Y I L M S K D R - - - K N N C G I A S V S S Y P L A
PACl Q N Y P K P Y S S L I N P S A F S M S K D G P V E
CA-B Q Y W G R F

homology between all the cysteine proteinases, residues 82 to 116 (Kamphius *et al.*, 1985; Baker, 1980). Substitutions in the more conserved regions are mostly of a conservative nature.

The least conserved region forms an extended surface loop and contains many negatively charged side chains in actinidin. Within this region of the pAC.1 and pAC.7 encoded proteins two Asp residues (position 97 and 104) are replaced by Asn, resulting in the loss of two negative charges. A further two negative charges in this region are lost in the pAC.1, but not the pAC.7 encoded protein, by Asp-86 → Gly and Asp-114 → Gly replacements. However, two of the charge losses are compensated by gains of negatively charged side chains through Ser_{100} → Asp and Gln_{105} → Glu replacements. Two replacements, Asp-97 → Glu and Val-100 → Leu, are conservative. Overall, this surface region, whose charge characteristics are in part responsible for the isoelectric points of the cysteine proteinases, has a net charge identical to that of actinidin in the case of the pAC.7 encoded protein, but that encoded by pAC.1 has two less negative side chains.

Within the less variable regions, two of the most conservative replacements are Val-160 → Ile and Val-175 → Ile, both in the vicinity of the active site (Baker, 1980) and showing conservation of hydrophobic residues between the cysteine proteinases. The moderately conservative substitutions Thr-41 → Val, Ser-44 → Val, Ala-148 → Ser, and Val-165 → Thr, are all in positions with considerable variability between enzymes but within more conserved regions. Two Asp → Asn substitutions, at positions 66 and 80, are within fairly well conserved areas, but the Asp found in the published amino acid sequence appears to be the exception since most other cysteine proteinases have Asn in these positions.

The replacement of Gln by His at position 146 is non-conservative, but it is interesting that the residue is associated with a cluster of aromatic

side chains (Phe-144, Tyr-147, Phe-152, Trp-184, and Trp-188), adding an additional ring structure.

Two non-conservative substitutions are close together in domain I, Asp-55 → Gly and Arg-58 → Gly, with the loss of a positive as well as a negative charge. Residue 58 in actinidin is adjacent to a deletion in papain, and gly is also found in three other cysteine proteinases at this position, therefore the overall structure of the protein may not be affected. However, the replacement of Asp-55 by Gly is difficult to reconcile with the proposed hydrogen-bonding of this residue to Gln-51 and Tyr-89 (Baker, 1980). The importance of the Asp side chain is underlined by its strong conservation in other cysteine proteinases.

The most extreme amino acid change is the replacement of Lys by Glu at position 181 in the pAC.1 encoded protein, resulting in a reversion of the charge of this side chain. This residue is at the interface between the two domains and forms part of a cluster of charged side chains contributed by both domains, which neutralise each other via a network of water molecules. The replacement of a positive by a negative charge may not be compatible with this proposed interaction.

6.5. Conclusions

The two cDNA clones characterised in detail have revealed coding information for two very similar proteins closely related to actinidin, with greater than 90% homology to the published amino acid sequence of actinidin.

Both clones have coding potential for polypeptide sequences additional to the mature protein, and confirm results from *in vitro* translation of fruit mRNA and immunoprecipitation, which suggested that actinidin is synthesised in the form of a larger precursor. However, only a portion of the amino-terminal extension appears to be encoded by clone pAC.1, since

no methionine codon is found in this region.

Both extensions are characterised by an excess of polar over non-polar amino acids. As suggested by the hydrophobicity profile, the putative cleavage sites are within polar regions, and are therefore probably on the surface of the protein, accessible to the proteolytic enzymes responsible for maturation of the enzyme.

It is surprising that the two clones show only very few differences at the nucleotide level, particularly as three of four nucleotide substitutions in the coding region result in non-conservative amino acid substitutions. The possibility that the observed differences may be the result of a cloning artefact is discussed below.

7. Actinidin gene expression and estimation of gene number

7.1. Introduction

Actinidin was above shown to be present mainly in the fruit of the plant, which also contained high levels of mRNA coding for a protein immunoprecipitable by actinidin antibodies. The low level of actinidin in young leaves was paralleled by low levels of mRNA coding for actinidin, since no immunoprecipitate was obtained from *in vitro* translation products with leaf RNA templates. It can be concluded from these results that actinidin is a fruit-specific protein, whose synthesis appears to be regulated at the transcriptional level, as has been shown to be the case for the majority of proteins analysed (Darnell, 1982). Therefore it was decided to study actinidin gene expression by estimating levels of actinidin-encoding mRNA in fruit at different stages during developments. By determining at which stage during fruit growth the message is first synthesised, it may be possible to distinguish between a function of the enzyme during fruit development or ripening.

Northern blot and dot blot experiments of RNA from fruit at different developmental stages allowed a comparative analysis of changes in levels of actinidin RNA. The size of the mRNA transcript was determined by comparison of its mobility on agarose gel electrophoresis with that of ribosomal RNA.

A Southern blot of Chinese gooseberry DNA was hybridised with the cDNA insert of clone pAC.1, providing an estimate of the number of genes coding for actinidin.

7.2. Results

7.2.1. Fruit growth

Chinese gooseberry fruits were harvested 7, 24, 42, 64, 82 and 104 days after anthesis. Growth was monitored by weight (Table A.4, Appendix), and the results are shown diagrammatically in Figure 7.1. The fruits reached half of their final weight after approximately 31 days from the time of anthesis, although it is possible that the last harvest on day 104 was before the maximum weight had been reached. However, climatic conditions would not have allowed further growth on the plant. After 104 days of development the fruits were still unripe and very firm. Storage at 10°C for up to 21 days resulted only in a slight degree of softening, but not ripening.

7.2.2. Northern blot analysis

A Northern blot, hybridised with ^{32}P-labelled cDNA insert of pAC.1, of total RNA from fruit at different developmental stages is shown in Figure 7.2.A. For comparison, an eight-fold amount of leaf RNA was electrophoresed in parallel (lane 1), as this was expected to contain a considerably lower level of hybridising transcript. All tracks show one main band hybridising with the cDNA probe. Its size was estimated by comparison with ribosomal RNA, visualised by ethidium bromide staining in a parallel lane (not shown), and was found to be 1.4kb. However, upon prolonged exposure of the autoradiograph, a further band at 3.7 Kb was seen in most lanes (Figure 7.2.B). This was most prominent in the lanes representing early developmental stages, but also in leaf RNA.
The amount of hybridising RNA was hardly detectable in 7 day-old fruit. Between 7 and 24 days after pollination only a slight increase occurred, but a significant amount was detectable 42 days after pollination, when

Figure 7.1. Fruit growth and RNA dot blot analysis

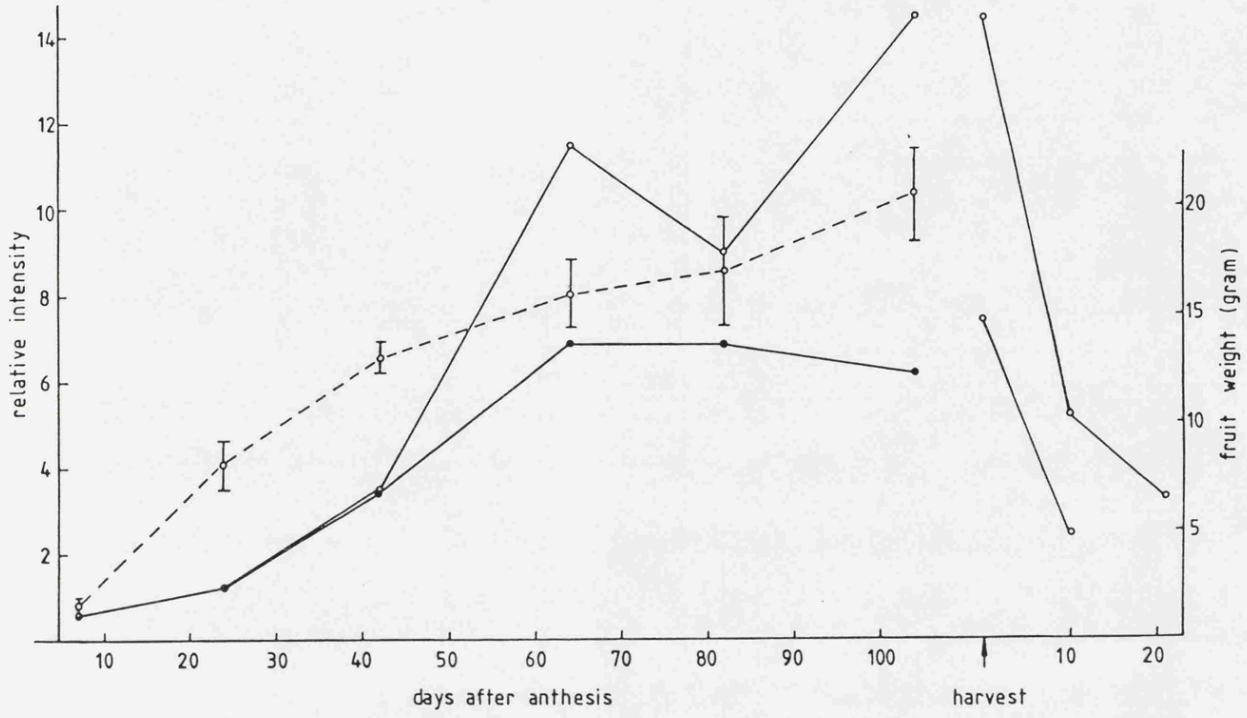
A) broken line: Plot of fruit weight
full circles: intensity profile of 0.3μg RNA
empty circles: " " " 0.9 " "
(intensity values plotted from Table A.5, Appendix)

B) Dot blots used for the determination of the intensity profile plotted above, by scanning with an Ultro Scan Laser Densitometer (LKB 2202)

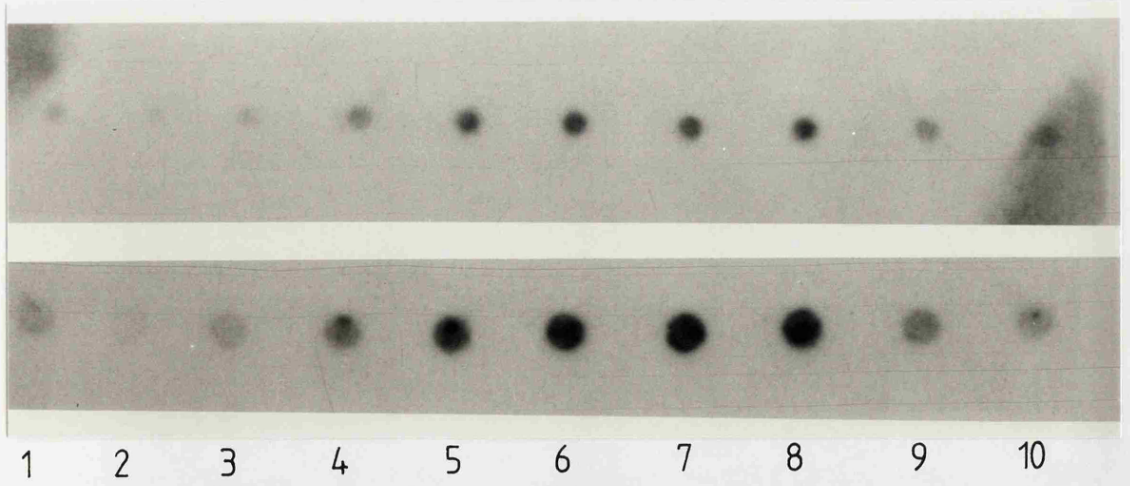
First row: 1 2.45 μg leaf RNA
2 0.3 μg RNA from 7 day-old fruit
3 " " " " 24 " "
4 " " " " 42 " "
5 " " " " 64 " "
6 " " " " 82 " "
7 " " " " 104 " "
8 " " " " 104 " "
9 " " " " fruit stored for 10 days
10 " " " " " " " 21 "

Second row: as first row, but using three-fold amounts of RNA

A


relative intensity
fruit weight (gram)
days after anthesis
harvest

B


1
2
3
4
5
6
7
8
9
10

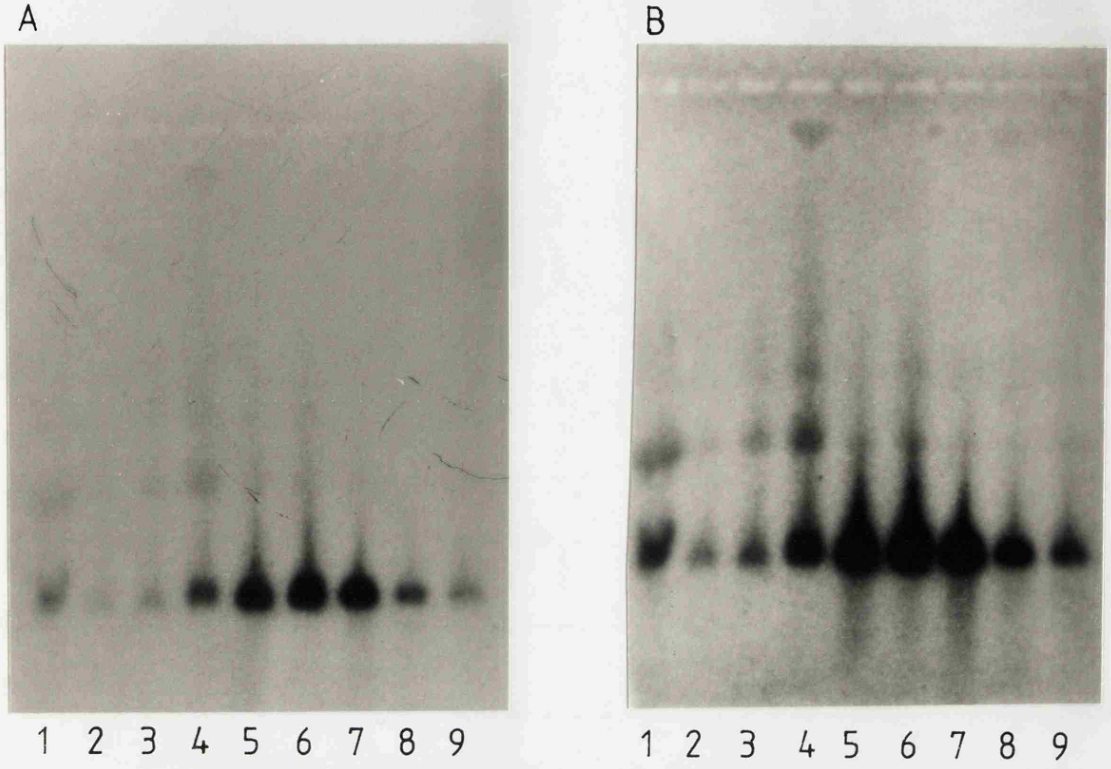


Figure 7.2. Northern Blot

RNA was electrophoresed on a formaldehyde gel, blotted onto nitrocellulose filter, and hybridised to ^{32}P-labelled cDNA insert of pAC.1 as described in Methods.

1 16.7μg leaf RNA
2 2.0μg RNA from 7 day-old fruit
3 " " " 24 " "
4 " " " 42 " "
5 " " " 64 " "
6 2.0μg RNA from 82 day-old fruit
7 " " " 104 " "
8 " " " fruit stored 10 days
9 " " " " " 21 "

A autoradiograph exposed for 1 day
B " " " 4 days

the fruit had reached more than half of its final weight. After 64 days of development, at about 4/5 of the final fruit weight, a near maximal level of hybridisable mRNA was present. After storage at 10°C for 10 days, a large fraction of hybridisable RNA had been lost, and not much remained after 21 days of storage.

Although an eight-fold concentration of leaf RNA showed a stronger signal than 7 day-old fruit RNA, an equivalent amount of leaf RNA would have resulted in a signal weaker than that of the young fruit, as demonstrated below.

Figure 7.1 shows the results of an attempt at quantifying hybridisable mRNA at the different developmental stages: autoradiographs of two series of dot blots (Figure 7.1.B), using a three-fold difference in total RNA, were scanned using a densitometer, and the areas under the peaks were integrated to give relative intensity values, which were plotted in Figure 7.1.A. It is clear from the results that the intensity values were not a linear function of the radioactive label hybridising to the dots, since a three-fold increase in mRNA did not result in three-fold higher intensity signal (Figure 7.1.A). Therefore these results can only be taken as an approximate guide to the difference between hybridisable RNA at the different stages. At the final harvest the amount of actinidin encoding RNA was found by this method to be 24-fold higher than that of the 7 day-old fruit. The non-linearity of this analysis, however, suggests that this may be an underestimate of the quantitative differences.

The intensity values obtained for an eight-fold amount of leaf RNA are indicated in Table A.5 (Appendix). After correction by a factor of 8, it was found that leaf RNA contained approximately one third of the amount of hybridisable RNA compared with RNA from 7 day-old fruit.

Probably a more accurate comparative analysis would have been achieved by the application of dot blot hybridisation to a serial dilution of RNA from

each fruit sample, in combination with scintillation counting of the dot blots.

7.2.3. Southern blot

Southern blots of DNA from two sources of plant material (the unidentified variety used for the quantification of actinidin in different parts of the plant, and the variety "Exbury" used for cDNA cloning) are shown in Figure 7.4. The DNA prepared from "Exbury" fruit gave rise to very faint hybridisation bands. Although similar amounts of DNA were used for both varieties, it is possible that the DNA isolated from fruit was contaminated by carbohydrate, which may have prevented efficient transfer of DNA to the nitrocellulose filter. However, those bands visible in the tracks containing the "Exbury" DNA corresponded exactly to the most strongly hybridising bands from the other variety, and therefore it is possible that the two varieties had the same hybridisation pattern. Restriction of DNA with BamHI resulted in only a few poorly resolved hybridising bands of high molecular weight, indicating that sites for this restriction enzyme are rare in the vicinity of the actinidin genes, even though the coding region contains one such site. Both EcoRI- and HindIII-digested DNA resulted in a number of hybridising fragments, approximately 14 in each case. The size of fragments is listed in Table 7.1., showing underlined those fragments which are visible in the "Exbury" DNA.

Considering that the actinidin genes may contain intervening sequences with restriction sites for EcoRI and HindIII, it is likely that only a proportion of the larger hybridising fragments represent whole genes.

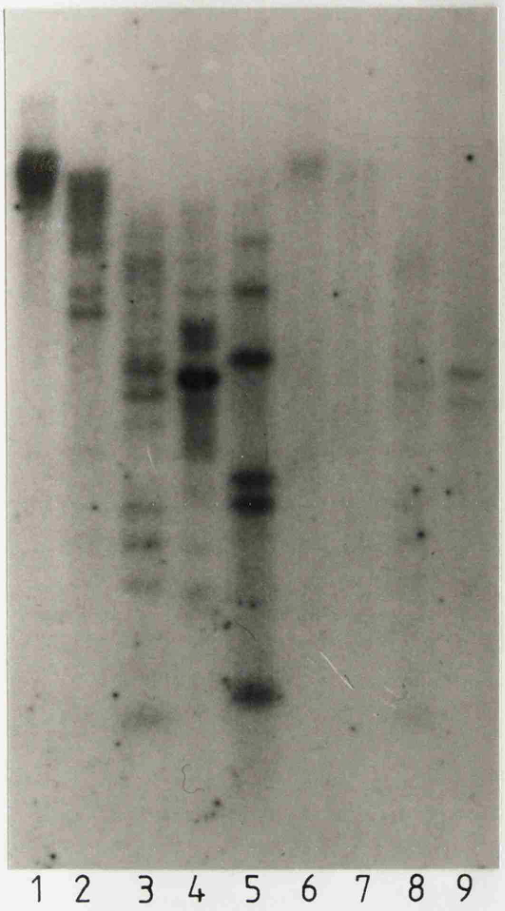


Figure 7.3. Southern blot

10 μg DNA was electrophoresed on agarose gel, blotted onto nitrocellulose filter, and hybridised to ^{32}P-labelled cDNA insert of pAC.1 at 65°C in 6XSSC, and washed under the same conditions.

1 undigested DNA
2 BamHI-digested DNA
3 EcoRI " "
4 HindIII " "
5 ^{32}P-labelled λHindIII markers
6 undigested DNA
7 BamHI-digested DNA
8 EcoRI " "
9 HindIII " "

Lanes 1-4: DNA from unidentified variety.
Lanes 6-9: " " variety Exbury used for cDNA cloning.

Table 7.1. Size distribution of restriction fragments hybridising on Southern blot

	Restriction enzymes used:	
	EcoRI	HindIII
size of fragments (kb)	9.9	13.0
	8.4	8.0
	7.4	6.8
	6.5	5.6
	6.0	5.4
	5.0	4.9
	4.5	4.1
	4.2	3.5
	3.7	3.0
	3.2	2.8
	1.9	2.6
	1.6	2.1
	1.2	1.5
	0.5	1.2

All fragments listed were identified in the unidentified variety obtained locally, and underlined fragments were visible in the tracks containing DNA from the variety "Exbury".

7.3. Conclusions

7.3.1. Transcript size

Size estimation by Northern blotting of the actinidin transcript has shown that the mRNA is approximately 1.4 kb long. This is about 250 bp in excess of the largest cDNA clone identified. Since pAC.1 apparently has a complete 3'untranslated region, this additional nucleotide sequence must be at the 5'end of the mRNA. As a short 5' leader of about 50 nucleotides may precede the coding region, the additional sequence has potential coding capacity for an extra 60 to 70 amino acids. Together with the 57 residues already identified by cDNA cloning, the size of the precursor peptide can be calculated at 117 to 127 amino acids. This is very similar indeed to the size of aminoterminal extensions of the other cysteine proteinases.

Prolonged exposure of the Northern blot on autoradiography revealed the presence of additional hybridising bands of about 3.7 kb. It is possible that these represent the unprocessed precursors of actinidin mRNA containing intervening sequences. Similar higher molecular weight bands were observed in Northern blots of papain mRNA and were also thought to represent unprocessed precursor (Mckee *et al.*, 1986).

7.3.2. Developmental changes in transcript levels

It was shown that actinidin-specific transcripts were present at low levels in very young fruit, and 7 day-old fruit contained approximately three times as much actinidin-specific message as young leaf tissue. Although it is not known what proportion of the total leaf mRNA population this level represents, it is likely to be a very small percentage, since leaf *in vitro* translation products could not be identified by immunoprecipitation.

A rapid increase in the amount of actinidin transcript was observed during early fruit development. It is interesting that the onset of this increased transcriptional activity coincides with a transition in fruit development from cell division to cell enlargement (see Introduction), and is preceded by a very gradual increase in the level of actinidin-specific transcripts. The results suggest that actinidin may have a function in early fruit development, rather than ripening, since gene expression is not specifically induced during ripening. However, the fact that mRNA accumulates early during development does not exclude the possibility that the protein is required at a later stage, particularly as it was shown above that the enzyme is synthesised as a larger precursor. Possible implications of the results with regard to the physiological function of actinidin are discussed below.

7.3.3. Southern blot

It was shown by Southern blotting that actinidin is encoded by a gene family. The number of genes cannot be estimated accurately, as some of the bands may represent fragments of the actinidin gene. It should be possible to obtain a more accurate estimate by independent hybridisations of the blot with several different small restriction fragments of the cDNA insert. If *A. chinensis* is indeed hexaploid, as its chromosome number suggests (Zhang and Beuzenberg, 1983), the discovery of an actinidin gene family is not surprising, and should contain at least six members.
The very strongly hybridising band at 4.1 kb is probably the result of several restriction fragments of identical size migrating together, and may reflect a recent gene duplication event.

8. Discussion

The objective of this study has been to examine, at the molecular level, some of the events associated with the presence of large quantities of actinidin in Chinese gooseberry fruit. Actinidin-related cDNAs were isolated to provide information on the structure of the genes and for use in the study of actinidin gene expression.

8.1. The approach

CDNA cloning was chosen as the method to obtain actinidin-encoding DNA sequences, because it seemed likely that a fruit-specific cDNA library would be enriched in actinidin-related sequences. The considerable effort involved in the adaptation of existing RNA extraction methods for the purification of RNA from fruit tissue was necessary as it seems unlikely that actinidin sequences could have been identified in a leaf cDNA or genomic library without a differential hybridisation probe. In addition, although it was shown that actinidin transcripts are present at low levels in leaf tissue, the frequency of actinidin clones in a leaf-specific cDNA library would have been very low indeed, requiring the screening of large numbers of recombinant colonies.

Differential screening of the fruit cDNA library, in combination with DNA dot blot analysis, resulted in the identification of 10 clones containing actinidin-related sequences, 1% of the clones in the library. This was a low yield, considering that actinidin-related RNA apparently represented a major fraction of total fruit RNA, but reflects the relatively small proportion of actinidin-encoding RNA recovered in the poly(A+) fraction of the RNA.

8.2. Structure of cDNAs

The two cDNAs characterised in detail show several characteristics typical of plant genes. The 3' untranslated regions of pAC.1 and pAC.7 contain three polyadenylation signals. Most plant genes, such as those of legumin and zein, have multiple polyadenylation sites (Lycett *et al.*, 1983; Geraghty *et al.*, 1981), sometimes consisting of overlapping signals, as also observed in pAC.1 and pAC.7. In an analysis of polyadenylation sites of several plant genes, Dean *et al.* (1986) showed that transcripts were polyadenylated at two or three sites. It would be interesting to investigate whether the other two polyadenylation sites of the actinidin transcripts are also used, and sequencing of the other eight clones identified may provide some information.

A G/T-rich region is present downstream from the second polyadenylation signal of pAC.1 and pAC.7. Although the significance of such a sequence is not clear, it has been found in 3'untranslated regions of other plant genes (Dean *et al.*, 1986; Werr *et al.*, 1985; Rosahl, 1986).

The coding strategy of pAC.1 does not conform to that observed for the majority of plant genes, which prefer A and T in the third codon position (Lycett *et al.*, 1983). Evolutionary aspects arising from the codon analysis are discussed below.

8.3. Sequence analysis

The protein partially encoded by pAC.1 differs from the published actinidin sequence more than the pAC.7 encoded protein, as it contains, in addition to the conservative substitutions also found in pAC.7, three non-conservative replacements, which may affect protein structure and/or function. These three substitutions are the only differences at the amino acid level between the two cloned proteins, and, surprisingly, are the result of only four nucleotide changes in the coding region. It is possible that the two sequences may represent two genes in the process of

divergence, following a recent duplication event. However, the differences would be expected to be more conservative in character, and show a greater number of silent mutations, considering the previously discussed selective constraint on important residues (see Introduction). In addition, the 3' untranslated regions, usually under less selective constraint than the coding region (Dunsmuir, 1985), differ by only three nucleotides plus an insertion/deletion of four bases at the polyadenylation site. These observations raise the possibility that the differences between pAC.1 and pAC.7 are due to cloning artefacts. A similar level of nucleotide differences was found between a genomic and a cDNA clone coding for aleurain, the cysteine proteinase from barley (Whittier *et al.*, 1987), and was attributed to errors during reverse transcription. Investigations into the fidelity of reverse transcriptase have shown misincorporations at frequencies between one in 600 and one in 6000 nucleotides (Gerard, 1984). The frequency of misincorporation in the case of the actinidin clones would be one in 280, not including the four nucleotide difference at the poly (A) tail which may have resulted from different processing of the mRNA precursor. This frequency seems very high, although, as discussed in chapter 4, the purification of mRNA was impeded by the presence of impurities co-purifying with poly(A) RNA, and the possibility, that traces of impurities present during reverse transcription may have reduced the fidelity of the enzyme, can perhaps not be excluded.

However, other reports in the literature support the view that pAC.1 and pAC.7 may indeed represent products of two different genes: patatin, the storage protein of potato tubers, consists of a family of very similar proteins, with 88% amino acid homology between two of its members (Rosahl *et al.*, 1986). More than half of the nucleotide substitutions in the coding region result in amino acid changes, and again more than half of

these are non-conservative. Patatin is an enzyme with lipolytic acyl hydrolase activity, and the authors suggest that some members of this family of proteins may have lost their enzyme activity as a result of non-conservative amino acid substitutions, following a reduction in the selective constraint on some members of the family. The 3' untranslated regions do not show a greater number of nucleotide differences than the coding regions, and highly conserved 3' untranslated regions have also been observed in several zein genes (Geraghty *et al.*, 1982). The α and α' subunits of conglycinin from soybean show even greater sequence conservation in their 3' untranslated than in their coding regions (Schuler *et al.*, 1982). Thus it is possible that the genes represented by pAC.1 and pAC.7 are in the process of divergence under reduced selective constraint on the pAC.1-encoded protein, which may have altered characteristics and activities. It should be possible to test whether different forms of actinidin are present in fruit of *A. chinensis*, by separating the proteins according to their different charge characteristics and investigating their catalytic properties. The fact that pAC.1 and pAC.7 were isolated by cDNA cloning shows that the gene(s) must be expressed, at least at the transcriptional level.

8.4. A family of actinidin genes

The two cDNA clones encode polypeptides with greater than 90% similarity to the published amino acid sequence of actinidin, and therefore represent additional members of the actinidin family of cysteine proteinases. The discovery, by Southern blotting, of a multigene family coding for actinidin is not surprising for two reasons: 1) All the other abundant plant cysteine proteinases, such as papain, bromelain, ficin and asclepain, consist of families with several, closely related, members. All seed storage proteins investigated to date are encoded by multigene

families (Higgins, 1984), as are other abundantly synthesised proteins such as barley α-amylase (Rogers and Milliman, 1984), the small subunit of ribulose 1,6 bisphospate carboxylase of petunia (Dean *et al.*, 1985), and the chlorophyll a/b binding protein, also of petunia (Dunsmuir, 1985).
2) *Actinidia chinensis* is probably hexaploid (Zhang and Beuzenberg, 1983), suggesting that most genes may be present in multiple copies as a result of polyploidisation, which is likely to have been followed by divergence of the individual members.
Two or more of the actinidin genes appear to be very similar indeed, as suggested by Southern blotting: The HindIII digest gave rise to a strongly hybridising band, likely to represent identical fragments from two or more adjacent gene copies resulting from a recent duplication event. The absence of a similar strongly hybridising band in the EcoRI digest suggests that these restriction sites have not been preserved. The two cDNAs described in this study may represent two such closely related genes.
A. chinensis is a polymorphic species, and is subdivided into several varieties, occurring naturally in different parts of its wide geographic range, and whose status as separate species was recently proposed (Liang and Ferguson, 1986). Although all varieties cultivated for their edible fruit, in New Zealand and elsewhere, are thought to be derived from the same variety, the history of origin of the cultivar used for cloning ("Exbury") is not known, and therefore its relationship to the cultivar used for amino acid sequencing (not identified in the publication; Carne and Moore, 1978) is unclear. It is possible that different varieties and cultivars have different alleles of the genes coding for actinidin, resulting in the observed differences between the sequences isolated from "Exbury" and published amino acid sequence. However, it is also possible that from similar sets of alleles of a multigene family, individual

members are differentially expressed in different cultivars. It would be interesting to examine in some detail the extent of diversity of this gene family, and as a first step it would be useful to analyse the remaining 8 clones isolated in this study, as they may represent different members of the gene family.

The expression of actinidin genes in different cultivars of *A. chinensis* could be studied by the use of gene specific probes derived from different genes. This approach has been used to demonstrate that members of the small subunit of ribulose 1,5-bisphosphate carboxylase from petunia and pea are differentially expressed (Dean *et al.*, 1985; Fluhr *et al.*, 1986).

8.5. Synthesis as a larger precursor

The two actinidin cDNAs encode polypeptides extending beyond the mature carboxy terminus of actinidin and, in addition, pAC.1 codes for a 57 residue portion of an amino-terminal extension. The discovery of a propeptide was not surprising as a similar extension has been found in papain and other cysteine proteinases. The analysis at the nucleotide level has confirmed the results from *in vitro* translation and immunoprecipitation experiments, which suggested that actinidin is synthesised in the form of a precursor 15 kDa larger than the mature protein. The size of the unprocessed precursor was estimated at 39 kDa, and this agrees very closely with the size predicted from the mRNA.

An alignment of the propeptide encoded by pAC.1 with amino-terminal sequences from other cysteine proteinases is shown in Figure 8.1. This reveals considerable homologies, suggesting that these portions of the cysteine proteinases are, like the mature enzyme domains, related by common ancestry. The overall level of homology is slightly less than that between the mature enzymes (Table 8.1), which could be the result of a less intense selective constraint during evolution. However, the fact

Table 8.1. Homologies of amino acid sequences of cysteine proteinases with that of actinidin and its deduced prosequence

	Protease domain	Prosegment
Pap	49.5% (109/220)	35.1 (20/57)
Ale	43.6% (96/220)	29.8 (17/57)
SB	41.2 (68/165)	-
CP-1	38.2 (84/220)	35.1 (20/57)
CP-2	38.6 (85/220)	31.6 (18/57)
Ca-B	30.0 (67/220)	-
Ca-H	40.0 (88/220)	-

Pap = Papain
Ale = Aleurain from barley
SB = Stem bromelain from pineapple
CP-1 = Cysteine proteinase 1 from slime mould
CP-2 = Cysteine proteinase 2 from slime mould
Ca-B = cathepsin B from rat
Ca-H = cathepsin H from rat

Figure 8.1. Alignment of the amino-terminal extension encoded by pAC.1 with 5'domains of other cysteine proteinases

Residues homologous with the pAC.1-derived sequence are boxed.
The amino-terminal portion not encoded by pAC.1 is aligned with the papain extension and homologies are circled.

PAC1 = Sequence encoded by pAC.1

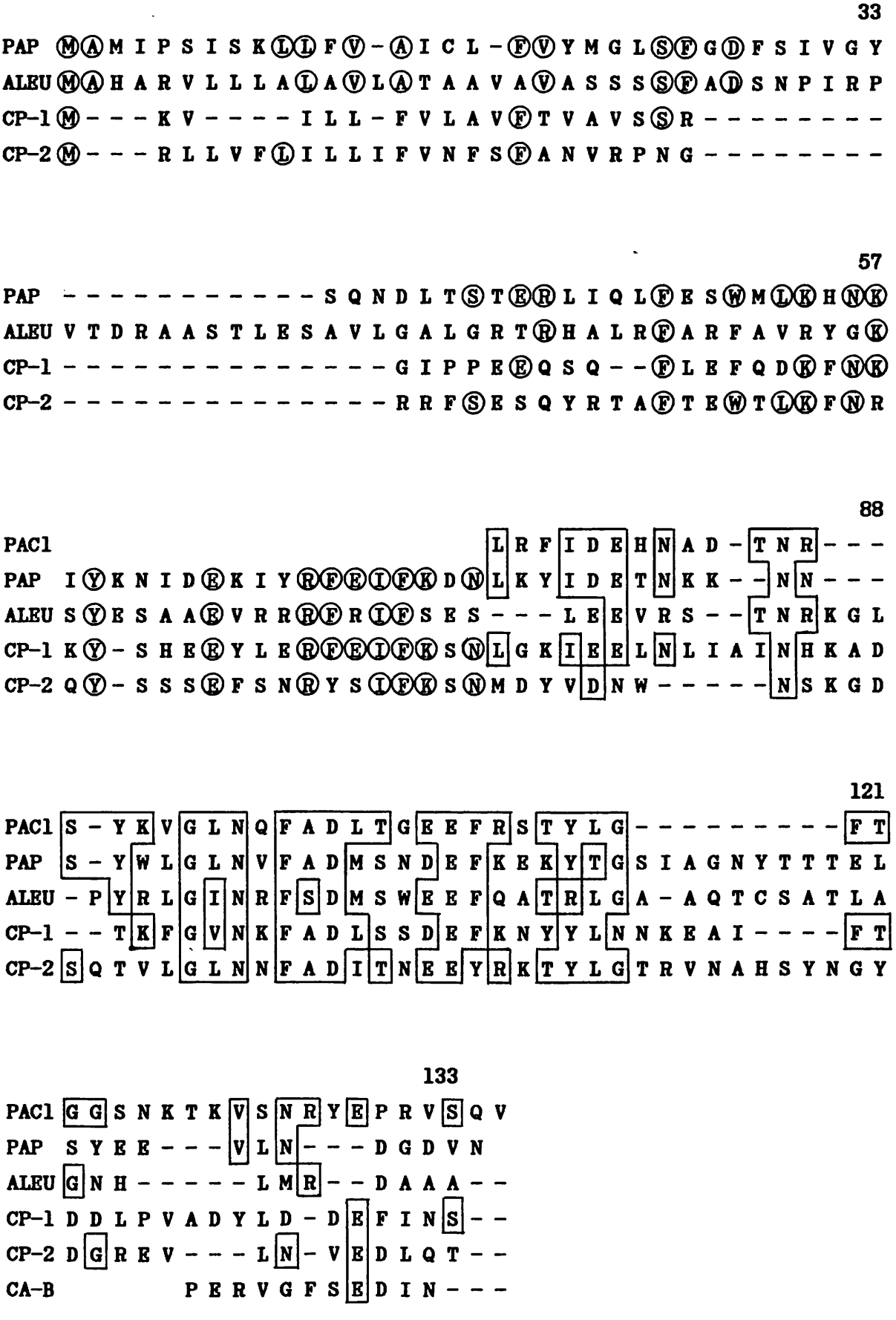
PAP = Papain from pawpaw

ALEU = Aleurain from barley

CP-1 = Cysteine proteinase 1 of slime mould

CP-2 = Cysteine proteinase 2 " " "

CA-B = Cathepsin B from rat


33
PAP MAMIPSISKLLFV-AICL-FVYMGLSFGDFSIVGY
ALEU MAHARVLLLALAVLATAAVAVASSSSFADSNPIRP
CP-1 M---KV----ILL-FVLAVFTVAVSSR---------
CP-2 M---RLLVFLILLIFVNFSFANVRPNG---------
57
PAP ----------SQNDLTSTERLIQLFESWMLKHNK
ALEU VTDRAASTLESAVLGALGRTRHALRFARFAVRYGK
CP-1 ----------GIPPEEQSQ--FLEFQDKFNK
CP-2 ----------RRFSESQYRTAFTEWTLKFNR
88
PAC1 LRFIDEHNAD-TNR---
PAP IYKNIDEKIYRFEIFKDNLKYIDETNKK--NN---
ALEU SYESAAEVRRRFRIFSES---LEEVRS--TNRKGL
CP-1 KY-SHEEYLERFEIFKSNLGKIEELNLIAINHKAD
CP-2 QY-SSSSEFSNRYSIFKSNMDYVDNW-----NSKGD
121
PAC1 S-YKVGLNQFADLTGEEFRSTYLG----------FT
PAP S-YWLGLNVFADMSNDEFKEKYTGSIAGNYTTTEL
ALEU -PYRLGINRFSDMSWEEFQATRLGA-AQTCSATLA
CP-1 --TKFGVNKFADLSSDEFKNYYLNNKEAI----FT
CP-2 SQTVLGLNNFADITNEEYRKTYLGTRVNAHSYNGY
133
PAC1 GGSNKTKVSNRYEPRVSQV
PAP SYEE---VLN---DGDVN
ALEU GNH-----LMR--DAAA--
CP-1 DDLPVADYLD-DEFINS--
CP-2 DGREV---LN-VEDLQT--
CA-B PERVGFSEDIN---

that it shows strongly conserved regions, and residues which are invariable between all the enzymes, suggests an important function, which may be similar for all the enzymes (North, 1986). Therefore it is possible that the precursor form of actinidin is an inactive zymogen, as has been proposed for that of papain (Cohen *et al.*, 1986).

The partial actinidin propeptide was shown to contain a large proportion of positively charged side chains. The papain propeptide, on the other hand, has a large number of negatively charged residues in the equivalent region. It is interesting that in both proteins the overall charge on the prosegment is the reverse of the net surface charge of the mature protein. These observations are consistent with a mechanism of inactivation by electrostatic interactions between the propeptide and the surface of the enzyme domain, such as that recently proposed for the aspartate proteinase, pepsinogen (James and Sielecki, 1986). Pepsinogen has a 44 amino acid propeptide containing many basic side chains, which interact with carboxylate groups of the enzyme domain, including those of the active site aspartyl residue. The overall shape of the inactivation domain of pepsinogen is that of a flattened disc, which fits neatly into the substrate binding cleft of pepsin, completely blocking access to the two catalytically important aspartyl residues. Activation of pepsinogen occurs spontaneously upon transfer from a neutral to an acidic environment, and is initiated by disruption of the electrostatic interactions leading to a conformational change in the propeptide. A portion of the propeptide is then cleaved by an intramolecular process giving rise to an enzymatically active intermediate. This is further converted to the mature enzyme by sequential proteolytic removal of the remaining propeptide, catalysed by mature pepsin.

It would be interesting to determine whether actinidin is activated in an acidic environment, and whether proteolytic processing involves an

intramolecular mechanism or requires other proteases, including mature actinidin.

As yet the precursor of actinidin has not been isolated, nor even detected, in fruit of *A.chinensis*, although it is possible that considerable amounts of it may be present in young fruit.

In a study of proteolytic enzymes in developing papaya fruit involving fractionation of fruit extract on hydroxylapatite, Skelton (1969) showed that 5 day-old fruit contained considerable amounts of papain. However, unlike the papain fractions isolated from older fruit, this had no proteolytic activity and was not activated by thiol compounds. Although the author does not consider this possibility, the papain of very young fruit may have been present in the form of an inactive precursor requiring proteolytic cleavage for activation. This could easily be tested by determining the M_r of the papain fraction by SDS-PAGE in combination with immuno-blotting (Burnette, 1981). Similarly, it would be interesting to analyse protein extracts from very young Chinese gooseberry fruit by SDS-PAGE. The antibodies available, which have been shown to react with the unprocessed precursor, can be used for the detection of actinidin protein by Western blotting.

If intact precursor can be isolated, it may be possible to study the conditions under which it is proteolytically processed, and to determine whether it is indeed an inactive zymogen.

As mentioned in the Introduction, cathepsin B is processed to the mature enzyme via an intermediate form of 39 kDa, which is believed to have limited enzyme activity and to be reponsible for the proteolytic maturation of insulin (Docherty *et al.*, 1984). Interestingly, it has been reported recently that insulin maturation takes place in post-Golgi vesicles which have an acid pH (Orci *et al.*, 1987). Since the cathepsin B precursor is believed to be sorted into the insulin secretory granules

during their formation in the Golgi apparatus (Docherty *et al.*, 1984), this may be an indication that cathepsin B is an example of a cysteine proteinase activated in a low pH environment.

8.6. Carboxy-terminal extension

Although the amino-terminal extension of actinidin is probably involved in inactivation of the protease activity, the function of the carboxy-terminal extension is difficult to assess. A carboxy-terminal extension of 6 amino acids was also discovered by cDNA cloning of cathepsin B, both from rat and human, and the first residue of this is identical to that of actinidin. Although this could indicate that the ancestral proteinase also had an extension, which requires the presence of a Gln residue for correct processing, other cysteine proteinases do not have this extension, and the differences between the two peptides are too great to conclude a common ancestry.

The carboxy-terminal extensions of cathepsin B, and of two unrelated lysosomal enzymes, cathepsin D and β-glucuronidase, have been implicated in subcellular targetting to the lysosomes. Although the subcellular location of actinidin has not been studied, it is possible that the function of its carboxy-terminal extension is targetting of the enzyme to the correct subcellular compartment. However, as described above, the mature carboxy-terminal peptide of actinidin crosses over to the amino-terminal domain and acts as a strap to hold the two domains together (Baker, 1980), and therefore it is equally possible that the transient extension is also involved in inactivation of the enzyme as it may be in close proximity to the active site cleft.

8.7. Subcellular location of actinidin

Although it is not known whether actinidin is compartmentalised, sequence analysis of the cDNA clones has shown that, unlike the mature form of actinidin, both the carboxy and amino termini have potential glycosylation sites. A large number of proteins destined for the lysosomes are glycosylated, as outlined in the Introduction. Therefore the potential for glycosylation of the actinidin precursor is circumstantial evidence that the protein is localised within lysosomal or related subcellular organelles, or possibly the vacuole.

If actinidin is indeed a compartmentalised enzyme, the initial translation product would be expected to possess an amino-terminal signal peptide for translocation into the lumen of the endoplasmic reticulum, similar to that of aleurain and the slime mould proteinases which are lysosomal enzymes. The size estimated for the actinidin mRNA allows for coding capacity of 60 to 70 amino acids in excess of the sequence determined from the longer cDNA, enough for a propeptide of similar dimension to those of other cysteine proteinases plus a signal sequence of 20 to 25 residues. However, this will have to be confirmed by sequence analysis of a full-length cDNA clone or of a genomic sequence.

It should be possible to verify the subcellular location of actinidin either by fractionation studies (Faye and Chrispeels, 1987), or immuno-cytochemical methods in combination with electron microscopy, using the antibodies to actinidin. In the studies of insulin maturation mentioned above, monoclonal antibodies to mature insulin which did not react with the prohormone, were used to identify the organelles containing the mature hormone (Orci *et al.*, 1987). If mature actinidin-specific antibodies can be produced it may be possible to distinguish between subcellular locations containing the mature enzyme and those containing precursor protein.

8.8. Evolutionary considerations

An analysis of the codon usage of pAC.1 has revealed slight differences in codon frequencies for some of the synonymous codons between the regions coding for the amino-terminal propeptide and the mature protein. The extension showed a striking preference for G and C residues in the third codon position, not observed in the remaining sequence. Rogers *et al.*(1985) have reported differences in the nucleotide composition between the sequences coding for the 5'domain and the enzyme domain of aleurain, and concluded that the amino-terminal portion may have been added to the enzyme later in evolution by a recombinational mechanism. Their hypothesis was supported by the finding that the position in aleurain corresponding to the mature amino terminus of other cysteine proteinases almost precisely coincides with the position of the third intron in the aleurain gene (Whittier *et al.*, 1987). The finding of inverted repeat structures flanking the first three exons and two introns has led the authors to speculate that this 5'domain evolved separately and was brought into juxtaposition with the remainder of the gene through a transposition event. The slight difference in coding strategy observed between the amino-terminal extension and mature protein coding regions of pAC.1 may be remnants of such an evolutionary event. Surprisingly, however, one of the arguments used by Whittier *et al.* to support the hypothesis of transposition is that only the mature enzyme domain of aleurain has homology with other known cysteine proteinases. In view of the increasing evidence for the synthesis of cysteine proteinases from diverse organisms in the form of larger precursors with homology to the 5'domain of aleurain, this argument no longer holds true. If a recombinational event has occurred in the evolution of the cysteine proteinases, this must have taken place before the divergence of plants and animals. The differences at the nucleotide level between regions coding for the amino-terminal and

mature protein domains may be preserved only in some of the descendant genes. It would be interesting to determine whether the genes coding for actinidin are also interupted by an intron at this junction.

8.9. Actinidin gene expression

The synthesis of actinidin was shown to be largely restricted to the fruit of the plant, and the difference in levels of actinidin between leaf and fruit was paralleled by the levels of actinidin mRNA. Thus actinidin gene expression seems to be fruit-specific and controlled at the level of transcription.

The results from the developmental study of actinidin gene expression show that actinidin transcripts are present at very low levels in young fruit. 5 day-old fruit contained only about three times more actinidin mRNA than leaf tissue. However, after an initial slow increase, the level of actinidin transcript rapidly increased early during fruit development. As mentioned in Chapter 7, this rapid increase coincided approximately with a change-over in development from cell division to cell enlargement in most parts of the fruit (see Introduction). Although this may not be a meaningful coincident, the expression of actinidin genes appears to parallel the increasing maturity of the fruit tissue. Assuming that a transition from cell division to cell enlargement indeed reflects an increase in differentiation of the fruit cells, it may be speculated that actinidin gene expression could be regulated simply in a fruit-specific manner, reaching maximal levels of transcription in fully differentiated fruit cells.

It is known from studies of other abundantly synthesised plant genes, that sequences 5' to the transcription start site are involved in the regulation of gene expression. For example, light regulation of the gene coding for the small subunit of ribulose 1,5-bisphosphate carboxylase

(rbcS) was shown by deletion mapping to be dependent on the presence of a 34 bp upstream sequence (Timko *et al.*, 1985). In addition, a longer sequence element was identified which was required for maximum transcriptional activation. Of the eight rbcS genes present in petunia, one was shown to be expressed at much higher levels than all the others (Dean *et al.*, 1985). Nucleotide sequencing revealed the presence in the 5' upstream region of this highly expressed gene of six blocks of sequence with homology to animal enhancer elements, whereas only two to four copies were present in the equivalent regions of the other genes. Similarities between sequence elements of 5'upstream regions from members of homologous gene families, for example the "legumin box" found in the legumin storage protein genes from three plant species, suggest that these elements are involved in tissue-specific gene expression (Baeumlein *et al.*, 1986). It would be interesting to analyse and compare the structure of the 5' upstream regions of the actinidin genes which may have similar conserved sequences. Since at least some of the genes must be highly expressed, and regulated in a tissue-specific manner, it would not be surprising to discover enhancer-like elements which modulate the levels of transcription. It may also be interesting to compare the regulatory regions to those of other plant cysteine proteinases, but also of other fruit-specific genes when sequences become available, such as the ripening genes identified in the tomato fruit (Grierson *et al.*, 1985). As has been observed in the case of the zein gene family, which contains a large number of members, it may be expected that the expression of the actinidin gene family is coordinately regulated. Recently, a short nucleotide sequence conserved in all zein genes was demonstrated to be the binding site for a nuclear protein factor, which is likely to be involved in the coordinated regulation of zein gene expression (Maier *et al.*, 1987). Interestingly, a similar sequence was also found in the maize sucrose

synthetase gene, which has a similar pattern of gene expression. An analysis of the 5'upstream sequences of the aleurain gene has revealed sequence homology with that of the α-amylase gene, whose expression is similarly controlled by gibberellic acid (Whittier *et al.*, 1987).
An analysis and comparison of regulatory sequences of the actinidin genes is likely to contribute to the increasing wealth of information on the mechanisms of gene regulation in plants, and the availability of cDNA clones coding for actinidin will be useful in the isolation of genomic sequences.

8.10. Function of actinidin

As yet nothing is known about the physiological function of actinidin. It seems reasonable to assume that the enzyme has an important role as it is present at such high levels in fruit. Since different levels of actinidin protein in fruit and leaves correlated with mRNA levels, it seems reasonable to assume that actinidin protein is also present in developing fruit, although this will have to be investigated experimentally. The probability that actinidin is synthesised in the form of an inactive precursor or zymogen raises the possibility that this may be stored in the fruit, possibly in some subcellular compartment separate from the active enzyme, until required. It is tempting to speculate that actinidin is not involved in the maturation process of the fruit itself but that it may have a role in the defence against predators, which in other cases is thought to be the function of proteinase inhibitors. The large amounts of actinidin in ripe fruit do not appear to deter mammalian predators, including man, and indeed the palatability of fruit is an adaptive property "designed" to ensure seed dispersal. Since unripe fruit is probably not palatable to most animals, the latter are unlikely to present a danger to developing Kiwi fruit. However, it is likely that insects

would cause damage by attempting to rear their offspring on the nourishing juices of the unripe fruit, possibly causing abscission of the fruit before the seeds have reached maturity. The presence of large amounts of highly active, non-specific proteinase might not be compatible with the survival of insects larvae. If inactive precursor was indeed found to be stored in separate subcellular organelles, damage to the cells by intruding predators would cause mixing of cell contents with those of the acidic vacuole, which may result in activation of the proteinase.
A study of the subcellular location of actinidin and of the conditions under which the putative inactive precursor is activated, although interesting in themselves, may also help illuminate the question of the physiological role of actinidin.

9. References

Arcus AC (1959): Proteolytic enzyme of *Actinidia chinensis*. Biochim Biophys Acta 33: 242-244

Baker EN (1977): Structure of actinidin: details of the polypeptide chain conformation and active site from an electron density map at 2.8 Å resolution. J Mol Biol 115: 263-277

Baker EN (1980): Structure of actinidin, after refinement at 1.7 Å resolution. J Mol Biol 141: 441-484

Baker EN, Boland MJ, Calder PC, Hardman MJ (1980): The specificity of actinidin and its relationship to the structure of the enzyme. Biochim Biophys Acta 616: 30-34

Baker EN (1981): Comparison of the refined structures of two sulphydryl proteases: actinidin and papain. In: Dodson G, Gluster JP,Jayre D (eds) Clarendon Press, Oxford pp339-349

Baeumlein H, Wobus U, Pustell J, Kafatos FC (1986): The legumin gene family: Structure of a B type gene of *Vicia faba* and a possible legumin gene specific regulatory element. Nucleic Acids Res 14: 2707-2720

Barrett AJ. Kirschke H (1981): Cathepsin B, cathepsin H, and cathepsin L. Meth Enzymol 80: 535-561

Baumgartner B, Chrispeels MJ (1977): Immunocytochemical localisation of vicilin-peptidohydrolase in mung bean cotyledons. Plant Physiol 59:86

Berger A, Schechter I (1970): Mapping the active site of papain with the aid of peptide substrates and inhibitors. Phil Trans Roy Soc Lond B 257: 249-264

Boland MJ, Hardman MJ (1972): Kinetic studies on the thiol protease from *Actinidia chinensis*. FEBS Lett 27: 282-284

Brocklehurst K, Baines BS, Kierstan MPJ (1981): Papain and other constituents of *Carica papaya* L. In: Wiseman, Topics in enzyme and fermentaiton biotechnology 5. Ellis Harwood Ltd, Chichester, pp262-335

Burnette WN (1981): Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulphate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal Biochem 112: 195-203

Carne AC, Moore CH (1978): The amino acid sequence of the tryptic peptides from actinidin, a proteolytic enzyme from the fruit of *Actinidia chinensis*. Biochem J 173: 73-83

Carotti A, Hansch C, Mueller MM, Blaney JM (1984): Actinidin hydrolysis of substituted phenyl hippurates - a quantitative structure activity relationship and graphics comparison with hydrolysis by papain. J Med Chem 27: 1401-1405

Chan SJ, Noyes BE, Argawal KL, Steiner DF (1979): Construction and selection of recombinant plasmids containing full-length complementary DNAs corresponding to rat insulin I and II. Proc Natl Acad Sci 76: 5036-5040

Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ (1979): Isolation of biologically active ribonucleic acid from sources enriched in ribocuclease. Biochemistry 18: 5294-5299

Cohen LW, Coghlan VM, Dihel LC (1986): Cloning and sequencing of papain-encoding cDNA. Gene 48: 219-227

Covey SN, Grierson D (1976): The measurement of plant polyadenylic acid by hybridisation with radioactive polyuridylic acid. Planta 131: 75-79

Croy RRD, Lycett GW, Gatehouse JA, Yarwood JN, Boulter D (1982): Cloning and analysis of cDNAs encoding plant storage protein precursors. Nature 295: 76-79

Dahms NM, Lobel P, Breitmeyer J, Chirgwin JM, Kornfeld S (1987): 46 Kd mannose 6-phosphate receptor: cloning, expression, and homology to the 215 Kd mannose 6-phosphate receptor. Cell 50: 181-192

Darnell JE (1982): Variety in the level of gene control in eukaryotic cells. Nature 297: 365-371

Dean C, van den Elzen P, Tamaki S, Dunsmuir P, Bedbrook (1985): Differential expression of the eight genes of the petunia ribulose bisphosphate carboxylase small subunit multi-gene family. EMBO J 4: 3055- 3061

Dean C, Tamaki S, Dunsmuir P, Favreau M, Katayama C, Dooner H, Bedbrook J (1986): mRNA transcripts of several plant genes are polyadenylated at multiple sites *in vivo*. Nucleic Acids Res 14: 2229-2240

De Vries SC, De Vos WM, Harmsen MC, Wessels JGH (1985): A shoot-specific mRNA from pea: nucleotide sequence and regulation as compared to light-induced mRNAs. Plant Mol Biol 4: 95-102

Docherty K, Hutton JC, Steiner DF (1984): Cathepsin B-related proteases in the insulin secretory granule. J Biol Chem 259: 6041-6044

Drenth J, Jansonius JN, Koekoek R, Swen HM, Wolthers BG (1968): Structure of papain. Nature 218: 929-932

Dunsmuir P (1985): The petunia chlorophyll a/b binding protein genes: a comparison of Cab genes from different gene families. Nucleic Acids Res 13: 2503-2518

Erickson AH, Blobel G (1983): Carboxy-terminal proteolytic processing during biosynthesis of the lysosomal enzymes β-glucuronidase and cathepsin D. Biochemistry 22: 5201-5205

Faye L, Chrispeels MJ (1987): Transport and processing of the glycosylated precursor of concanavalin A in jack-bean. Planta 170: 217-224

Figura K von, Hasilik A (1986): Lysosomal enzymes and their receptors. Ann Rev Biochem 55: 167-193

Fluhr R, Moses P, Morelli G, Coruzzi G, Chua N-H (1986): Expression dynamics of the pea rbcS multigene family and organ distribution of the transcripts. EMBO J 5: 2063-2071

Fong D, Calhoun DH, Hsieh W-T, Lee B, Wells RD (1986): Isolation of a cDNA clone for the human lysosomal proteinase cathepsin B. Proc Natl Acad Sci 83: 2909-2913

Geraghty D, Peifer MA, Rubenstein I, Messing J (1981): The primary structure of a plant storage protein: zein. Nucleic Acids Res 9: 5163-5174

Gerard GE (1982): The error rate of cloned M-MLV reverse transcriptase during DNA synthesis. Focus 8 (3): 12

Glazer AN, Smith EL (1971): Papain and other plant sulphydryl proteolytic enzymes. Enzymes 3: 501-546

Godfrey T, Reichelt J (1983): Industrial enzymology. Macmillan Ltd, Oxford

Goding J (1983): Monoclonal antibodies: principles and practice. Academic Press Ltd, London

Goto K, Murachi T, Takahashi N (1976): Structural studies on stem bromelain isolation, characterisation and alignment of the cyanogen bromide fragments. FEBS Lett 62: 93-95

Goto K, Takahashi N, Murachi T (1980): Structural studies on stem bromelain - cyanogen bromide cleavage and amino acid sequence of carboxyl-terminal half of the molecule. Int J Pept Protein Res 15: 335-341

Graf L, Craik CS, Patthy A, Roczniak S, Fletterick RJ, Rutter WJ (1987): Selective alteration of substrate specificity by replacement of aspartic acid-189 with lysine in the binding pocket of trypsin. Biochemistry 26: 2616-2623

Grierson D, Slater A, Speirs, J Tucker GA (1985): The appearance of polygalacturonase mRNA in tomatoes: one of a series of changes in gene expression during development and ripening. Planta 163: 263-271

Gruenbaum Y, Naveh-Many T, Cedar H, Razin A (1981): Sequence specificity of methylation in higher plant DNA. Nature 292: 860-862

Gubler U, Hoffman B (1983): A simple and very efficient method for generating cDNA libraries. Gene 25: 263-269

Hall TC, Ma Y, Buchbinder BU, Pyne JW, Sun SM, Bliss FA (1978): Messenger RNA for Gl protein forFrench bean seeds: cell-free translation and product characterisation. Proc Natl Acad Sci 75: 3196-3200

Hames BD, Rickwood D (eds) Gel electrophoresis of proteins: a practical approach. IRL Press Ltd, Oxford

Hammerton RW, Ho T-HD (1986): Hormonal regulation of the development of protease and carboxypeptidase activities in barley aleurone layers. Plant Physiol 80: 692-697

Harris TJR, Lowe PA, Lyons A, Thomas PG, Eaton MAW, Millican TA, Patel TP, Bose CC, Carey NIH, Doel MT (1982): Molecular cloning and nucleotide sequence of cDNA coding for calf preprochymosin. Nucleic Acids Res 10: 2177-2187

Heidecker G, Messing J, Gronenborn B (1980): A versatile primer for DNA sequencing in the M13 mp2 cloning system. Gene 10: 68-73

Heijne G von (1985): Signal sequences. The limits of variation. J Mol Biol 184: 99-105

Heinemann U, Pal GP, Hilgenfeld R, Saenger W (1982): Crystal and molecular structure of the sulphydryl protease calotropin DI at 3.2 Å resolution. J Mol Biol 161: 591-606

Higgins TJV (1984): Synthesis and regulation of major proteins in seeds. Ann Rev Plant Physiol 35: 191-221

Hopping ME (1976): Effect of exogenous auxins, gibberellins, and cytokinins on fruit development in Chinese gooseberry (*Actinidia chinensis* Planch.). NZ J Bot 14: 69-75

Hudson L, Hay FC (1980); Practical immunology (2nd ed) Blackwell Scientific Publications, Oxford

Husain SS, Lowe G (1969): Completion of the amino acid sequence of papain. Biochem J 114: 279-288

Husain SS, Lowe G (1970): The amino acid sequence around the active-site cysteine and histidine residues, and the buried cystein residue in ficin. Biochem J 117: 333-340

James MNG, Sielecki AR (1986): Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8A resolution. Nature 319: 33-38

Kamphius IG, Kalk KH, Swarte MBA, Drenth J (1984): Structure of papain refined at 1.65 Å resolution. J Mol Biol 179: 233-252

Kamphius IG, Drenth J, Baker EN (1985): Thiol proteases: Comparative studies base on the high resolution structures of papain and actinidin, and on amino acid sequence information for cathepsins B and H, and stem bromelain. J Mol Biol 182: 317-329

Kaulen H, Schell J, Kreuzaler F (1986): Light-induced expression of the chimeric chalcone synthase-NPTII gene in tobacco cells. EMBO J 5: 1-8

Kawamura M, Wadano A, Miura K (1984): Purification and some properties of cathepsin-like thiol protease from pupae of the blowfly *Aldrichina grahami*. Comp Biochem Physiol 78B: 279-286

Kirby KS (1964): Isolation and fractionation of nucleic acids. Progr Nucl Acid Res Mol Biol 3: 1-31

Laemmli UK (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. Nature 227: 680-685

Lathe R (1985): Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. J Mol Biol 183: 1-12

Liang CF, Ferguson AR (1986): The botanical nomenclature of the Kiwifruit and related taxa. NZ J Bot 24: 183-184

Lichtenstein C, Draper J (1985): Genetic engineering of plants. In: Dover (ed) DNA cloning vol II. IRL Press, Oxford pp 67-119

Liu T-Y, Elliott SD (1971): Streptococcal proteinase. In: Boyer P.D. (ed) The Enzymes III, Acad Press

Lomedico PT, Gubler U, Hellmann CP, Dukovich M, Giri JG, Pan Y-C, Collier K, Semionow R, Chua AO, Mizel SB (1984): Cloning and expression of murine interleukin-1 cDNA in *Escherichia coli*. Nature 312:458-62

Lowe G: The cysteine proteinases. Tetrahedron 32:291-302 (1976).

Lycett GW, Delauney AJ, Croy RRD(1983): Are plant genes different? FEBS Lett 153: 43-46

Lynn KR, Bockbank WJ, Clevette NA (1980): Multiple forms of the asclepains cysteinyl proteases from milkweed. Biochim Biophys Acta 612: 119-125

Lynn KR, Yagouchi M, Roy C (1980): Homologies of the N-terminal sequences of asclepains and papain. Biochim Biophys Acta 624: 579-580

Magnuson MA, Nikodem VM (1983): Molecular cloning of a cDNA sequence for rat malic enzyme. J Biol Chem 258: 12712-12717

Maier U-G, Brown JWS, Toloczyki C, Feix G (1987): Binding of a nuclear factor to a consensus sequence in the 5' flanking region of zein genes from maize. EMBO J 6: 17-22

Maniatis T, Fritsch EF, Sambrook J (1982): Molecular cloning. A laboratory manual. Cold Spring Harbour Laboratory, Cold Spring Harbour NY

McDowall MA (1970): Anionic proteinase from *Actinidia chinensis*. Eur J Biochem 14:214-221

Noyes BE, Mevarech M, Stein R, Agarwal KL (1979): Detection and partial

Negoro T, Momma T, Fukazawa C (1985): A cDNA clone encoding a glycinin Ala subunit precursor of soybean. Nucleic Acids Res 13: 6719-6731

Neurath H (1984): Evolution of proteolytic enzymes. Science 224: 350-357

Nielssen SS, Liener JE (1984): Degradation of the major storage protein of *Phaseolus vulgaris* [cultivar Improved Tendergreen] during germination: Role of endogenous proteases and protease inhibitors. Plant Physiol 74: 494-498

North MJ (1986): Homology within the N-terminal extension of cysteine proteinases. B J Letters 238: 623-624 (1986)

Biochemistry 26: 2077-2082

Paterson BM, Roberts BE, Kuff EL (1977): Structural gene identification and mapping by DNA.mRNA hybrid-arrested cell-free translation. Proc Natl Acad Sci 74: 4370-4374

Pears CJ, Mahbubani HM, Williams JG (1985): Characterisation of two highly diverged but developmentally co-regulated cysteine proteinase genes in *Dictyostelium discoideum.* Nucleic Acids Res 13: 8853-8866

Pelham HRB, Jackson RJ (1976): An efficient mRNA - dependent translation system from reticulocyte lysates. Eur J Biochem 67: 247-256

Rath VL, Fletterick RJ (1987): Protein structure and design. Cell 49: 583-586

Roberts BE, Paterson BM (1973): Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. Proc Natl Acad Sci 70: 2330-2334

Rogers JC, Milliman C (1984): Coordinate increase in major transcripts from the high pI α-amylase multigene family in barley aleurone cells stimulated with gibberellic acid. J Biol Chem 259: 12234-12240

Rogers JC, Dean D, Heck GR (1985): Aleurain: a barley thiol protease closely related to mammalian cathepsin H. Proc Natl Acad Sci USA 82: 6512-6516

Rosahl S (1986): Isoliering, Charakterisierung und funktionelle Analyse eines Patatin-Gens aus *Solanum tuberosum*. Doctoral Thesis, Koeln

Rosahl S, Schmidt R, Schell J, Willmitzer L (1986): Isolation and characterisation of a gene from *Solanum tuberosum* encoding patatin, the major storage protein of potato tubers. Mol Gen Genet 203: 214-220

Ryan CA (1973): Proteolytic enzymes and their inhibitors in plants. Ann Rev Plant Physiol 24: 173-196

Sanger F, Nicklen S, Coulson AR (1977): DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74:5463-5467

San Segundo B, Chan SJ, Steiner DF (1985): Identification of cDNA clones encoding a precursor of rat liver cathepsin B. Proc Natl Acad Sci USA 82:2320-2324

Schimke RT, Rhoads RE, McKnight GT (1974): Assay of ovalbumin mRNA in reticulocyte lysate. Meth Enzymol 30: 694-701

Schmid R (1978): Reproductive anatomy of *Actinidia chinensis*. Bot Jahrb f System Pflanzengesch Pflanzengeogr 100: 149-195

Schopf JW (1978): The evolution of the earliest cells. In: Piel G, Flannagan (eds): Evolution. A Scientific American Book. Freemann, San Francisco

Schroeder CA, Fletscher WA (1967): The Chinese gooseberry in New Zealand. Econ Bot 21: 81-92

Schuler MA, Schmitt ES, Beachy RN (1982): Closely related families of genes code for the α and α' subunits of the soybean 7S storage protein complex. Nucleic Acids Res 10: 8225-8244

Slightom JL, Drony RF, Klassy RC, Hoffman LM (1985): Nucleotide sequence from phaseolin cDNA clones: the major storge proteins from *Phaseolus vulgaris* are encoded by two unique gene families. Nucleic Acids Res 13: 6483-6498

Skelton GS (1969): Development of proteolytic enzymes in growing papaya fruit. Phytochemistry 8: 5760

Stroud RM (1974): A family of protein-cutting proteins. Scientific American 231: 74-88

Swank RT, Munkres KD (1971): Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dodecyl sulfate. Anal Biochem 39: 462-477

Takio K, Towatari T, Katunuma N, Teller DC, Titani K (1983): Homology of amino acid sequences of rat liver cathepsin B and H with that of papain. Proc Natl Acad Sci 80: 3666-3670

Taylor JM (1979): The isolation of eukaryotic messenger RNA. Ann Rev Biochem 48: 681-717

Taylor WR (1986): The classification of amino acid conservation. J Theor Biol 119: 205-218

Thomas PS (1980): Hybridisation of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci 77: 5201-5205

Timko MP, Kausch AP, Castresana C, Fassler J, Herrera-Estrella L, Broeck van den G, Montague van M, Schell J, Cashmore AR (1985): Light regulation of plant gene expression by an upstream enhancer-like element. Nature 318: 579-582

Veitch JH (1904): Recently introduced trees, shrubs, etc. from Central China. J Roy Hort Soc Lond 28: 57-70

Voller A, Bidwell DE, Bartlett A (1979): The enzyme linked immunosorbent assay (ELISA). Dynatech Europe Guernsey

Walburg G, Larkins BA (1986): Isolation and characterisation of cDNAs encoding oat 12S globulin mRNAs. Plant Mol Biol 6: 161-169

Warburg O, Christian W (1942): Isolierung und Kristallisierung des Gaerungsferments Enolase. Biochem Z 310: 384-421

Werr W, Frommerer WB, Mans C, Starlinger P (1985): Structure of the sucrose synthase gene on chromosome 9 of *Zea mays* L. EMBO J 4: 1373-1380

Whittier RF, Dean DA, Rogers JC (1987): Nucleotide sequence analysis of alpha amylase and thiol protease genes that are hormonally regulated in barley aleurone cells. Nucleic Acids Res 15:2515-2535

Willenbrock F, Brocklehurst K (1985): A general framework of cysteine proteinase mechanism deduced from studies on enzymes with structurally different analogous catalytic-site residues asp-158 and -161 (papain and actinidin) gly-196 (cathepsin B) and asn-165 (cathepsin H). Biochem J 227: 521-528

Willenbrock F, Brocklehurst K (1984): Natural structural variation in enzymes as a tool in the study of mechanism exemplified by a comparison of the catalytic site structure and characteristics of cathepsin B and papain. Biochem J 222: 805-814

Williams JG, North MJ, Mahbubani H (1985): A developmentally regulated cysteine proteinase in *Dictyostelium discoideum.* EMBO J 4: 999-1006

Zhang J, Beuzenberg EJ (1983): Chromosome numbers in two varieties of *Actinidia chinensis* Planch. NZ J Bot 21: 353- 355

10. Appendix

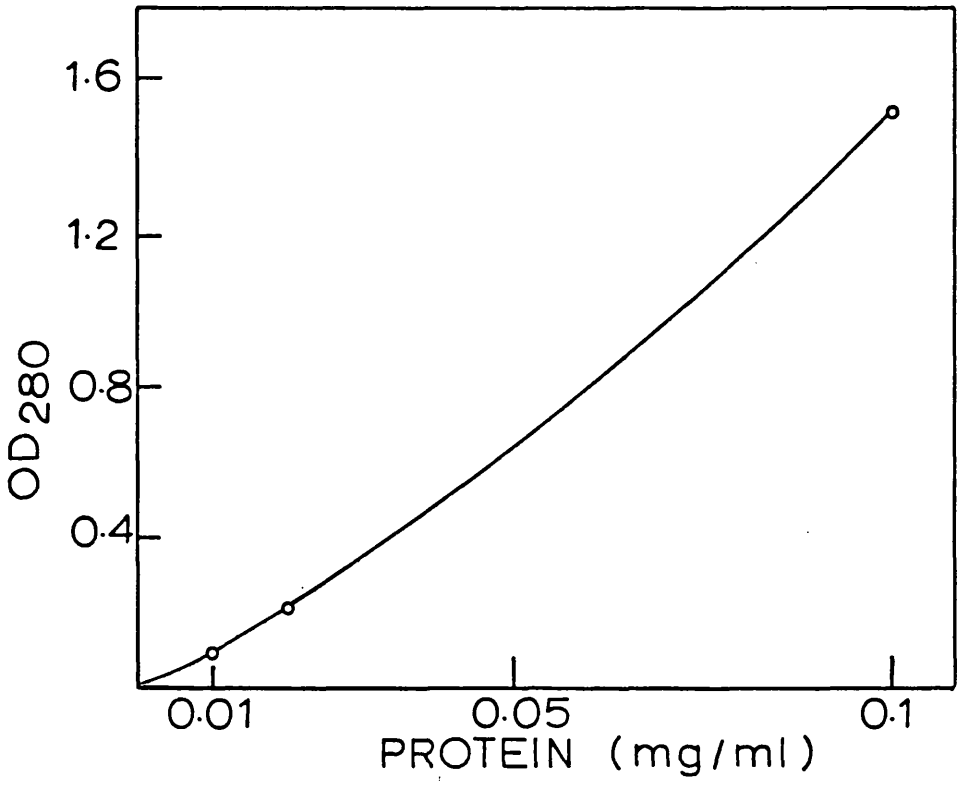


Figure A.1. Papain hydrolysis of casein: standard curve

The assay of casein hydrolysis was carried out as described in Methods.

Table A.1.

Estimation of protein content and protease activity of column fractions

fraction	total protein (mg)	protease activity (OD_{280})	activity per mg protein	activity equiv. to mg papain
25	9.7	1.99	0.22	0.14
26	15.8	3.00	0.19	0.20
27	17.8	3.71	0.21	0.25
28	20.2	4.85	0.24	0.33
29	21.8	6.36	0.29	0.43
30	16.2	4.76	0.29	0.32
31	18.5	6.12	0.33	0.42
32	10.2	4.07	0.39	0.28
33	4.2	2.23	0.53	0.15
34	1.2	1.03	0.89	0.07
Total	135.6	38.12		2.59

Table A.2. ELISA of plant extracts: OD_{450} values

	actinidin		leaf		apex		root		fruit		fruit	
	−	+	−	+	−	+	−	+	−	+	−	+
A	0.32	1.42	0.21	0.91	0.13	0.51	0.67	0.68	0.19	1.15	0.16	1.13
B	0.32	1.37	0.21	0.87	0.21	0.45	0.58	0.57	0.20	1.12	0.12	1.14
C	0.26	1.37	0.26	0.72	0.18	0.34	0.51	0.49	0.17	0.98	0.17	1.04
D	0.27	1.30	0.23	0.52	0.18	0.31	0.36	0.34	0.10	0.94	0.21	0.96
E	0.28	1.17	0.26	0.38	0.16	0.24	0.23	0.23	0.18	0.71	0.19	0.69
F	0.28	0.86	0.26	0.28	0.16	0.17	0.21	0.18	0.20	0.50	0.20	0.49
G	0.28	0.60	0.27	0.25	0.15	0.19	0.18	0.20	0.22	0.37	0.22	0.38
H	0.29	0.28	0.21	0.23	0.20	0.21	0.20	0.20	0.21	0.23	0.23	0.23

− non-immune antibody
+ actinidin-specific antibody

A to H: two-fold serial dilutions.
Leaf, apex and root extracts from 0.017 g/ml.
Fruit extract from a thousand-fold dilution of 0.017 g/ml.

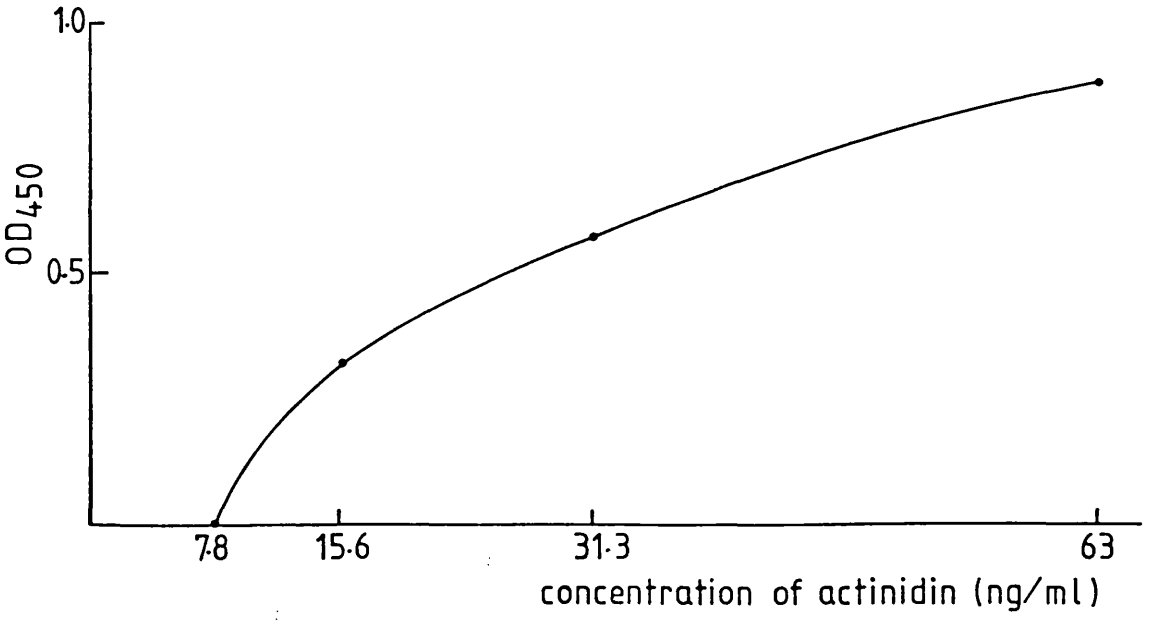


Figure A.2. ELISA: Standard curve of actinidin

Table A.3.1. Summary of results from ELISA of plant extracts

	Tissue	dilution	µg actinidin/ g dry weight	µg actinidin/ g fresh weight
A)	leaf	1:4	6.1	
		1:8	6.7	
		1:16	7.6	
		average:	6.8 ± 0.8	0.69
	apex	1:2	1.3 1.4	
		1:4	1.6 2.0	
		1:8	1.9 2.4	
		1:16	2.9 3.4	
		average:	2.1 ± 0.7	0.24
	root	not measurable		
	fruit (unripe)	1:4,000	18,500	
		1:8,000	21,800	
		1:16,000	24,900	
		1:32,000	24,000	
		average:	22,300 ± 2,800	3,800
B)	leaf	1:4	5.6	
		1:8	6.7	
		1:16	5.3	
		average:	5.9 ± 0.8	0.61
	apex	1:2	1.4	
		1:4	1.8	
		1:8	2.9	
		average:	2.0 ± 0.8	0.23
	root	not measurable		
	fruit (unripe)	1:8,000	21,600 22,600	
		1:16,000	26,900 25,400	
		1:32,000	29,800 28,800	
		average:	25,900 ± 3,300	4,300

Parts A and B represent two independent assays with different extracts.

Table A.3.2. Ratios of fresh to dry weight of different parts of the plant

leaf	8.84
apex	9.74
root	9.71
fruit (unripe)	5.93

Table A.4. Weight of *A. chinensis* var. "Exbury" fruit at different developmental stages

	Days after anthesis					
	7	24	42	64	82	104
weight	1.5	6.5	13.5	17.0	22.0	18.0
(gram)	1.8	11.0	12.5	16.0	14.5	22.0
	1.2	8.2	13.5	14.0	18.5	23.0
		7.8	13.5	17.0	17.0	21.0
		7.5	13.5	14.0	13.0	21.0
		7.5	12.0	18.0	18.0	22.5
		8.2			15.5	19.0
		8.5			12.5	22.0
					17.5	20.0
					17.0	17.0

Averages

7 days	1.5 ± 0.3
24 days	8.2 ± 1.3
42 days	13.1 ± 0.7
64 days	16.0 ± 1.7
82 days	16.7 ± 2.5
104 days	20.6 ± 2.0

Table A.5. Intensity data of dot blot

Age of fruit (days)	relative intensity 0.3 μg RNA	0.9μg RNA
7	0.62	0.55
24	1.27	1.25
42	3.53	3.44
64	6.84	11.31
82	6.83	8.95
104	6.17	14.45
104	7.54	14.40
10 days storage	2.41	5.18
21 days storage	n.d.	3.30
leaf RNA	2.45 μg RNA	7.35μg RNA
	1.26	1.99

Dot blots were hybridised with ^{32}P-labelled cDNA insert of pAC.1, autoradiographed and scanned with an Ultro Scan Laser Densitometer (LKB 2202), and peak areas calculated using a recording integrator.

Molecular Analysis of Actinidin

Uta M. Praekelt

Abstract

Actinidin, the 23.6 kDa cysteine proteinase of Chinese gooseberry (*Actinidia chinensis*), is present at high concentration in fruits. A fruit-specific cDNA library was established and screened by differential hybridisation and using a synthetic oligonucleotide. Two of ten actinidin clones identified were characterised by sequence analysis. The two very similar cDNAs code for proteins with approximately 90% sequence homology to the published amino acid sequence of actinidin, as well as an additional 25 amino acids following the mature carboxyl terminus. The larger clone in addition has coding potential for 57 residues of an amino-terminal extension with considerable homology to amino-terminal sequences of other cysteine proteinases. From size determinations of both mRNA (1.4 kb) and immunoprecipitated *in vitro* translation product (39 kDa) it was estimated that actinidin is synthesised as a precursor approximately 15 kDa larger than the mature protein. Features of the prosegment primary sequence are considered with regard to a possible mechanism of inactivation of the proteinase, by analogy with other proteolytic zymogens. The presence of three potential glycosylation sites, one within the carboxy-terminal and two in the amino-terminal extension are consistent with subcellular location of the enzyme within membrane-bound organelles. Results from a Southern blot show that actinidin is encoded by a multigene family of up to ten members. Actinidin gene expression, both at the level of mRNA and protein, is largely restricted to the fruit of *A. chinensis*, where the level of actinidin mRNA accumulates early during development.