

Molecular analysis of actinidin, the cysteine proteinase of *Actinidia chinensis*

Uta M. Praekelt,¹ Raymond A. McKee² and Harry Smith

Department of Botany, University of Leicester, University Road, Leicester LE1 7RH, UK; (present addresses: ¹Leicester Biocentre, University of Leicester, University Road, Leicester LE1 7RH, UK; ²AFRC Institute of Food Research, Colney Lane, Norwich NR4 7UA, UK)

Received 28 July 1987; accepted in revised form 21 October 1987

Key words: *Actinidia chinensis*, cDNA clones, cysteine proteinase, developmental expression, organ-specific, precursor

Abstract

We have isolated and sequenced two very similar cDNA clones of 1 145 and 809 bp length, from a fruit-specific library of *Actinidia chinensis*, the larger encoding all 220 amino acids of actinidin, showing 91% homology to the published amino acid sequence. Both cDNAs code for an additional 25 amino acids following the mature carboxy terminus of actinidin. The larger clone has coding potential for 57 residues of an amino-terminal extension with considerable homology to amino-terminal sequences of other cysteine proteinases. From size determination 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. Both proteolytic cleavage sites are located on the surface of the molecule as illustrated by the hydropathy profile of the deduced amino acid sequence. 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 suggest 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 the plant, where the level of actinidin mRNA accumulates early during development.

Introduction

Actinidin, the cysteine proteinase from the fruit of *Actinidia chinensis* (Chinese gooseberry or Kiwifruit), is a well characterised endopeptidase (M_r 23.6 kDa) with wide substrate specificity and pH range [1, 25] whose amino acid sequence and tertiary structure have been determined [2, 3, 6]. It belongs to a family of homologous enzymes, related through common ancestry, from other plants, such as papain from pawpaw, *Carica papaya* [10],

bromelain from pineapple, *Ananas comosus* [14], ficin from fig, *Ficus glabrata* [23], and aleurain from barley, *Hordeum vulgare* [29], but also from widely different organisms including mammals and slime mould (*Dictyostelium*). In spite of considerable amino acid sequence divergence, structurally and functionally important residues are highly conserved between all the enzymes, suggesting similar folding patterns [20]. The strong conservation of tertiary structure has been demonstrated for actinidin and papain whose polypeptide backbones are virtually

superimposable although their primary sequences are only 52% homologous [4]. However, with the exception of aleurain, whose role in the hydrolysis of storage proteins during seed germination has been established [37], no clear function has been identified for the other plant enzymes, present at high concentrations in the fruit or latex of the plants, and in many cases consisting of several isozymes. The cysteine proteinases are highly active enzymes and, as expected, some have been shown to be localised in membrane-bound organelles, including mammalian cathepsin B [9] and the cysteine proteinases from *Dictyostelium* [27]. Consistent with their location, signal peptides were identified by analysis at the DNA level of the two *Dictyostelium* proteinases [27], and of aleurain, which is secreted [26]. In addition, cDNA cloning of these enzymes, but also of papain [7] and cathepsin B from rat [32], revealed the presence of an extra 100 to 120 amino acids N-terminal to the region of homology with actinidin and to the mature N-termini of papain and cathepsin B. It is not known whether they are proteolytically removed from aleurain or the *Dictyostelium* enzymes. However, the 39-kDa precursor form of cathepsin B has been isolated and is known to have limited enzyme activity [9]. Therefore it seems likely that at least some of the plant cysteine proteinases are synthesised in the form of inactive precursors or zymogens [7, 26].

Since actinidin is a highly active proteinase and accounts for 50% of soluble fruit protein, the plant cell itself must in some way be protected against this activity. In analogy with other proteinases, it may be expected that actinidin is synthesised in the form of a zymogen and located within membrane-bound organelles, the lysosomes or plant vacuole. We decided to address this question by studying actinidin at the molecular level and have isolated and sequenced two actinidin-encoding cDNAs. Together with data from Northern blots and *in vitro* translations of mRNA the results show that actinidin is indeed synthesised as a precursor approximately 15 kDa larger than the mature protein, with both amino- and carboxy-terminal extensions. Results are also presented of an analysis of actinidin gene expression in developing fruit, and of an estimation of gene copy number by Southern blotting revealing the presence of a gene family.

Materials and methods

Materials

Chinese gooseberry fruit was purchased from local markets as a source of actinidin, and seeds from these fruits were used for growing plants. Nearly-mature and developing fruit, variety "Exbury", for cDNA cloning and Northern blot analysis were obtained from the Royal Horticultural Society's Gardens at Wisley, Surrey, and stored frozen in liquid nitrogen. Crystalline actinidin for raising antibodies was obtained as a gift from C.H. Moore, Palmerston North, New Zealand. The synthetic oligonucleotide, a 32-fold degenerate mixture of 15 nucleotides, was purchased from J. Kyte, Biochemistry Department, University of Leicester.

Methods

Plant extracts for actinidin measurements were prepared by grinding freeze-dried material to a fine powder in liquid nitrogen, homogenising in buffer containing 0.8% NaCl, 0.02% KH_2PO_4 , 0.29% $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 0.02% KCl and 0.05% Tween 20, and centrifuged to remove debris.

Antibodies were raised in rabbits by immunisation followed by two booster injections over 6 weeks with a total of 2 mg actinidin. The antiserum was prepared as described by Jurd [19]. Enzyme-linked immunosorbent assay was carried out as described by Voller *et al.* [35].

RNA was purified by a phenol method adapted from Covey and Grierson [8] and obtained from S. Curson, Norwich. Tissue was ground frozen in liquid nitrogen with pestle and mortar, and mixed with 1 cm³/g of frozen buffer containing 1% triisopropyl naphthalene sulphonate, 6% 4-aminosalicylate, 50 mM Tris-HCl pH 8.3 and 6% phenol mixture prepared from phenol with 0.1% 8-hydroxyquinoline, dissolved in 50 mM Tris-HCl pH 8.3, followed by 2.5 g sodium diethylthiocarbamate for each 100 g tissue. After rapid thawing a further 1 cm³/g of liquid buffer was added and the slurry extracted three times with phenol:chloroform:isoamylalcohol (50:50:1). Nucleic acids were precipitated with 0.2 M lithium acetate pH 6.0 and

2.5 volumes ethanol at -20°C , pelleted and resuspended in sterile distilled water (SDW). RNA was precipitated by addition of an equal volume 6 M lithium acetate pH 6.0, after which it was left at 4°C for 2 h, and collected by centrifugation at 10000 r.p.m. in a Sorvall HB-4 rotor for 10 min at 4°C . The pellet was washed twice with 3 M lithium acetate, dissolved in SDW and ethanol precipitated. Fruit RNA was prepared from skinned and deseeded fruit, and both lithium acetate precipitation and phenol extraction were repeated several times alternately to remove impurities. *In vitro* translation was carried out using rabbit reticulocyte lysate as advised by the manufacturers and incorporating [^{35}S]-methionine. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Hames [17] and immunoprecipitation as described by Lomedico *et al.* [22]. cDNA was synthesised using poly(A) RNA fractionated twice on oligo(dT) and followed the method of Gubler and Hoffman [16]. First-strand cDNA for use as probe in colony hybridisation was synthesised in the presence of [^{32}P] α -dATP. Double-stranded cDNA was C-tailed as described in the Leicester University cloning course manual [5], but using a reaction mixture containing 5% dideoxy CTP. cDNA was annealed to G-tailed pUC9 vector DNA [34] and transformed into *E. coli* JM83. Colony hybridisations were carried out by standard methods [25] at 35°C in $6 \times \text{SSC}$ overnight for end-labelled synthetic oligomer and at 65°C overnight in $6 \times \text{SSC}$ for first-strand cDNA probes, with washes at the same temperatures. cDNA inserts were restriction-mapped, subcloned into M13 mp18 and M13 mp19 [35], and sequenced on both strands by the chain termination method of Sanger *et al.* [31].

All other methods not described here followed the instructions in the Maniatis *et al.*'s Manual [24].

Results

Immuno-detection of actinidin in different parts of the plant

Plant extracts were initially subjected to rocket immunoelectrophoresis (results not shown) but only fruit extracts contained levels of actinidin detectable

by this method. The more sensitive enzyme-linked immunosorbent assay (ELISA) revealed traces of actinidin in other parts of the plant. However, as shown in Table 1, fruit contained a 3800-fold level of the protein compared with leaf, and 11 500 times more than apex tissue. Actinidin thus appears to be specifically synthesised in the fruit of the plant.

In vitro translation and immunoprecipitation

The *in vitro* translation products of total fruit RNA (Fig. 1A, lane 2) contained an abundant protein of

Table 1. Levels of actinidin in different parts of the plant determined by ELISA.

	Actinidin ($\mu\text{g/g}$ fresh wt.)	Actinidin ($\mu\text{g/g}$ dry wt.)
Leaf	0.65	6.4
Apex	0.23	2.1
Root	<0.03	<0.3
Fruit	4100	24100

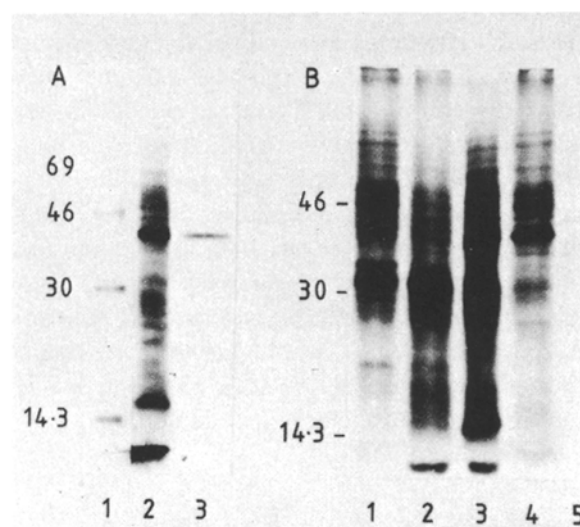


Fig. 1. Autoradiographs of SDS-polyacrylamide gels of *in vitro* translations. A. Lane 1, protein standards; lane 2, total RNA from mature fruit translated *in vitro*; lane 3, immunoprecipitate of fruit translation products. B. Templates for *in vitro* translation were: lane 1, poly(A) RNA from apex; lane 2, total RNA from apex; lane 3, total RNA from mature fruit; lane 4, poly(A) RNA from fruit; lane 5, immunoprecipitate of translation products from fruit poly(A) RNA. Size markers are in kDa.

approximately 39 kDa, which was immunoprecipitated by actinidin-specific antibodies (Fig. 1A, lane 3), suggesting that mRNA encoding a protein antigenically related to actinidin is very abundant in fruit. However, the relative abundance of the immunoprecipitated protein was much reduced in translation products directed by oligo(dT)-fractionated mRNA (Fig. 1B, lane 4), suggesting differential affinities of mRNA species for the oligo(dT) column. The major protein synthesised *in vitro* by fruit RNA was not present at high levels in translation products of young leaf (apex) RNA (Fig. 1B, lanes 1 and 2). This difference was subsequently exploited for the construction of a differential hybridisation probe. Assuming that the immunoprecipitate is indeed related to actinidin this must be synthesised as a 15 kDa larger precursor.

cDNA synthesis and screening of library

A library containing approximately 1000 recombinants was initially screened using a 32-fold degenerate, mixed synthetic oligomer of 15 nucleotides, corresponding to amino acids 190 to 194 of actinidin. However, colony hybridisation experiments using end-labelled synthetic oligomer were repeatedly unsuccessful. Therefore, the library was screened with a differential probe of labelled first-strand cDNA from fruit and apex. The fruit-specific probe was made more specific by priming the reverse transcription with the synthetic oligonucleotide instead of oligo(dT). Plasmid DNA from 90 colonies hybridising strongly with the fruit probe was immobilised on nitrocellulose and hybridised with end-labelled synthetic oligomer, and in this way ten positive clones were identified.

Sequence analysis of two cDNA clones

The majority of the cDNA inserts were between 500 and 600 bp long, but two clones had inserts of 1145 bp (pAC.1) and 809 bp (pAC.7). Their restriction maps (Fig. 2A) indicate coding and non-coding regions as well as the area of homology with actinidin, only partially encoded by pAC.7. Both

cDNAs extend into the poly(A) tail (Fig. 2B) and therefore represent the complete 3' untranslated regions. pAC.1 and pAC.7 are very similar indeed, showing four nucleotide differences in the coding, and three in the non-coding region, whilst pAC.1 has four additional nucleotides at the 3' end preceding the poly(A) tail. The differences at the nucleotide level result in the absence of two restriction sites in pAC.7. Three of the four nucleotide differences in the coding region result in amino acid substitution.

The 3' untranslated regions have three consensus polyadenylation signals, one of these 26 bp (pAC.1) or 22 bp (pAC.7) upstream of the polyadenylation site (Fig. 2B).

The sequence in pAC.1 which lies 5' to the region homologous with actinidin does not encode a methionine residue and therefore does not represent a full-length cDNA. In fact, since the size of the mRNA was determined by Northern blotting to be 1.4 kb (see below), it can be estimated that the complete cDNA would be extended by 200 to 250 bp.

Analysis of the deduced amino acid sequences

pAC.1 encodes a polypeptide of 302 amino acids of M_r 32709 Da. Residues 58 to 277 of the deduced sequence correspond to the whole of the known amino acid sequence of actinidin, showing 20 amino acid differences (Fig. 3). The deduced sequence of pAC.7 (PAC.7) covers residues 62 to 220 of actinidin, and within this region shows all but three of the differences encoded by pAC.1.

When compared to the published amino acid sequence of actinidin the deduced sequences show mainly conservative substitutions in regions affecting tertiary structure and catalytic function. A large proportion of differences, some non-conservative, are clustered in the region of poorest homology between all cysteine proteinases, containing most of the charged side chains affecting the pI of the proteins. The overall charge in this region of PAC.7 is identical to that of the published sequence, but the pAC.1-encoded protein (pAC.1) differs by having lost two negative charges. Two non-conservative replacements, asp-55 by gly and lys-181 by glu, are only found in PAC.1. Lys-181 is highly conserved between

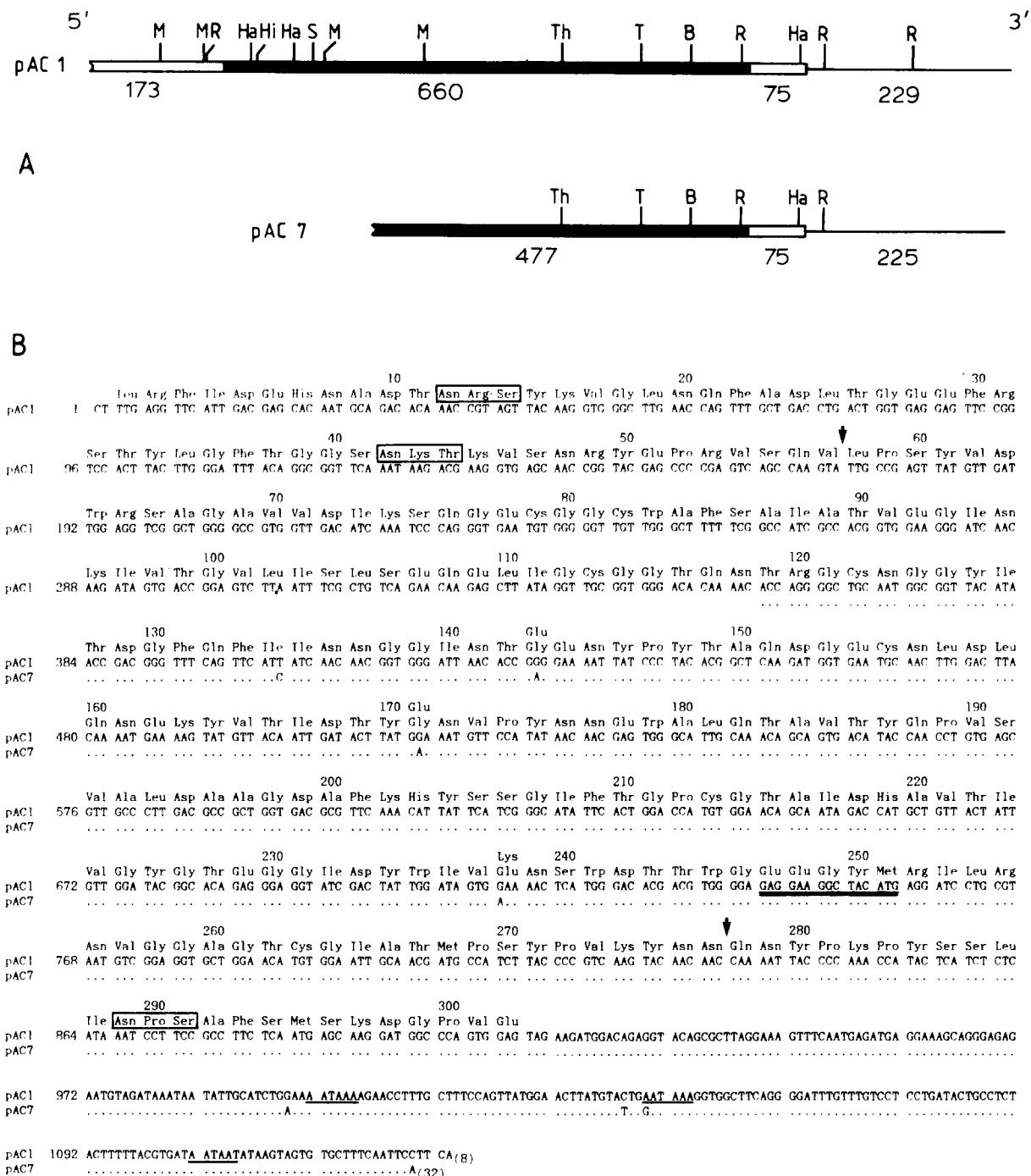


Fig. 2. A. Restriction maps of the two cDNA clones, pAC.1 and pAC.7. Black boxes indicate coding regions corresponding to mature actinidin, blank boxes code for sequences not present in the mature protein. Lines represent untranslated regions. Numbers refer to nucleotides in each region. Abbreviations: B = BamHI, Ha = HaeIII, Hi = HincII, M = MspI, R = RsaI, S = Sau3A, T = TaqI. **B.** Nucleotide sequence and deduced amino acid sequence of pAC.1 and pAC.7. Identical nucleotides in pAC.7 are indicated by dots below the sequence of pAC.1. Amino acids differing in pAC.7 are shown above the sequence deduced from pAC.1. Polyadenylation consensus signals are underlined, and the region corresponding to the synthetic oligonucleotide is heavily underlined. Arrows indicate probable proteolytic cleavage sites. Potential glycosylation sites are boxed.

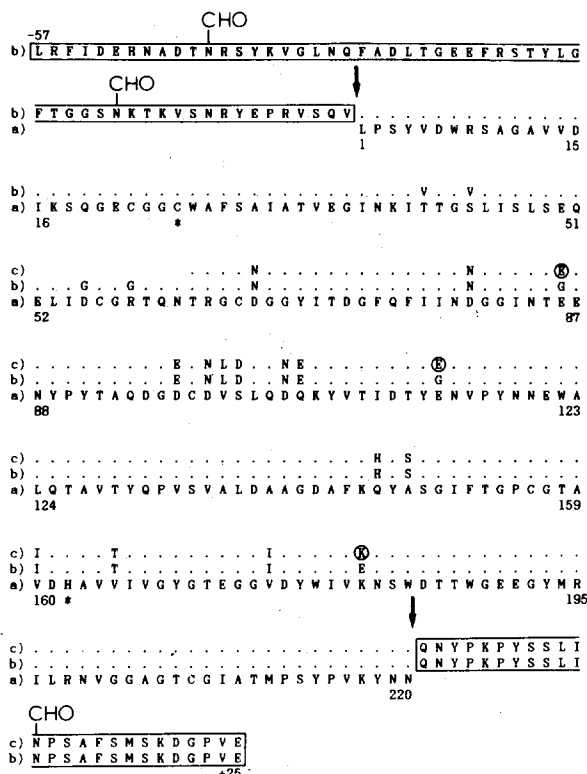


Fig. 3. Amino acid sequence deduced from cDNA clones aligned with published sequence. a) amino acid sequence determined by Carne and Moore [6] and amended by Baker [3]; b) sequence deduced from pAC.1; c) sequence deduced from pAC.7 (dots represent homologies with published sequence); asterisks denote the active site cysteine and histidine residues; arrows indicate possible proteolytic cleavage sites; boxed residues are not present in mature actinidin; potential glycosylation sites are marked by "CHO"; circles indicate residues in pAC.7 which differ from pAC.1 but are identical to the published sequence.

all the cysteine proteinases and has been implicated in electrostatic interaction with positively charged side chains across the active site cleft, therefore its replacement by glu may not be compatible with this role. Surprisingly, both cDNAs code for an additional 25 amino acids at the carboxy terminus of mature actinidin. It has an M_r of 2754 Da and contains a striking excess of Pro and Ser residues (16% and 20%, respectively). A potential glycosylation site, characterised by the canonical sequence Asn-X-Ser, is indicated in Fig. 3.

The amino-terminal 57 residue extension of PAC.1 has an M_r of 6508 Da, however, as it contains no met residue the identified sequence must

represent only part of a larger segment. It is characterised by a large proportion of polar residues, as illustrated by the hydropathy profile (Fig. 4), and is particularly rich in positively charged amino acids (16% of total amino acids, compared with 5% in the mature protein). There are two sites in this region which may be glycosylated (Fig. 3).

An alignment of this polypeptide with amino-terminal regions of papain [7], aleurain [29], and two cysteine proteinases from *Dictyostelium* [27] shows considerable homologies (between 30 and 35%), clustered in regions of strong conservation (Fig. 5). Based on these similarities in the actinidin precursor peptide can be predicted to contain a total of 120 to 140 amino acids.

Northern blot analysis of RNA at different developmental stages

Northern blots of total RNA from fruit at different developmental stages show a single band of 1.4 kb which hybridises to the cDNA probe (Fig. 6). Since the cDNA insert of pAC.1 is 1.15 kb long, the mRNA probably has additional coding potential for up to 70 amino acids, taking into account a short untranslated leader at the 5' end.

The developmental changes in levels of mRNA coding for actinidin are illustrated in Fig. 6. Actinidin mRNA was present at low levels in young

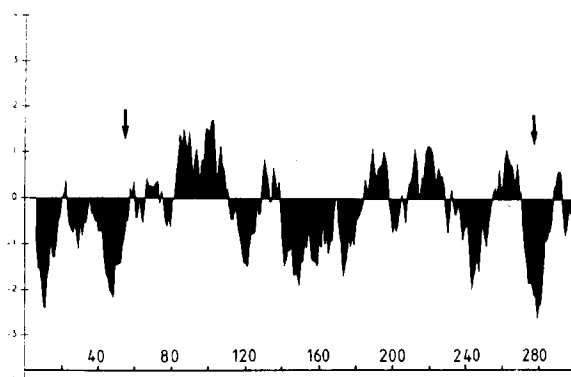


Fig. 4. Hydropathy plot of the amino acid sequence deduced from pAC.1. Numbers on the x-axis refer to amino acid residues, and those of the y-axis to relative hydrophobicity of moving window averages of 11 residues [21]. Arrows indicate proteolytic cleavage sites.

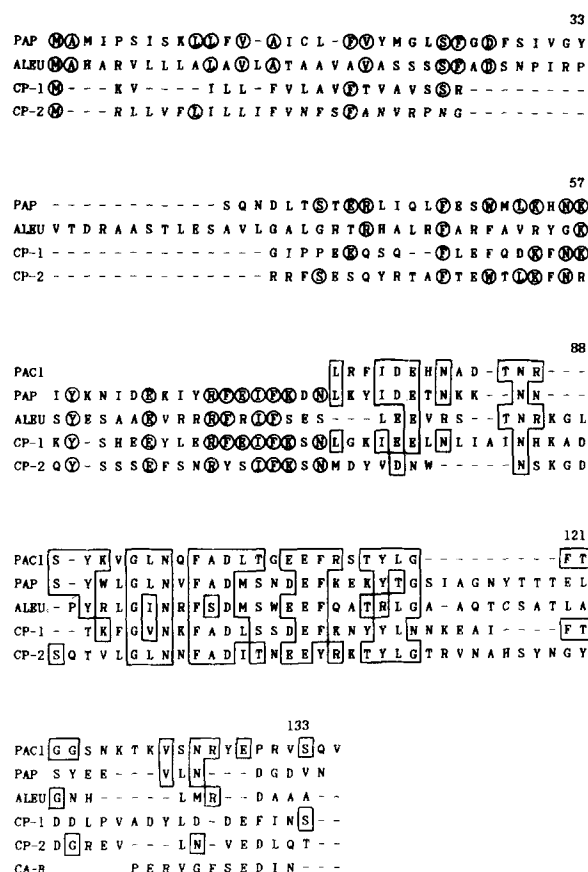


Fig. 5. Alignment of amino-terminal extensions of cysteine proteinases. PAC.1 = sequence deduced from pAC.1, PAP = papain, ALEU = aleurain, CP-1 and CP-2 = cysteine proteinase 1 and 2 (respectively) from slime mould, CA-B = cathepsin B. Boxed residues are homologous with the pAC.1-encoded sequence, and circled residues are homologous with the papain extension. Numbering is based on the papain sequence.

fruit, but increased rapidly when the fruit had reached about half its mature size. A maximum level appeared to be attained at about 64 days after pollination, less than half way to maturity, reached at approximately 160 days [28]. After harvest and during storage of the fruit at 10 °C the amount of hybridisable actinidin mRNA rapidly decreased. Although an 8-fold excess of young leaf RNA (Fig. 6, lane 1) showed a faint hybridising band of identical size, the proportion of actinidin-encoding mRNA must be at least as low as in undeveloped fruit.

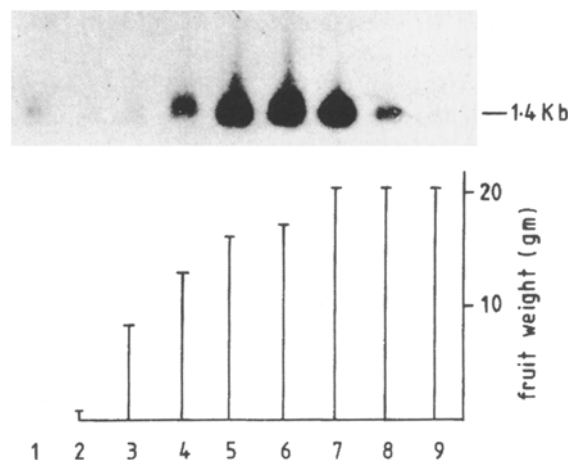


Fig. 6. Northern blot of total RNA from leaf (lane 1, 16 µg), and fruit (lanes 2–9, 2 µg) at different developmental stages. Fruit was harvested 7, 24, 42, 64, 82 and 102 days after pollination (lanes 2–7, respectively), and stored at 10 °C for 10 and 21 days (lanes 8 and 9, respectively). The weight of fruit (average of 10 fruits) at each stage is indicated below the Northern blot.

Southern blot

A Southern blot analysis of EcoRI- and HindIII-digested leaf DNA shows in each track approximately 14 fragments, hybridising at a low stringency to the cDNA probe (Fig. 7). One fragment of 4.1 kb of the HindIII-digested DNA hybridised particularly strongly. This could be due to perfect matching with the probe, or could reflect recent gene duplication giving rise to multiple, identical-sized, fragments. Since the whole of the pAC.1 cDNA was used as a probe, the lower hybridising bands must represent gene fragments, and therefore the total number of actinidin genes probably does not exceed 10.

Discussion

Two cDNAs encoding proteins with more than 90% homology to actinidin were isolated from a fruit-specific library of *Actinidia chinensis*. An analysis of their deduced amino acid sequences suggests that they represent isoforms if not isozymes of actinidin, since structurally important residues are

conserved. The pAC.1-encoded protein is likely to differ from actinidin to a greater extent than PAC.7 and it probably has a higher pI.

The existence of a family of closely related cysteine proteinases in *A. chinensis*, estimated at up to 10 members by Southern blot analysis, comes as no surprise as it is paralleled by the presence of protein families in all the other cases of abundant plant cysteine proteinases, but also of unrelated abundant plant proteins such as the seed storage proteins.

The small number of nucleotide differences between the two cDNA sequences, in the non-coding as well as the coding region, with three out of four substitutions non-silent in the latter, raises the possibility of a cloning artefact, such as that observed with a cDNA clone of aleurain [37]. However, examples of other multigene families, e.g. patatin, show many non-conservative differences between individual members [30], and in the case of soybean conglycinin, the 3' untranslated regions are even more highly conserved than the coding regions [33]. Thus it is possible that the two actinidin cDNAs rep-

resent two genes in the process of divergence following a recent duplication event. Due to the presence of several gene copies, selective constraint may not operate equally on all members of the family. It would be interesting to examine whether all of the genes encode functional enzymes.

As expected, actinidin was shown to be synthesised in the form of a larger precursor. Although only a portion of the prosegment was identified by cDNA cloning, the size of the precursor estimated at 39 kDa by immunoprecipitation of *in vitro* translation products agrees with that predicted from the size of the mRNA, and is similar to that of other cysteine proteinases.

The significance of the carboxy-terminal extension is not clear, although both human and rat cathepsin B have carboxy-terminal extensions of six amino acids, and in both cases the first residue is identical to that of the actinidin extension. Carboxy-terminal extensions of unrelated enzymes, such as cathepsin D and β -glucuronidase, have been implicated in intracellular transport to the lysosomes [11, 12].

The considerable sequence homology of the actinidin prosegment with those of other cysteine proteinases indicates that these extensions may have structural and functional similarities [26]. There is strong evidence that papain is synthesised in the form of a zymogen which becomes activated by the removal of the amino-terminal extension [7], and the precursor of cathepsin B is thought to have very limited proteolytic activity [9]. Recently the mechanism of inactivation of the protease activity in the aspartate proteinase pepsinogen has been shown to involve electrostatic interactions of a highly charged, transitory amino-terminal peptide with charged residues at and near the active site of the enzyme [18]. Both the amino-terminal extensions of papain and actinidin are highly hydrophilic, papain containing an excess of negative charges while actinidin has more positively charged residues in this region. The overall charge on the prosegment is in both cases the reverse of the net surface charge of the mature proteins. These observations are consistent with a mechanism of inactivation by electrostatic interactions between the prosegment and the surface of the mature enzyme domain. It would be interesting to de-

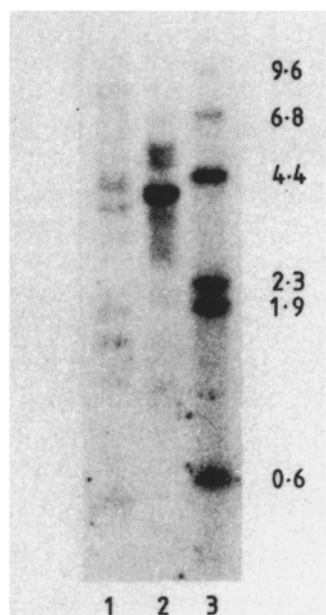


Fig. 7. Southern blot of leaf DNA. Lane 1, DNA digested with EcoRI; lane 2, DNA digested with HindIII; lane 3, λ -DNA digested with HindIII and radioactively labelled. The blot was hybridised with ^{32}P -labelled, nick-translated pAC.1 insert at 65 °C in $6 \times \text{SSC}$ and washed under the same conditions.

termine whether the actinidin precursor can cleave itself by an intramolecular process upon transfer to an acidic environment as has been described for pepsinogen, or whether mature actinidin or some other protease in the fruit of *A. chinensis* carries out this conversion. The presence of several potential glycosylation sites on the actinidin precursor is consistent with sequestration of the enzyme to lysosomal organelles. Correct targeting may be achieved partly through recognition of the precursor by receptors of the mannose-6-phosphate pathway [13]. Actinidin was shown to be specifically synthesised in the fruit of the plant with very low levels detected in other tissues. The difference in the levels of actinidin between fruit and young leaf tissue was due to different amounts of actinidin-encoding mRNA. It will be interesting to identify promoter sequences of the actinidin genes responsible for this organ-specific expression, and the availability of cDNA clones will be useful in the isolation of genomic sequences.

As yet nothing is known about the physiological function of actinidin. The results of the developmental study of actinidin gene expression show that actinidin mRNA accumulates to very high levels early during fruit development, and therefore does not belong to the class of "ripening-specific" mRNAs described for the tomato fruit [15]. The enzyme appears to be required during fruit development, and it is tempting to speculate that it is involved in the defence against invading predators, such as larvae or fungi, and is released from storage vesicles upon damage to the tissue.

Acknowledgements

We wish to thank C.H. Moore for the gift of crystalline actinidin, and H.A. Baker of Wisley Gardens for a generous supply of Kiwifruit. We are grateful to S. Adams, A. Greenland and L. and I. Eperon for much helpful advice during cDNA cloning and sequencing, and to S. Rosahl for helpful discussions.

This work was supported by an SERC Biotechnology Directorate Research Studentship to U.M.P.

References

1. Arcus AC: Proteolytic enzyme of *Actinidia chinensis*. *Biochim Biophys Acta* 33: 242–244 (1959).
2. Baker EN: 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 (1977).
3. Baker EN: Structure of actinidin, after refinement at 1.7 Å resolution. *J Mol Biol* 141: 441–484 (1980).
4. Baker EN: Comparison of the refined structures of two sulphhydryl proteases: actinidin and papain. In: Dodson G, Gluster JP, Jayre D (eds) Clarendon Press, Oxford (1981) pp 339–349.
5. Boulnois GJ (ed.): *Gene Cloning and Analysis: a Laboratory Manual*. Blackwell Scientific Publications Ltd, Oxford (in press).
6. Carne AC, Moore CH: The amino acid sequence of the tryptic peptides from actinidin, a proteolytic enzyme from the fruit of *Actinidia chinensis*. *Biochem J* 173: 73–83 (1978).
7. Cohen LW, Coghlan VM, Dihel LC: Cloning and sequencing of papain-encoding cDNA. *Gene* 48: 219–227 (1968).
8. Covey SN, Grierson D: The measurement of plant polyadenylic acid by hybridisation with radioactive polyuridylic acid. *Planta* 131: 75–79 (1976).
9. Docherty K, Hutton JC, Steiner DF: Cathepsin B-related proteases in the insulin secretory granule. *J Biol Chem* 259: 6041–6044 (1984).
10. Drenth J, Jansonius JN, Koekoek R, Wolthers BG (1968): The structure of papain. *Adv Protein Chem* 25: 79–115 (1971).
11. Erickson AH, Blobel G: Carboxy-terminal proteolytic processing during biosynthesis of the lysosomal enzymes β -glucuronidase and cathepsin D. *Biochemistry* 22: 5201–5205 (1983).
12. Faust PL, Kornfeld S, Chirgwin JM: Cloning and sequencing cDNA for human cathepsin D. *Proc Natl Acad Sci USA* 82: 4910–4914 (1985).
13. Figura K von, Hasilik A: Lysosomal enzyme and their receptors. *Ann Rev Biochem* 55: 167–193 (1986).
14. Goto K, Murachi T, Takahashi N: Structural studies on stem bromelain: isolation, characterisation and alignment of the cyanogen bromide fragments. *FEBS Lett* 62: 93–95 (1976).
15. Grierson D, Slater A, Speirs J, Tucker GA: The appearance of polygalacturonase mRNA in tomatoes: one of a series of changes in gene expression during development and ripening. *Planta* 163: 263–271 (1983).
16. Gubler Y, Hoffman B: A simple and very efficient method for generating cDNA libraries. *Gene* 25: 263–269 (1983).
17. Hames BD: An introduction to polyacrylamide gel electrophoresis. In: Hames BD, Rickwood D (eds) *Gel Electrophoresis of Proteins: a Practical Approach*. IRL Press Ltd, Oxford (1981) pp 1–91.
18. James MNG, Sielecki AR: Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 Å resolution. *Nature* 319: 33–38 (1986).

19. Jurd RD: Immuno-electrophoresis. In: Hames BD, Rickwood D (eds) *Gel Electrophoresis of Proteins: a Practical Approach*. IRL Press, Oxford (1981) pp 229–248.
20. Kamphius IG, Drenth J, Baker EN: Thiol proteases. Comparative studies based 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 (1985).
21. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105–132 (1982).
22. Lomedico PT, Gubler U, Hellman CP, Dukovich M, Giri JG, Pan Y-CE, Collier K, Semionow R, Chua AO, Mizel SB: Cloning and expression of murine interleukin-1 cDNA in *Escherichia coli*. *Nature* 312: 458–462 (1984).
23. Lowe G: The cysteine proteinases. *Tetrahedron* 32: 291–302 (1976).
24. Maniatis T, Fritsch EF, Sambrook J: *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbour Laboratory, Cold Spring Harbour NY (1982).
25. McDowall MA: Anionic proteinase from *Actinidia chinensis*. *Eur J Biochem* 14: 214–221 (1970).
26. North MJ (1986): Homology within the N-terminal extension of cysteine proteinases. *B J Letters* 238: 623–624 (1986).
27. 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 (1985).
28. Pratt HK, Reid MS: Seasonal patterns in fruit growth and maturation, ripening, respiration and the role of ethylene. *J Sci Fd Agric* 25: 747–757 (1974).
29. Rogers JC, Dean D, Heck GR: Aleurain: a barley thiol protease closely related to mammalian cathepsin H. *Proc Natl Acad Sci USA* 82: 6512–6516 (1985).
30. Rosahl S, Schmidt R, Schell J, Willmitzer L: Isolation and characterisation of a gene from *Solanum tuberosum* encoding patatin, the major storage protein of potato tubers. *Mol Gen Genet* 203: 214–220 (1986).
31. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain termination inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467 (1977).
32. San Segundo B, Chan SJ, Steiner DF: Identification of cDNA clones encoding a precursor of rat liver cathepsin B. *Proc Natl Acad Sci USA* 82: 2320–2324 (1985).
33. Schuler MA, Schmitt ES, Beachy RN: 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 (1982).
34. Vieira J, Messing J: The pUC plasmids, an M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259–268 (1982).
35. Vieira J, Messing J: A new pair of M13 vectors for selecting either DNA strand of double-digested restriction fragments. *Gene* 19: 269–276 (1982).
36. Voller A, Bidwell DE, Bartlett A: The enzyme-linked immunosorbent assay (ELISA). Dynatech Europe Guernsey (1979).
37. Whittier RF, Dean DA, Rogers JC: Nucleotide sequence analysis of alpha amylase and thiol protease genes that are hormonally regulated in barley aleurone cells. *Nucleic Acids Res* 15: 2515–2535 (1987).