Biochem. J. (1991) 274, 231–236 (Printed in Great Britain)

Purification and characterization of thaumatopain, a cysteine protease from the arils of the plant *Thaumatococcus daniellii*

Maggie CUSACK,* Andrew G. STEPHEN, Roy POWLS and Robert J. BEYNON†
Proteolysis Research Group, Department of Biochemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K.

Aqueous extracts of the aril of the seed of *Thaumatococcus daniellii* contain, in addition to the intensely sweet protein thaumatin, a cysteine protease that we have termed thaumatopain. Thaumatopain has been purified by ion-exchange chromatography from arils, and is a monomeric protein of M_r 30000. The protease strongly resembles papain in proteolytic activity, pH optima, susceptibility to inhibitors of cysteine proteases and in *N*-terminal sequence. The protease has also been identified in crude aril extracts by affinity labelling with iodo[^{14}C]acetate. Thaumatopain is responsible for the cysteine protease activity previously attributed to thaumatin. Thaumatin is digested by thaumatopain at neutral to alkaline pH values.

INTRODUCTION

The intensely sweet protein thaumatin is derived from the west African shrub *Thaumatococcus daniellii* (benth) (Van der Wel & Loeve, 1972; Higginbotham, 1979; Van der Wel & Ledeboer, 1989). Thaumatin constitutes about 50% of the dry mass of the aril of *Thaumatococcus daniellii*. It is a protein of M_r 22000, possessing eight disulphide bridges, and is a member of a family of related proteins (Cusack & Pierpoint, 1988). The role for thaumatin is unknown, but the acute sensory properties may aid in seed dispersal. Alternatively, it may function as a storage protein for the germinating embryo. The latter role would require hydrolysis of thaumatin in order to release amino acids for utilization by the embryo. In this respect, it was of interest to note the claims by Van der Wel & Bel (1980) that purified thaumatin possessed intrinsic proteolytic activity and was able to effect autodigestion. The proteolytic activity appeared after treatment with reducing agents, raising the possibility that thaumatin exhibited cysteine protease activity. However, the sequence of thaumatin revealed that histidine, a residue thought to be essential for the catalytic activity of 'classical' cysteine proteases, was not present. The purported cysteine protease activity of thaumatin was thus enigmatic.

We demonstrated that the cysteine protease activity of thaumatin preparations was also inhibitable by the highly specific cysteine protease inhibitor E-64 such that the inhibitor could be used as an active-site titrant for the protease. When thaumatin preparations were titrated with E-64, the cysteine protease activity was inhibited completely at an equivalence point equal to 1/100 of the molar quantity of thaumatin, indicating that the cysteine protease activity was attributable to a trace amount of a contaminating protein (Cusack *et al.*, 1987). In a separate investigation, we have been using limited proteolysis to study the relationship between structure and sensory properties of thaumatin (Shamil *et al.*, 1990). As part of this study, we have examined further the cysteine protease activity in order to eliminate it from thaumatin preparations, and to assess it as a potential new probe of thaumatin structure. We report here the purification and properties of the protease, which we have termed thaumatopain (a contraction of '*Thaumatococcus* papain') to indicate its origin and strong resemblance to the archetypical cysteine protease, papain.

MATERIALS AND METHODS

Materials

Arils from the ripe fruit of *Thaumatococcus daniellii*, a gift from Tate & Lyle Group Research and Development (Reading, Berks., U.K.), were gathered in the Ivory Coast and were shipped frozen to the U.K. Z-Phe-Arg-7-amido-4-methylcoumarin (ZFRNMec), E-64 and cystatin were from Sigma, Poole, Dorset, U.K. Papain was from BCL, Lewes, Sussex, U.K. Iodo[^{14}C]acetic acid (CFA 269; 2 GBq/ml) and Na^{125}I (IMS 30) were from Amersham International, Amersham, Bucks., U.K., and [^{3}H]formaldehyde (NET 099; 2 GBq/ml) was from NEN–Dupont (U.K.), Stevenage, Herts., U.K. Iodogen was from Pierce Chemical Co., Cambridge, U.K. Optiphase 'Safe' was from Pharmacia–LKB, Bromma, Sweden. X-ray film (NIF RX 100) was from Fuji. All other chemicals were of analytical grade.

Purification of thaumatopain from arils

Whole arils (250 g) were homogenized at 4 °C in 500 ml of water with a Polytron homogenizer. After a low-speed centrifugation at 1400 *g* for 10 min (4 °C) to remove cell debris, the supernatant was centrifuged at 74000 *g* for 2 h. The supernatant (470 ml) was stored at −20 °C until required, with no loss of thaumatopain protease activity.

The aril extract (0.5–1.0 ml) was filtered through a 0.25 μm-pore filter and applied to a 1 ml column of MonoS (Pharmacia–LKB) equilibrated in 20 mM-Hepes buffer, pH 7.4 (Buffer A). Thaumatopain was retained under these conditions, and the column was subsequently developed at 1.0 ml/min by a two-stage gradient with Buffer B (20 mM-Hepes, pH 7.4, containing 1 M-NaCl); the first stage comprising 0–14% B over 14 min, followed by the second stage (14–100% buffer B over 6 min). Fractions (0.5 or 1.0 ml) were collected and assayed for thaumatopain activity with ZFRNMec or ^{125}I-insulin B chain as substrates. Active fractions were subsequently pooled, diluted and rechromatographed under the same conditions to eliminate low levels of contaminating thaumatin from the preparation.

Assay of protease activity

ZFRNMec. Cysteine protease activity was routinely assayed with the fluorogenic substrate ZFRNMec. The enzyme samples

Abbreviation used: ZFRNMec, Z-Phe-Arg-7-amido-4-methylcoumarin.

* Present address: Department of Geology and Applied Geology, University of Glasgow, Glasgow G12 8QQ, U.K.
† To whom correspondence should be addressed.

were preincubated with 25 mM-2-mercaptoethanol in 100 mM-Hepes, pH 7.5, for 10 min in a final volume of 100 μl to activate the cysteine protease. Subsequently, a solution (2.0 ml) of 5 μM-ZFRNMec in 100 mM-Hepes, pH 7.5, was added, and the formation of product was monitored fluorimetrically in a Perkin–Elmer 3000 recording fluorimeter (excitation at 350 nm, emission 460 nm). The fluorimeter was calibrated with standard solutions of the product, 7-amino-4-methylcoumarin.

^{125}I-Insulin B chain. The performic acid-oxidized B chain of bovine insulin was iodinated with Iodogen (1 mg/ml iodination solution) to a specific radioactivity of 60000 d.p.m./μg. After activation of the fractions with 25 mM-2-mercaptoethanol in 100 mM-Hepes, pH 7.5, for 20 min, the assay was conducted in 100 mM-Hepes, pH 7.5 (200 μl), containing 5 μg of ^{125}I-insulin B chain. At suitable times throughout the digestion, samples (80 μl) were removed and added to 80 μl of 2% (w/v) casein as a coprecipitant, followed by 160 μl of 25% (w/v) trichloroacetic acid. After centrifugation at 10000 *g* for 3 min, a portion of the supernatant (160 μl) containing trichloroacetic acid-soluble digestion products was removed for determination of γ-radioactivity. Another portion of the digestion mixture was removed to establish the specific radioactivity of the substrate at the time of assay.

[^{3}H]Casein. True proteolytic activity towards a large protein was determined with [^{3}H]casein, prepared by the procedure of Maurizi (1987). Digestion reactions were conducted in a final volume of 1.0 ml and were initiated by addition of 2 μg of [^{3}H]casein. At suitable times, samples (100 μl) of the reaction mixture were removed and added to 100 μl of 10% trichloroacetic acid. The mixture was centrifuged at 10000 *g* for 2 min, and the radioactivity in 100 μl of the supernatant was determined by scintillation counting.

pH-optima determination

The optimal pH for activity of thaumatopain was determined for ZFRNMec and [^{3}H]casein. Thaumatopain was preincubated in a series of buffers spanning the range pH 3–9.5, and containing 30 mM-2-mercaptoethanol, for a period of 15 min. After this period, substrate ZFRNMec was added as described previously. The buffers used were citrate (pH 3.0–5.5), acetate (pH 4.0–5.0), pyrophosphate (6.0–7.0), phosphate (pH 6.5–7.5), Hepes (pH 7.0–8.0) and ethanolamine (pH 9.0–9.5). The final concentration of each buffer species was 50 mM, and the ionic strength was maintained at 0.2 by addition of the appropriate amount of NaCl. The buffer compositions were calculated by using a computer program that makes appropriate adjustments to the thermodynamic pK_a, taking into account temperature and ionic strength (Beynon, 1988).

SDS/PAGE

This was conducted by the method of Studier (1973), on 17.5%-acrylamide gels. Gels were stained for protein with Fast Green, or were electroblotted overnight on to nitrocellulose (0.2 μm; Schleicher and Schuell) at 80 mA in a buffer comprising 25 mM-Tris/HCl, 192 mM-glycine, 20% (v/v) methanol and 0.1% SDS, pH 8.3.

Peptide fragment analysis by h.p.l.c.

The fragmentation pattern of the performic acid-oxidized insulin B chain was compared when the peptide was incubated with thaumatopain or papain. Insulin B chain (500 μg) was incubated with preactivated proteases (5 μg of papain, 6 μg of thaumatopain) in 100 mM-Hepes buffer, pH 7.5. At suitable times, samples were removed and adjusted to 1% (v/v) trifluoroacetic acid and separated on reverse-phase h.p.l.c. on a

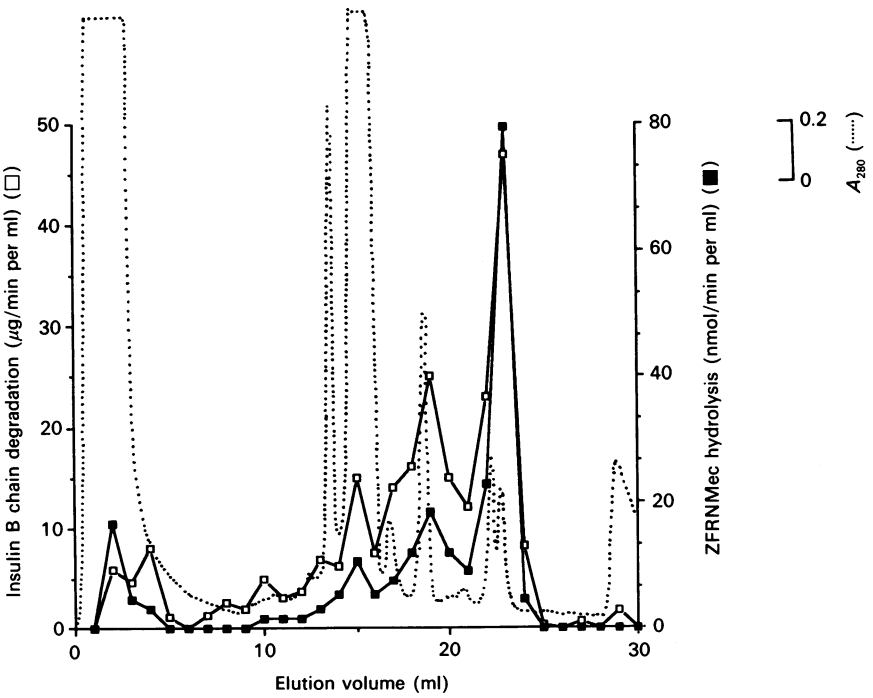

Fig. 1. MonoS cation-exchange chromatography of crude aril extract

A sample of crude aril extract (1 ml; 8.7 mg of protein) was applied to a 1 ml column of MonoS equilibrated in 20 mM-Hepes, pH 7.5, and eluted with a NaCl gradient from buffer A (20 mM-Hepes, pH 7.5) to buffer B (20 mM-Hepes/1.0 M-NaCl, pH 7.5). The A_{280} of the eluate was monitored (dotted line, representing an absorbance range of 0–2.0). The gradient programme was 0–6 ml, 0% B; 6–20 ml, 6–24% B; 20–26 ml, 14–100% B, 26–30 ml, 100% B. The gradient profile has been omitted for clarity. Fractions (1 ml) were collected and assayed for proteolytic activity towards ^{125}I-insulin B chain (□) and ZFRNMec (■) as described in the Materials and methods section.

3 μm Spherisorb ODS-2 column (125 mm × 4 mm) equilibrated in 0.1 % trifluoroacetic acid, and was subsequently developed with a linear gradient to 0.1 % trifluoroacetic acid/60 % (v/v) acetonitrile. The flow rate was 1.0 ml/min throughout. The eluted peptides were monitored at 206 nm, to identify the full complement of peptides. Other analyses were conducted with ^{125}I-insulin B chain as substrate, and the reverse-phase-h.p.l.c. eluate was collected for γ-radiation counting to identify iodinated peptides. The insulin B chain is iodinated approximately equally at both tyrosine residues (R. J. Beynon, unpublished work).

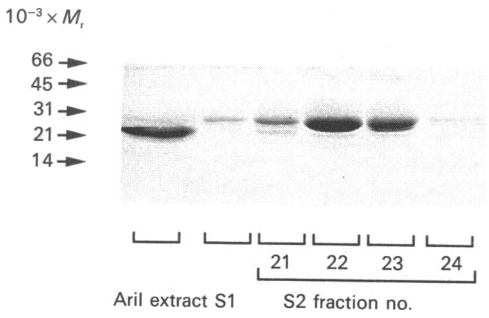

Fig. 2. SDS/PAGE of purified thaumatopain

Aril extract (8 μg of protein) and fractions 21–24 (2.5, 7.4, 6.7 and 0.6 μg of protein respectively) of the material rechromatographed on MonoS were separated by reducing SDS/PAGE (17.5 % gel) and stained for protein with Fast Green. Key: S1, first MonoS eluate; S2, second MonoS eluate.

Sequence analysis

Thaumatopain (twice chromatographed on MonoS) was subjected to *N*-terminal sequence analysis by the S.E.R.C.-supported facility at Aberdeen University. The protein (100 pmol) was sequenced on an Applied Biosystems 477A sequenator.

RESULTS AND DISCUSSION

Purification of thaumatopain

We had previously observed that commercial preparations of thaumatin were contaminated by trace levels of a protease (Cusack *et al.*, 1987). Aqueous extracts of arils of *Thaumatococcus daniellii* consists primarily of the thaumatins (M_r approx. 22000), but contain, in addition, a number of other proteins that are present at much lower levels. Initial experiments established that the proteolytic activity (assayed towards the fluorogenic peptide derivative ZFRNMec or ^{125}I-insulin B chain) was also present in aril extracts and could bind to MonoS cation exchanger. When aril extracts were applied to MonoS, a linear salt gradient resolved the multiple thaumatin species, but the protease was eluted as a broad peak at a higher ionic strength than the thaumatins. This peak was sharpened by application of a two-phase gradient, such that the proteolytic activity was eluted in the second sharp phase (Fig. 1). A similar elution profile was observed irrespective of whether ZFRNMec or ^{125}I-insulin B chain was used as substrate. At this stage, the active fractions consisted primarily of a peak (V_e = 23 ml) which on SDS/PAGE consisted of a major band of M_r approx. 30000 and a small amount of thaumatin. A second peak of activity (V_e = 19 ml) was associated with a thaumatin peak. Re-chromatography of

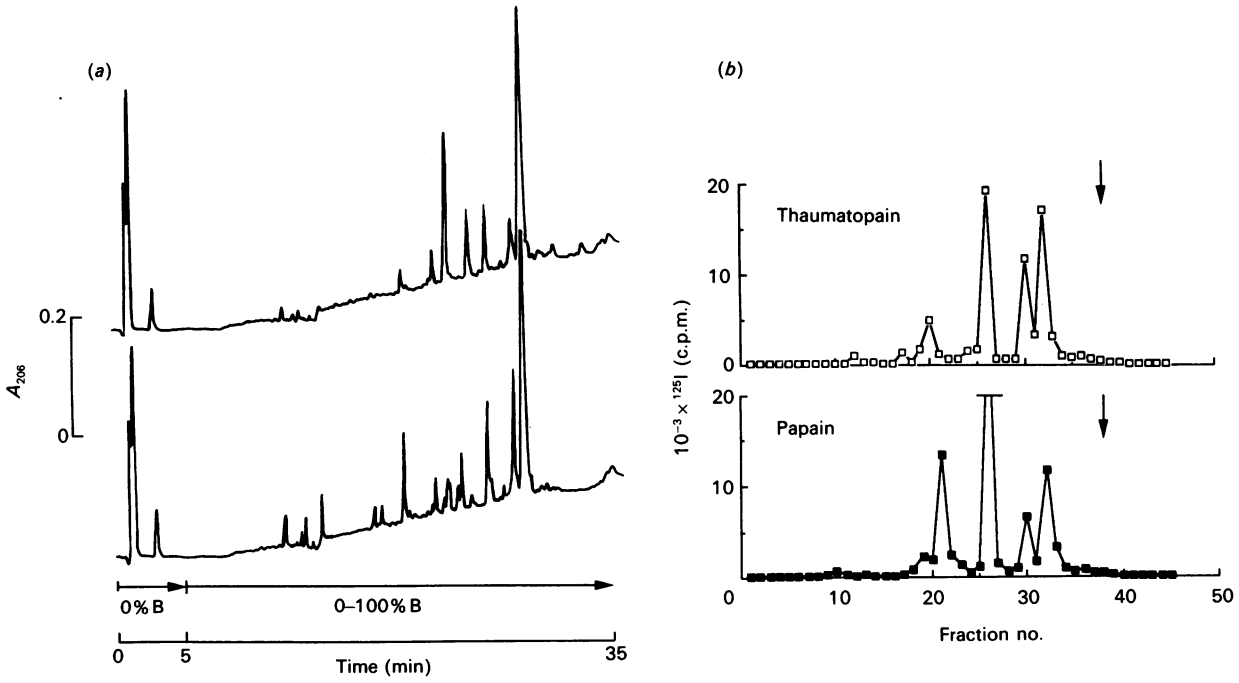

Fig. 3. Degradation of insulin B chain by thaumatopain

Insulin B chain (500 μg) was partially digested in a final volume of 1.0 ml of 100 mM-Hepes, pH 7.5, with 6 μg of thaumatopain (panel *a*, upper trace) or 5 μg of papain (panel *a*, lower trace), each of which was previously activated by incubation for 20 min in the presence of 25 mM-2-mercaptoethanol. At suitable times, samples were removed, adjusted to 1 % (v/v) trifluoroacetic acid and separated on reverse-phase h.p.l.c. The sample was applied in 0.1 % trifluoroacetic acid and the column was developed with a gradient of 0–60 % acetonitrile (in 0.1 % trifluoroacetic acid throughout) at a gradient of 2 % acetonitrile/min. The eluate was monitored at 206 nm (panel *a*). In parallel assays (panel *b*), incubations were conducted with 7.5 μg of ^{125}I-insulin B chain, and the h.p.l.c. eluate was fractionated to identify iodinated peptides. In this incubation, ^{125}I-insulin B chain, which would have been eluted in fraction 38 (indicated by the arrow), was digested to completion (60 min).

the pooled active fractions on MonoS under the same conditions yielded a preparation that was pure on SDS/PAGE (Fig. 2). After two cycles of ion-exchange chromatography, the recovery of activity was 52%, with a 22-fold purification. The M_r of this material by SDS/PAGE was 29800 ± 400 (mean $\pm$ S.D., $n = 4$). The protease activity was partially resolved into two components on both MonoS separations. These two peaks, which can be completely resolved by hydrophobic chromatography on phenyl-Superose (A. G. Stephen, unpublished work) have very similar proteolytic activities. All studies described here were therefore conducted with the pooled material from the second MonoS column.

Catalytic properties of thaumatopain

Preliminary experiments (M. Cusack & R. J. Beynon, unpublished work) established that thaumatopain was insensitive to classical inhibitors of serine, aspartic or metallo- proteinases. However, little or no proteolytic activity could be detected unless the thaumatopain-containing samples were first incubated with a reducing agent such as 2-mercaptoethanol (optimally for 10 min with 25 mM reductant). This observation, together with the complete inhibition of thaumatopain by reagents such as iodoacetate (Fig. 5) and E-64 confirmed that thaumatopain was a true cysteine protease. Inhibition of thaumatopain by E-64 was rapid, and this inhibitor could be used as an active-site titrant for thaumatopain. The concentration of thaumatopain in aril extracts was 0.29 mg/ml (assuming $M_r = 30000$), compared with a protein content of approx. 10 mg/ml. By contrast, the pure thaumatopain (protein concn. 0.96 mg/ml), although chromatographically homogeneous, contained 0.48 mg of active thaumatopain/ml. It is likely that the preparation contains some inactive thaumatopain. Active-site titration with cystatin revealed an apparent thaumatopain concentration approximately twice that determined with E-64; we attribute this discrepancy to the ability of the inactive fraction of thaumatopain to bind and thus to sequester cystatin. Cystatin can bind cysteine proteases that have been inactivated by modification of the active-site cysteine residue, and indeed is purified by affinity chromatography on immobilized carboxymethyl-papain (Anastasi *et al.*, 1983). Inhibition of thaumatopain by cystatin confirms that the enzyme resembles papain and ananain in properties more than stem bromelain or clostripain (Anastasi *et al.*, 1983).

The second-order rate constant for inhibition of thaumatopain by E-64 was determined [by non-linear curve fitting (Beynon, 1985) of the exponential inactivation of the enzyme during continuous fluorogenic assay] to be $3.6 \times 10^4\ M^{-1} \cdot s^{-1}$ (0.22 μM-E-64; mean of two determinations), a value similar in magnitude to that obtained for many other cysteine proteases that are papain-like, and higher than that obtained for bromelain and clostripain (Barrett *et al.*, 1982; Rowan *et al.*, 1988). Under the conditions of this experiment, the second-order rate constant was independent of inhibitor concentration ($3.4 \times 10^4\ M^{-1} \cdot s^{-1}$ at 0.15 μM-E-64).

The proteolytic activity towards ZFRNMec in crude aril extracts is optimally active at neutral pH values. In the absence of reducing agents, little or no activity can be measured at any pH value between 3.5 and 9.5. Thus all of the proteolytic activity across the whole of this pH range is attributable to a cysteine protease. Purified thaumatopain is optimally active at neutral pH values. ZFRNMec was hydrolysed optimally at pH 6.5, with substantial diminution of activity at higher or lower pH values, whereas [^{3}H]casein digestion was optimally active at neutral pH, with a slow decline of activity at higher or lower pH values. This behaviour is similar to that of papain (Kimmel & Smith, 1954).

Specificity of thaumatopain

Thaumatopain and papain were compared in terms of the pattern of hydrolysis of insulin B chain. After 20 min digestion, 5 μg of papain had degraded 190 μg of insulin B chain, and in the same time interval, 6 μg of thaumatopain had degraded 205 μg of the peptide. When the products were resolved by reverse-phase h.p.l.c., the numbers of peptides and their elution positions were broadly similar, irrespective of the proteinase used (Fig. 3). Additionally, digestion of the radioiodinated insulin B chain with papain or thaumatopain indicated a similar pattern of elution of radiolabelled peptides (those that contain Tyr-16 or Tyr-26). By several criteria we consider that thaumatopain and papain are similar in catalytic activity and primary specificity.

N-Terminal sequence of thaumatopain

This similarity between the two proteinases is reflected in the *N*-terminal sequences. Unmodified thaumatopain (the eluate from the second MonoS chromatographic separation) was sequenced to a total of 35 residues, which showed strong similarity to many other plant cysteine proteases (Fig. 4). The cycles 22 and 25 are both blank, but, by analogy with other members of this family, it is likely that these are cysteine residues which would not be identified during sequencing of the unmodified protein. In

```
                          1                   2                   3
                1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4

Thaumatopain    N L P N S V D W W K K G A V A A V K N Q R X X G S X X A F S S I K T S

Papain          - 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 V V T I
pplII           - L P E N V D W R K K G A V T P V R H Q G S C G S C W A F S A V A T V
pplV            - L P E S V D W R A K G A V T P V K H Q G Y C E S C W A F S T V A T V
Chymopapain     - Y P Q S I D W R A K G A V T P V K N Q G A C G S C W A F S T I A T V
Actinidin       - 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 I A T V
vigcp           - V P A S V D W R K K G A V T D V K D Q G Q C G S C W A F S T I V A V
Tomato          - L P E S I D W R E K G V L V G V K D Q G S C G S C W A F S A V A A M
Aleurian        - 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 T G A L
Bromelain       A 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 A A I A T V
```

Fig. 4. *N*-Terminal sequence analysis of thaumatopain

Thaumatopain was subjected to automated pulsed liquid-phase sequencing. The *N*-terminal sequence is aligned with those of the plant cysteine proteases papain (reference numbering system; Cohen *et al.*, 1986), papaya protease III (ppIII; Dubois *et al.*, 1988), papaya protease IV (ppIV; Ritonja *et al.*, 1989*b*), chymopapain (Jacquet *et al.*, 1989), actinidin (Carne & Moore, 1978; Baker, 1980), mung-bean seed protease (vigcp; Akasofu *et al.*, 1989), tomato cysteine protease (tomato; Schaffer & Fischer, 1988), barley (aleurain; Rogers *et al.*, 1985) and stem bromelain (Ritonja *et al.*, 1989*a*). Unassigned residues are indicated by X.

particular, Cys-25 is the nucleophilic centre of the cysteine proteases. Residue 23 is a glycine residue, and, by analogy with papain, is a major contributor to the S_1 substrate-binding pocket. This Gly-25 is conserved in most other cysteine proteases, an exception being papaya proteinase IV, in which this residue is glutamic acid (Ritonja *et al.*, 1989*b*), a substitution that is reflected in the inability of this enzyme to accommodate with ease any residue other than a glycine at the S_1 position (Buttle *et al.*, 1990). The overall identity between thaumatopain and papain is 18 residues in a total of 30 identified amino acids; a further four amino acids constitute conservative substitutions. By all criteria, thaumatopain should be considered as a 'papain' from this species.

Relationship of thaumatopain to the proteolytic activity in crude thaumatin preparations

Thaumatin had previously been claimed to possess cysteine protease activity (Van der Wel & Bel, 1980). We had shown that the cysteine protease activity in commercial preparations of thaumatin (Talin) could be suppressed by an amount of E-64 corresponding to 1/100th of the amount of protein in the preparation, and thus was likely to be a contaminating protease in the preparation (Cusack *et al.*, 1987). It was of interest, however, to extend those studies, as the original claim of thaumatin activity was made with preparations of 'pure' thaumatins. First, we attempted active-site labelling of the cysteine protease activities in crude aril extracts with iodo[^{14}C]acetate (Fig. 5). A portion of aril extract was treated with iodo[^{14}C]acetate to inhibit the cysteine protease completely. At the same time, a parallel incubation was conducted, except that the cysteine protease activity was inhibited with E-64 before radiolabelling. Both preparations were then resolved by SDS/PAGE and subjected to autoradiography. Faint labelling was evident in both preparations, coincident with the large band of thaumatins (accounting for over 95% of the protein in the aril extract). A second, heavily labelled, band was coincident with thaumatopain, but only in the sample that had not been exposed to E-64. We interpret this experiment to mean that all of the cysteine protease activity in aril extracts was due to thaumatopain, although there was slight non-specific labelling of thaumatins. It is possible that treatment with the reducing agent had exposed one pair of cysteine residues such that slight labelling occurred. However, the extent of labelling must also be seen in the context of the relative masses of the two proteins in the aril extract (thaumatin, > 8.5 mg/ml; thaumatopain, 0.3 mg/ml).

Although thaumatopain itself chromatographed as a very cationic protein, in the first separation of aril extract there was some evidence for cysteine protease activity that was coincident with thaumatin peaks (Fig. 1; fractions 15 and 19), identified by Western blotting with an antiserum to thaumatin I (results not shown). Accordingly, these peak fractions, in addition to fraction 23 (thaumatopain) were titrated with E-64 and separated by SDS/PAGE. Fraction 19, which showed substantial activity towards ZFRNMec and ^{125}I-insulin B chain, contained about

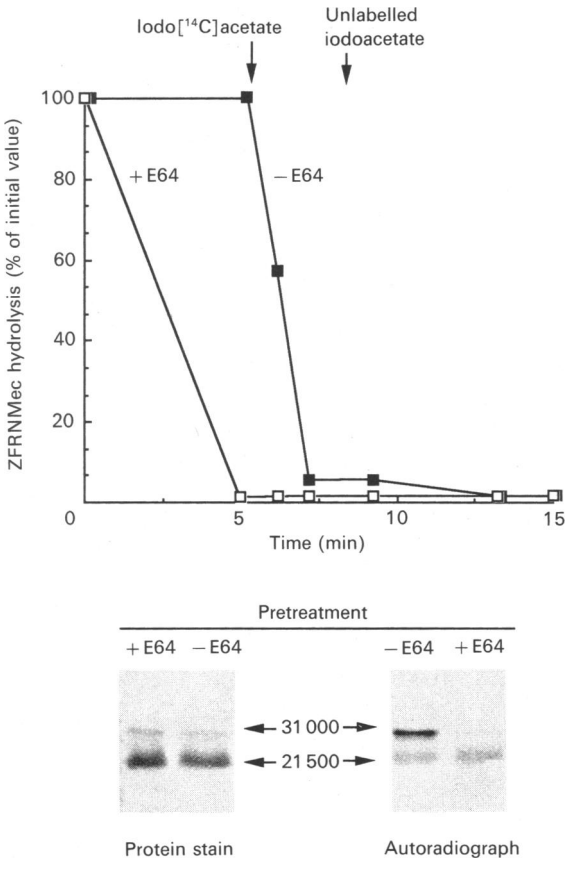

Fig. 5. Active-site labelling of cysteine protease activity in aril extract

Two samples of aril extract (100 μl; 85 μg of protein, containing approx. 2 nmol of thaumatopain) were incubated with 30 mM-2-mercaptoethanol for 15 min in 1 ml of 10 mM-pyrophosphate buffer, pH 6.0. E-64 (100 nmol) was added to one sample (□, +E-64); 100 μl of water was added to the other (■, −E-64); 5 min later, iodo[^{14}C]acetate (50 nmol; 2.5 μCi) was added to each reaction. After a further 4 min incubation, unlabelled iodoacetate (10 μmol) was added to each reaction. The activity in each preparation was monitored throughout by assaying 20 μl of the reaction mixture at various times by using ZFRNMec. Samples from both reactions were freeze-dried and resolved by SDS/PAGE. The gel was electroblotted on to nitrocellulose, and the proteins on the membrane (M_r values indicated) were revealed by Fast Green staining ('Protein stain') before autoradiography on Fuji RX100 film for 5 days ('Autoradiograph').

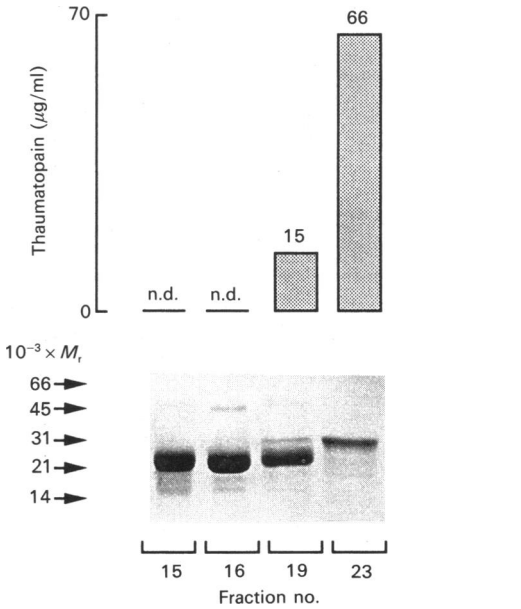

Fig. 6. Active-site titration of cysteine protease activity in chromatographic fractions derived from aril extracts

Aril extract (8.7 mg of protein) was separated on MonoS as in Fig. 1. Fractions containing cysteine protease activity were titrated with E-64 to quantify the cysteine protease (n.d., not detected), and additionally were resolved on SDS/PAGE and stained for protein.

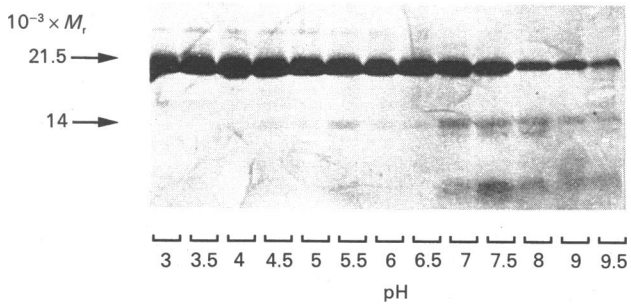

Fig. 7. Degradation of thaumatin by thaumatopain

Thaumatopain I (2 nmol; 50 μg) was incubated with 0.1 nmol of thaumatopain in buffers of final concentration 25 mM and ionic strength 0.1 M, maintained by NaCl, ranging from pH 3.0 to 9.5; 2-mercaptoethanol (80 mM) was included in the incubation. After incubation for 4 h at 25 °C, each sample was prepared for SDS/PAGE (17.5% gel) before proteins were made visible by silver staining.

one-fifth of the amount of thaumatopain as fraction 23, roughly in parallel with the relative activities of the two peaks. On SDS/PAGE the 30 kDa thaumatopain is clearly evident in fraction 19, although the most abundant protein is thaumatin (Fig. 6). There was substantially less activity in fractions 15 and 16. There may be some association between thaumatopain and thaumatin that is stable to the conditions of cation-exchange chromatography. Alternatively, a modified form of thaumatopain (for example, rendered more anionic by deamidation) could have co-eluted with this thaumatin fraction. Despite a recent communication that draws attention to a weak cysteine protease motif in thaumatins (Skern *et al.*, 1990), we consider that the lack of histidine in thaumatins, together with our own enzymological data, favour the view that the thaumatins are not proteolytically active (Beynon & Cusack, 1990).

It has proved difficult to demonstrate that thaumatin is a 'good' substrate for thaumatopain. Thaumatin is refractory to proteolysis (Shamil *et al.*, 1990), and digestion of thaumatin I by thaumatopain was only possible at pH values greater than 7.5, under highly reducing conditions (80 mM-2-mercaptoethanol; Fig. 7). Under these conditions the substrate was extensively degraded, and there was little evidence for accumulation of intermediates. This differs from the degradation pattern seen with other proteases such as trypsin or thermolysin, which elicit a limited digestion, but parallels the digestion with chymotrypsin or subtilisin (Shamil *et al.*, 1990). Whether the degree of digestion is consistent with digestion *in vivo* is open to conjecture, as it is difficult to emulate these conditions. Between them, thaumatin and thaumatopain constitute over 99% of the total protein in an aqueous aril extract. When intact arils were incubated in a buffered reducing environment (40 mM-2-mercaptoethanol, 50 mM-phosphate, pH 8.0), the thaumatins were extensively degraded in an E-64-inhibitable process. There was no digestion in the non-reduced control incubations (A. G. Stephen, unpublished work). Thus, conditions in the concentrated jelly-like aril, where the thaumatopain concentration could be as high as 1–2 mg/ml, might favour degradation of these unusual sweet proteins. Some post-translational modification of thaumatins is seen during fruit development (Mackenzie *et al.*, 1985), and it is possible that limited proteolysis by thaumatopain could contribute in part to the formation of these variants.

Received 11 June 1990/13 August 1990; accepted 22 August 1990

This work was supported by a grant from the A.F.R.C. (CRG 19). A. G. S. holds a University of Liverpool Research Studentship, M. C. was in receipt of an SERC–CASE Award with Tate & Lyle Group R&D. We acknowledge helpful discussions with Alan Barrett and Dave Buttle (Strangeways Research Laboratories, Cambridge). We are also grateful to Alan Barrett for providing the aligned sequences of the plant cysteine proteases. We are grateful to Bryan Dunbar and John Fothergill (SERC Sequencing Facility, Department of Biochemistry, University of Aberdeen) for useful discussion and for conducting the sequence analysis.

REFERENCES

Akasofu, H., Yamauchi, D., Mitsuhashi, W. & Minamikawa, T. (1989) Nucleic Acids Res. **17**, 6733
Anastasi, A., Brown, M. A., Kembhavi, A., Nicklin, M. J. H., Sayers, C. A., Sunter, D. C. & Barrett, A. J. (1983) Biochem. J. **211**, 129–138
Baker, E. N. (1980) J. Mol. Biol. **141**, 441–484
Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirshke, H., Knight, C. G., Tamai, M. & Hanada, K. (1982) Biochem. J. **201**, 189–198
Beynon, R. J. (1985) Comput. Appl. Biosci. **1**, 111–115
Beynon, R. J. (1988) Comput. Appl. Biosci. **4**, 487–490
Beynon, R. J. & Cusack, M. (1990) Nature (London) **334**, 498
Buttle, D. J., Ritonja, A., Pearl, L. H., Turk, V. & Barrett, A. J. (1990) FEBS Lett. **260**, 195–197
Carne, A. & Moore, C. H. (1978) Biochem. J. **173**, 73–83
Cohen, L. W., Coghlan, V. M. & Dihel, L. C. (1986) Gene **48**, 219–227
Cusack, M. R. & Pierpoint, W. S. (1988) Phytochemistry **27**, 3817–3821
Cusack, M., Beynon, R. J. & Rodgers, P. B. (1987) Biochem. Soc. Trans. **15**, 880
Dubois, T., Kleinschmidt, T., Schnek, A. G., Looze, Y. & Braunitzer, G. (1988) Biol. Chem. Hoppe–Seyler **369**, 741–754
Higginbotham, J. D. (1979) in Developments in Sweeteners, vol. 1 (Hough, C. A. M., Parker, K. J. & Vlitos, A. J., eds.), pp. 87–123, Applied Science Publishers, London
Jacquet, A., Kleinschmidt, T., Schnek, A. G., Looze, Y. & Braunitzer, G. (1989) Biol. Chem. Hoppe–Seyler **370**, 425–434
Kimmel, J. R. & Smith, E. L. (1954) J. Biol. Chem. **207**, 515–531
Mackenzie, A., Pridham, J. B. & Saunders, N. (1985) Phytochemistry **11**, 2503–2506
Maurizi, M. J. (1987) J. Biol. Chem. **262**, 2696–2703
Ritonja, A., Rowan, A. D., Buttle, D. J., Rawlings, N. D., Turk, V. & Barrett, A. J. (1989*a*) FEBS Lett. **247**, 419–424
Ritonja, A., Buttle, D. J., Rawlings, N. D., Turk, V. & Barrett, A. J. (1989*b*) FEBS Lett. **258**, 109–112
Rogers, J. C., Dean, D. & Heck, G. R. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 6512–6516
Rowan, A. D., Buttle, D. J. & Barrett, A. J. (1988) Arch. Biochem. Biophys. **267**, 262–270
Schaffer, M. A. & Fischer, R. L. (1988) Plant Physiol. **87**, 431–436
Shamil, S., Cusack, M. R. & Beynon, R. J. (1990) J. Sci. Food Agric. **53**, 73–84
Skern, T., Zorn, M., Blaas, D., Kuechler, E. & Sommegruber, W. (1990) Nature (London) **334**, 26
Studier, F. W. (1973) J. Mol. Biol. **79**, 237–248
Van der Wel, H. & Bel, W. J. (1980) Eur. J. Biochem. **104**, 413–418
Van der Wel, H. & Loeve, K. (1972) Eur. J. Biochem. **31**, 221–225
Van der Wel, H. & Ledeboer, A. M. (1989) in The Biochemistry of Plants, vol. 15 (Marcus, A., ed.), pp. 379–391, Academic Press, London