Structure of chymopapain M the late-eluted chymopapain deduced by comparative modelling techniques and active-centre characteristics determined by pH-dependent kinetics of catalysis and reactions with time-dependent inhibitors: the Cys-25/His-159 ion-pair is insufficient for catalytic competence in both chymopapain M and papain

Mark P. THOMAS,*|| Christopher M. TOPHAM,†‡ Devanand KOWLESSUR,*¶ Geoffrey W. MELLOR,*†† Emrys W. THOMAS,§ David WHITFORD* and Keith BROCKLEHURST*‡‡

*Laboratory of Structural and Mechanistic Enzymology, Department of Biochemistry, Queen Mary & Westfield College, University of London, Mile End Road, London E1 4NS, U.K., †Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, U.K., ‡St. Luke's Institute of Cancer Research, Department of Chemistry, University College Dublin, Dublin 4, Ireland, and §Department of Biological Sciences, University of Salford, Salford M5 4WT, U.K.

1. Chymopapain M, the monothiol cysteine proteinase component of the chymopapain band eluted after chymopapains A and B in cation-exchange chromatography, was isolated from the dried latex of *Carica papaya* and characterized by kinetic and chromatographic analysis. This late-eluted chymopapain is probably a component of the cysteine proteinase fraction of papaya latex discovered by Schack [(1967) *Compt. Rend. Trav. Lab. Carlsberg* **36**, 67–83], named papaya peptidase B by Lynn [(1979) *Biochim. Biophys. Acta* **569**, 193–201] and partially characterized by Polgár [(1981) *Biochim. Biophys. Acta* **658**, 262–269] and is the enzyme with unusual specificity characteristics (papaya proteinase IV) that Buttle, Kembhavi, Sharp, Shute, Rich and Barrett [Biochem. J. (1989) **261**, 469–476] claimed to be a previously undetected cysteine proteinase eluted from a cation-exchange column near to the early-eluted chymopapains. 2. A study of the time-dependent chromatographic consequences of thiol-dependent proteolysis of the components of papaya latex is reported. 3. Chymopapain M was isolated by (i) affinity chromatography followed by separation from papain using cation-exchange f.p.l.c. on a Mono S HR5/5 column and (ii) cation-exchange chromatography followed by an unusual variant of covalent chromatography by thiol–disulphide interchange. 4. The existence in chymopapain M of a nucleophilic interactive Cys/His catalytic-site system analogous to those in papain (EC 3.4.22.2) and other cysteine proteinases was deduced from the characteristic shape of the pH–second-order rate constant (k) profiles for its reactions with 2,2'-dipyridyl disulphide and ethyl 2-pyridyl disulphide. 5. Analysis of the pH– k data for the reactions of chymopapain M with the 2-pyridyl disulphides and

with 4,4'-dipyrimidyl disulphide permits the assignment of molecular pK_a values of 3.4 and 8.7 to the formation and subsequent dehydration of the Cys-S[−]/His-Im⁺H state of the catalytic site and reveals three other kinetically influential ionizations with pK_a values 3.4, 4.3 and 5.6. 6. The pH-dependences of k_{cat}/K_m for the hydrolysis of *N*-acetyl-L-Phe-Gly-4-nitroanilide at 25.0 °C and 1.0 M catalysed by chymopapain M and papain were determined. For both enzymes, little catalytic activity (5–7 % of the maximal) develops consequent on formation of the catalytic site Cys-S[−]/His-Im⁺H ion-pair state (across pK_a 3.4 for both enzymes). For papain, full expression of k_{cat}/K_m for the uncharged substrate requires only the additional hydronic dissociation with pK_a 4.2. By contrast, full expression of k_{cat}/K_m for chymopapain M requires additional hydronic dissociation with pK_a values of 4.3 and 5.6. 7. To provide a structural basis for the present, and future, mechanism-targeted kinetic studies of chymopapain M, a model of its three-dimensional structure was built by using comparative modelling techniques. Superpositioning of the model of the structure of chymopapain M and the crystal structure of papain demonstrates the presence in chymopapain M of many of the kinetically and mechanistically relevant structural features present in the papain molecule and reveals a number of interesting differences, which are discussed. The additional kinetically influential side chain with pK_a 3.4 that is unique to chymopapain M is assigned to the carboxy group of Glu-23 which is close to the catalytic site and can engage in hydrogen-bonding interactions with Arg-65, a cationic residue unique to chymopapain M which is Gly in papain, actinidin (EC 3.4.22.14) and papaya proteinase Ω (EC 3.4.22.30).

INTRODUCTION

The natural variation in both structure and chemical behaviour that exists in the cysteine proteinases provides opportunities for

investigation of the consequences of different electrostatic effects (Pickersgill et al., 1988, 1989; Brocklehurst et al., 1989; Kowlessur et al., 1989a) and different binding interactions (Brocklehurst et al., 1988a; Kowlessur et al., 1989b) for active-centre

Abbreviations used: PDB, Protein Data Bank (Brookhaven); PPA, papaya peptidase A; PPB, papaya peptidase B; PPIII, papaya proteinase III; PPIV, papaya proteinase IV; PP Ω , papaya proteinase Ω ; 2-Py-S-S-2-Py, 2,2'-dipyridyl disulphide; Et-S-S-2-Py, ethyl 2-pyridyl disulphide; Bz-L-Arg-4NA, *N*^α-benzoyl-L-arginine 4-nitroanilide; N-Ac-L-Phe-Gly-4NA, *N*-acetyl-L-phenylalanylglycine 4-nitroanilide; Boc, *t*-butoxycarbonyl; Boc-Gly-4NPE, *t*-butoxycarbonylglycine 4-nitrophenyl ester; K_m (acyl), the steady-state dissociation constant, $(k_{-1} + k_{-2})/k_{+1}$ of the adsorptive complex, ES, of the three-step acyl-enzyme model; K_a , the analogous equilibrium dissociation constant, k_{-1}/k_{+1} .

|| Present address: Department of Chemistry, University of York, Heslington, York, YO1 5DD, U.K.

¶ Present address: Laboratory of Neurochemistry, NIMH, Bethesda, MD, U.S.A.

†† Present address: The Wellcome Research Laboratories, The Wellcome Foundation Ltd., Langley Court, South Eden Park Road, Beckenham, Kent BR3 3BS, U.K.

‡‡ To whom correspondence and requests for reprints should be addressed.

chemistry and catalytic mechanism [see Brocklehurst et al. (1987a) for an in-depth review, Brocklehurst (1987), Baker & Drenth (1987) and Polgár (1989) for shorter reviews and Topham et al. (1991), Patel et al. (1992a,b) and Mellor et al. (1993a–c) for recent structure-related kinetic studies and for references to recent crystallographic, spectroscopic and molecular-biological studies]. As part of a programme of co-ordinated structural and structure-related kinetic studies on the cysteine proteinases, we are continuing to develop the use of specially designed time-dependent inhibitors [e.g. Mellor et al. (1993b,c)] and variations found in this enzyme family, e.g. papain (EC 3.4.22.2), actinidin (EC 3.4.22.14) and ficin (EC 3.4.22.3), the cysteine proteinases from *Carica papaya*, *Actinidia chinensis* and *Ficus glabrata* respectively (Kowlessur et al., 1989a,b; Patel et al., 1992a, 1993) and the calpain II 80 kDa/18 kDa heterodimer (Mellor et al., 1993a).

The latex of *Carica papaya* is a particularly rich source of cysteine proteinase variants [see e.g. Brocklehurst et al. (1981, 1987a) for reviews]. Aqueous extracts of papaya latex contain a multiplicity of cysteine proteinases which are separable by ion-exchange chromatography (Brocklehurst et al., 1981; Baines and Brocklehurst, 1982) and may be obtained in fully active forms by hydron-activated covalent chromatography by thiol–disulphide interchange (Brocklehurst et al., 1985a). (IUPAC recommends the use of the term ‘hydron’ for H^+ when a particular isotope is not specified; ‘proton’ is then reserved for $^1H^+$.) The aqueous extracts contain, as a minor component, the well-characterized enzyme papain (the least basic cysteine proteinase component), papaya proteinase Ω (PP Ω) (EC 3.4.22.30) (Brocklehurst et al., 1985b) [originally called papaya peptidase A (PPA) (see Schack, 1967), the most basic cysteine proteinase component] and a complex mixture of cysteine proteinases of intermediate basicity which are usually referred to as the chymopapains (see Baines et al., 1986). The most widely recognized types of chymopapain are the ‘A’ and ‘B’ types, both of which contain non-essential thiol groups in addition to the essential thiol group that is part of the catalytic site (Brocklehurst and Salih, 1983; Baines et al., 1986). Initially the designations ‘A’ and ‘B’ were used for the earlier- and later-eluted forms of chymopapain respectively [Kunimitsu and Yasunobu (1967); see Brocklehurst et al. (1987a) for a review of the chymopapain literature up to 1986]. When it became clear that the positions of some chymopapain forms in ion-exchange elution profiles vary with the source of the papaya latex and its treatment in the laboratory (Brocklehurst et al., 1985b,c; Baines et al., 1986), the development of a more incisive method of recognizing ‘A’ and ‘B’ types of chymopapain became imperative. To that end therefore we introduced a method of distinguishing catalytic sites characteristic of chymopapain A from those characteristic of chymopapain B. This involves kinetic analysis of the reactions of the essential thiol group with a two-hydronic-state reactivity probe such as 2,2’-dipyridyl disulphide (2-Py-S-S-2-Py) (Brocklehurst and Salih, 1983; Brocklehurst et al., 1985b,c), and Khan and Polgár (1983) suggested that this method may be the simplest and most reliable way in which to distinguish between different chymopapain forms.

There have been a number of reports of other cysteine proteinase components of papaya latex within the chymopapain band between papain and PP Ω in addition to the multiple forms of chymopapains A and B. Appraisal of these literature reports together with the new experimental work reported in the present paper suggests that the major (and possibly only) additional cysteine proteinase component is the late-eluted proteinase of the chymopapain band which is highly selective for substrates with Gly at P_1 (see below). By contrast with chymopapains A and B, each of which contains a non-essential thiol group in addition to

the catalytic-site thiol group, the late-eluted enzyme contains only the essential catalytic-site thiol group. We have called this enzyme chymopapain M, the designation M being used to denote monothiol chymopapain (Topham et al., 1990a). Evidence is presented in the Results and discussion section that chymopapain M is probably the major component of the cysteine proteinase fraction discovered by Schack (1967) and named papaya peptidase B (PPB) by Lynn (1979) and partially characterized by Polgár (1981), and is the enzyme with unusual specificity characteristics that Buttle et al. (1989) called papaya proteinase IV (PPIV). Buttle et al. (1989) claimed that PPIV was a previously undetected cysteine proteinase which is eluted from a cation-exchange column near to early-eluted chymopapains. In fact, when due care is taken to prevent proteolysis, this enzyme is eluted as the last component of the chymopapain band, as is the PPB fraction.

Chymopapain M (PPIV) is of considerable interest because of its narrow substrate specificity which is largely restricted to cleavage of peptide bonds with Gly at P_1 (Buttle et al., 1990a) and because, unlike most other papain-like cysteine proteinases, it is not inhibited by chicken cystatin or human cystatins A or C (Buttle et al., 1990b). The strict specificity of chymopapain M (PPIV), marketed by Novabiochem, initially as Endopeptidase Gly-C and currently as Proteinase Gly-C, has found application in the preparation of protein fragments and in peptide mapping and sequencing (Buttle et al., 1990a; Ritonja et al., 1990, 1991).

There have been few kinetic studies and no substantive reports on the three-dimensional structure and active-centre characterization of chymopapain M. As part of our programme of studies on structure–function relationships in the cysteine proteinase family of enzymes, we are increasing the availability of models of three-dimensional structures by using a comparative modelling approach incorporated in the COMPOSER suite of programs (Sutcliffe et al., 1987a,b; Blundell et al., 1988; Topham et al., 1990b) and refinement by energy minimization [for stem bromelain, Topham et al. (1990c); for PP Ω , Topham et al. (1991) and see also Pickersgill et al. (1991) for the crystal structure of PP Ω determined by X-ray crystallography]. The present paper reports a model of the three-dimensional structure of chymopapain M, the isolation of the enzyme from papaya latex, clarification of its ion-exchange elution behaviour relative to that of other chymopapains and its active-centre characteristics revealed by analysis of the pH-dependent kinetics of catalysis and the reaction of its catalytic-site (and only) thiol group with reactivity probes. Preliminary accounts of some of this work have appeared (Topham et al., 1990a; Thomas et al., 1993).

Hydronic dissociation to provide the (Cys-25)-S[−]/(His-159)-Im⁺H ion-pair state of chymopapain M (pK_a 3.4) is clearly separated from another hydronic dissociation of pK_a 4.3 which is necessary for production of substantial catalytic competence. Evidence has been accumulating over the past few years for the modulation of cysteine proteinase catalytic-site geometry by combinations of specific binding interactions and electrostatic effects (Brocklehurst et al., 1988a,b, 1989; Kowlessur et al., 1989a,b). Only recently has hydronic dissociation leading to the formation and dehydration of the catalytic site ion-pair state of papain been clearly distinguished from additional ionizations that modulate its reactivity (see Mellor et al., 1993b,c). Chymopapain M provides a clear example of another cysteine proteinase in which pH-dependent perturbation of the catalytic-site ion-pair is clearly separated from its formation by hydron loss from the catalytic site itself. In addition, chymopapain M is unique among the cysteine proteinases thus far investigated in that the reactivity of its catalytic-site ion-pair (generated across pK_a 3.4) is modulated by hydronic dissociation of another group of pK_a 3.4 in

addition to modulation across pK_a values 4 and 5–6 seen with other members of the family.

The pK_a values of the kinetically influential ionizations of chymopapain M seen in reactions with disulphide probes are now established as 3.4, 3.4, 4.3, 5.6 and 8.7, of which 3.4 and 8.7 characterize ion-pair formation and subsequent dehydronation and 4.3, 5.6, 8.7 and the additional pK_a 9.9 are seen in the pH-dependence of $k_{cat.}/K_m$ for the catalysed hydrolysis of *N*-acetyl-L-phenylalanylglycine 4-nitroanilide (N-Ac-L-Phe-Gly-4NA). The kinetically influential ionizations of chymopapain M are compared with those of papain, and the pH-dependence of $k_{cat.}/K_m$ for the papain-catalysed hydrolysis of N-Ac-L-Phe-Gly-4NA is reported.

MATERIALS AND METHODS

Chemical reagents, substrates, inhibitors and reactivity probes

The syntheses of ethyl 2-pyridyl disulphide (Et-S-S-2-Py) (Brocklehurst et al., 1988b), N-Ac-L-Phe-Gly-4NA (Kowlessur et al., 1990) and 4,4'-dipyrimidyl disulphide (Mellor et al., 1993b) and the purification of 2-Py-S-S-2-Py (Salih et al., 1987) have been described previously. Other commercially available chemicals were obtained from Sigma, Poole, Dorset, U.K. or from BDH, Poole, Dorset, U.K.

Buffers

Buffers of constant ionic strength [Biochemists' Handbook (1961) pp. 30–40, E & F. N. Spon Ltd., Colchester] were prepared by using BDH Analar-grade buffer salts, and stock solutions, made up triple strength, contained 3 mM EDTA. Dilution 1:3 provided reaction mixtures with 1.0 M in experiments conducted with conventional spectrophotometers. In stopped-flow experiments one of the syringes contained a reactant (usually the time-dependent inhibitor) in buffer of 1.0 M and the other syringe contained the other reactant (usually the enzyme) in 0.1 M KCl.

Chromatographic materials

Sephadex G-25, S-Sepharose Fast Flow, Sepharose 4B, activated CH-Sepharose 4B and prepacked HR5/5 Mono S and S-Sepharose High Performance Hi Load 26/10 columns were supplied by Pharmacia (GB) Ltd.

Synthesis of covalent chromatography gel

Sepharose–glutathione 2-pyridyl disulphide gel for covalent chromatography by thiol–disulphide interchange was prepared from Sepharose 4B as described by Brocklehurst et al. (1973, 1974).

Synthesis of the affinity-chromatography gel for the isolation of chymopapain M

In the present work a new relatively simple two-step method of synthesis of Boc-Gly-L-Phe-NHCH₂CN was developed. De-protection of the amino group and coupling to activated CH-Sepharose 4B were then performed as described by Buttle et al. (1989).

To a stirred mixture of Gly-L-Phe (1.34 g, 6 mmol) and triethylamine (1.7 ml, 12.6 mmol) in 40 ml of methanol/dichloromethane (1:1, v/v) was added 2-(*t*-butoxycarbonylimino)-2-phenylacetonitrile (BOC-ON, Aldrich; 1.55 g, 6.3 mmol) [see Itoh et al. (1975) for a description of the synthesis of this compound and of its use in provision of contaminant-free

Boc-amino acids in high yield] and stirring was continued at room temperature (approx. 22 °C) overnight. The solvent was removed *in vacuo* and the residue partitioned between ethyl acetate and pH 10 buffer. The aqueous phase was acidified and extracted into ethyl acetate. The organic layer was dried, concentrated to give a clear oil which was crystallized from ethyl acetate/cyclohexane (1:1, v/v) to yield small white crystals (1.4 g, 73 %) of Boc-Gly-L-Phe. Boc-Gly-L-Phe (1 g, 3.1 mmol) was dissolved in acetonitrile (10 ml) to which was then added *N*-hydroxysuccinimide (393 mg, 3.4 mmol) and 1-(3,3-dimethylaminopropyl)-3-ethylcarbodi-imide (891 mg, 4.65 mmol) and the mixture was stirred for 45 min at room temperature. To this was added a solution of aminoacetonitrile hydrochloride (301 mg, 3.25 mmol) in acetonitrile (10 ml) containing triethylamine (627 mg, 860 μ l, 62 mmol) and the mixture was stirred overnight. The solvent was removed *in vacuo* and the residue taken up in ethyl acetate and washed with 0.1 M HCl (twice), sodium hydrogen carbonate, pH 11 (twice) and water (twice). The organic layer was dried and concentrated to yield a white solid (770 mg). This was recrystallized from methanol/ethyl acetate to yield the product (Boc-Gly-L-Phe-NHCH₂CN; 359 mg; 32 % yield; m.p. 162–164 °C). The n.m.r. spectrum (Bruker WH 400-MHz Spectrometer, University of London Intercollegiate Research Service at Queen Mary & Westfield College) was identical with that reported by Buttle et al. (1989).

Before attachment of Gly-Phe-NHCH₂CN to the activated Sepharose gel, the Boc group was removed by dissolving Boc-Gly-Phe-NHCH₂CN (30 mg) in 1 ml of 25 % (v/v) trifluoroacetic acid in dichloromethane containing 10 % (v/v) anisole and keeping the mixture at 0 °C for 30 min. Solvent was removed by rotary evaporation at approx. 35 °C and the residue was dissolved in 1.5 ml of methanol to which 1.5 ml of 0.1 M NaHCO₃, pH 8.0, was then added. Activated CH-Sepharose 4B (3 g dry weight) was kept overnight in 75 ml of 1 mM HCl at 4 °C and washed with 600 ml of 1 mM HCl followed by 300 ml of 0.1 M NaHCO₃ (pH 8.0) in a sintered-glass funnel. The swollen washed activated gel was suspended in 30 ml of 0.1 M NaHCO₃ (pH 8.0) and the solution of Gly-Phe-NHCH₂CN was added.

The mixture was agitated gently overnight at room temperature and the Sepharose–6-aminohexanoyl-Gly-L-Phe-aminoacetonitrile affinity gel was isolated in a sintered-glass funnel. After washing with aq. 50 % (v/v) methanol and then with water, the gel was suspended in 30 ml of 0.1 M ethanolamine adjusted to pH 9.0 with HCl and shaken for 4 h at room temperature. Finally the gel was collected, washed with water and stored at 4 °C in 50 mM sodium phosphate buffer, pH 6.8, containing 1 mM EDTA in water/ethanediol (2:1, v/v).

Purification and characterization of papain

This was carried out as described by Mellor et al. (1993b).

Isolation of chymopapain M

Chymopapain M was isolated from the spray-dried latex of *Carica papaya* (Powell and Scholefield Ltd., Liverpool, U.K.) by the two methods detailed below. All solutions used in chromatographic procedures were passed through 0.22 μ m Durapore membrane filters (Millipore U.K. Ltd) before use.

Isolation by affinity chromatography followed by separation from papain using cation-exchange f.p.l.c.

The procedure used is similar to, but not identical with, that described by Buttle et al. (1989). A column of Sepharose–

6-aminohexanoyl-Gly-L-Phe-aminoacetonitrile (4 ml bed volume, column diameter 16 mm) was washed first with 12 ml of 50 mM sodium citrate buffer, pH 4.5, prepared in a water/ethanediol mixture (2:1, v/v) and subsequently with 12 ml of 50 mM sodium phosphate buffer, pH 6.8, containing 1 mM EDTA and prepared in the same water/ethanediol solvent. Spray-dried papaya latex (0.5 g) was added to 10 ml of the sodium phosphate buffer containing also 40 mM cysteine, the mixture was stirred at room temperature for 10 min and residual solid was removed by filtration through a 0.22 μ m filter. A flow rate of 0.5 ml/min was set and a portion of the filtrate (0.5 ml) was applied to the column at room temperature. The column was washed with the sodium phosphate buffer, devoid of cysteine, until the A_{280} reading returned to its baseline value. The column was then washed with the sodium citrate buffer to elute a mixture of the other chymopapains (A and B) and PP Ω until the A_{280} reading again returned to its baseline value. In each procedure washing with 8 ml of buffer produced baseline readings of A_{280} . One bed volume (4 ml) of the sodium citrate buffer now containing 1.5 mM 2-Py-S-S-2-Py was applied to the column, the flow was stopped and the column was allowed to incubate in the 2-Py-S-S-2-Py solution overnight at room temperature. Elution with the buffer containing 2-Py-S-S-2-Py (12 ml) was then carried out and 1 ml fractions were collected. The 4–6 ml of eluate that contained chymopapain M-S-S-2-Py was concentrated to 1 ml by using an Amicon stirred ultrafiltration cell and a Diaflo PM10 membrane and this solution was stored at 4 °C until required. The column was re-equilibrated with the sodium phosphate buffer in preparation for chromatography of a subsequent portion of latex solution.

The protein fraction eluted by washing with the sodium citrate buffer can be used as a source of chymopapains A and B and PP Ω or discarded. The fraction eluted in 2-Py-S-S-2-Py solution contains a mixture of papain-S-S-2-Py and chymopapain M-S-S-2-Py. These were separated by cation-exchange f.p.l.c. using a Mono S HR5/5 column pre-equilibrated in 0.1 M sodium acetate buffer, pH 5.0, containing 1 mM EDTA. After application of the enzyme-S-S-2-Py mixture, the column was washed with the same buffer (1 ml/min) until A_{280} returned to its baseline value. A gradient (39 mM Na⁺/ml; 0.1 M Na⁺ → 0.8 M Na⁺ in 18 ml) was then applied to the column followed by a final wash providing a change from 0.8 M Na⁺ to 1.1 M Na⁺ in 2 ml; 1 ml fractions were collected.

Papain-S-S-2-Py with catalytic activity towards both Bz-L-Arg-4NA and t-butoxycarbonylglycine 4-nitrophenyl ester (Boc-Gly-4NPE) after activation by treatment with 20 mM cysteine is eluted at pH 5 at about 0.2 M Na⁺ well separated from chymopapain M-S-S-2-Py (with catalytic activity towards Boc-Gly-4NPE but not towards Bz-L-Arg-4NA after activation with cysteine), which is eluted at about 0.4 M Na⁺ [see Figure 1(b) of Buttle et al. (1989) for a typical elution profile]. This procedure on the scale described produces sufficient chymopapain M for about three stopped-flow kinetic runs of reactions with 2-pyridyl disulphide reactivity probes or up to about 30 initial-rate determinations using Boc-Gly-4NPE as substrate.

Isolation by cation-exchange chromatography followed by covalent chromatography by thiol-disulphide interchange

Papaya latex (15 g) was added to 100 ml of 0.1 M sodium acetate buffer, pH 5.0, containing 1 mM EDTA and saturated with 2-Py-S-S-2-Py. The mixture was stirred for 30 min at room temperature and then dialysed overnight against 3 × 5 litres of the same buffer but without the 2-Py-S-S-2-Py. The protein solution was clarified by centrifugation for 15 min at 10000 g

and 4 °C and was then applied to a column (30 cm × 2.5 cm) of S-Sepharose Fast Flow. Ion-exchange chromatography was carried out with a start buffer of 0.1 M sodium acetate, pH 5.0, containing 1 mM EDTA, and finish buffer comprising the start buffer containing, in addition, 1 M NaCl.

After loading the protein solution on to the column, the start buffer was pumped through until the A_{280} reading returned to its baseline value. An elution gradient from 0.1 M Na⁺ to 0.8 M Na⁺ was then applied at a flow rate of 1.5 ml/min for conventional chromatography or a flow rate of 4 ml/min for f.p.l.c. Fractions of 12 ml were collected, the number of which varied according to the volume of eluent used (about 150–300 fractions and 2–4 litres of eluent). These volumes are sufficient to collect all of the cysteine proteinase components from papain to PP Ω and then to permit A_{280} to return to its baseline value. The fractions were analysed by measurement of A_{280} and catalytic activity towards both Bz-L-Arg-4NA and Boc-Gly-4NPE in the presence of 20 mM L-cysteine to regenerate the active enzymes from the mixed disulphide derivatives. Appropriate fractions (see the Results and discussion section) identified by their positions in the elution profile as derivatives of papain, chymopapains A and B and PP Ω were either discarded or used to provide fully active samples of the enzymes by covalent chromatography (Brocklehurst et al., 1973, 1974) as required. All of these enzymes exhibited substantial catalytic activity towards both of the substrates used to evaluate the fractions. The fraction rich in the chymopapain M derivative is eluted after chymopapain B and papaya lysozyme and, after activation by cysteine in the assay mixtures, discriminates markedly between the P₁ Arg and P₁ Gly substrates in favour of the latter. Chymopapain M was then freed from the contaminant with activity towards Bz-L-Arg-4NA (possibly a minor chymopapain B component) by covalent chromatography by the procedure described below.

The fractions containing chymopapain M-S-S-2-Py were pooled and brought to 85% saturation with respect to (NH₄)₂SO₄. The mixture was stirred for 30 min at room temperature and then centrifuged for 15 min (10000 g, 4 °C). The pellet was dissolved in 0.1 M Tris/HCl buffer, pH 8.0, containing 40 mM L-cysteine and 1 mM EDTA to convert the enzyme 2-pyridyl mixed disulphides into active enzymes. The solution was acidified to about pH 5 by addition of 1 M HCl before (NH₄)₂SO₄ precipitation to avoid evolution of ammonia. The enzymes were precipitated with 85% satd. (NH₄)₂SO₄ and isolated by centrifugation as described above. The pellet was dissolved in 300 ml of either Tris/HCl buffer, pH 8.0, or sodium acetate buffer, pH 4.0, in each case containing 1 mM EDTA and 0.3 M NaCl and mixed with Sepharose–glutathione-2-pyridyl disulphide gel (200 g wet weight) equilibrated in either pH 4.0 or pH 8.0 buffer. This mixture was stirred at room temperature for 30 min, isolated by suction on a sintered funnel and washed with 200 ml of the application buffer. The filtrate and washings (about 500 ml) were brought to 85% saturation with (NH₄)₂SO₄ and stirred at room temperature for 30 min. When the application was carried out at pH 8.0, the solution was acidified to pH 5.0 before the (NH₄)₂SO₄ precipitation. Chymopapain M was isolated by centrifugation as described above and the pellet was dissolved in the minimum volume of 0.1 M sodium acetate buffer, pH 5.0, containing 1 mM EDTA and saturated with 2-Py-S-S-2-Py. The reversibly blocked form of the enzyme was stored in this state at 4 °C until required for kinetic studies. A preparation on the scale described provides sufficient chymopapain M for about 500 stopped-flow kinetic runs of reactions with disulphide reactivity probes or up to about 5000 initial-rate determinations using Boc-Gly-4NPE as substrate. Chymopapain M protein was quantified by using A_{280} and $\epsilon_{280} = 38\,500\text{ M}^{-1}\cdot\text{cm}^{-1}$ (Buttle et al., 1989), total

thiol content by spectroscopic titration at 343 nm [$\Delta\epsilon_{343} = 8.08 \times 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$ (Stuchbury et al. (1975))] with 2-Py-S-S-2-Py at pH 8.0, and intact catalytic-site content by an analogous titration at pH 4.0 [see Brocklehurst et al. (1982) for a description of 2-Py-S-S-2-Py as a titrant of intact cysteine proteinase catalytic sites].

Investigation of the effects of proteolysis on the form of the ion-exchange elution profile of papaya latex

Papaya latex (40 g) was added with stirring at room temperature to 200 ml of 0.1 M sodium acetate buffer, pH 5.0, containing 1 mM EDTA and saturated with 2-Py-S-S-2-Py. The mixture was dialysed overnight at 4 °C against 4 × 5 litres of the acetate buffer devoid of 2-Py-S-S-2-Py. Three 18 ml samples (a–c) of the stirred contents of the dialysis sac (total volume 260 ml) were removed and L-cysteine was added to the remainder to provide a concentration of 100 mM. This mixture was stirred at 4 °C for 5 days, during which time a further eleven 18 ml samples (d–n) were withdrawn at intervals of 0.5, 1, 2, 4, 8, 16, 24, 72, 96 and 120 h.

Sample (a) was clarified by centrifugation at 4 °C and 23 000 *g* for 30 min and then subjected to cation-exchange f.p.l.c. using a 26/10 S-Sepharose HiLoad column at room temperature. The start buffer was 0.1 M sodium acetate buffer, pH 5.0, containing 1 mM EDTA and the finish buffer was the start buffer containing also 1 M NaCl. A gradient was applied from 0.1 M Na⁺ to 0.8 M Na⁺ over 600 ml (approx. 10 column vol.) at a flow rate of 4 ml/min, followed by an increase to 1.1 M Na⁺ over a further 60 ml. The first 644 ml eluted were collected in 92 × 7 ml fractions which were evaluated by measuring values of A_{280} and catalytic activity towards both Bz-L-Arg-4NA and Boc-Gly-4NPE in the presence of 13 mM L-cysteine.

Sample (b) was stored at 4 °C for 6 weeks before clarification by centrifugation and analysis by f.p.l.c. Sample (c) was brought to 90% saturation with respect to (NH₄)₂SO₄ and stirred for 30 min at room temperature. The precipitate was isolated by centrifugation as described above and then redissolved in 20 ml of the acetate buffer by stirring for 1 h at room temperature. This solution was dialysed overnight against 2 × 500 ml of the acetate buffer devoid of 2-Py-S-S-2-Py, and subjected to analysis by f.p.l.c. as described above. Samples (d)–(n) were subjected to (NH₄)₂SO₄ precipitation, centrifugation, dissolution in acetate buffer (which for these samples was saturated with 2-Py-S-S-2-Py) dialysis and f.p.l.c. as described for sample (c).

Evaluation of catalytic activity in ion-exchange chromatography fractions

Activity towards Bz-L-Arg-4NA was evaluated by spectral analysis at 410 nm, essentially as described for the ficin-catalysed hydrolysis (Malthouse and Brocklehurst, 1976). Reaction mixtures in potassium phosphate buffer, pH 6.0, 1.0 M, contained 250 μM substrate and 13 mM L-cysteine at 25 °C. Enzyme and L-cysteine were preincubated in the buffer for 15 min before the reaction was started by addition of substrate. Stock solutions of substrate (1.5 mM) were prepared by dissolving 65 mg of substrate in 1.5 ml of dimethyl sulphoxide followed by addition of water to make 100 ml. Activity of fractions towards Boc-Gly-4NPE was evaluated by spectral analysis at 400 nm of reaction mixtures in sodium acetate buffer, pH 4.5, 1.0 M, containing 1 mM EDTA and 13 mM L-cysteine. This concentration of L-cysteine was sufficient to transform enzyme-S-S-2-Py to active enzyme even at pH 4.5. The stock solution of the substrate was 20 mM in acetonitrile, and 0.1 ml of this was used to start the

reaction after preincubation of enzyme and L-cysteine for 15 min. It was found to be convenient to assay fractions for activity towards Boc-Gly-4NPE at 400 nm even at pH 4.5. The relatively low pH obviates problems of alkaline hydrolysis, and by using 400 nm to detect the small fraction of product existing as 4-nitrophenolate anion rather than a wavelength around 345 nm which detects undissociated 4-nitrophenol, it was possible to quantify the activity of the fractions without the need for excessive dilution before assay.

Demonstration of the lack of inhibition of chymopapain M by chicken cystatin

This was performed by using activated chymopapain M (0.1 μM) preincubated with chicken cystatin (1.0 μM) for 20 min at 25 °C and Boc-Gly-4NPE (1.0 mM) as substrate as described above.

Study of the oxidation of chymopapain M

A solution of activated activator-free enzyme (15 ml of about 2 μM in 0.1 M KCl containing 1 mM EDTA) was kept at 25.0 °C in a clear-glass stoppered bottle. A duplicate solution was kept under the same conditions in a stoppered bottle shielded from light by aluminium foil. Samples were removed immediately after preparation of the solution and at intervals of 1 h over a 9 h period and used to fill one of the 1 ml syringes of the SF machine. The enzyme was allowed to react with 2-Py-S-S-2-Py (250 μM in phosphate buffer, pH 6.0, in the other SF syringe); four or five progress curves were overlaid on the visual display unit and the value of ΔA_{343} was computed. The remaining enzyme solution was returned to the stock bottle in the thermostat bath.

pH-dependent kinetics

Instrumentation and software

Stopped-flow kinetic studies were performed with an Applied Photophysics SF.17MV stopped-flow spectrophotometer, kinetics workstation and data-acquisition and analysis software. Monochromator entrance and exit slit widths were set at 1 mm. Kinetics of catalysis were performed by using either a Kontron Uvikon 810 spectrophotometer or a Cary 118 spectrophotometer. Closely similar results were obtained with each instrument. Kinetic data ([S], (*v*) pairs) were fitted to rate equations by weighted non-linear regression analysis using either the AR computer program (1988 PC version) from the BMDP statistical software package (Dixon et al., 1988) or Sigmaplot 4.1 (Jandel Scientific) and either a Compaq Deskpro 386/20e computer or a Tandon MCS 486/33 computer and a Hewlett-Packard Colour ProPlotter. pH-dependence studies of the type reported in the present paper involving a multiplicity of reactive states require a rapid method of evaluating a series of kinetic models differing in the number and reactivity of the reactive states. These were evaluated by using a multitasking application program (SKETCHER) written in ANSI C running under RISCOS on an Acorn Archimedes microcomputer (Brocklehurst et al., 1990; Topham et al., 1991) which was developed from an earlier program (QUATRO) written in BBC BASIC for use with a BBC Micro Model B (Brocklehurst and Brocklehurst, 1988). SKETCHER permits estimation of characterizing parameters in the generic set of equations for the various models by means of interactive manipulation of calculated curves. Values of the parameters thus obtained are generally close to the final mean values obtained by weighted non-linear regression performed by using the AR computer program from the statistical software

package BMDP and a Compaq Deskpro 386/20c computer or a Tandon MCS 486/33 computer, particularly when one or more of the parameters are known from an independent experiment which obviates problems deriving from multiple local minima. As demonstrated in the present paper, information from independent experiments is essential for a unique solution for systems with overlapping ionizations. Rate equations for reactions in a variety of hydronic states were written down and generated within the Archimedes by using the simple general expression, eqn. (1)

$$k = \sum_{i=1}^n \frac{A_i}{1 + \sum_{j=1}^n B_{ij}} \quad (1)$$

Eqn. (1) relates an experimentally determined pH-dependent rate constant, k , and macroscopic acid dissociation constants, K , where n is the number of reactive hydronic states; the numerator A_i is a generalized pH-independent rate constant ($\bar{k}_{x_{H^+}, i}$) where i , which may take values from 1 to n , specifies the particular hydronic state. The denominator of eqn. (1) provides for the pH-dependent variation in each of the contributions to k from reaction in a particular hydronic state, i.e. reaction associated with an individual pH-independent rate constant ($\bar{k}$). Each term generated by j in the denominator comprises a product of $[H^+]$ and one or more acid dissociation constants each raised to a particular power. The construction of specific rate equations for particular kinetic models from eqn. (1) involves determining expressions for B_{ij} by using the two information matrices described by Brocklehurst et al. (1990) and Topham et al. (1991). pH- k data and the associated theoretical curves were displayed by using the Sigmaplot 4.1 software package, a Tandon MCS 486/33 computer and a Hewlett-Packard Colour Pro Plotter.

Stopped-flow kinetic analysis of reactions of chymopapain M with disulphide reactivity probes

All reactions were carried out at 25 °C and 1.0 M in solutions containing 1 mM EDTA under pseudo-first-order conditions with $[\text{disulphide}] \gg [\text{enzyme}]$. In all cases the disulphide was in solution in the appropriate buffer and the enzyme was in solution in 0.1 M KCl. Reactions of 2-Py-S-S-2-Py and Et-S-S-2-Py were monitored at 343 nm [$\Delta\epsilon_{343} = 8080 \text{ M}^{-1} \cdot \text{cm}^{-1}$ (Stuchbury et al., 1975)] with $[\text{enzyme}] = 2\text{--}3 \mu\text{M}$, $[2\text{-Py-S-S-2-Py}] = 45\text{--}500 \mu\text{M}$ and $[\text{Et-S-S-2-Py}] = 30\text{--}1150 \mu\text{M}$. Reactions of 4,4'-dipyrimidyl disulphide were monitored at the isosbestic point of pyrimidine-4-thione, 312 nm, with $[\text{enzyme}] = 2\text{--}3 \mu\text{M}$ and $[\text{probe}] = 30\text{--}350 \mu\text{M}$. Although unnecessary for the determination of first-order rate constants, the absorbance changes at 312 nm can be used at pH values ≥ 2 to calculate the concentration of product formed by using $\Delta\epsilon_{312} = 7763\text{--}317$ [the value of ϵ_{312} for pyrimidine-4-thione at pH ≥ 2) = $7446 \text{ M}^{-1} \cdot \text{cm}^{-1}$ (Mellor et al., 1993b)]. First-order rate constants (k_{obs}) were obtained by fitting the absorbance (A)- t data collected by the Acorn Archimedes microcomputer of the SF machine either to an equation for a single-exponential process, i.e. $A = P_1 e^{-P_2 t} + P_3$, where $P_1 = A_\infty - A_0$, $P_2 = k_{\text{obs}}$, and $P_3 = A_\infty$, or to an equation for a single exponential with a zero-order component, i.e. $A = P_1 e^{-P_2 t} + P_3 t + P_4$, where $P_1 = A_\infty - A_0$ for the single-exponential component, $P_2 = k_{\text{obs}}$, P_3 is the slope of the zero-order component of the A versus t trace in s^{-1} , and P_4 is the extrapolated value of the ordinate intercept, i.e. A_∞ for the single-exponential component. Reactions of the 2-pyridyl disulphides were evaluated by using the single-exponential equation. Reactions of

4,4'-dipyrimidyl disulphide exhibited small but noticeable deviations from a single-exponential process, particularly at pH values > 7.5 (see Mellor et al., 1993b), and for these reactions the fit was improved by incorporation of a zero-order component. Values of k were calculated from $k = k_{\text{obs}}/[\text{disulphide}]$.

Kinetics of catalysis

Initial rates (v) of release of 4-nitroaniline from N-Ac-L-Phe-Gly-4NA at 25 °C in buffer, 1.0 M, containing 8.3% (v/v) *NN*-dimethylformamide in a 3 ml reaction volume with $[E] = 0.1\text{--}0.3 \mu\text{M}$ and $[S] = 20\text{--}200 \mu\text{M}$ were measured at 410 nm and quantified by using $\epsilon_{410} = 8.8 \times 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$ (Erlanger et al., 1961). This value is essentially pH-independent at pH values above 3; the value of the pK_a of the 4-nitroanilinium cation is 1.0 (Bascombe and Bell, 1959). When corrections are significant they are made by using

$$\epsilon_{410} = 8.8 \times 10^3 / (1 + 10 [H^+]) \text{ M}^{-1} \cdot \text{cm}^{-1}$$

Model building

The model of the three-dimensional structure of chymopapain M was built by using comparative modelling techniques (Sutcliffe et al., 1987a,b; Blundell et al., 1988; Topham et al., 1990a-d, 1993) based on a procedure similar to that described in detail for PPQ by Topham et al. (1991). The model was built from the amino acid sequence of chymopapain M (Ritonja et al., 1989) and the three currently known high-resolution crystal structures of the homologous enzymes, papain [0.16 nm (Kamphuis et al., 1985)], actinidin [0.17 nm (Baker, 1980)], and PPQ [0.18 nm (Pickersgill et al., 1991)] by using the COMPOSER suite of programs and refined manually by interactive graphics on an Evans and Sutherland PS390 before energy minimization using the Tripos forcefield within the SYBYL 5.5 package. Computations were performed on an Iris-4D workstation.

The conserved framework and all but one of the structurally variable regions were constructed from the structures of the three homologous enzymes. Their co-ordinates are available from the Brookhaven Protein Data Bank (PDB) codes: papain, 9PAP; actinidin, 2ACT; PPQ, 1PPO). The loop fragment used to model the region 150-152 (EGS) was taken from residues 86-88 (PMA) of xylose isomerase (PDB code 3 XIA). This was necessary because Ser-152 aligns with a conserved *cis*-proline residue (152, papain numbering) in the three starting structures. All of the loops in the model were selected by using the fragment-ranking procedure described by Topham et al. (1993). Arg-65 was built in the same extended positive ϕ conformation as that of Gly-65 in papain and the analogous residues (both Gly) in actinidin and PPQ.

Comparison of the three-dimensional structure of papain and the model of the three-dimensional structure of chymopapain M by interactive computer graphics

Superposition of the model of chymopapain M and the structure of papain (PDB code 9PAP) and measurements of bond angles and hydrogen-bond lengths were carried out by using the Insight II molecular graphics package running on a Silicon Graphics Iris Indigo XS24 Workstation. The polypeptide backbones were aligned by using 138 residues that are identical in the two proteins, comprising 2096 atoms in total, with a root-mean-square deviation of 0.116 nm.

Figure 1 Fractionation of the latex of *Carica papaya* by cation-exchange f.p.l.c. on S-Sepharose Fast Flow: Identification of chymopapain M as the late-eluted chymopapain with specificity for P₁ Gly substrates

To prevent thiol-dependent proteolysis, the papaya latex was dissolved in a saturated solution of 2-Py-S-S-2-Py to convert thiol groups into mixed disulphides. The various cysteine proteinase fractions (protected as E-S-S-2-Py) are indicated as PAP (papain), CA (chymopapain A), CB (chymopapain B), CM (chymopapain M) and PPQ. The fractions containing CM are the only ones with high activity towards Boc-Gly-4NPE and low activity towards Bz-L-Arg-4NA. The activity towards Bz-L-Arg-4NA (presumably the result of a small amount of late-eluted contaminant B-type chymopapain) is removed by covalent chromatography (see the text), and rechromatography of the non-bonded fraction by f.p.l.c. produced a single peak of chymopapain M with activity towards Boc-Gly-4NPE and N-Ac-L-Phe-Gly-4NA but devoid of activity towards Bz-L-Arg-4NA. —, A_{280} ; —, activity towards Bz-L-Arg-4NA; ---, activity towards Boc-Gly-4NPE.

RESULTS AND DISCUSSION

Detection and isolation of chymopapain M

To extend our earlier work on the chymopapains [see Topham et al. (1990d) and references therein] we fractionated papaya latex by cation-exchange f.p.l.c. (Figure 1). In view of the fact that papaya cysteine proteinases are known to undergo thiol-dependent proteolysis (Baines et al., 1986), the latex was dissolved in a saturated solution of 2-Py-S-S-2-Py which converts the thiol groups into mixed disulphides (E-S-S-2-Py) and thus reversibly inactivates the various cysteine proteinase components. We screened the protein peaks in Figure 1 for catalytic activity (after cysteine activation) not only with Bz-L-Arg-4NA but also with Boc-Gly-4NPE, expecting to find the enzyme PPIV with high specificity for Gly at P₁ reported by Buttler et al. (1989) to occur near to the early-eluted chymopapain fractions. Figure 1 shows the presence in the latex of papain, chymopapains A and B and PPQ, all of which exhibit catalytic activity towards both Bz-L-Arg-4NA and Boc-Gly-4NPE, and papaya lysozyme (see Brocklehurst et al., 1981). We were surprised to find that PPIV as described by Buttler et al. (1989) was apparently not present. Instead of this early-eluted P₁ Gly-specific proteinase, however, there was a substantial amount of a late-eluted protein with high activity towards Boc-Gly-4NPE and little or no activity towards Bz-L-Arg-4NA. This enzyme was incompletely resolved from another protein on the trailing edge of the papaya lysozyme band with activity towards both Boc-Gly-4NPE and Bz-L-Arg-4NA. In an attempt to isolate the cysteine proteinase component(s) eluted after papaya lysozyme in Figure 1, fractions 104–110 were pooled and applied to Sepharose–glutathione-2-pyridyl disulphide gel at pH 4. At this pH all cysteine proteinases with intact catalytic sites thus far investigated react with the electrophilic sites of the gel, and other components pass through. This type of hydron-activated covalent chromatography [see Brocklehurst et

al. (1985a) for a review] permits the preparation of fully active cysteine proteinases. Its basis is the specific reaction of catalytic site $-S^-/-Im^+H$ ion-pairs with activated electrophilic sites produced by hydronation (Gel-S-S-2-Py ^+H). The unexpected result of using this procedure in the present work was that the P₁ Gly-specific proteinase failed to react with the covalent chromatography gel. The procedure did remove the cysteine proteinase component with activity towards both Bz-L-Arg-4NA and Boc-Gly-4NPE (a late-eluted chymopapain B) and rechromatography by f.p.l.c. of the fraction that did not bond to the covalent chromatography gel produced a single peak of protein devoid of activity towards Bz-L-Arg-4NA but with high activity towards Boc-Gly-4NPE. The same result was obtained when the initial part of the covalent chromatography procedure was performed at pH 8 instead of at pH 4. The failure of the late-eluted cysteine proteinase to bond to the gel therefore was not due to an unusually high value of the pK_a characteristic of the formation of the ion-pair state from unreactive $-SH/-Im^+H$. Application of the covalent chromatography procedure at pH 8 usually results in the isolation of any thiol-containing molecule. The failure of covalent chromatography at either pH 4 or pH 8 to isolate the late-eluted P₁ Gly-specific cysteine proteinase compels the view that some feature of the gel ligand hinders the approach of the catalytic site of the enzyme to the disulphide site of the gel, an observation that is in accord with the P₁ Gly specificity.

The P₁ Gly-specific cysteine proteinase that had been freed from late-eluted chymopapain B by covalent chromatography was shown to exhibit two other key properties of 'PPIV'. Thus this enzyme (a) binds to the affinity-chromatography gel, Sepharose-6-aminohexanoyl-Gly-L-Phe-aminoacetonitrile, and the eluate from the affinity gel exhibits elution behaviour in ion-exchange chromatography identical with that reported by Buttler et al. (1989) and that of enzyme prepared in the present work directly by this affinity chromatography procedure, and (b) is not inhibited by chicken cystatin.

It would seem reasonable to suppose therefore that the P₁ Gly-specific enzyme (chymopapain M of Figure 1, see below) is the same enzyme that Buttler et al. (1989) named PPIV, were it not for one apparent difference. Thus Buttler et al. (1989) had reported PPIV to be eluted in ion-exchange chromatography close to the position expected for early-eluted forms of chymopapain, whereas the P₁ Gly-specific cysteine proteinase of Figure 1 is a late-eluted component (indeed the last) of the chymopapain band. The answer to this apparent difficulty is that the conclusion about elution behaviour reached by Buttler et al. (1989) is based on an inappropriate comparison of the elution positions of the various protein components of papaya latex observed under different experimental conditions. When the mixture of PPIV and papain was isolated by their affinity-chromatography procedure, proteolysis was inhibited. When the elution position of inhibited PPIV is compared with the elution profile of the whole latex in which proteolysis has not been inhibited (see e.g. Buttler and Barrett, 1984), it does appear that PPIV is eluted early in the chymopapain region. Thus it appears that 'PPIV' may seem to be eluted early only when the behaviour of non-proteolysed 'PPIV' is compared with the limited proteolysis products of the chymopapain band.

To evaluate this possibility the changes in the positions of the protein components of papaya latex as proteolysis was allowed to occur progressively were monitored and some of the elution patterns are shown in Figure 2. When thiol-dependent proteolysis is inhibited (sample a, see the Materials and methods section), the P₁ Gly-specific proteinase is the last-eluted component of the chymopapain band. As thiol-dependent proteolysis is allowed to occur (Figures 2b–2d), samples f, g and i, the complex chymo-

Figure 2 Changes in the elution behaviour of the components of the latex of *Carica papaya* consequent on progressive proteolysis: clarification of the elution behaviour of chymopapain M (PPIV)

—, A_{280} ; ---, Bz-L-Arg-4NA activity ($\Delta A_{412}/\text{min}$); ----, Boc-Gly-4NPE activity ($\Delta A_{400}/\text{min}$). (a) Chromatography of the latex solution with all cysteine proteinase components reversibly inhibited as E-S-S-2-Py (sample a; see the Materials and methods section for a description of samples a, f, g and i); this elution profile is analogous to that shown in Figure 1 except that the resolution is not as great (because of the smaller number of fractions collected over a shorter period of time with a steeper salt gradient to provide more rapid analysis); the CM fractions occur as a late-eluted chymopapain. (b)–(d) show changes in the elution profile resulting from addition of L-cysteine to promote proteolysis (see the Materials and methods section for details); (b) elution profile after 2 h of proteolysis (sample f); (c) elution profile after 4 h of proteolysis (sample g); (d) elution profile after 16 h of proteolysis (sample i). Chymopapain M-S-S-2-Py is eluted in fractions 37–42 in (a); in (b)–(d) changes in the position of the components of the chymopapain band occur; the consequence of this is that, if the elution position of chymopapain M-S-S-2-Py is compared with the elution profile of proteolysed latex [as was done by Buttle et al. (1989)], chymopapain M would appear to be an early eluted chymopapain.

papain band moves to the right in the elution profile (i.e. is gradually eluted later). As proteolysis is allowed to occur progressively, further changes in the elution profile are observed (see e.g. Figures 2c and 2d). Neither storage of the uninhibited latex (sample b) nor $(\text{NH}_4)_2\text{SO}_4$ precipitation (sample c) resulted in changes to the elution profile. Extensive proteolysis (samples j–n, storage in L-cysteine solution for 24–120 h) produced further changes in the elution profile including changes in the relative heights of the peaks within the ‘chymopapain’ band and splitting of the PP Ω band. The particularly important conclusion from these proteolysis studies is confirmation that when the elution position of inhibited chymopapain M is compared with the elution profile of the whole latex in which proteolysis has not been inhibited [which, although inappropriate, is the comparison

that was made by Buttle et al. (1989)] chymopapain M would then appear to be eluted early in the chymopapain region even though in reality it is a late-eluted chymopapain.

It should be stressed that chymopapain M does not become an early-eluted chymopapain: it appears to be an early-eluted chymopapain only when an inappropriate comparison is made of the elution positions of the various protein components of papaya latex observed under different experimental conditions.

Cysteine proteinase components of the chymopapain band: chymopapains A, B and M

The results of the experiments on the elution behaviour of the cysteine proteinase components of the chymopapain band discussed above, together with other evidence recorded in the literature over a number of years, now permit clarification of the types of chymopapain present in papaya latex. The presence of A-type and B-type chymopapains is well established (Brocklehurst and Salih, 1983). The evidence presented by Brocklehurst et al. (1984) leads to the conclusion that chymopapain S (Khan and Polgár, 1983) is chymopapain A. Chymopapain C (Claggett et al., 1974) is almost certainly the very basic proteinase PP Ω , which was shown to be present sometimes in samples of commercial chymopapain prepared by salt fractionation (Robinson, 1975). It may now be concluded that the late-eluted monothiol cysteine proteinase, chymopapain M, is the only major cysteine proteinase component present in the chymopapain band in addition to chymopapains A and B. Chymopapain M is probably the major component of the protein fraction noted by Schack (1967) and studied by Lynn (1979) (who named it PPB), Lynn and Yaguchi (1979) and Polgár (1981). It is particularly important to note that chymopapain M is the same enzyme that Buttle et al. (1989) claimed to be an undiscovered cysteine proteinase when they named it PPIV. Had they compared its elution position with those of the other cysteine proteinases under conditions such that they too were protected from proteolysis, PPIV would have been recognized as PPB, i.e. a late-eluted chymopapain. Mixtures of chymopapain M and one of the other (bismercapto) chymopapains, either native or as proteolysis products, could be responsible for the apparent existence of three types of thiol group per ‘molecule’ noted in some chymopapain preparations (Baines et al., 1986).

We prefer the name chymopapain M for the late-eluted component of the chymopapain band, not only because this enzyme is eluted from ion-exchange columns as part of the complex chymopapain band, but also because both PPB and PPIV depend upon the use of names other than PP Ω for the most basic cysteine proteinase component of papaya latex. This enzyme was originally named papaya peptidase A (PPA) by Schack (1967) and was subsequently given a variety of names (see Topham et al., 1991). Because the number and nature of the cysteine proteinase components eluted from ion-exchange columns between papain and the most basic enzyme vary with the source of the papaya latex and its treatment, we suggested that the most basic enzyme might usefully be called papaya proteinase Ω (PP Ω) as it is most certainly the last cysteine proteinase component to be eluted (Brocklehurst et al., 1985b,c). This name has since been used by laboratories reporting results of structural and kinetic investigations of the enzyme (e.g. Dubois et al., 1988a,b; Jacquet et al., 1989; Pickersgill et al., 1990, 1991; Sumner et al., 1993). The other names for chymopapain M, i.e. PPB and PPIV, depend upon naming PP Ω either PPA (Schack, 1967) or papaya proteinase III, PPIII (Buttle and Barrett, 1984). Because neither PPA nor PPIII is now widely used, continued use of PPB or PPIV seems unhelpful.

Oxidation of chymopapain M

In view of earlier reports (Lynn, 1979; Buttle et al., 1989) that chymopapain M (i.e. PPB or PPIV; see above) undergoes relatively rapid loss of activity in the absence of a reducing agent, it was necessary to undertake minimal characterization of this process (as described in the Materials and methods section) in connection with kinetic study of catalysis. The decrease in the concentration of the catalytic-site thiol group with time was monitored by observing the decrease in the amplitude of the increase in A_{343} consequent on reaction with 2-Py-S-S-2-Py. The increase in A_{343} recorded on the visual display unit of the SF machine obeyed good first-order kinetics for all samples. This provides no evidence that activity loss might be due to protein unfolding or proteolysis in addition to oxidation of the thiol group. Either of these additional processes might be expected to produce deviation from a single exponential as a result of loss of the enhanced thiol reactivity characteristic of the native conformation of the active centre. When chymopapain M is stored at 25.0 °C in 0.1 M KCl containing 1 mM EDTA at an initial concentration of 2 μ M (the conditions under which the enzyme is kept during a series of kinetic studies of catalysed substrate hydrolysis), oxidation occurs with $k_{\text{obs.}} = 3.08 \times 10^{-5} \text{ s}^{-1}$. This corresponds to $t_{1/2} = 2.25 \times 10^4 \text{ s} \equiv 62.5 \text{ h}$. Exposure of the solution to daylight does not affect the rate of oxidation.

Characterization of the kinetically influential ionizations of chymopapain M by stopped-flow kinetic analysis of its reactions with thiol-specific reactivity probes

A major feature of the catalytic sites of those cysteine proteinases that have been subjected to detailed study is an interactive system comprising the side chains of a Cys residue and a His residue of which a thiolate/imidazolium ion-pair is a plausible component. In the case of the much-studied enzyme, papain, these catalytic-site residues are Cys-25 and His-159. The S atom of Cys-25 becomes transiently acylated during catalysis and the side chain of His-159 is considered to play a number of roles involving maintenance of the thiolate anion in neutral and weakly acidic media, general acid catalysis of leaving-group departure from an anionic tetrahedral intermediate during acylation and general base catalysis of the hydrolysis of the acyl-enzyme intermediate. Calculations by Arad et al. (1990) suggest that the nucleophilic sulphur atom can approach the carbonyl carbon atom of a substrate but the potential is repulsive until accompanied by (His-159)-Im⁺H...amide N hydron transfer.

A particularly clear way of revealing the existence and characteristics of a nucleophilic Cys/His interactive system is by kinetic

analysis of its pH-dependent reactions with 2-pyridyl disulphides such as 2-Py-S-S-2-Py (I) and Et-S-S-2-Py (II) augmented by analogous kinetic data from reaction with 4,4'-dipyrimidyl disulphide (III) (see Mellor et al., 1993b,c). 2-Pyridyl disulphides are particularly effective two-hydronic-state electrophiles (Brocklehurst, 1982), and reactions of papain with simple members of the series such as (I) and (II) that do not contain key substrate-analogous binding sites are characterized by a striking rate maximum at pH approx. 4, a rate minimum at pH 6–7 and a plateau of higher reactivity at high pH. The high reactivity of these reagents, the reactions of which (specifically with thiol groups) occur on the stopped-flow time scale, permits reliable kinetic study of reactions of cysteine proteinases over a wide range of pH. The particularly high reactivity at pH values around 4 derives from the coexistence of significant concentrations of the activated (hydronated) inhibitor (R-S-S-2-Py⁺H) and the nucleophilic interactive Cys-25/His-159 system (hereafter designated -S⁻/-Im⁺H ion-pair). Analysis of the pH-second-order rate constant (k) profile provides pH-independent rate constants ($\bar{k}$) for each of the reactive ionization states, and macroscopic pK_a values characteristic of their formation and loss by further hydronic dissociation.

Augmentation of the pH-dependent kinetic data obtained with 2-pyridyl disulphide probes [such as (I) and (II)] with analogous data obtained with 4,4'-dipyrimidyl disulphide (III) permits the pK_a values characteristic of the formation and subsequent dehydronation of the Cys-S⁻/His-Im⁺H ion-pair state to be distinguished from those characteristic of other kinetically influential ionizations of the enzyme molecule (Mellor et al., 1993b,c). Identification of the pK_a values of the ion-pair state of the enzyme is facilitated by the lower basicity of (III) (pK_a 0.91) than that of 2-pyridyl disulphides such as (I) (pK_a 2.45), combined with retention of high reactivity across the pH range. The lower basicity greatly decreases the intervention of reaction with the activated (hydronated) form of the pyrimidyl disulphide and provides a simpler form of pH-dependence. Evaluation of papain by combined use of 2-pyridyl disulphide and 4,4'-dipyrimidyl disulphide probes demonstrated the existence of five kinetically influential ionizations with pK_a values 3.4, 4.0, 5.0, 8.3 and 10.0, of which 3.4 and 8.3 relate to the formation and subsequent dehydronation of the (Cys-25)-S⁻/(His-159)-Im⁺H state.

In the present work the existence in chymopapain M of a nucleophilic interactive catalytic-site system analogous to that in papain has been deduced and the kinetically influential ionizations identified by the data shown in Figures 3(a)–3(c). The assignment of this system to a Cys/His couple by structural analysis is presented below. The existence of the nucleophilic interactive system is revealed by the characteristic shape of the pH- k profiles in Figures 3(a) and (b) for the reactions of chymopapain M with the two 2-pyridyl disulphide probes (I) and (II), i.e. the striking rate maximum in acidic media, the rate minimum in neutral media and the plateau of higher reactivity at high pH. The values of the characterizing parameters needed to fit the data to pH-dependent rate equations are compared with the corresponding values for the analogous reactions of papain in Table 1. The following points are particularly noteworthy: (i) in acidic media chymopapain M is characterized by four kinetically influential ionizations (pK_a values 3.4, 3.4, 4.3 and 5.6), whereas papain is characterized by one fewer (pK_a values 3.4, 4.0 and 5.0); (ii) the major component of the complex bell-shaped curve has a rate maximum characterized by two pK_a values, both of value 3.4, which is symptomatic of positively co-operative hydron binding and is responsible for the narrowness of the bell (see Dixon, 1973); (iii) in alkaline media the increase in k towards the plateau value at high pH is characterized by pK_a 8.7; for

Figure 3 pH-dependences of k for the reactions of the catalytic-site thiol group of chymopapain M at 25 °C and / 0.1 M with (a) 2-Py-S-S-2-Py, (b) Et-S-S-2-Py and (c) 4,4'-dipyrimidyl disulphide

For (a) and (b) the continuous lines are theoretical for eqn. (2) which relates to reaction in six hydronic states with the following values of the characterizing parameters (molecular pK_a values and pH-independent rate constants, $\tilde{k}$) with values of $\tilde{k}_1$ – $\tilde{k}_6$ in $M^{-1} \cdot s^{-1}$:

$$k = \frac{\tilde{k}_1}{1 + \frac{K_I}{[H^+]} + \frac{K_{II}}{[H^+]^2} + \frac{K_{III}}{[H^+]^3} + \frac{K_{IV}}{[H^+]^4} + \frac{K_{V}}{[H^+]^5} + \frac{K_{VI}}{[H^+]^6}} + \frac{\tilde{k}_2}{1 + \frac{[H^+]^2}{K_I K_{II}} + \frac{[H^+]}{K_{II}} + \frac{K_{III}}{[H^+]^2} + \frac{K_{IV}}{[H^+]^3} + \frac{K_{V}}{[H^+]^4} + \frac{K_{VI}}{[H^+]^5}} + \frac{\tilde{k}_3}{1 + \frac{[H^+]^3}{K_I K_{II} K_{III}} + \frac{[H^+]^2}{K_{II} K_{III}} + \frac{[H^+]}{K_{III}} + \frac{K_{IV}}{[H^+]} + \frac{K_{V}}{[H^+]^2} + \frac{K_{VI}}{[H^+]^3}} + \frac{\tilde{k}_4}{1 + \frac{[H^+]^4}{K_I K_{II} K_{III} K_{IV}} + \frac{[H^+]^3}{K_{II} K_{III} K_{IV}} + \frac{[H^+]^2}{K_{III} K_{IV}} + \frac{[H^+]}{K_{IV}} + \frac{K_{V}}{[H^+]} + \frac{K_{VI}}{[H^+]^2}} + \frac{\tilde{k}_5}{1 + \frac{[H^+]^5}{K_I K_{II} K_{III} K_{IV} K_{V}} + \frac{[H^+]^4}{K_{II} K_{III} K_{IV} K_{V}} + \frac{[H^+]^3}{K_{III} K_{IV} K_{V}} + \frac{[H^+]^2}{K_{IV} K_{V}} + \frac{[H^+]}{K_{V}} + \frac{K_{VI}}{[H^+]}} + \frac{\tilde{k}_6}{1 + \frac{[H^+]^6}{K_I K_{II} K_{III} K_{IV} K_{V} K_{VI}} + \frac{[H^+]^5}{K_{II} K_{III} K_{IV} K_{V} K_{VI}} + \frac{[H^+]^4}{K_{III} K_{IV} K_{V} K_{VI}} + \frac{[H^+]^3}{K_{IV} K_{V} K_{VI}} + \frac{[H^+]^2}{K_{V} K_{VI}} + \frac{[H^+]}{K_{VI}}}$$

chymopapain M there is no evidence from the reactivity probe data for the additional ionization of pK_a about 10 seen in reactions of papain (see Mellor et al., 1993b), although a pK_a of 9.9 is seen in the pH-dependence of k_{cat}/K_m (see below).

The pH- k profile shown in Figure 3(c) for the reaction of chymopapain M with 4,4'-dipyrimidyl disulphide (III) is simpler than those in Figures 3(a) and 3(b), as is the case for papain. For the papain reaction the pH- k profile is of double sigmoid form (Mellor et al., 1993b,c) and exhibits only the pK_a values of 3.4 and 8.3 characteristic of the (Cys-25)-S⁻/(His-159)-Im⁺H ion-pair system. The pH- k profile for the reaction of chymopapain M with (III) is slightly more complex, with reactivity dependent on three of the five enzyme ionizations (pK_a values 3.4, 5.6 and 8.7) seen with the 2-pyridyl disulphide probes. The additional ionization with pK_a 3.4 and the other ionization with pK_a 4.3 do not significantly influence the reaction of chymopapain M with (III). The simplest interpretation of the data in Figures 3(a) and 3(b) taken together with those in Figure 3(c) is that the ion-pair state of chymopapain M is formed by hydronic dissociation with pK_a 3.4 and transformed into the simple thiolate anion state (-S⁻/Im) by hydronic dissociation with pK_a 8.7, and that the reactivity of the ion-pair is modulated by the state of ionization of other groups associated with pK_a values 3.4, 4.3 and 5.6. This is similar to the situation deduced for papain but with the noteworthy difference that, in chymopapain M, there is an additional kinetically influential ionization operative in acidic media with a pK_a value (3.4) similar to that characteristic of the formation of the nucleophilic ion-pair state. The assignment of this additional regulatory ionization by structural analysis is described below.

The pH-dependence of k_{cat}/K_m for the hydrolysis of N-Ac-L-Phe-Gly-4NA catalysed by chymopapain M and by papain: evidence that for both enzymes the existence of the catalytic-site Cys-S⁻/His-Im⁺H ion-pair is an insufficient condition for catalytic competence

To relate the kinetically influential ionizations of chymopapain M detected by using the disulphide time-dependent inhibitors as

for (a) $pK_I = 2.45$ (the pK_a value of 2-Py-S-S-2-Py⁺H), $pK_{II} = pK_{III} = 3.4$, $pK_{IV} = 4.3$, $pK_V = 5.6$, $pK_{VI} = 8.7$, $\tilde{k}_1 = 8.0 \times 10^3$, $\tilde{k}_2 = 9.0 \times 10^4$, $\tilde{k}_3 = 8.0 \times 10^3$, $\tilde{k}_4 = 8.0 \times 10^3$, $\tilde{k}_5 = 50$, $\tilde{k}_6 = 7.75 \times 10^2$; the inset in (a) shows the part of the pH- k profile at pH values greater than 7.5 on an expanded scale; for (b) $pK_I = 2.8$, $pK_{II} = pK_{III} = 3.5$, $pK_{IV} = 4.3$, $pK_V = 5.6$, $pK_{VI} = 8.7$, $\tilde{k}_1 = 5.0 \times 10^2$, $\tilde{k}_2 = 1.32 \times 10^5$, $\tilde{k}_3 = 5.0 \times 10^3$, $\tilde{k}_4 = 2.5 \times 10^3$, $\tilde{k}_5 = 1.0 \times 10^3$, $\tilde{k}_6 = 1.7 \times 10^4$; the pK_a of Et-S-S-2Py⁺H is 2.9 and the first two enzyme pK_a values are each 3.4; the proximity of these numerical values results in perturbation leading to lower and higher values respectively of the observed values (i.e. 2.8 and 3.5 respectively). The broken lines in (a) and (b) correspond to contributions to k of the various hydronic states provided by the individual terms of eqn. (2). The reaction with 2-Py-S-S-2-Py (a) shows the kinetically influential ionization with pK_a 5.6 particularly clearly and the reaction with Et-S-S-2-Py (b) shows the ionization with pK_a 8.7 particularly clearly. For (c) the continuous line is theoretical for eqn. (3) which relates to reaction in three reactive hydronic states with the following values of the characterizing parameters with values of $\tilde{k}_1$ – $\tilde{k}_3$ in $M^{-1} \cdot s^{-1}$:

$$k = \frac{\tilde{k}_1}{1 + \frac{[H^+]}{K_I} + \frac{K_{II}}{[H^+]} + \frac{K_{III}}{[H^+]^2}} + \frac{\tilde{k}_2}{1 + \frac{[H^+]^2}{K_I K_{II}} + \frac{[H^+]}{K_{II}} + \frac{K_{III}}{[H^+]}} + \frac{\tilde{k}_3}{1 + \frac{[H^+]^3}{K_I K_{II} K_{III}} + \frac{[H^+]^2}{K_{II} K_{III}} + \frac{[H^+]}{K_{III}}} \quad (3)$$

$pK_I = 3.4$, $pK_{II} = 5.6$, $pK_{III} = 8.7$, $\tilde{k}_1 = 2.1 \times 10^3$, $\tilde{k}_2 = 4.5 \times 10^3$, $\tilde{k}_3 = 3.5 \times 10^4$; the broken lines correspond to contributions to k of the various hydronic states provided by the individual terms of eqn. (3). Reaction with the hydronated form of 4,4'-dipyrimidyl disulphide (pK_a 0.91) at low pH is negligible.

Table 1 Comparison of the characterizing parameters of the reactions of chymopapain M and papain with 2-Py-S-S-2-Py, Et-S-S-2-Py and 4,4'-dipyrimidyl disulphide in various ionization states at 25.0 °C and / 0.1 M

Except for the reaction of chymopapain M with 4,4'-dipyrimidyl disulphide, pK_i is or approximates to the pK_a value of the monohydrated form of the disulphide; in the reactions of Et-S-S-2-Py the value of pK_i (2.8) is perturbed slightly from its measured value (2.9) by the proximity of this pK_a value and the first two enzyme pK_a values (3.4), which are both perturbed to 3.5. pH-independent rate constants (k_1 – k_6) ($M^{-1} \cdot s^{-1}$) are given in parentheses.

Enzyme	Disulphide	Macroscopic pK_a values					
		pK_i	pK_{II}	pK_{III}	pK_{IV}	pK_V	pK_{VI}
Chymopapain M	4,4'-Dipyrimidyl disulphide	3.4 (2.1×10^3)	5.6 (4.5×10^3)	8.7 (3.5×10^4)	—	—	—
Papain	4,4'-Dipyrimidyl disulphide	0.91 (1.0×10^2)	3.4 (2.05×10^3)	8.3 (3.95×10^4)	—	—	—
Chymopapain M	2-Py-S-S-2-Py	2.45 (8.0×10^3)	3.4 (9.0×10^4)	3.4 (8.0×10^3)	4.3 (8.0×10^3)	5.6 (50)	8.7 (7.75×10^2)
Papain	2-Py-S-S-2-Py	2.45 (1.0×10^3)	3.45 (3.9×10^4)	4.0 (5.5×10^3)	5.0 (5.0×10^2)	10.0 (1.7×10^3)	—
Chymopapain M	Et-S-S-2-Py	2.8 (5.0×10^2)	3.5 (1.32×10^5)	3.5 (5.0×10^3)	4.3 (2.5×10^3)	5.6 (1.0×10^3)	8.7 (1.7×10^4)
Papain	Et-S-S-2-Py	2.8 (10)	3.5 (5.5×10^4)	4.2 (1.4×10^4)	5.3 (5.2×10^3)	10.0 (6.2×10^3)	—

reactivity probes to the requirements for catalytic competence, it is necessary to analyse the pH-dependence of k_{cat}/K_m . This kinetic parameter (Bender and Kézdy, 1965; Brocklehurst et al., 1968; Fersht, 1977) is now widely used in the characterization of enzymes and the study of mechanism. In terms of the conventional three-step acyl-enzyme model, $k_{cat}/K_m = k_{+2}/K_m$ (acyl) $\equiv k_{+1}k_{+2}/(k_{-1} + k_{+2})$ and when quasi-equilibrium obtains around the adsorptive complex, ES, $k_{cat}/K_m \approx k_{+2}/K_s$, where $K_s = k_{-1}/k_{+1}$. The pH-dependence of k_{cat}/K_m for substrates and k for time-dependent inhibitors reveals free reactant state pK_a values provided that the value of k_{cat}/K_m (or k) is substantially smaller than that of the on-rate constant for adsorptive complex-formation (see Brocklehurst, 1994). This condition corresponds to quasi-equilibrium around ES, i.e. the condition $k_{+2} \ll k_{-1}$. Before the present work there had been no reports of the pH-dependence of k_{cat}/K_m for reactions catalysed by chymopapain M (PPB or PPIV) although some kinetic studies (indicated below) had been reported.

Chymopapain M is probably the unnamed enzyme with proteolytic activity towards casein observed by Schack (1967) to be present in commercial dried papaya latex (18 % of the latex) in addition to PPA (14 % of the latex, now known as PPQ). This enzyme was studied by Lynn (1979), who named it PPB. Digestion of the insulin B chain produced bond scissions comparable with those produced by papain, 'chymopapain' and PPQ. Profiles of v at arbitrary fixed substrate concentration against pH in the range 5–8 were obtained for the hydrolysis of both Z-Gly-4NPE and its L-Lys analogue and exhibited maximal activity around pH 6. Catalytic activity could not be detected towards either the L-Tyr analogue or Bz-L-Arg-4NA. The activity of the PPB preparation towards the substrate with L-Lys at P_1 may have been due to contamination by a late-eluted chymopapain B.

In the present work the chymopapain M-catalysed hydrolysis of N-Ac-L-Phe-Gly-4NA was shown to obey Michaelis-Menten kinetics at pH 6.0 by the linearity of a plot of $[S]_0/v$ versus $[S]_0$ for values of $[S]_0$ in the range 10–90 μM . Values of K_m and k_{cat} were obtained by fitting the data to the Michaelis-Menten equation as described in the Materials and methods section and using the catalytic-site concentration determined by titration with 2-Py-S-S-2-Py. The values of the parameters ($\pm$ S.E.M.) $K_m = 3.2 \pm 1.7$ mM, $k_{cat} = 4.7 \pm 2.5$ s $^{-1}$ and $k_{cat}/$

$K_m = 1497 \pm 24$ M $^{-1} \cdot s^{-1}$ are similar to those for the corresponding papain-catalysed hydrolysis, $K_m = 1.5 \pm 0.6$ mM, $k_{cat} = 3.2 \pm 1.2$ s $^{-1}$ and $k_{cat}/K_m = 2088 \pm 55$ M $^{-1} \cdot s^{-1}$ (Kowlessur et al., 1990). In the paper by Kowlessur et al. (1990) the concentration range of N-Ac-L-Phe-Gly-4NA is given correctly on p. 654, column 2, paragraph 2, line 5 as approx. 20–220 μM but is printed incorrectly on p. 654, column 2, paragraph 3, line 9 as 20–240 mM. N-Ac-L-Phe-Gly-4NA, the use of which as a substrate for chymopapain M has not been reported previously, was chosen to meet the specificity requirements predicted from the model of the three-dimensional structure of the enzyme (see below) and to provide both a chromophoric leaving group and better stability towards non-catalysed hydrolysis than that provided by reactive ester substrates. To ensure that the condition $[S]_0 \ll K_m$ was obeyed at other pH values, v was determined at each pH at several values of $[S]_0$ in the range 20–100 μM . For various values of $[S]_0$ in this range the values k_{cat}/K_m calculated from $v/[E] = (k_{cat}/K_m)[S]$ (where under these conditions $[S] = [S]_0$) were essentially identical at a given pH value.

Because it was necessary to perform a large number of kinetic runs throughout the day, two methods were used to obviate the problem of irreversible enzyme oxidation. In one method the concentration of active enzyme remaining at a particular time was calculated by using the value of the rate constant for oxidation determined as described above as 3.08×10^{-5} s $^{-1}$. In the other method a sample of activated activator-free enzyme was allowed to react with 2-Py-S-S-2-Py immediately after the catalytic-site concentration had been determined. This provides a sample of reversibly oxidized enzyme (enzyme-S-S-2-Py) which is converted into active enzyme in the assay mixture by incubation with 13 mM L-cysteine for 20 min before the reaction is started by addition of substrate. Each of the two methods gave closely similar results for the values of k_{cat}/K_m (Figure 4a).

The pH-dependence of k_{cat}/K_m for the chymopapain M-catalysed hydrolysis of N-Ac-L-Phe-Gly-4NA is shown in Figure 4(a) and the form of this pH-dependence is compared with that for the analogous papain-catalysed hydrolysis in Figure 4(c) (the data for the papain catalysis being shown in Figure 4b). The most striking feature of both pH- (k_{cat}/K_m) profiles (Figures 4a and 4b) is that very little catalytic activity (about 5–7 % of the maximum) develops consequent on formation of the catalytic

Figure 4 pH-dependences of k_{cat}/K_m for the hydrolysis of N-Ac-L-Phe-Gly-4NA at 25.0 °C and / 0.1 M catalysed by (a) and (c) chymopapain M and (b) and (c) papain

For (a) the open circles correspond to data obtained by calculating the concentration of active chymopapain M by allowing for oxidation of the enzyme as described in the Materials and methods section and closed squares correspond to data obtained by using enzyme stored as E-S-S-Py with thiol-promoted activation immediately before addition of substrate in the assay. For both (a) and (b) the continuous lines are theoretical for eqn. (4), where $k \equiv k_{\text{cat}}/K_m$, which relates to catalysis in three active hydronic states bounded by two inactive hydronic states at low and high pH with the following values of the characterizing parameters with values of $\bar{k}_{\text{cat}_1}/K_{m_1} - \bar{k}_{\text{cat}_4}/K_{m_4}$ (shown as $\bar{k}_1 - \bar{k}_4$, severally) in $\text{M}^{-1} \cdot \text{s}^{-1}$:

$$k = \frac{\bar{k}_1}{1 + \frac{[\text{H}^+]}{K_1} + \frac{K_{\text{II}}}{[\text{H}^+]} + \frac{K_{\text{II}} K_{\text{III}}}{[\text{H}^+]^2} + \frac{K_{\text{II}} K_{\text{III}} K_{\text{IV}}}{[\text{H}^+]^3}} + \frac{\bar{k}_2}{1 + \frac{[\text{H}^+]^2}{K_1 K_{\text{II}}} + \frac{[\text{H}^+]}{K_{\text{II}}} + \frac{K_{\text{III}}}{[\text{H}^+]} + \frac{K_{\text{III}} K_{\text{IV}}}{[\text{H}^+]^2}} + \frac{\bar{k}_3}{1 + \frac{[\text{H}^+]^3}{K_1 K_{\text{II}} K_{\text{III}}} + \frac{[\text{H}^+]^2}{K_{\text{II}} K_{\text{III}}} + \frac{[\text{H}^+]}{K_{\text{III}}} + \frac{K_{\text{IV}}}{[\text{H}^+]}} + \frac{\bar{k}_4}{1 + \frac{[\text{H}^+]^4}{K_1 K_{\text{II}} K_{\text{III}} K_{\text{IV}}} + \frac{[\text{H}^+]^3}{K_{\text{II}} K_{\text{III}} K_{\text{IV}}} + \frac{[\text{H}^+]^2}{K_{\text{III}} K_{\text{IV}}} + \frac{[\text{H}^+]}{K_{\text{IV}}}} \quad (4)$$

for (a) $pK_1 = 4.3$, $pK_{\text{II}} = 5.6$, $pK_{\text{III}} = 8.7$, $pK_{\text{IV}} = 9.9$, $\bar{k}_1 = 1.05 \times 10^3$, $\bar{k}_2 = 1.4 \times 10^3$, $\bar{k}_3 = 6.0 \times 10^2$, $\bar{k}_4 = 0$ and for (b) $pK_1 = 3.4$, $pK_{\text{II}} = 4.2$, $pK_{\text{III}} = 8.3$, $pK_{\text{IV}} = 10.0$,

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

Figure 5 Alignment of the sequences of chymopapain M (Chymo M), PPQ and actinidin to that of papain

The numbering of the residues in single-letter code relates to the chymopapain M sequence. Conserved residues are shown as capitals. Lower-case letters indicate non-conserved residues and, where there are insertions in the other sequences, they too are shown as lower-case letters. Deletions are indicated by —. The residues Glu-23 and His-127, which were considered to be candidates for one of the kinetically influential ionizations with pK_a 3.4, and Arg-65, all of which are unique to chymopapain M, are shown in bold type and doubly underlined. The papain sequence is that of Kamphuis et al. (1984) as modified by Cohen et al. (1986) on the basis of the cDNA sequence which identifies residues 47, 118 and 135 (numbering common to papain and chymopapain M) as Glu rather than Gln. The alignment of the PPQ sequence is as reported by Dubois et al. (1988a). The alignment of the actinidin sequence is as reported by Dubois et al. (1988a) but taking account of the sequence differences discussed by Varughese et al. (1992). They discussed 17 differences between the amino acid sequence of actinidin (Carne and Moore, 1978) and its cDNA sequence (Podivinsky et al., 1989) and made decisions about the assignments based on electron density in the crystal structure of an actinidin–E64 complex.

site Cys-S⁻/His-Im⁺H ion-pair state (pK_a 3.4 for both enzymes). The data for the papain catalysis (Figure 4b) establish the conclusion that ion-pair formation is an insufficient condition for catalytic competence in papain. This was originally suggested by a re-analysis (Mellor et al., 1993c) of the data for the papain-catalysed hydrolysis of nine other substrates reported by Polgár and Halasz (1978). The data in Figure 4(a) demonstrate that this conclusion holds also for chymopapain M, although the two profiles differ, particularly in two respects. First, full expression of k_{cat}/K_m for the papain catalysis requires only hydronic

$\bar{k}_1 = 1.0 \times 10^2$, $\bar{k}_2 = 1.95 \times 10^3$, $\bar{k}_3 = 1.5 \times 10^2$, $\bar{k}_4 = 0$; the broken lines correspond to contributions to k (i.e. k_{cat}/K_m) of the various hydronic states provided by the individual terms ($\bar{k}_1 - \bar{k}_4$) of eqn. (4). The forms of pH-dependence of k_{cat}/K_m for the reactions catalysed by chymopapain M [continuous line as in (a)] and by papain [broken line identical with the continuous line in (b)] are compared in (c).

Figure 6 Three-dimensional arrangement of the protein backbone and selected side chains of chymopapain M

The backbone is shown in green with the exception of the backbone atoms of residues Gly-66 and Asp-158, which are shown in magenta. These two residues, the carbonyl oxygens of which are shown in magenta, are in the S_1 – S_2 intersubsite region. The side chains of the catalytic site and S_1 -subsite residues (Gln-19, Glu-23, Cys-25, Arg-65, Asp-158, His-159, Asn-175 and Trp-181) are shown in red. The side chains of the residues that line the S_2 subsite (Tyr-61, Tyr-67, Gln-68, Ser-69, Val-133, Val-157, Ser-209 and Tyr-211) are shown in blue. The three disulphide bridges (22–63, 56–95 and 153–204) are shown in yellow. The side chains of Glu-135 and Lys-156, which may have an electrostatic influence on enzyme activity, are shown in white.

Figure 7 Structural similarity between chymopapain M and papain

The colour scheme used is as described in the legend to Figure 6. The major difference between the two structures can be observed on the left-hand side where chymopapain M has a four-residue insertion relative to papain. Differences in the S_1 and S_2 subsites are shown in more detail in Figure 8.

dissociation with pK_a 4.2, in addition to hydronic dissociation with pK_a 3.4 needed to generate ion-pairs. This is in accord with our analysis (Mellor et al., 1993c) of the data of Polgár and Halasz (1978), where hydronic dissociation with pK_a 5.0 in addition to the pK_a values of 3.4 and 4.0 is required for maximal activity only for cationic substrates (with Arg at P_1). By contrast, full expression of k_{cat}/K_m for the chymopapain M catalysis requires hydronic dissociation with pK_a values of 4.3 and 5.6. There is no evidence for a kinetically significant pK_a of 3.4 in the pH - k_{cat}/K_m profile. The other difference is that the value of the pK_a for the conversion of the ion-pair state to uncomplicated thiolate anion Cys-S⁻/His-Im is 8.7 for chymopapain M, as reflected also in reactivity probe kinetics, as against 8.3 in papain. These two differences result in a pH - k_{cat}/K_m profile for the chymopapain M catalysis shifted somewhat to higher pH values than the profile for the corresponding papain catalysis (see Figure 4c).

Towards a structural interpretation of the kinetic behaviour of chymopapain M

Use of the sequence alignment

To provide a structural basis for the present and future mechanism-targeted kinetic studies on chymopapain M, a model of its three-dimensional structure was built by comparative modelling techniques as described in the Materials and methods section. A necessary prerequisite was the sequence alignment for chymopapain M, papain, PPΩ and actinidin shown in Figure 5, which suggested the presence in chymopapain M of some structural features of kinetic significance. Most of these, such as the ubiquitous Cys-25/His-159 catalytic-site pair, are more usefully discussed below with reference to the model of the three-dimensional structure. The sequence alignment itself, however, was particularly useful in revealing the potential candidates for

the additional kinetically influential ionization with pK_a 3.4 unique to chymopapain M. The value of this pK_a suggests the side chain of an Asp or a Glu or of a substantially perturbed His. Inspection of the sequences of the four cysteine proteinases (Figure 5) leads to the conclusion that there are only two possible candidates for the additional ionization with pK_a 3.4, namely Glu-23 and His-127. His-127 is unique to chymopapain M (the analogous residues in the other three enzymes being Asn in papain, Lys in PPΩ and Tyr in actinidin) and is the only His residue in chymopapain M other than the conserved His (His-159; see the model of the three-dimensional structure in Figure 6) that is part of the catalytic site. Similarly Glu-23 is the only residue with a carboxy side chain (i.e. Glu or Asp) unique to chymopapain M, the analogous residue in the other enzymes being Gly in all cases.

Use of the model of the three-dimensional structure

The model of the three-dimensional structure of chymopapain M, built as part of the present work (Figure 6), and the superposition of this and the crystal structure of papain (Figures 7 and 8) demonstrate the presence in chymopapain M of many of the kinetically and mechanistically relevant structural features present in the papain molecule, and reveal a number of interesting differences. Thus common features (chymopapain M numbering with papain numbering in parentheses where it differs) include: (i) the three disulphide bonds Cys-22–Cys-63, Cys-56–Cys-95 and Cys-153–Cys-204 (Cys-153–Cys-200) involved in the maintenance of the folded structure; (ii) the Trp-shielded nucleophilic/acid-base and oxyanion hole systems of the catalytic site, Trp-181 (Trp-177), Cys-25, His-159, Asn-179 (Asn-175) and Gln-19; (iii) the potential electrostatic influences close to the catalytic site described by Brocklehurst et al. (1983), Pickersgill et al. (1989) and Kowlessur et al. (1989a), Glu-135, Lys-156 and Asp-158; (iv) the hydrogen-bond donor/acceptor systems considered to bind the non-scissile P_1 - P_2 amide bond of substrates within the active-centre cleft, the backbone NH of Gly-66 and the backbone C=O of Asp-158; (v) many of the residues in or near to the hydrophobic pocket of the S_2 subsite (the principal binding site of papain; see Figure 8), Tyr-61, Tyr-67, Val-133, Val-157, Ala-160 and Ser-209 (Ser-205). The presence of a Cys-His interactive system, the catalytic device used by all cysteine proteinases studied to date to cleave the scissile bond of substrates, was predicted by the characteristic shape of the pH - k profiles for the reactions with 2-Py-S-S-2-Py and Et-S-S-2-Py shown in Figures 3(a) and 3(b).

Differences in the S_2 -subsite region (Figure 8), which might have implications for recognition characteristics of substrates and inhibitors, are that Pro-68 and Trp-69 in papain are Gln-68 and Ser-69 in chymopapain M. Differences elsewhere that are particularly noteworthy relate to the additional kinetically influential side chain with pK_a 3.4 that is unique to chymopapain M. The model of the three-dimensional structure of the enzyme (Figure 6) shows that one of the potential candidates (His-127) is remote from the active-centre region [S(Cys-25)–Cα (His-127), 2.3 nm (23 Å); S(Cys-25)–ND1 (His-127), 2.5 nm (25 Å); S(Cys-

Figure 8 Three-dimensional arrangement of residues forming the catalytic site, the S_1 and S_2 subsites and the S_1 - S_2 intersubsite regions of chymopapain M and papain

—, Chymopapain M; ----, papain. The differences are in the replacement of Gly-23 and Gly-65 of papain by Glu and Arg respectively (in the S_1 subsite) and of Pro-68 and Trp-69 (papain) by Gln-68 and Ser-69 (chymopapain M) in the S_2 subsite. The identities of the other numbered residues are shown in the legend to Figure 6.

Figure 9 Arrangement of residues forming the catalytic site and the S_1 subsite of chymopapain M

The hydration state of the side chains is shown for the enzyme at pH 7.0 with the aspartate, glutamate and cysteine side chains negatively charged and the arginine and histidine side chains positively charged.

25)–NE2 (His-127) 2.6 nm (26 Å)] whereas the other (Glu-23) is close to the catalytic site where it can engage in hydrogen-bonding interactions with Arg-65 (see Figures 9 and 10) a cationic residue unique to chymopapain M which is Gly in the other three enzymes (see Figure 5). The possibility that a Glu-23–Arg-65 interaction might exclude most amino acid side chains in substrates from binding in the S_1 -subsite was suggested by Buttle et al. (1990a) and this is now clearly seen to be an appropriate conclusion from inspection of the modelled structure here reported.

The possibilities for hydrogen-bonding interactions between Arg-65 and Glu-23 are shown in Figures 9 and 10. The modelled disposition of the two ionic side chains corresponds to the 'side-on staggered' structure discussed by Mitchell et al. (1992) in connection with their survey and analysis of Arg–Asp interactions in proteins. As they pointed out, the interaction between the Arg side chain and a carboxylate group is one of the most important in protein structure and is characterized by its high directionality. This survey showed that in intramolecular interactions a 'side-on' geometry with the formation of hydrogen bonds between the NE and NH of the Arg side chain and the oxygen atoms of the Asp side chain is favoured. Figure 10 shows that a similar conformation could exist for the Glu-23–Arg-65 interactions in chymopapain M, namely side-on with a staggered arrangement of atoms. In this conformation side-chain interactions can exist between OE1 (23) and NH21 (65) (0.29 nm), between OE1 and HE (0.22 nm) and between OE2 and HE (0.28 nm). In addition, another potential hydrogen bond could exist between the backbone NH of Arg-65 and the backbone carbonyl O of Glu-23

Figure 10 Potential hydrogen-bonding interactions involving Arg-65 and Glu-23 of chymopapain M

Distances are shown in nm. Potential hydrogen bonds between the side chains are those between OE1 and NH21 (0.29 nm), between OE1 and HE (0.22 nm) and between OE2 and HE (0.28 nm). The other potential hydrogen bond (0.24 nm) is between the backbone NH of Arg-65 and the backbone carbonyl O of Glu-23.

(0.24 nm). An analogous hydrogen bond could exist in papain (0.22 nm), where both residues are Gly, and this hydrogen bond may well be a common feature of the papain subfamily of cysteine proteinases.

Concluding comments

The work presented in the present paper demonstrates (i) the importance of using reversibly inhibited cysteine proteinases during their isolation and manipulation during protein chemical procedures and (ii) a multiplicity of kinetically influential ionizations in both chymopapain M and papain, a phenomenon that may be general to the cysteine proteinase family. The combined use of 2-pyridyl disulphides and 4,4'-dipyrimidyl disulphide as thiol-specific reactivity probes differing in their susceptibility to activation by hydration permits the ionizations pertaining to formation and subsequent dehydration of the ion-pair state of the catalytic sites of these enzymes to be distinguished from the other kinetically influential ionizations. The surprising fact that the development of catalytic activity in both chymopapain M and papain is not determined solely or mainly by formation of the Cys-S[−]/His-Im⁺H ion-pair state, but rather requires additional dehydration, may also be a general facet of cysteine proteinase mechanism. The identification of at least some of the structural features of chymopapain M responsible for kinetically observed characteristics emphasizes the value of the combined use of comparative modelling techniques and mechanism-targeted kinetic analysis in structure–function studies on enzyme families.

We thank Dr. Richard Pickersgill for providing the co-ordinates of the structure of PP2 in advance of their release via the Brookhaven PDB, the Science and Engineering Research Council for Postdoctoral Research Assistantships for M. P. T., D. K. and G. W. M. and an Earmarked Studentship for M. P. T., Hasu Patel for valuable technical assistance, and Rita Dobrin and Joy Smith for the efficient production of the typescript.

REFERENCES

- Arad, D., Langridge, R. and Kollman, P. A. (1990) *J. Am. Chem. Soc.* **112**, 491–502
- Baines, B. S. and Brocklehurst, K. (1982) *J. Protein Chem.* **1**, 119–139
- Baines, B. S., Brocklehurst, K., McKee, R., O'Driscoll, M., Salih, E. and Smith, H. (1986) *Biochem. Soc. Trans.* **14**, 1226–1227
- Baker, E. N. (1980) *J. Mol. Biol.* **141**, 441–484
- Baker, E. N. and Drenth, J. (1987) in *Biological Macromolecules and Assemblies 3* (Jurnak, F. and McPherson, A., eds.), pp. 313–368, John Wiley and Sons, New York
- Bascombe, K. N. and Bell, R. P. (1959) *J. Chem. Soc.* 1096–1104
- Bender, M. L. and Kézdy, F. J. (1965) *J. Am. Chem. Soc.* **87**, 4954–4955
- Blundell, T., Carney, D., Gardiner, S., Hayes, F., Howlin, B., Hubbard, T., Overington, J., Singh, D. A., Sibanda, L. and Sutcliffe, M. (1988) *Eur. J. Biochem.* **172**, 513–520
- Brocklehurst, K. (1982) *Methods Enzymol.* **87**, 427–469
- Brocklehurst, K. (1987) in *Enzyme Mechanism* (Page, M. I. and Williams, A., eds.), pp. 140–158, Royal Society of Chemistry, London
- Brocklehurst, K. (1994) *Protein Eng.* **7**, 291–299
- Brocklehurst, K. and Brocklehurst, S. M. (1988) *Biochem. J.* **256**, 556–558
- Brocklehurst, K. and Salih, E. (1983) *Biochem. J.* **213**, 559–560
- Brocklehurst, K., Crook, E. M. and Wharton, C. W. (1968) *FEBS Lett.* **6**, 565–571
- Brocklehurst, K., Carlsson, J., Kierstan, M. P. J. and Crook, E. M. (1973) *Biochem. J.* **133**, 573–544
- Brocklehurst, K., Carlsson, J., Kierstan, M. P. J. and Crook, E. M. (1974) *Methods Enzymol.* **34**, 531–544
- Brocklehurst, K., Baines, B. S. and Kierstan, M. P. G. (1981) *Top. Enzyme Ferment. Biotechnol.* **5**, 262–335
- Brocklehurst, K., Mushiri, S. M., Patel, G. and Willenbrock, F. (1982) *Biochem. J.* **201**, 101–104
- Brocklehurst, K., Willenbrock, S. J. F. and Salih, E. (1983) *Biochem. J.* **211**, 701–708
- Brocklehurst, K., Baines, B. S., Salih, E. and Hatzoulis, C. (1984) *Biochem. J.* **221**, 553–554
- Brocklehurst, K., Carlsson, J. and Kierstan, M. P. J. (1985a) *Top. Enzyme Ferment. Biotechnol.* **10**, 146–188
- Brocklehurst, K., Salih, E., McKee, R. and Smith, H. (1985b) *Biochem. J.* **228**, 525–527
- Brocklehurst, K., McKee, R., Salih, E. and Smith, H. (1985c) *J. Protein Chem.* **4**, 103–127
- Brocklehurst, K., Willenbrock, F. and Salih, E. (1987a) *New Compr. Biochem.* **16**, 39–158
- Brocklehurst, K., Kowlessur, D., O'Driscoll, M., Patel, G., Quenby, S., Salih, E., Templeton, W., Thomas, E. W. and Willenbrock, F. (1987b) *Biochem. J.* **244**, 173–181
- Brocklehurst, K., Brocklehurst, S. M., Kowlessur, D., O'Driscoll, M., Patel, G., Salih, H., Templeton, W., Thomas, E., Topham, C. M. and Willenbrock, F. (1988a) *Biochem. J.* **256**, 543–555
- Brocklehurst, K., Kowlessur, D., Patel, G., Templeton, W., Quigley, K., Thomas, E. W., Wharton, C. W., Willenbrock, F. and Szawelski, R. J. (1988b) *Biochem. J.* **250**, 761–772
- Brocklehurst, K., O'Driscoll, M., Kowlessur, D., Phillips, I. R., Templeton, W., Thomas, E. W., Topham, C. M. and Wharton, C. W. (1989) *Biochem. J.* **257**, 309–310
- Brocklehurst, S. M., Topham, C. M. and Brocklehurst, K. (1990) *Biochem. Soc. Trans.* **18**, 598–599
- Buttle, D. J. and Barrett, A. J. (1984) *Biochem. J.* **223**, 81–88
- Buttle, D. J., Kembhavi, A. A., Sharp, S. L., Shute, R. E., Rich, D. H. and Barrett, A. J. (1989) *Biochem. J.* **261**, 469–476
- Buttle, D. J., Ritonja, A., Pearl, L. H., Turk, V. and Barrett, A. J. (1990a) *FEBS Lett.* **260**, 195–197
- Buttle, D. J., Ritonja, A., Dando, P. M., Abrahamson, M., Shaw, E. N., Wikstrom, P., Turk, V. and Barrett, A. J. (1990b) *FEBS Lett.* **262**, 58–60
- Carne, A. and Moore, C. H. (1978) *Biochem. J.* **173**, 73–83
- Claggett, J. A., Tokuda, S. and Englehard, W. E. (1974) *Proc. Soc. Exp. Biol. Med.* **145**, 1250–1257
- Cohen, L. W., Coghlan, V. M. and Dihel, L. C. (1986) *Gene* **48**, 221–227
- Dixon, H. B. F. (1973) *Biochem. J.* **131**, 149–154
- Dixon, W. J. (chief ed.), Brown, M. B., Engelman, L., Frane, J. W., Hill, M. A. and Jennrich, R. I. (1988) *BMDP Statistical Software Manual*, vol. 1, pp. 389–417, University of California Press, Berkeley, CA
- Dubois, T., Jacquet, A., Schnek, A. G. and Looze, Y. (1988a) *Biol. Chem. Hoppe-Seyler* **369**, 733–740
- Dubois, T., Kleinschmidt, T., Schnek, A. G., Looze, Y. and Braunitz, G. (1988b) *Biol. Chem. Hoppe-Seyler* **369**, 741–754
- Erlanger, B. F., Kokowsky, N. and Cohen, W. (1961) *Arch. Biochem. Biophys.* **95**, 271–278
- Fersht, A. R. (1977) *Enzyme Structure and Mechanism*. Freeman, New York
- Itoh, M., Haglwara, D. and Kamiya, T. (1975) *Tetrahedron Lett.* **49**, 4393–4394
- Jacquet, A., Kleinschmidt, T., Dubois, T., Schnek, A. G., Looze, Y. and Braunitz, G. (1989) *Biol. Chem. Hoppe-Seyler* **370**, 819–829
- Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A. and Drenth, J. (1984) *J. Mol. Biol.* **179**, 233–256
- Kamphuis, I. G., Drenth, J. and Baker, E. N. (1985) *J. Mol. Biol.* **182**, 317–329
- Khan, I. U. and Polgár, L. (1983) *Biochim. Biophys. Acta* **760**, 350–356
- Kowlessur, D., O'Driscoll, M., Topham, C. M., Templeton, W., Thomas, E. W. and Brocklehurst, K. (1989a) *Biochem. J.* **259**, 443–452
- Kowlessur, D., Topham, C. M., Thomas, E. W., O'Driscoll, M., Templeton, W. and Brocklehurst, K. (1989b) *Biochem. J.* **258**, 755–764
- Kowlessur, D., Thomas, E. W., Topham, C. M., Templeton, W. and Brocklehurst, K. (1990) *Biochem. J.* **266**, 653–660
- Kunimitsu, D. K. and Yasunobu, K. T. (1967) *Biochim. Biophys. Acta* **139**, 405–417
- Lynn, K. R. (1979) *Biochim. Biophys. Acta* **569**, 193–201
- Lynn, K. R. and Yaguchi, M. (1979) *Biochim. Biophys. Acta* **581**, 363–364
- Malthouse, J. P. G. and Brocklehurst, K. (1976) *Biochem. J.* **159**, 221–234
- Mellor, G. W., Sreedharan, S. K., Kowlessur, D., Thomas, E. W. and Brocklehurst, K. (1993a) *Biochem. J.* **290**, 75–83
- Mellor, G. W., Thomas, E. W., Topham, C. M. and Brocklehurst, K. (1993b) *Biochem. J.* **290**, 289–296
- Mellor, G. W., Patel, M., Thomas, E. W. and Brocklehurst, K. (1993c) *Biochem. J.* **294**, 201–210
- Mitchell, J. B., Thornton, J. M., Singh, J. and Price, S. L. (1992) *J. Mol. Biol.* **226**, 251–262
- Patel, M., Kayani, I. S., Mellor, G. W., Sreedharan, S., Templeton, W., Thomas, E. W., Thomas, M. and Brocklehurst, K. (1992a) *Biochem. J.* **281**, 553–559
- Patel, M., Kayani, I. S., Templeton, W., Mellor, G. W., Thomas, E. W. and Brocklehurst, K. (1992b) *Biochem. J.* **287**, 881–889
- Patel, M., Thomas, M. P., Kayani, I. S., Mellor, G. W., Thomas, E. W. and Brocklehurst, K. (1993) *Biochem. Soc. Trans.* **21**, 216S
- Pickersgill, R. W., Goodenough, P. W., Sumner, I. G. and Collins, M. E. (1988) *Biochem. J.* **254**, 235–238
- Pickersgill, R. W., Sumner, I. G., Collins, M. E. and Goodenough, P. W. (1989) *Biochem. J.* **257**, 310–312
- Pickersgill, R. W., Sumner, I. G. and Goodenough, P. W. (1990) *Eur. J. Biochem.* **190**, 443–444
- Pickersgill, R. W., Rizkallah, P., Harris, G. W. and Goodenough, P. W. (1991) *Acta Crystallogr.* **B47**, 766–771
- Podivinsky, E., Forster, R. L. S. and Gardner, R. C. (1989) *Nucleic Acids Res.* **17**, 8363
- Polgár, L. (1981) *Biochim. Biophys. Acta* **658**, 262–269
- Polgár, L. (1989) *Mechanism of Protease Action*. CRC Press, Boca Raton, FL
- Polgár, L. and Halasz, P. (1978) *Eur. J. Biochem.* **88**, 513–521
- Ritonja, A., Buttle, D. J., Rawlings, N. D., Turk, V. and Barrett, A. J. (1989) *FEBS Lett.* **258**, 109–112
- Ritonja, A., Križaj, I., Meško, P., Kopitar, M., Lučovnik, P., Štrukelj, B., Pungertar, J., Buttle, D. J., Barrett, A. J. and Turk, V. (1990) *FEBS Lett.* **267**, 13–15
- Ritonja, A., Čolić, A., Dolenc, I., Ogrinc, T., Podobnik, M. and Turk, V. (1991) *FEBS Lett.* **283**, 329–331
- Robinson, G. W. (1975) *Biochemistry* **14**, 3695–3700
- Salih, E., Malthouse, J. P. G., Kowlessur, D., Jarvis, M., O'Driscoll, M. and Brocklehurst, K. (1987) *Biochem. J.* **247**, 181–193
- Schack, P. (1967) *Compt. Rend. Trav. Lab. Carlsberg* **36**, 67–83
- Stuchbury, T., Shipton, M., Norris, R., Malthouse, J. P. G., Brocklehurst, K., Herbert, J. A. L. and Suschitsky, H. (1975) *Biochem. J.* **151**, 417–432
- Sumner, I. G., Vaughan, A., Eisenthal, R., Pickersgill, R. W., Owen, A. J. and Goodenough, P. W. (1993) *Biochim. Biophys. Acta* **1164**, 243–251
- Sutcliffe, M. J., Haneef, I., Carney, D. and Blundell, T. L. (1987a) *Protein Engng.* **1**, 377–384
- Sutcliffe, M. J., Hayes, F. R. F. and Blundell, T. L. (1987b) *Protein Engng.* **1**, 385–392
- Thomas, M. P., Mellor, G. W. and Brocklehurst, K. (1993) *Biochem. Soc. Trans.* **21**, 217S
- Topham, C. M., Overington, J., Thomas, M. P., Kowlessur, D., Thomas, E. W. and Brocklehurst, K. (1990a) *Biochem. Soc. Trans.* **18**, 934–935
- Topham, C. M., Thomas, P., Overington, J. P., Johnson, M. S., Eisenmenger, F. and Blundell, T. L. (1990b) *Biochem. Soc. Symp.* **57**, pp. 1–9
- Topham, C. M., Overington, J., Kowlessur, D., Thomas, M., Thomas, E. W. and Brocklehurst, K. (1990c) *Biochem. Soc. Trans.* **18**, 579–580
- Topham, C. M., Overington, J., O'Driscoll, M., Salih, E., Thomas, M., Thomas, E. W. and Brocklehurst, K. (1990d) *Biochem. Soc. Trans.* **18**, 933–934
- Topham, C. M., Salih, E., Frazao, C., Kowlessur, D., Overington, J. P., Thomas, M., Brocklehurst, S. M., Patel, M., Thomas, E. W. and Brocklehurst, K. (1991) *Biochem. J.* **280**, 79–92
- Topham, C. M., McLeod, A., Eisenmenger, F., Overington, J. P., Johnson, M. S. and Blundell, T. L. (1993) *J. Mol. Biol.* **229**, 194–220
- Varughese, K. I., Ying, S., Cromwell, D., Hasnain, S. and Xuong, N. (1992) *Biochemistry* **31**, 5172–5176