Crystal Structure of Glycyl Endopeptidase from *Carica papaya*: A Cysteine Endopeptidase of Unusual Substrate Specificity‡

Bernard P. O'Hara,§ Andrew M. Hemmings,§,|| David J. Buttle,⊥,# and Laurence H. Pearl*,§

Section of Structural Biochemistry, Department of Biochemistry and Molecular Biology, University College London, Gower Street, London WC1E 6BT, U.K., and Department of Biochemistry, Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 4RN, U.K.

Received May 22, 1995⊗

ABSTRACT: Glycyl endopeptidase is a cysteine endopeptidase of the papain family, characterized by specificity for cleavage C-terminal to glycyl residues only and by resistance to inhibition by members of the cystatin family of cysteine proteinase inhibitors. Glycyl endopeptidase has been crystallized from high salt with a substrate-like inhibitor covalently bound to the catalytic Cys 25. The structure has been solved by molecular replacement with the structure of papain and refined at 2.1 Å to an *R* factor of 0.196 (R_{free} = 0.258) with good geometry. The structure of the S_1 substrate binding site of glycyl endopeptidase differs from that of papain by the substitution of glycines at residues 23 and 65 in papain, with glutamic acid and arginine, respectively, in glycyl endopeptidase. The side chains of these residues form a barrier across the binding pocket, effectively excluding substrate residues with large side chains from the S_1 subsite. The constriction of this subsite in glycyl endopeptidase explains the unique specificity of this enzyme for cleavage after glycyl residues and is a major component of its resistance to inhibition by cystatins.

Many tropical plants, on damage, secrete a protective latex rich in proteolytic enzymes. One of the most studied of these is the latex of the papaya or paw-paw (*Carica papaya*) which contains a variety of cysteine proteinases of which papain (EC 3.4.22.2) is the best described. Papaya latex contains at least three other cysteine proteinases: chymopapain (EC 3.4.22.6), caricain (also known as papaya peptidase A, papaya proteinase III, and papaya proteinase Ω) (EC 3.4.22.30), and glycyl endopeptidase (also known as papaya proteinase IV, chymopapain M, and commercially as Proteinase Gly-C [Calbiochem-Novabiochem U.K. (EC 3.4.22.25)]. Glycyl endopeptidase has a M_r of 23 316 and is 67% identical to papain in amino acid sequence (Ritonja *et al.*, 1989).

The papain family of cysteine proteinases (Rawlings & Barrett, 1993) also contains enzymes of mammalian origin, known to be important in lysosome function and implicated in several pathological conditions (Barrett *et al.*, 1988). Of the plant enzymes, crystal structures have previously been obtained for papain (Kamphuis *et al.*, 1984), caricain (Pickersgill *et al.*, 1991), actinidin (EC 3.4.22.14) (Baker, 1980) from the kiwi fruit *Actinidia chinensis*, and calotropin (Heinemann *et al.*, 1982) from the madar plant *Calotropis gigantea*. Of the nonplant papain homologues, detailed structural information is available for human cathepsin B (Musil *et al.*, 1991) and cruzain (McGrath *et al.*, 1995) from *Trypanasoma cruzi*, the causal agent of Chaga's disease. These studies clearly demonstrate the conservation of the overall structure of this family of molecules.

Most members of the papain family of cysteine proteinases show primary specificity for hydrophobic residues in the S_2 and S_3 subsites with relatively little specificity toward residues in the S_1 subsite. In contrast, glycyl endopeptidase displays a strong preference for protein and small peptide substrates with a glycine residue in the S_1 subsite (Buttle *et al.*, 1990c; Buttle, 1994). Glycyl endopeptidase is also unusual in its resistance to inhibition by the small protein inhibitors, the cystatins, and unique among the papain family of cysteine proteinases in its ability to irreversibly inactivate family 2 cystatins (Barrett *et al.*, 1988) such as chicken cystatin and cystatin-C by cleavage of a glycyl bond near the N-terminus of these inhibitors (Buttle *et al.*, 1990b). Glycyl endopeptidase is also unusually resistant to inactivation by iodoacetate and iodoacetamide; k_2 for the inactivation of glycyl endopeptidase by iodoacetate is about 450 times slower than for the equivalent inactivation of papain, at near-neutral pH (Buttle *et al.*, 1990b).

Glycyl endopeptidase has now been crystallized with a specific peptide inhibitor, Z-Leu-Val-Gly-diazomethane, covalently bound to the active site cysteine. The crystal structure of this complex, refined at 2.1 Å, provides for the first time a clear structural basis for understanding the unusual properties of this enzyme.

MATERIALS AND METHODS

Protein Purification. Glycyl endopeptidase was purified from dried papaya latex (J. E. Siebel Sons Co. Inc., Chicago) using a single active site-directed affinity chromatography step following selective inactivation of other cysteine pro-

‡ The refined coordinates of glycyl endopeptidase have been deposited in the Brookhaven Protein Databank under the name 1GEC.

* Corresponding author.
§ University College London.
|| Present address: Department of Chemistry, University of East Anglia, Norwich NR4 7TJ, U.K.
⊥ Strangeways Research Laboratory.
Present address: Institute for Bone and Joint Medicine, Department of Human Metabolism and Clinical Biochemistry, University of Sheffield Medical School, Beech Hill Road, Sheffield S10 2RX, U.K.

⊗ Abstract published in *Advance ACS Abstracts,* September 15, 1995.

0006-2960/95/0434-13190$09.00/0

teinases with iodoacetate, as described elsewhere (Buttle, 1994). Briefly, incubation of latex with iodoacetate (0.5 mM) in buffer containing cysteine (4 mM, pH6.8) for 30 min at 4 °C resulted in complete loss of hydrolytic activity against Bz-Arg-pNA (Bachem, Bubendorf) while retaining most of the activity against Boc-Ala-Ala-Gly-pNA (Bachem, Bubendorf). This material was then applied to a Sepharose-Ahx-Gly-Phe-$NHCH_2CN$ affinity column (Buttle *et al.*, 1989), unbound material was washed off in pH 6.8 buffer, and the bound glycyl endopeptidase was eluted in citrate buffer (pH 4.5) containing 10 mM $HgCl_2$. The eluted protein was dialyzed extensively against Tris-HCl (20 mM, pH 7.2) containing EDTA (50 mM). The purified mercurial enzyme was activated with cysteine and incubated at 20 °C for 20 min with a molar equivalent of Z-Leu-Val-Gly-CHN_2 (Hall *et al.*, 1992), producing complete loss of enzyme activity. Contamination of the final material by other papaya cysteine proteinases was less than 0.5% in total, as determined by radial immunodiffusion assay (Buttle *et al.*, 1990a).

Crystallization. The purified protein was concentrated in a stirred cell concentrator over a PM 10 membrane (Diaflo, Amicon) to a final concentration of 25 mg/mL. A large range of conditions was examined for crystallization, which was finally achieved using NaCl as precipitant in acidic buffers. Stellate clusters of fine needles were initially obtained in hanging drop experiments with 4 μL droplets containing glycyl endopeptidase at 18 mg/mL, in sodium acetate buffer (100 μM, pH 5.6) and NaCl (3.1 M) equilibrated against 1 mL reservoirs containing 3.1 M NaCl and 100 mM sodium acetate buffer, pH 5.6. Although initial conditions were identified by vapor diffusion methods, only needle clusters could be obtained by this method despite considerable variation of the crystallization conditions. Using microbatch methods under paraffin oil (Chayen *et al.*, 1990), however, well formed single crystals of trigonal habit and dimension 0.2 mm × 0.2 mm × 0.4 mm could be grown from 2.0 μL volume experiments containing 21.7 mg/mL glycyl endopeptidase, 3.2 M NaCl, and 50 mM sodium acetate buffer, pH 5.6.

X-ray Data Collection and Processing. Microbatch droplets containing crystals were swollen to 10 μL with the addition of a harvesting solution (4.2 M NaCl, 50 mM sodium acetate, pH 5.6), and the suspended crystals were drawn into 1.0 mm diameter glass capillaries. Data were collected initially to 2.3 Å resolution on a Enraf Nonius FAST Area Detector using Cu Kα radiation (λ = 1.5418 Å) and processed using the MADNES software package (Pflugrath & Messerschmidt, 1992). Subsequently data to 2.1 Å were collected on an 18 cm MAR Research Image Plate detector and integrated using the MOSFLM package (Leslie, 1994). For both datasets, intensities were merged, scaled, and truncated using the ROTAVATA/AGROVATA and TRUNCATE programs of the CCP4 (CCP4, 1994) suite.

Glycyl endopeptidase crystallizes in a trigonal lattice with cell dimensions a = 55.78 Å, c = 64.35 Å. Specific volume calculations suggest three molecules in the unit cell, with a solvent content of approximately 50% by volume. The 00l reflections showed a clear pattern of systematic absence for $l \neq 3$ indicating a space group of $P3_1$ or $P3_2$. Data collection statistics are given in Table 1.

Structure Solution and Refinement. Molecular replacement using the refined structure of papain (Kamphuis *et. al.*, 1984) was performed using normalized structure factors with the programs ALMN and TFFC from the CCP4 suite. A rotation function using the refined structure of papain gave a clear solution at 9.96 times the root mean square (rms) function value. The rotated search model was used in a T2 translation function (Tickle, 1985) in both possible space groups. The function for $P3_1$ gave a clear solution at 9.88 × rms, while the translation function calculated in $P3_2$ was essentially featureless at 3.0 × rms. The reasonableness of the packing of the rotated and translated search model in the $P3_1$ cell was verified by visual inspection.

The transformed papain search model was changed to the sequence of glycyl endopeptidase and refined against data to 2.3 Å resolution (native-1 in Table 1). Refinement employed simulated annealing and conjugate gradient refinement in X-PLOR (Brunger, 1992a) using geometric restraints based on Engh and Huber (1991). Difference maps were calculated with coefficients $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ and σ_A-weighted according to Read (1991) to minimize model bias and used in repeated cycles of manual adjustment and simulated annealing refinement with the resolution extended to 2.1 Å (native-2 in Table 1). Examination of electron density maps and manual adjustments to the model were made using O (Jones *et al.*, 1991). The final model including the bound inhibitor, and 118 solvent molecules making at least one geometrically reasonable hydrogen bond to the protein, has an R factor of 0.196 for all reflections in the range 8−2.1 Å. A free R factor (Brunger, 1992b) calculated on 5% of the reflections in this range omitted from refinement was 0.258. The results of this refinement are summarized in Table 2.

The geometric parameters of the model have been analyzed with PROCHECK (Laskowski *et al.*, 1993) and are all inside or better than the expected deviations from ideality at 2.1 Å resolution. A single residue, Arg 65, has main-chain torsion angles within a "forbidden" region of the Ramachandran plot but is in unambiguous electron density.

Table 1: Data Collection Statistics for Glycyl Endopeptidase[a]

dataset	d_{min} (Å)	N unique	% complete	multiplicity	R_{merge}
native-1	2.3	8304	83.3	1.5	0.042
native-2	2.1	14026	97.4	2.7	0.050

[a] Data collection statistics: d_{min}, spacing of highest angle diffraction data; N unique, number of unique reflections after data reduction; % complete, fraction of possible reflections actually measured; multiplicity, total number of observations/number of symmetry unique reflections. R_{merge}, $\Sigma\Sigma_i|I(h) - I(h)_i|/\Sigma\Sigma_i I(h)$, where $I(h)$ is the mean intensity after rejection of outliers.

Table 2: Refinement Statistics[a]

no. of protein atoms (non-hydrogen)	1631
no. of ligand atoms	30
no. of water molecules	118
R factor	0.196 (all data)
R free	0.258
RMS deviation in bond lengths	0.016 Å
RMS deviation in bond angles	2.2°
RMS deviation in improper torsions	1.9°

[a] R factors are for 12 725 reflections (all data) in the resolution range 8−2.1 Å. The free R factor (Brunger, 1992b) was calculated from a random selection constituting approximately 5% of the data. RMS deviations are from ideal values derived from Engh and Huber (1991).

FIGURE 1: Stereopair of the Cα trace of glycyl endopeptidase (white) superimposed on the Cα trace of papain (PDB code P9PAP) (black). Figure produced with MOLSCRIPT (Kraulis, 1991).

```
                10         20         30         40         50
papain      IPEYVDWRQK GAVTPVKNQG SCGSCWAFSA VVTIEGIIKI RTGNLNEYSE
chymopapain YPQSIDWRAK GAVTPVKNQG ACGSCWAFST IATVEGINKI VTGNLLELSE
caricain    LPENVDWRKK GAVTPVRHQG SCGSCWAFSA VATVEGINKI RTGKLVELSE
glep        LPESVDWRAK GAVTPVKHQG YCGSCWAFST VATVEGINKI KTGNLVELSE
                                  1 *

                60         70         80         90        100
papain      QELLDCDRRS YGCNGGYPWS ALQLVAQYGI HYRNTYPYEG VQRYCRSREK
chymopapain QELVDCDKHS YGCKGGYQTT SLQYVANNGV HTSKVYPYQA KQYKCRATDK
caricain    QELVDCERRS HGCKGGYPPY ALEYVAKNGI HLRSKYPYKA KQGTCRAKQV
glep        QELVDCDLQS YGCNRGYQST SLQYVAQNGI HLRAKYPYIA KQQTCRANQV
                    3   123

               110        120        130        140        150
papain      GPYAAKTDGV RQVQPYNEGA LLYSIANQPV SVVLEAAGKD FQLYRGGIFV
chymopapain PGPKVKITGY KRVPSNCETS FLGALANQPL SVLVEAGGKP FQLYKSGVFD
caricain    GGPIVKTSGV GRVQPNNEGN LLNAIAKQPV SVVVESKGRP FQLYKGGIFE
glep        GGPKVKTNGV GRVQSNNEGS LLNAIAHQPV SVVVESAGRD FQNYKGGIFE

               160        170        180        190        200
papain      GPCGNKVDHA VAAVGYGP.. ..NYILIKNS WGTGWGENGY IRIKRGTGNS
chymopapain GPCGTKLDHA VTAVGYGTSD GKNYIIIKNS WGPNWGEKGY MRLKRQSGNS
caricain    GPCGTKVDHA VTAVGYGKSG GKGYILIKNS WGTAWGEKGY IRIKRAPGNS
glep        GSCGTKVDHA VTAVGYGKSG GKGYILIKNS WGPGWGENGY IRIRRASGNS
                42*2

               210
papain      YGVCGLYTSS FYPVKN..
chymopapain QGTCGVYKSS YYPFKGFA
caricain    PGVCGLYKSS YYPTKN..
glep        PGVCGVYRSS YYPIKN..
                4
```

FIGURE 2: Alignment of the amino acid sequences of the papaya cysteine proteinases, papain, chymopapain, caricain, and glycyl endopeptidase (glep). Asterisks under the sequences indicate the catalytic Cys and His residues, and numbers under the sequences indicate residues involved in the S1, S2, S3, and S4 substrate specificity subsites. The glutamate (E) at position 23 and the arginine (R) at position 65 which are unique to glycyl endopeptidase are in bold type.

RESULTS

Structure of Glycyl Endopeptidase. As expected from the 67% sequence identity between them, the overall structure of glycyl endopeptidase is extremely similar to papain, and the two proteins can be superimposed with a root mean square difference of only 0.64 Å for 212 common Cα positions (Figure 1). The only major backbone differences in the two structures are a relative insertion of four residues in the turn at residue 168 in glycyl endopeptidase and a small difference in the conformation of the loop around residues 191−195 (papain numbering). The longer loop at 168 is found in most other members of the papain family, and its absence can therefore be considered as a deletion mutation in papain (Figure 2).

Glycyl Endopeptidase−Inhibitor Interactions. The inhibitor Z-Leu-Val-Gly-CHN_2 is bound in the active-site cleft of the enzyme, with the methylene of the diazomethane group covalently attached to Sγ of the catalytic Cys 25 at the end of the long helix. Overall the peptide backbone of the inhibitor is in an extended conformation similar to that observed in complexes of papain with elongated peptide-like inhibitors (Figure 3), with the inhibitor side chains occupying the S_4, S_3, S_2, and S_1 subsites (defined by homology with papain) (Table 3 and Figure 4).

The phenyl ring of the amino-terminal benzyloxycarbonyl blocking group (Z) of the inhibitor is in van der Waals contact with the side chains of Val 157 and Ser 209 in the S_4 subsite. This is in contrast to the benzyloxycarbonyl blocking group in the papain-Z-Gly-Phe-Gly-CMK complex (Drenth *et al.*, 1976) which binds on the opposite side of the cleft. The carbonyl oxygen of this group makes no direct interaction with the enzyme but is hydrogen bonded to a well ordered solvent molecule bound by the side and main chain of Gln 68. The inhibitor leucine and valine residues occupy hydrophobic pockets on opposite faces of the cleft, in the S_3 and S_2 subsites, respectively. The side chain of the valine in the S_2 subsite makes van der Waals contacts with the side chains of Ala 160 and Val 133 and with the Cα of Asp 158, all of which residues are highly conserved in other papain family proteinases. The peptide nitrogen and carbonyl groups of the valine in the S_2 subsite hydrogen bond to the peptide carbonyl and nitrogen groups of the totally conserved Gly 66, in a short antiparallel β-sheet interaction. The side chain of the inhibitor leucine residue in the S_3 subsite makes van der Waals interactions with the side chains of Tyr 61 and Tyr 67, both highly conserved in papaya proteinases, and also with the side chain of Arg 65. The presence of Arg 65 makes the S_3 pocket in glycyl endopeptidase slightly more restricted than in other papain family proteinases where this residue is glycine. This may be responsible for the decreased preference for large aromatic substrate residues in the S_3 subsite observed for glycyl endopeptidase (Buttle, 1994). The guanidinium head group of Arg 65 makes a hydrogen bonding contact with the peptide carbonyl oxygen of the leucine in the S_3 subsite.

The S_1 subsite in papain and other members of the papain family is a wide and unrestricted pocket which exerts relatively little influence on the substrate specificity. The S_1 subsite is bounded by the catalytic Cys 25 on one side and the highly conserved glycines at 23 and 65 on the other side; however, in the amino acid sequence of glycyl endopeptidase (Ritonja *et al.*, 1989) glycines 23 and 65 are replaced by glutamic acid and arginine, respectively. In the crystal structure of glycyl endopeptidase, the S_1 subsite is extremely restricted by the side chains of these two residues, and the Cα of the inhibitor glycine residue is in van der Waals contact with the Cα and Cβ atoms of Glu 23 and the Cα, Cβ, and Cγ atoms of Arg 65. Modeling of any other amino acid type at this position in the inhibitor results in severe steric clashes with Glu 23 and Arg 65 which cannot be relieved without major changes in the backbone conformation of the protein. The presence of these large side chains

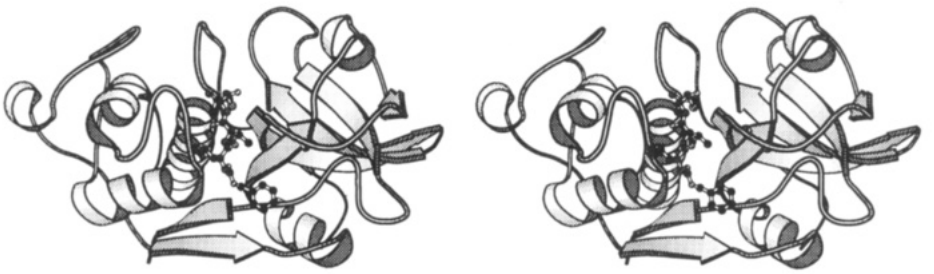

FIGURE 3: Stereopair of the glycyl endopeptidase inhibitor complex. α-Helices are shown as helical ribbons and β-strands as arrows. The covalent inhibitor is shown as a ball-and-stick model.

Table 3: Inhibitor Interactions

inhibitor group	type of interaction[a]	site	enzyme group
methylene	covalent		Cys 23 (side chain)
Gly Cα	vdw	S1	Glu 23 (side chain)
Gly Cα	vdw	S1	Arg 65 (side chain)
Gly N	H bond		solvent
Gly O	H bond		solvent
Val (side chain)	vdw	S2	Ala 160 (side chain)
Val (side chain)	vdw	S2	Val 133 (side chain)
Val (side chain)	vdw	S2	Asp 158 (main chain)
Val (main chain)	H bond	S2	Gly 66
Leu (side chain)	vdw	S3	Tyr 61 (side chain)
Leu (side chain)	vdw	S3	Tyr 67 (side chain)
Leu (side chain)	vdw	S3	Arg 65 (side chain)
Leu (main chain)	H bond	S3	Arg 65 (side chain)
Z (side chain)	vdw	S4	Val 157 (side chain)
Z (side chain)	vdw	S4	Ser 209 (side chain)
Z (main chain)	water bridge	S4	Gln 68 (main and side chain)

[a] vdw are van der Waals interactions; H bond are hydrogen bonding interactions.

in the S_1 subsite and the tightness of their interaction with the substrate satisfactorily account for the observed specificity of glycyl endopeptidase for glycine in the S_1 subsite.

Although the side chain carboxyl of Glu 23 does not interact directly with the methylene carbonyl in the inhibitor complex, it would be close enough to the scissile peptide of a productively bound substrate to have some influence on the electronics of the hydrolytic reaction, as suggested from molecular modeling and enzymological studies (Thomas *et al.*, 1994). The steric crowding and the anionic state of Glu 23 at the S_1 subsite could account for the significantly slower rate of inactivation of glycyl endopeptidase by iodoacetamide and iodoacetate than papain (Buttle *et al.*, 1990b).

Model building studies (Buttle *et al.*, 1990c; Thomas *et al.*, 1994) had suggested that the side chain functional groups of Glu 23 and Arg 65 would interact with each other, probably via charge assisted hydrogen bonding, commonly observed between these amino acid side chains in other proteins (Singh & Thornton, 1992). However, in the experimentally determined structure these side chains make no direct interaction with each other. This may be due to the presence of the inhibitor or may be a result of crystal packing, as the guanidinium head group of Arg 65 is in contact with the main chain of the inhibitor and with a symmetry related molecule in the crystal lattice.

In papain, the glycine at position 65 has a backbone conformation ($\phi = 98$, $\psi = 174$) normally only acceptable for glycine and is constrained in that conformation by hydrogen bond interactions in the secondary structure of the protein. In glycyl endopeptidase the backbone conformation at residue 65 ($\phi = 83$, $\psi = 165$) is extremely similar despite the presence of an arginine, for which this region of the Ramachandran plot is normally considered "forbidden". Nonetheless, the large side chain of Arg 65 is accommodated without any steric clash with adjacent residues, by the tight bend conformation of Asn 64 preceding it and the totally conserved Gly 66 in a fully extended conformation immediately following. The main chains of papain and glycyl endopeptidase in this region are virtually superimposable.

The Structural Basis of Cystatin Resistance. Cystatins are small disulfide bonded proteins which form a tight inhibitory complex with most members of the papain family. Glycyl endopeptidase is unusually resistant to inhibition by cystatins and irreversibly inactivates two family 2 cystatins, chicken cystatin and cystatin C, by cleavage of a Gly-Ala or Gly-Gly bond in the reactive site of the inhibitor (Buttle *et al.*, 1990b). Resistance to family 1 cystatins does not involve peptide bond hydrolysis, however, and has been assumed to be structural in origin. To further understand the structural basis of family 1 cystatin resistance, we have modeled a putative glycyl endopeptidase−stefin B complex by superimposing the structure of glycyl endopeptidase onto papain in the papain−stefin B complex structure (Stubbs *et al.*, 1990).

Three regions of the stefin B structure are involved in binding to papain: the N-terminal strand containing MetI6-MetI7-SerI8-GlyI9-AlaI10, and two hairpin loops, GlnI53-ValI54-ValI55-AlaI56-GlyI57 and ProI103-HisI104-GluI105-AsnI106 (Stubbs *et al.*, 1990). Modeled interactions between stefin B and glycyl endopeptidase are reasonably favourable in the S_4, S_3, and S_2 subsites, with MetI6, MetI7, and SerI8 in the N-terminal strand making few bad contacts. Inhibitor residues GlyI9 and AlaI10, however, clash extremely badly with the side chains of Arg 65 and Glu 23 (Figure 5), and this unfavorable interaction would be sufficient to prevent binding of the inhibitory amino-terminal strand in the substrate binding site of glycyl endopeptidase. Interactions with stefin B might be further disfavored by the steric clash of the hairpin loop (GlnI53 to AsnI59) with Glu 23 and Tyr 21 of glycyl endopeptidase (Figure 5).

DISCUSSION

The presence of the bulky side chains of glutamate at residue 23 and arginine at residue 65 confer a very tight substrate specificity on glycyl endopeptidase, making it a much less proteolytically destructive peptidase than the other cysteine proteinases occurring in papaya latex. On the other hand, Glu 23 and Arg 65 render glycyl endopeptidase resistant to cystatins which efficiently inhibit the other papaya cysteine proteinases and even allow it to inactivate family 2 cystatins by cleavage of a peptide bond in the reactive site.

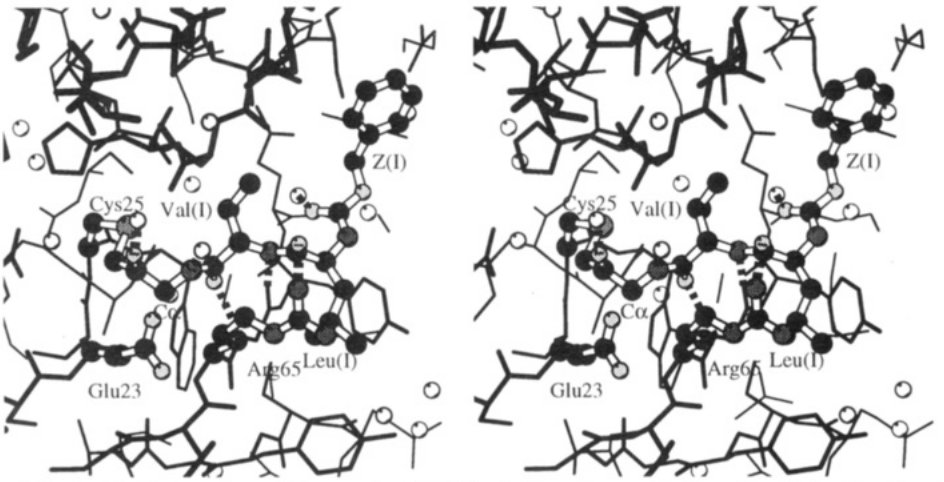

FIGURE 4: Stereopair of the detailed interactions of the covalent inhibitor in the substrate binding site of glycyl endopeptidase. The inhibitor and the enzyme residues Cys 25, Glu 23, and Arg 65 are highlighted as ball-and-stick models. Bound solvent molecules are shown as white spheres. Dashed lines indicate hydrogen-bonding interactions.

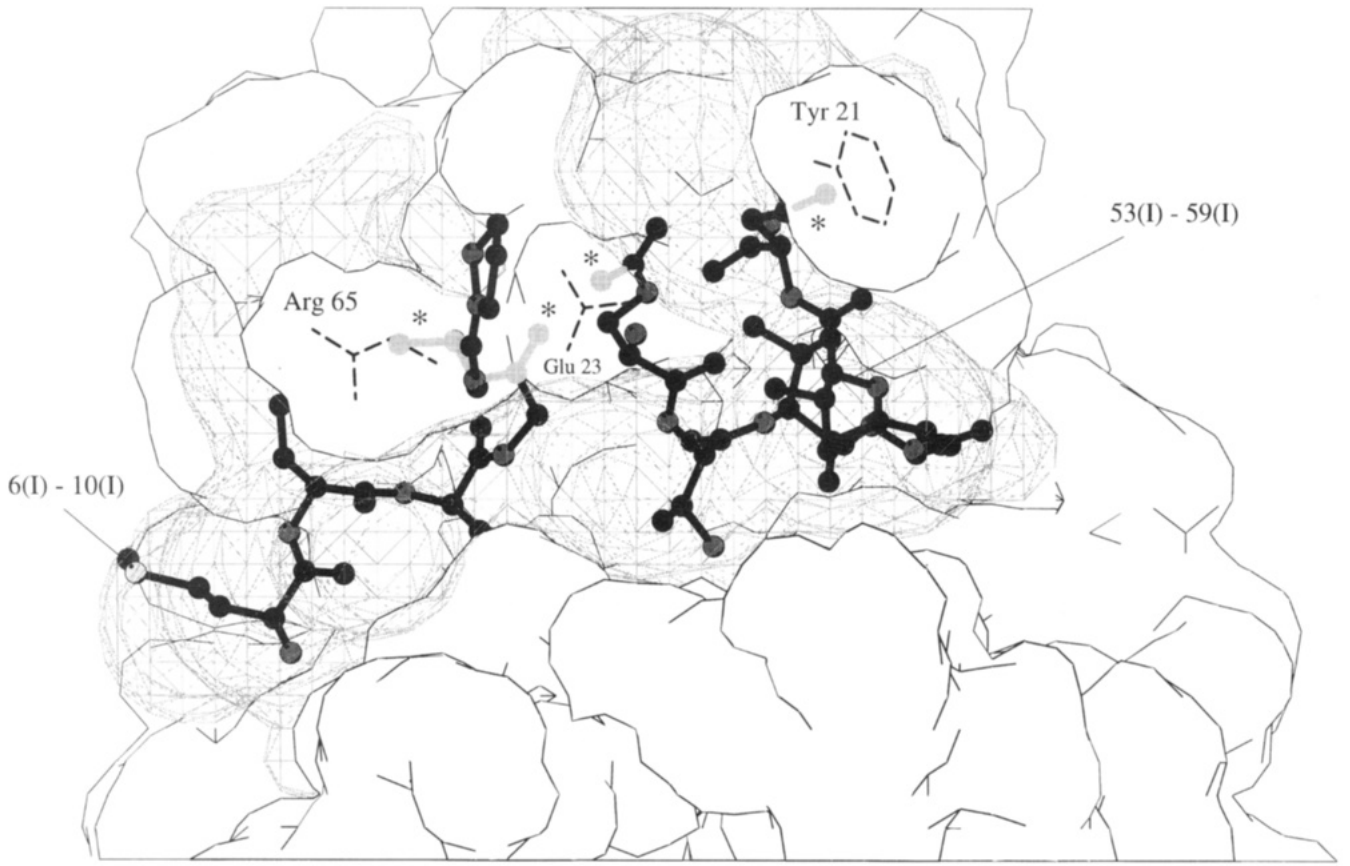

FIGURE 5: Model of the glycyl endopeptidase−stefin B interaction. The van der Waals surface of glycyl endopeptidase is depicted by the sketched white surface with the wire-cage contours outlining the volume of the protein's groove. The two loops of the inhibitor [6(I)−10(I) and 53(I)−59(I)] that have been modeled into the binding groove are shown in ball-and-stick representation. The key clashes of stefin B with glycyl endopeptidase are highlighted by asterisks and occur at Arg 65, Glu 23, and Tyr 21. The plot was produced using SURFNET (Laskowski, 1991).

Secretion of latex by tropical plants on injury probably serves to protect the site of the wound from invasion by insects and their larvae and from fungi and molds. Certainly the latex is rich in degradative enzymes of many sorts, including chitinases which would digest insect and arthropod exoskeletons. The proteinases may serve to cause more injury to an invading insect and thereby protect the fruit from further damage. Glycyl endopeptidase, with its unusual resistance to cysteine proteinase inhibitors and ability to inactivate some of them, is a sophisticated weapon in the tropical plant's defenses.

ACKNOWLEDGMENT

We are grateful to the Department of Crystallography at Birkbeck College for use of X-ray diffraction equipment. We thank Ms. Ruth Feltell for excellent technical assistance at the Strangeways Laboratory and Dr. Graham Elliot for helpful discussions. We are very grateful to Dr. Magnus Abrahamson for the gift of Z-Leu-Val-Gly-CHN_2 and to Dr. Roman Laskowski for assistance with figures. D.J.B. is an Arthritis and Rheumatism Council (U.K.) Research Fellow.

REFERENCES

Baker, E. N. (1980) *J. Mol. Biol. 141*, 441−484.

Barrett, A. J., Buttle, D. J., & Mason, R. W. (1988) *ISI Atlas Sci.: Biochem. 1*, 256−260.

Brunger, A. T. (1992a) *X-PLOR Version 3.1. A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Brunger, A. T. (1992b) *Nature 255*, 472−474.

Buttle, D. J. (1994) *Methods Enzymol. 244*, 539−555.

Buttle, D. J., Dando, P. M., Coe, P. F., Sharp, S. L., Shepherd, S. T., & Barrett, A. J. (1990a) *Biol. Chem. Hoppe-Seyler 371*, 1083−1088.

Buttle, D. J., Ritonja, A., Dando, P. M., Abrahamson, M., Shaw, E. N., Wikstrom, P., Turk, V., & Barrett, A. J. (1990b) *FEBS Lett. 262*, 58−60.

Buttle, D. J., Ritonja, A., Pearl, L. H., Turk, V., & Barrett, A. J. (1990c) *FEBS Lett. 260*, 195−197.

Chayen, N. E., Shaw Stewart, P. D., Maeder, D. L., & Blow, D. M. (1990) *J. Appl. Crystallogr. 23*, 297−302.

Collaborative Computational Project, No. 4. (CCP4) (1994) *Acta Crystallogr. A D50*, 760−763.

Drenth, J., Kalk, K. H., & Swen, H. M. (1976) *Biochemistry 15*, 3731−3738.

Engh, R. A., & Huber, R. (1991) *Acta Crystallogr. Sect. A, 47*, 392−400.

Hall, A., Abrahamson, M. Grubb, A., Trojnat, J., Kania, P., Kaspryzkowska, R., & Kasprzykowski, F. (1992) *J. Enzyme Inhib. 6*, 113−123.

Heinemann, U., Pal, G. P., Hilgenfeld, R., & Saenger, W. (1982) *J. Mol. Biol. 161*, 591−606.

Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A, 47*, 110−119.

Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A., & Drenth, J. (1984)- *J. Mol. Biol. 179*, 233−256.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Laskowski, R. A. (1991) SURFNET, Department of Biochemistry and Molecular Biology, University College London, U.K.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−290.

Leslie, A. G. W. (1994) MOSFLM Users Guide, MRC-LMB, Cambridge, U.K.

McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S. & Fletterick, R. J. (1995) *J. Mol. Biol. 247*, 251−259.

Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber. R., Popovic, T., Turk, V., Towatari, T., Katanuma, N., & Bode, W. (1991) *EMBO J. 10*, 2321−2330.

Pickersgill, R. W., Rizkallah, P., Harris, G. W., & Goodenough, P. W. (1991) *Acta Crystallogr. B47*, 766−771.

Plugrath, J. W., & Messerschmidt, A. (1992) Munich Area Detector New ECC System Max-Planck-Institut für Biochemie, 8033 Martinsried, West Germany.

Rawlings, N. D., & Barrett, A. J. (1993) *Biochem. J. 290*, 205−218.

Read, R. J. (1991) *Acta Crystallogr. Sect. A, 42*, 140−149.

Ritonja, A., Buttle, D. J., Rawlings, N. D., Turk, V., & Barrett, A. J. (1989) *FEBS Lett. 258*, 109−112.

Singh, J., & Thornton, J. M. (1992) *Atlas of Protein Side-Chain Interactions*, Vol. I and II, IRL Press, Oxford, U.K.

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., & Turk, V. (1990) *EMBO J. 9*, 1939−1947.

Thomas, M. P., Topham, C. M., Kowlessur, D., Mellor, G. W., Thomas, E. W., Whitford, D., & Brocklehurst, K. (1994) *Biochem. J. 300*, 805−820.

Tickle, I. J. (1985) in *Proceedings of the Daresbury Study Weekend* (Machin, P. A., Ed.) pp 22−26, SERC, Daresbury, U.K.

BI9511447