A Reinvestigation of Residues 64–68 and 175 in Papain

EVIDENCE THAT RESIDUES 64 AND 175 ARE ASPARAGINE

By S. S. HUSAIN AND G. LOWE

The Dyson Perrins Laboratory, South Parks Road, University of Oxford, Oxford OX1 3QY, U.K.

The tryptophan-containing peptides were isolated from the chymotryptic digest of *S*-carboxymethylated papain. Residue 175, which is strongly hydrogen-bonded to the active-site histidine residue in the tertiary structure of papain, is asparagine, confirming the work of Kimmel, Rogers & Smith (1965). Its function is probably to maintain the orientation and tautomeric state of the imidazole ring of histidine-159. The amino acid sequence predicted from the electron-density map of papain for residues 64–68 was confirmed, but residue 64 is asparagine, not aspartic acid. This residue, which is about 10 Å from the thiol group of the active-site cysteine-25, cannot therefore be a site of electrostatic attraction for substrates of basic amino acids.

The X-ray crystallographic structure of α-chymotrypsin revealed that residue 102 is hydrogen-bonded to the active-site histidine-57, which in turn is probably hydrogen-bonded to the active-site serine-195 (Matthews, Sigler, Henderson & Blow, 1967; Sigler, Blow, Matthews & Henderson, 1968). The discovery that residue 102 is aspartic acid rather than asparagine has led to the proposal of a 'charge-relay system' made up of these three residues in which transfer of electron density from the buried aspartic acid-102 is relayed through hydrogen bonds and the polarizable imidazole ring of histidine-57 to the oxygen of serine-195 at the surface of the molecule. In this way the oxygen of serine-195 becomes a more powerful nucleophile and a possible mechanism of action for α-chymotrypsin based on this 'charge-relay system' has been proposed (Blow, Birktoft & Hartley, 1969). The bacterial proteinase subtilisin also contains this 'charge-relay system' (Alden, Wright & Kraut, 1970) and it seems that the bacterial and mammalian serine proteinases have a similar catalytic mechanism.

The plant proteinase papain contains in its active site cysteine-25 and histidine-159, the functional groups of which are in close proximity and probably hydrogen-bonded (Husain & Lowe, 1968*a*,*b*). The X-ray crystallographic structure of papain has confirmed this (Drenth, Jansonius, Koekoek, Swen & Wolthers, 1968) and also revealed that the side-chain amide of asparagine-175 is hydrogen-bonded to the N-3 position of histidine-159 (Drenth, Jansonius, Koekoek, Sluyterman & Wolthers, 1970).

Again, therefore, a hydrogen-bonded system of three residues is found. Although the evidence for the presence of asparagine-175 in papain appears to be adequate (Kimmel, Rogers & Smith, 1965) and the need to increase the nucleophilicity of the thiol group unnecessary, it has been suggested that this residue may in fact be aspartic acid, giving rise to a 'charge-relay system' similar to that in α-chymotrypsin (Perutz, 1970). A reinvestigation of this residue was therefore desirable.

The main differences between the structure of papain as revealed by X-ray analysis (Drenth *et al.* 1968) and the tentative amino acid sequence (Light, Frater, Kimmel & Smith, 1964) have now been clarified (Drenth *et al.* 1970; Husain & Lowe, 1969; Lowe, 1970), but the sequence of residues 64–68 (160–164 in the tentative sequence of Light *et al.* 1964) deduced from the electron-density map and the amino acid composition (Drenth *et al.* 1968) differs from that proposed by Light *et al.* (1964). Since aspartic acid-64 is only 10 Å from the active-site thiol group of cysteine-25 and has tentatively been considered to be involved in the binding of substrates containing side chains (Drenth *et al.* 1968), a reinvestigation of this sequence was also desirable.

Papain contains only five tryptophan residues and since two of these, residues 69 and 177, were close to the residues required for reinvestigation, papain after carboxymethylation and denaturation was digested with α-chymotrypsin and the tryptophan-containing peptides were isolated.

EXPERIMENTAL

Reduction and carboxymethylation of papain. Twice-crystallized papain (0.27g), prepared from granular papaya latex (we are grateful to the Wallerstein Co., Mariners Harbor, Staten Island, N.Y., U.S.A. for a generous gift of this material) by the method of Kimmel & Smith (1954), was dissolved in 1.5M-tris–HCl buffer, pH8.6 (15ml) and irreversibly inhibited with iodoacetic acid (18.6mg). Guanidinium chloride (14.3g) and EDTA (47mg) followed by mercaptoethanol (0.4ml) were added. After 4h at 25°C, iodoacetic acid (1.07g) in M-NaOH (5ml) was added. After 20min at 25°C the solution was exhaustively dialysed against deionized water.

Digestion with α-chymotrypsin. The carboxymethylated protein suspended in water (40ml) was digested with α-chymotrypsin (8mg) in a Radiometer pH-stat. The pH was maintained at 8.0 by addition of 0.1M-NaOH. After 6.5h at 30°C the digestion was stopped by lowering the pH to 5.0 with M-HCl.

Chromatography of the digest on Sephadex G-25. The digest was concentrated to 2ml in a rotary evaporator, the insoluble residue centrifuged off and the clear solution applied to a column (140cm×1cm) of Sephadex G-25 (fine grade). The peptides were eluted with deionized water at a flow rate of 15ml/h and fractions (1ml) were collected. Peptides containing aromatic residues were located by measuring E_{280}. Tryptophan-containing peptides were detected with Ehrlich reagent. The chromatogram is shown in Fig. 1. Peaks 1–8 gave a strong Ehrlich-positive reaction; peaks 7 and 8 were pure peptides.

Paper chromatography. Peaks 1–6 were purified by chromatography on Whatman no. 3MM paper with butanol–acetic acid–water (40:6:15, by vol.). After development for 15–20h a guide strip was cut and the peptides were located with ninhydrin and Ehrlich reagents. Tryptophan-containing peptides were cut out and eluted with water.

Amino acid analysis. Peptides (0.1μmol) in constant-boiling HCl (1ml) in a sealed evacuated tube were kept at 110°C for 24h. The hydrolysate was evaporated to dryness in a rotary evaporator and the residue, in 0.1M-HCl (1ml), was applied to a Technicon AutoAnalyzer with the buffer system of Thomson & Miles (1964).

Carboxypeptidase A digestion. The peptides (0.1μmol) were dissolved in 0.1M-sodium phosphate buffer, pH7.8 (0.5ml) and a solution of carboxypeptidase A treated with di-isopropyl phosphorofluoridate [5μg; Sigma (London) Chemical Co., London S.W.6, U.K.], solubilized with 0.05ml of 0.5% Na$_2$CO$_3$) was added. The solutions were incubated at 30°C and the reaction was stopped by acidifying to pH2.5.

Leucine aminopeptidase digestion. The peptides (0.1μmol) were dissolved in 0.1M-tris–HCl buffer, pH8.5 (0.5ml, containing 10mM-MgCl$_2$) and a suspension of leucine aminopeptidase in 5mM-tris–HCl buffer, pH8.0, containing 5mM-MgCl$_2$ (20μl, Worthington Biochemical Corp., Freehold, N.J., U.S.A.) was added. The solutions were incubated at 35°C for 24h and the digestion was stopped by acidifying to pH2.5.

Edman degradation. Subtractive Edman degradation was performed as described by Konigsberg (1967). After
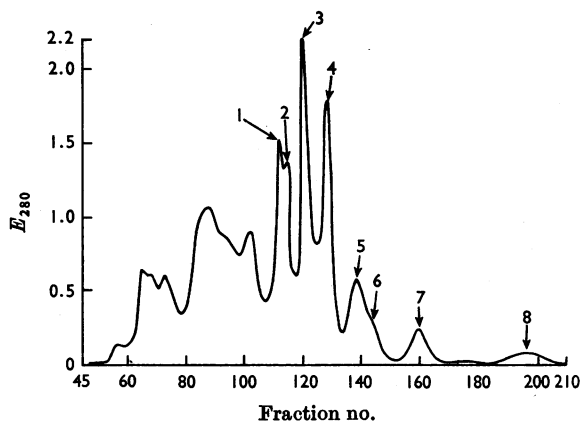

Fig. 1. Gel filtration on a column (140cm × 1cm) of Sephadex G-25 of the chymotryptic digest of carboxy-methylated papain. The tryptophan-containing peptides isolated from peaks 1–8 and their position in the amino acid sequence of papain are as follows.

Peak 1 } Peak 2 } Asn-Gln-Gly-Ser-CMCys-Gly-Ser-CMCys-Trp	Residues 18–26	Peptide 1
Peak 3 { Ile-Pro-Glu-Tyr-Val-Asp-Trp	Residues 1–7	Peptide 2
{ Gly-Thr-Gly-Trp-Gly-Glu-Asn-Gly-Tyr	Residues 178–186	Peptide 3
Peak 4 Gly-CMCys-Asn-Gly-Gly-Tyr-Pro-Trp	Residues 62–69	Peptide 4
Peak 5 Gly-Thr-Gly-Trp	Residues 178–181	Peptide 5
Peak 6 Asn-Ser-Trp	Residues 175–177	Peptide 6
Peak 7 Ser-Trp	Residues 176–177	Peptide 7
Peak 8 Asn-Ser-Trp-Gly-Thr-Gly-Trp	Residues 175–181	Peptide 8

Table 1. *Amino acid analysis of peptide 4 and residual peptides after successive Edman degradations*

	Peptide 4	After 1st cycle	After 2nd cycle	After 3rd cycle	After 4th cycle	After 5th cycle	After 6th cycle
CMCys*	0.8	0.9	>0.05	>0.05	>0.05	>0.05	>0.05
Asp	1.0	1.0	1.0	0.3	0.3	0.2	0.2
Pro	1.0	0.9	1.0	1.0	0.9	1.0	1.0
Gly	3.0	2.0	2.1	2.0	1.4	0.4	0.4
Tyr	1.0	0.9	1.0	0.9	0.9	0.9	0.6
Trp	+	+	+	+	+	+	+

each cycle, a sample was removed and prepared for amino acid analysis. 'Dansyl'–Edman degradation was performed as described by Gray (1967).

Electrophoresis. The electrophoretic mobility of a peptide relative to aspartic acid was measured in pyridine–acetic acid–water (25:1;225, by vol.), pH 6.5, and compared with the data of Offord (1966). (We are grateful to Dr Offord for assistance with this experiment.)

Peptide 1. Peaks 1 and 2 (Fig. 1) gave the same Ehrlich-positive peptide on paper chromatography, but it was not completely free of contaminating peptides. The amino acid analysis, however, indicated that the peptide arose from residues 18–26 in papain.

Peptides 2 and 3. Peak 3 (Fig. 1) gave two Ehrlich-positive peptides on paper chromatography. Peptide 2 had the following analysis: Asp (1.9), Glu (1.1), Pro (1.1), Val (1.0), Ile (1.0), Tyr (0.9), Trp (present). Peptide 3 had the analysis: Asp (1.0), Thr (1.0), Glu (1.0), Gly (4.0), Tyr (0.9), Trp (present).

Peptide 4. Peak 4 (Fig. 1) after paper chromatography gave a pure Ehrlich-positive peptide. Its amino acid analysis and that of the residual peptides after successive Edman degradations were as shown in Table 1. Digestion with carboxypeptidase A for 24 h gave 1.0 mol of tryptophan/mol of peptide; no other amino acid was released. Digestion with leucine aminopeptidase gave CMCys* (1.0), Asn (0.8) and Gly (3.1).

Peptide 5. Peak 5 (Fig. 1) after paper chromatography gave a pure Ehrlich-positive peptide with the analysis: Thr (0.9), Gly (2.0), Trp (present).

Peptide 6. Peak 6 (Fig. 1) after paper chromatography gave a pure Ehrlich-positive peptide, with the analysis: Asp (1.0), Ser (0.9), Trp (present). Digestion with carboxypeptidase A for 24 h released 1.0 mol of tryptophan/mol of peptide. Digestion with leucine aminopeptidase gave 1.0 mol each of Asn, Ser and Trp/mol of peptide.

Peptide 7. Peak 7 (Fig. 1) was a pure Ehrlich-positive peptide, with the analysis: Ser (1.0), Trp (present).

Peptide 8. Peak 8 (Fig. 1) was a pure Ehrlich-positive peptide, with the analysis: Asp (1.0), Thr (1.0), Ser (1.0), Gly (2.0), Trp (present). Digestion with carboxypeptidase A for 1 h released 1.0 mol of tryptophan/mol of peptide. Digestion with leucine aminopeptidase gave 1.0 mol each of Thr, Asn and Ser, Gly (1.9) and Trp (2.0)/mol of peptide.

RESULTS AND DISCUSSION

In order to isolate the tryptophan-containing peptides in the chymotryptic digest of carboxy-

* Abbreviation: CMCys, carboxymethylcysteine.

methylated papain, advantage was taken of the affinity of Sephadex G-25 for such peptides in de-ionized water. The separation of eight tryptophan-containing peptides was achieved, of which two (peaks 7 and 8, Fig. 1) were completely pure; five more were obtained pure by paper chromatography. The amino acid analysis allowed each peptide to be identified from the known sequence of the enzyme (Husain & Lowe, 1969) and accounted for all five tryptophan residues. Peptides 4, 6 and 8 (derived from peaks 4, 6 and 8) were selected for further study.

Peptides 6 and 8 were shown to contain an asparagine residue and no aspartic acid by digestion with leucine aminopeptidase. 'Dansyl'–Edman degradation revealed that this was the N-terminal residue in each peptide. Thus asparagine-175 was confirmed and the three residues hydrogen-bonded in the catalytic site of papain are therefore cysteine-25, histidine-159 and asparagine-175. The principal function of asparagine-175 seems to be to maintain the imidazole group of histidine-159 in a fixed orientation and tautomeric state. Since the hydrogen bond of the thiol group to the 'pyridine-type' nitrogen atom of the imidazole ring is expected to be relatively weak, without this restricting influence the imidazole group would have much more rotational freedom. Aligning the imidazole and thiol groups for hydrogen bonding will clearly assist catalysis by making the entropy of activation less negative. In this way the imidazole group can act as a general base towards the thiol group by using a pre-existing, albeit weak, hydrogen bond (cf. Wang & Parker, 1967). The catalytic mechanism would then be expected to follow the course already proposed (Lowe, 1970).

The amino acid analysis of peptide 4 and the residual peptides after six cycles of subtractive Edman degradation, together with the release of tryptophan by carboxypeptidase A, confirmed the sequence deduced from the electron-density map and the known amino acid composition for residues 62–69 (Drenth *et al.* 1968). The peptide, however, had an electrophoretic mobility of 0.24 relative to aspartic acid at pH 6.5, strongly suggesting that residue 64 was asparagine rather than aspartic acid

(Offord, 1966). Digestion with leucine amino-peptidase confirmed this, releasing carboxymethyl-cysteine, asparagine and glycine with no evidence of aspartic acid. Thus residue 64 in the amino acid sequence of papain (Husain & Lowe, 1969) must be asparagine. This residue, which is some 10 Å from the thiol group of cysteine-25, may still be involved in binding, but not because of any electrostatic attraction for substrates of basic amino acids.

REFERENCES

Alden, R. A., Wright, C. S. & Kraut, J. (1970). *Phil. Trans. R. Soc.* B, **257**, 119.

Blow, D. M., Birktoft, J. J. & Hartley, B. S. (1969). *Nature, Lond.*, **221**, 337.

Drenth, J., Jansonius, J. N., Koekoek, R., Sluyterman, L. A. Æ. & Wolthers, B. G. (1970). *Phil. Trans. R. Soc.* B, **257**, 231.

Drenth, J., Jansonius, J. N., Koekoek, R., Swen, H. M. & Wolthers, B. G. (1968). *Nature, Lond.*, **218**, 929.

Gray, W. R. (1967). *Methods in Enzymology*, vol. 11, p. 469. Ed. by Hirs, C. H. W. New York: Academic Press Inc.

Husain, S. S. & Lowe, G. (1968a). *Biochem. J.* **108**, 855.

Husain, S. S. & Lowe, G. (1968b). *Biochem. J.* **108**, 861.

Husain, S. S. & Lowe, G. (1969). *Biochem. J.* **114**, 279.

Kimmel, J. R., Rogers, H. J. & Smith, E. L. (1965). *J. biol. Chem.* **240**, 266.

Kimmel, J. R. & Smith, E. L. (1954). *J. biol. Chem.* **207**, 515.

Konigsberg, W. (1967). *Methods in Enzymology*, vol. 11, p. 461. Ed. by Hirs, C. H. W. New York: Academic Press Inc.

Light, A., Frater, R., Kimmel, J. R. & Smith, E. L. (1964). *Proc. natn. Acad. Sci. U.S.A.*, **52**, 1276.

Lowe, G. (1970). *Phil. Trans. R. Soc.* B, **257**, 265.

Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. (1967). *Nature, Lond.*, **214**, 652.

Offord, R. E. (1966). *Nature, Lond.*, **211**, 591.

Perutz, M. F. (1970). *Phil. Trans. R. Soc.* B, **256**, 205.

Sigler, P. B., Blow, D. M., Matthews, B. W. & Henderson, R. (1968). *J. molec. Biol.* **35**, 143.

Thomson, A. R. & Miles, B. J. (1964). *Nature, Lond.*, **203**, 483.

Wang, J. H. & Parker, L. (1967). *Proc. natn. Acad. Sci. U.S.A.* **58**, 2451.