Biochem. J. (1966) **101**, 232

Histidine Sequences in the Active Centres of some 'Serine' Proteinases

By L. B. SMILLIE and B. S. HARTLEY
Department of Biochemistry, University of Alberta, Edmonton, Canada, and Medical Research Council Laboratory of Molecular Biology, Cambridge

(*Received* 7 *March* 1966)

1. A comparison of the diagonal 'maps' of chymotrypsin A and 'tosylphenylalanyl chloromethyl ketone'-inhibited chymotrypsin A showed that His-57 is alkylated specifically by this substrate analogue. 2. From peptic digests of chymotrypsinogen A and B, trypsin and elastase it was demonstrated by the diagonal electrophoretic technique that a common di-histidine cystine-bridged structure is present in all four enzymes. 3. The sequences of these peptides were determined and show that the positions of the two histidine residues relative to the disulphide bond are a common feature. Thus His-40 of chymotrypsin A is only two residues removed from CyS-42, and His-57 is adjacent to the other half of this bridge, CyS-58. 4. Considerable variation in sequence occurs about His-40, where the aromatic residues 39 and 41 of the chymotrypsins and trypsin are replaced by alanine and threonine in elastase. There is a remarkable similarity in sequence following CyS-42 and preceding CyS-58 in all four enzymes.

Many types of evidence have implicated histidine in the catalytic activity of chymotrypsin (Hartley, 1960; Koshland, Strumeyer & Ray, 1962). Recent elucidation of the complete amino acid sequence of chymotrypsin A (Hartley, 1964*a*) and the determination of its disulphide bridges (Brown & Hartley, 1963, 1966; Keil, Prusík & Šorm, 1963) have shown that the two histidine residues, His-40 and His-57, are brought close to each other by the CyS-42 to CyS-58 disulphide bridge. Considerable similarities in sequence have been observed around the serine residue that forms part of the active centre in trypsin, chymotrypsin and elastase (Dixon, Kauffman & Neurath, 1958; Naughton, Sanger, Hartley & Shaw, 1960; Hartley, 1961). We have used the diagonal electrophoretic technique of Brown & Hartley (1966) to discover whether such homologies of sequence are also present around the histidine residues of these enzymes.

Of the group of enzymes that we have studied bovine chymotrypsin B has similar activity to chymotrypsin A (Enenkel & Smillie, 1963*b*) but shows several known sequence differences (Smillie & Hartley, 1965), bovine trypsin has markedly different specificity to bovine chymotrypsin, and porcine elastase differs from all these both in specificity and species origin. Brief reports of many of our experimental findings have already been made (Smillie & Hartley, 1964*a*,*b*, 1965; Hartley, Brown, Kauffman & Smillie, 1965).

MATERIALS

Enzymes. Bovine trypsin, chymotrypsin A and carboxypeptidase A and porcine pepsin were crystalline products from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Crystalline subtilisin B (Hunt & Ottesen, 1956) was from Novo Terapeutisk Laboratorium, Copenhagen, Denmark. Bovine chymotrypsinogen B was prepared by the method of Enenkel & Smillie (1963*a*).

Porcine elastase. This was the three-times-crystallized product obtained from Worthington Biochemical Corp. or was equivalent material prepared from the Trypsin 1-300 product of Nutritional Biochemicals Corp., Cleveland, Ohio, U.S.A., by the method of Lewis, Williams & Brink (1956). It was further purified by a batchwise treatment with DEAE-Sephadex A-50 to remove an inactive acidic fraction that represents some 20–50% of the 'crystalline elastase' (Lewis *et al.* 1956). For this purpose, 2g. of the crystalline material was suspended in 100ml. of 0·05M-tris–HCl buffer, pH 8·8, and dissolved by adjustment to pH 10·4 with N-NaOH. Approx. 700ml. of the tris–HCl buffer, pH 8·8, was added and the pH adjusted to 9·4 with N-HCl. To the elastase solution was added 250ml. of settled DEAE-Sephadex A-50 previously equilibrated and washed with 0·05M-tris–HCl buffer, pH 8·8. The suspension was stirred at 5° for 4hr. and filtered on a sintered-glass funnel under vacuum. The filtrate was freeze-dried, dissolved in 100ml. of mN-NH_3, dialysed at 2° for 24hr. against two 10l. changes of dilute NH_3 and again freeze-dried. Activity measurements against elastin dyed with Congo Red (Naughton & Sanger, 1961) showed a threefold increase in specific activity of the filtrate over the crystalline elastase. When chromatographed on CM-cellulose according to the method

of Naughton & Sanger (1961) 70% of the filtrate protein was recovered under a single peak with elastase activity.

*TPCK.** This was a gift from Miss B. A. Jeffry and had been prepared by the method of Schoellmann & Shaw (1963).

METHODS

Amino acid composition. Quantitative determination of the amino acid composition of peptides was made on the 10 cm. and 50 cm. columns of the Beckman automatic amino acid analyser equipped with the accelerated system (Spackman, Stein & Moore, 1958). About 0·01–0·10 μmole of peptide was hydrolysed in 0·1–0·2 ml. of constant-boiling HCl for 16–24 hr. in a sealed evacuated tube. For quantitative determination of the peaks on the chart, the usual triangular approximation method was used for peaks with an extinction greater than 0·1. For smaller peaks, the weighing technique of Yamasaki *et al.* (1963) was found most satisfactory. Accuracy and reproducibility with 0·01 μmole standards were better than 10%. In the recoveries reported, no correction was made for destructive losses during hydrolysis.

A few peptide hydrolysates were examined by two-dimensional electrophoresis and chromatography. Electrophoresis was carried out on Whatman no. 1 paper at pH 2·0 (formic acid–acetic acid–water; 1:4:45, by vol.) for 20 min. in solvent-cooled tanks as described by Ryle, Sanger, Smith & Kitai (1955). The strip containing the amino acids was then stitched at right angles to a sheet of Whatman no. 4 paper and descending chromatography was carried out overnight with 2-methylbutan-2-ol–pyridine–acetic acid–water (200:10:0·3:190, by vol.). Amino acids were located by dipping the paper in the cadmium–ninhydrin reagent of Heilmann, Barollier & Watzke (1957) and heating at 60° for 20 min. The amino acids could be determined quantitatively by cutting out the spots, eluting with 2 ml. of 95% (v/v) ethanol at 5° for 2 hr. with occasional shaking and reading the extinction at 510 mμ. Recoveries were $\pm$20%. This method was used only for the simplest peptides and the results were always confirmed from other evidence.

Pepsin digestions of proteins. These were always carried out in 5% (v/v) formic acid. Protein was weighed out into a glass-stoppered tube or flask and enough 5% formic acid and pepsin were added to give a 0·5% protein solution and a final pepsin protein weight ratio 1:10. After solution, the digest was incubated at 37° for 16 hr.

Enzymic digestion of peptides. Tryptic, chymotryptic and subtilisin digests of peptides were carried out in 0·2 M-*N*-ethylmorpholine–acetic acid buffer, pH 8·0, with a 1:50 or 1:100 molar ratio at 37° for 4–16 hr. depending on the extent of digestion desired. The final concentration of peptide was about 0·20–0·50 μmole/ml. Fresh enzyme solutions were prepared in water for each digestion.

For carboxypeptidase digestion, the peptide (0·02 μmole) in 0·025 ml. of water was added to 0·1 ml. of 0·5% NH_4HCO_3. Then carboxypeptidase (0·025 mg. in 0·05 ml.) was added and incubated for 16 hr. at 37°. Blanks (without peptide) were incubated simultaneously. The carboxypeptidase solution was prepared by adding 0·005 ml. of suspension (20 mg./ml.) to 0·1 ml. of water at 0°, clarifying by adding 0·01 ml. of N-NaOH and readjusting to neutrality with 0·01 ml. of N-HCl and 0·075 ml. of 0·5% NH_4HCO_3. The liberated amino acids were identified by electrophoresis at pH 6·5 and 1·8 as described by Ambler (1963).

N-Terminal groups of peptides. These were determined by their reaction with DNS chloride (Gray & Hartley, 1963*a*). This method, coupled with the Edman procedure (designated below as the 'dansyl'–Edman procedure) as described by Gray & Hartley (1963*b*), was extremely valuable for determining the sequences of the peptides isolated in this work. Approx. 0·2 μmole of peptide in water or volatile buffer was evaporated to dryness in a 3 ml. glass-stoppered test tube. The peptide was dissolved in 0·2 ml. of water and 0·1 ml. of pyridine. A 0·02 ml. sample was transferred to a 30 mm. × 6 mm. Durham fermentation tube (A. Gallenkamp and Co. Ltd., London, E.C.2) and the remainder frozen and stored at −10°. To the Durham tube was added 0·02 ml. of 0·1 M-$NaHCO_3$, the contents were thoroughly mixed and the solution was evaporated to dryness in a desiccator under vacuum. Deionized water (0·02 ml.) and 0·02 ml. of DNS chloride (2·5 mg./ml. in acetone) were added, the solutions mixed, and the tubes sealed with Parafilm and incubated at 37° for 3 hr. About half of the mixture was transferred to a separate Durham tube and evaporated to dryness under vacuum. Constant-boiling HCl (0·02 ml.) was added, and the tube was sealed without vacuum and incubated at 105° for 16–20 hr. The tube was opened, the contents were evaporated to dryness and, after solution in N-NH_3, about half was used for identification of the DNS-amino acid by electrophoresis at pH 4·40 (Gray & Hartley, 1963*a*).

The phenyl isocyanate degradation was essentially as described by Gray & Hartley (1963*b*) except that the reaction with the reagent was at 37° for 2–3 hr. Anhydrous trifluoroacetic acid was employed for 1 hr. at room temperature instead of acetic acid–anhydrous HCl. After removal of the trifluoroacetic acid under vacuum, the degraded peptide was again dissolved in water and pyridine and a further sample taken for reaction with DNS chloride. In none of the *N*-terminal degradations described in this work was it necessary to repurify the peptide at each step. Unequivocal identification of the *N*-terminal amino acid was obtained after as many as seven or eight degradations. When more than one DNS derivative was observed after electrophoresis at pH 4·40 or when the fluorescent spots became faint, the series was discontinued.

High-voltage paper electrophoresis. This was carried out in an apparatus similar to that described by Michl (1951) and by Ryle *et al.* (1955). The buffer systems and coolants at pH 6·5, 3·5 and 2·0 were as described by Ambler (1963) except that the toluene was 8% (v/v) with respect to pyridine. The papers used, time of runs and location of peptides were essentially the same as described by Ambler (1963) except that the cadmium–ninhydrin reagent of Heilmann *et al.* (1957) was used instead of 0·25% ninhydrin in acetone.

Diagonal electrophoretic procedure. For isolating the paired cysteic acid peptides containing histidine, the diagonal procedure of Brown & Hartley (1966) was employed. A peptic digest of the protein (40–50 mg.) was applied to a sheet of Whatman 3MM paper and submitted to

* Abbreviations: TPCK, chloromethyl 2-phenyl-L-1-toluene-*p*-sulphonamidoethyl ketone ('tosylphenylalanyl chloromethyl ketone'); DNS, 1-dimethylaminonaphthalene-5-sulphonyl; in amino acid sequences $CySO_3H$, Asn and Gln refer to cysteic acid, asparagine and glutamine residues respectively, and Asx and Glx to residues that could be either aspartic acid or asparagine and either glutamic acid or glutamine respectively.

high voltage ionophoresis at either pH 6·5 or 2·0. After being dried at 30–40°, a 3 cm. strip from this sheet was exposed on a glass rack in a desiccator to performic acid fumes generated from a mixture of 98–100% (v/v) formic acid and 30% (v/v) H_2O_2 (19:1, v/v) in the bottom of the desiccator. After 2 hr. at room temperature the strip was dried over NaOH pellets under vacuum for 1 hr. The strip was stitched to a full sheet of Whatman 3MM paper and submitted to electrophoresis at right angles to the original direction under the same conditions as before. After location of the histidine cysteic acid peptides by staining with cadmium–ninhydrin reagent and the Pauly reagent (Dent, 1947), the corresponding bands from the original electrophoretogram were cut out and oxidized, and the cysteic acid peptides were isolated and purified by electrophoresis.

RESULTS

TPCK-chymotrypsin A

TPCK-chymotrypsin A, prepared by the method of Schoellman & Shaw (1963), and native chymotrypsin A were subjected to pepsin digestion and the diagonal 'mapping' procedure at pH 6·5. Fig. 1 shows a comparison of the diagonal 'maps' of the native and inhibited enzyme. Peptides A1 and A2 arise from the CyS-42 to CyS-58 disulphide bridge of chymotrypsin A (Brown & Hartley, 1966) and contain the sequences around His-40 (peptide A1) and His-57 (peptide A2). Both peptides therefore give a characteristic red colour with the Pauly reagent (Dent, 1947). The only detectable difference between the diagonal 'maps' of peptic digests of chymotrypsin A and TPCK-chymotrypsin A lies in these cystine-bridged histidine peptides. The electrophoretic mobility of peptide A1 is unchanged and it still reacts with Pauly reagent. But a new more acidic peptide (A2′) replaces A2 and no longer reacts with Pauly reagent.

The diagonal 'map' can be used as a guide to the selective purification of cysteic acid peptides (Brown & Hartley, 1966). Thus the band containing the cystine peptide parent of A1 and A2′ was cut out of a preparative pH 6·5 electrophoretogram and run at pH 2·0 at 80 v/cm. for 45 min. The histidine peptide was detected by its reaction with Pauly reagent, cut out, oxidized with performic acid and run at pH 6·5. The two cysteic acid peptides A1 and A2′ were located by ninhydrin staining, eluted and subjected to qualitative amino acid analysis by two-dimensional electrophoresis and chromatography. Peptide A1 had the composition expected for residues 40–46. The composition of peptide A2′ was that expected for residues 55–64 but was devoid of histidine. Instead, a new amino acid derivative was present migrating in the approximate position of glutamic acid in both directions of the two-dimensional system (see the Methods section). Subsequent work by Stevenson & Smillie (1965) has shown that this new derivative is 3-carboxymethylhistidine. It may be concluded that TPCK reacts preferentially with N-3 of His-57 of chymotrypsin A.

Chymotrypsinogen B

The pH 6·5 diagonal 'map' of a peptic digest of chymotrypsinogen B is shown in Fig. 2. Although it differs in many respects from the corresponding 'map' of chymotrypsinogen A (Fig. 1), two Pauly-positive cysteic acid peptides, B1 and B2, are found in similar positions to the histidine peptides A1 and A2 from chymotrypsin A. Since no other peptides were vertically in line with these two, it was clear that they were disulphide-bridged in the native protein. The corresponding band on the original

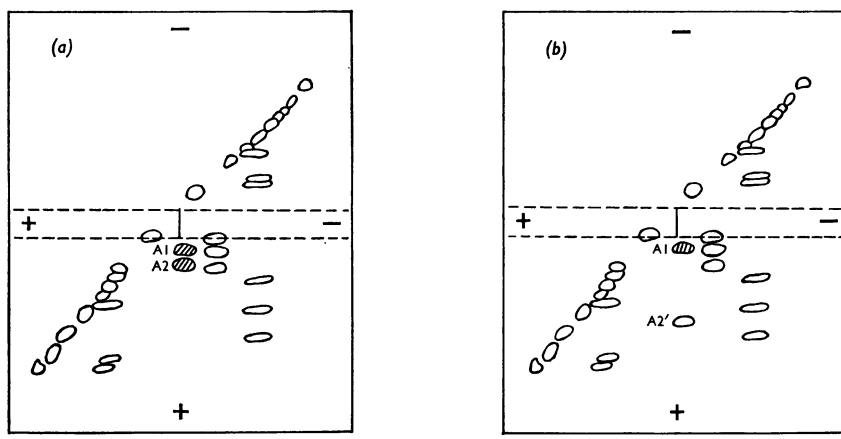

Fig. 1. Diagonal peptide 'map' of peptic digest of (*a*) chymotrypsin A and (*b*) TPCK-chymotrypsin A. Electrophoresis was at pH 6·5 in both dimensions. The origin is the vertical line in the middle of the sewn strip. Peptides staining for histidine are hatched.

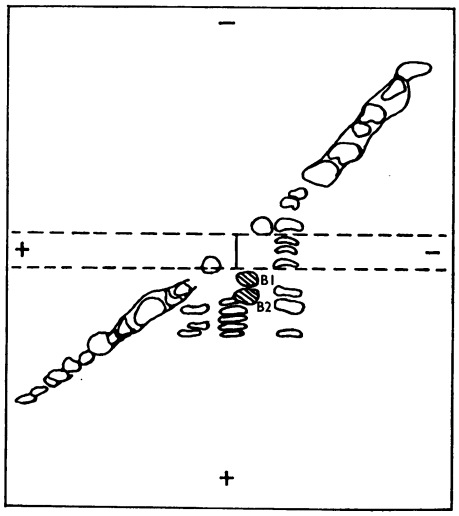

Fig. 2. Diagonal peptide 'map' of peptic digest of chymotrypsinogen B. Electrophoresis was at pH 6·5 in both dimensions. The origin is the vertical line in the middle of the sewn strip. Peptides staining for histidine are hatched.

electrophoretogram was cut out, run at pH 2·0, oxidized with performic acid and run again at pH 6·5. The two peptides B1 and B2 were eluted and subjected to amino acid analysis. Peptide B2 was found to be contaminated with peptide B1 and was further purified by electrophoresis at pH 3·5. The final amino acid analysis of both peptides is shown in Table 1.

Peptide B1. Carboxypeptidase digestion of peptide B1 gave a good yield of leucine and a small amount of serine. A chymotryptic digest (ratio 1:50; 16 hr. at 37°) when subjected to electrophoresis at pH 6·5 gave three bands, one basic and two acidic. The slower acidic band was Pauly-positive and had the composition of the original peptide B1. The most acidic peptide, B1B (Pauly-negative), and the basic peptide, B1A (Pauly-positive), had the compositions and *N*-terminal sequences indicated in Table 2. The sequence of B1 is therefore: His-Phe-CysO_3H-Gly-Gly-Ser-Leu.

Peptide B2. A subtilisin digest (ratio 1:100; 5 hr. at 37°) was subjected to electrophoresis at pH 6·5. The neutral band was cut out and subjected to electrophoresis at pH 2·0. One acidic (B2L) and 11 neutral peptides (B2A–B2K) were observed. Only peptides B2G, B2H, B2I, B2J, B2K and B2L were recovered in adequate yield for analysis of amino acid composition and *N*-terminal sequence. The results are shown in Table 2. The sequence of peptide B2 is therefore: Ala-Ala-His-CysO_3H-Gly-Val-Thr-Thr-Ser-Asp.

Table 1. *Amino acid composition and other properties of peptides isolated from diagonal 'maps' of peptic digests of chymotrypsinogen B, trypsin and elastase*

Protein	Peptide	Amino acid composition															Other details
		Ala	Arg	Asp	CysO_3H	Glu	Gly	His	Ile	Leu	Lys	Phe	Pro	Ser	Thr	Tyr	Val	
Chymotrypsinogen B	B1	—	—	—	1·04	—	2·08	0·90	—	1·04	—	0·95	—	0·98	—	—	—	*N*-Terminal His; *C*-terminal Leu
	B2	1·77	—	0·97	1·01	—	1·24	0·88	—	—	—	—	—	1·24	1·87	—	1·05	*N*-Terminal Ala
Trypsin	T1	1·73	—	—	1·05	1·04	0·86	1·14	1·04	—	1·0	—	—	1·10	—	1·05	—	*N*-Terminal Ala; *C*-terminal (Ile,Glu)
	T2	2·0	—	—	1·07	0·86	1·07	1·18	1·04	—	1·15	—	—	1·93	—	0·89	0·75	*N*-Terminal Val
	T3	—	—	—	0·80	—	2·17	0·60	—	1·03	—	0·89	—	1·08	—	—	—	*N*-Terminal His; *C*-terminal Leu
	T4	—	—	0·99	1·04	—	2·97	1·08	—	0·99	—	0·99	—	2·02	—	0·90	—	*N*-Terminal Asp; *C*-terminal Leu
	T5	—	—	—	0·92	—	2·20	0·27	—	1·03	—	0·82	—	1·03	—	—	—	Pauly-positive; ninhydrin-negative
Elastase	E1	1·75	1·03	1·14	1·06	1·11	—	1·02	—	—	—	—	—	—	—	—	0·87	*N*-Terminal Ala
	E2	1·76	1·01	1·38	0·87	1·03	—	1·05	—	—	—	—	—	—	1·06	—	0·87	*N*-Terminal Thr
	E3	1·03	—	—	1·03	—	2·03	1·24	—	0·97	—	—	—	—	1·73	—	—	Fluorescent; no *N*-terminus
	E4	0·95	—	—	1·11	—	1·94	1·05	—	1·17	—	—	—	—	1·79	—	—	Fluorescent; ninhydrin-negative; no *N*-terminus
	E5	—	—	—	1·05	—	—	0·62	—	0·95	—	—	—	—	—	—	—	*N*-Terminal His
	E6	—	0·97	1·08	0·85	—	—	—	—	—	—	—	—	—	—	—	1·11	*N*-Terminal CysO_3H; acidic at pH 6·5
	E7	1·17	—	—	1·10	—	1·70	0·97	—	1·34	—	—	—	—	1·73	—	—	*N*-Terminal Ala
	E8	0·94	—	2·05	1·12	—	—	—	—	—	—	—	0·80	1·13	—	0·81	—	*N*-Terminal Ala; neutral at pH 2·0

Trypsin

The pH 2·0 diagonal 'map' for trypsin is shown in Fig. 3. In addition to the histidine cysteic acid peptides T1, T2, T3, T4 and T5 lying off the diagonal, there were two Pauly-positive peptides on the diagonal that are presumably two varieties of a single histidine sequence not adjacent to a disulphide bond. This view is consistent with the presence of three histidine residues in trypsin and the sequences reported by Walsh, Kauffman, Sampath Kumar & Neurath (1964). The most basic of these two peptides was difficult to purify and was not investigated further. The one with lower mobility was purified by electrophoresis at pH 6·5 and 3·5 and had the amino acid composition corresponding to residues 79–82 of the sequence given by Walsh & Neurath (1964).

Table 2. *Amino acid sequence of histidine cysteic acid peptides from chymotrypsinogen B*

The symbols →→··· indicate *N*-terminal analyses by the 'dansyl'–Edman procedure; ···← represents *C*-terminal analysis with carboxypeptidase.

Peptide	Sequence
B1	His-Phe-$CySO_3H$-Gly-Gly-Ser-Leu
B1A	His-Phe
B1B	$CySO_3H$-Gly-Gly-Ser-Leu
B2	Ala-Ala-His-$CySO_3H$-Gly-Val-Thr-Thr-Ser-Asp
B2G	Thr
B2H	Gly-Val-Thr
B2I	Ala-Ala-His-$CySO_3H$
B2J	Ala-Ala-His-$CySO_3H$-Gly-Val
B2K	Ala-Ala-His-$CySO_3H$-Gly-Val-Thr-Thr
B2L	Ser-Asp

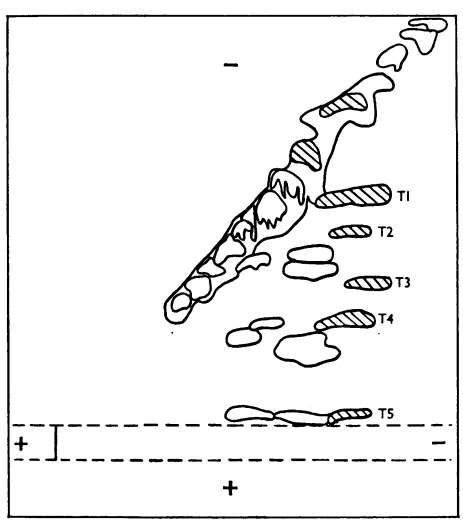

Fig. 3. Diagonal peptide 'map' of peptic digest of trypsin. Electrophoresis was at pH 2·0 in both dimensions. The origin is the vertical line on the sewn strip. Peptides staining for histidine are hatched.

The histidine cysteic acid peptides T1, T2, T3 and T4 were both Pauly- and ninhydrin-positive. Peptide T5 was Pauly-positive but ninhydrin-negative. After purification by electrophoresis at pH 6·5 and 3·5, samples were taken for amino acid analysis with the results shown in Table 1. Some *N*- and *C*-terminal analyses are also reported in Table 1. Peptides T1 and T2 were recovered in approximately equal yields and are clearly two varieties of the same sequence. Peptide T4 was recovered in considerably better yield than peptides T3 and T5 and is a larger variety of the latter two peptides. Peptide T5, which appears to have the same composition as peptide T3, but which did not stain with ninhydrin and was neutral at pH 2·0, also gave a low yield of histidine, which is apparently *N*-terminal. In fact, it has been a general observation in this work that peptides with *N*-terminal histidine give low yields of histidine on amino acid analysis (Table 1), perhaps as a consequence of modification during performic acid oxidation or high-voltage electrophoresis.

Peptide T1. Carboxypeptidase digestion of peptide T1 gave good yields of glutamine and isoleucine in approximately equal amounts with smaller yields of glycine and serine. The peptide (0·4 μmole in 0·5 ml.) in 0·2 M-*N*-ethylmorpholine–acetate buffer, pH 8·0, was digested with both trypsin and chymotrypsin simultaneously (ratio of each enzyme to peptide 1:100) at 37° for 3 hr. After electrophoresis at pH 2·0, five bands, T1A–T1E, were observed, which on analysis of amino acid composition and *N*-terminal sequence gave the results shown in Table 3. It was concluded that peptide T1 had the sequence: Ala-Ala-His-$CySO_3H$-Tyr-Lys-Ser-Gly-Ile-Gln.

Peptide T2. From its amino acid analysis (Table 1) and the identification of its *N*-terminal sequence as Val-Ser, it was concluded that peptide T2 had the sequence: Val-Ser-Ala-Ala-His-$CySO_3H$-Tyr-Lys-Ser-Gly-Ile-Gln.

Peptide T4. Carboxypeptidase digestion of this peptide gave a molar yield of leucine with no other amino acid. Chymotryptic digestion of the peptide (ratio 1:40; 16 hr. at 37°) yielded one basic peptide, T4A, one neutral peptide, T4B, and two acidic peptides, T4C and T4D, on electrophoresis at pH 6·5. These peptides were analysed for amino acids and *N*-terminal sequence with the results shown in Table 3. The sequence of peptide T4 was concluded

Table 3. *Amino acid sequence of histidine cysteic acid peptides from trypsin*

The symbols are defined in Table 2.

Peptide	Sequence
T1	Ala-Ala-His-$CySO_3H$-Tyr-Lys-Ser-Gly-Ile-Gln
T1A	Lys
T1B	Lys-Ser-Gly-Ile-Gln
T1C	Ala-Ala-His-$CySO_3H$-Tyr-Lys
T1D	Ser-Gly-Ile-Gln
T1E	Ala-Ala-His-$CySO_3H$-Tyr
T4	Asn-Ser-Gly-Tyr-His-Phe-$CySO_3H$-Gly-Gly-Ser-Leu
T4A	His-Phe
T4B	Asn-Ser-Gly-Tyr
T4C	His-Phe-$CySO_3H$-Gly-Gly-Ser-Leu
T4D	$CySO_3H$-Gly-Gly-Ser-Leu

to be: Asn-Ser-Gly-Tyr-His-Phe-$CySO_3H$-Gly-Gly-Ser-Leu.

Peptides T3 and T5. From their amino acid composition and *N*- and *C*-terminal residues it was clear that these were smaller varieties of peptides T4 and they were not examined further.

Elastase

A pH 2·0 diagonal 'map' of a pepsin digest of purified elastase is shown in Fig. 4. In addition to the four Pauly-positive peptides E1, E2, E3 and E4 indicated, there was a fifth Pauly-positive peptide lying off the diagonal and apparently paired to a Pauly-negative and neutral (at pH 2·0) peptide near the origin of the electrophoretogram in the second dimension. This band was not further investigated at this point (but see below). When examined under ultraviolet light before staining with ninhydrin or for histidine, peptides E3 and E4 showed strong fluorescence. Peptide E1 stained red with the cadmium–ninhydrin reagent, peptide E2 was yellow, peptide E3 a weak yellow and peptide E4 gave little or no reaction with cadmium–ninhydrin reagent. All four peptides were Pauly-positive.

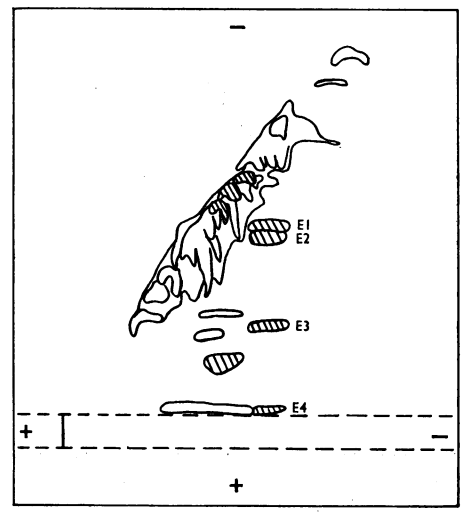

Fig. 4. Diagonal peptide 'map' of peptic digest of elastase. Electrophoresis was at pH 2·0 in both dimensions. The origin is the vertical line on the sewn strip. Peptides staining for histidine are hatched.

Peptides E1–E4 were eluted without further purification and analysed for amino acid composition and *N*-terminal residues (Table 1). Peptides E1 and E2 are evidently varieties of the same sequence with an extra threonine residue in the *N*-terminal position of peptide E2. Peptides E3 and E4 had the same composition and gave no detectable end group. Since they were fluorescent, it seemed possible that they contained oxidized *N*-terminal tryptophan, and the two forms, peptides E3 and E4, represented different states of oxidation of the same peptide. Further support for this view came from the presence of at least two abnormal compounds on pH 2·0 electrophoretograms of hydrochloric acid hydrolysates of these peptides. Such compounds have previously been observed on electrophoretograms of acid hydrolysates of other known oxidized tryptophan-containing peptides (B. S. Hartley, unpublished work).

Peptides E1 and E2. When subjected to *N*-terminal sequence analysis, these peptides gave the results shown in Table 4.

Peptide E3. Carboxypeptidase digestion yielded molar quantities of leucine and a trace of threonine.

Table 4. *Amino acid sequence of cysteic acid histidine peptides from peptic digest of elastase*

The symbols are defined in Table 2.

Peptide	Sequence	Other details
E1	Ala-Ala-His-$CySO_3H$-Val-Asx-Arg-Glx	
E2	Thr-Ala-Ala-His-$CySO_3H$-Val-Asx-Arg-Glx	
E3	Trp?-(Ala,His,Thr,$CySO_3H$,Gly,Gly)-Thr-Leu	Fluorescent
E3A	Gly-Thr-Leu	
E3B, E3C, E3D, E3E	Trp?-(Ala,His,Thr,$CySO_3H$,Gly,Gly,Thr,Leu)	Fluorescent
E3F	Trp?-(Ala,His,Thr,$CySO_3H$)	Fluorescent

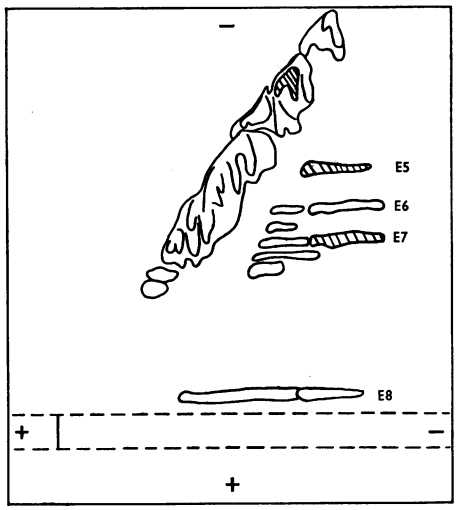

Fig. 5. Diagonal peptide 'map' of chymotryptic plus peptic digest of elastase. Electrophoresis was at pH 2·0 in both dimensions. The origin is the vertical line on the sewn strip. Peptides staining for histidine are hatched.

Digestion with subtilisin (ratio 1:100; 3hr. at 37°) and electrophoresis at pH 6·5 and 3·5 gave a series of peptides E3A–E3F. Peptide E3A was neutral at pH 6·5 whereas peptides E3B–E3F were acidic. The compositions and *N*-terminal analyses of each of these are shown in Table 4.

Because of the difficulties involved in determining the sequence of a peptide with a presumed oxidized *N*-terminal tryptophan, alternative approaches were explored for the isolation of this peptide. Attempts to purify the intact di-histidine cystine peptide before oxidation were unsuccessful because of extremely poor recoveries of the peptide during elution from the electrophoretograms. Finally, a combined chymotryptic plus peptic digestion was adopted in the expectation that a peptide corresponding to peptide E3 but without *N*-terminal tryptophan would be recovered.

Chymotryptic plus peptic digestion of elastase. A 200mg. portion of purified elastase was dissolved in 20ml. of 5% (v/v) formic acid and incubated at 37° for 5hr. to destroy elastase activity. After freeze-drying, 20ml. of 0·5% ammonium hydrogen carbonate, containing 4mg. of chymotrypsin, was added. After incubation for 16hr. at 37° and freeze-drying, the preparation was digested with pepsin (ratio 1:10) under the usual conditions. A pH 2·0 diagonal 'map' of this digest is illustrated in Fig. 5. Two peptides, E5 and E7, stained for histidine but were not pairs since they were not directly in line with one another. Instead, peptide E5 was paired with peptide E8 (ninhydrin-positive, Pauly-negative and neutral at pH 2·0), whereas peptide E7 was paired with peptide E6 (ninhydrin-positive and Pauly-negative). Peptides E5, E6 and E7 were isolated without further purification in the usual way. Peptide E8 was purified by electrophoresis at pH 6·5. The results of amino acid analysis and end-group determination are given in Table 1. Each of these peptides was analysed from the *N*-terminal end by the 'dansyl'–Edman procedure with the results shown in Table 5. Peptide E6 was shown to be acidic at pH 6·5 so that the aspartate residue was not present as the amide.

As indicated above, it was clear that peptide E5 was paired with peptide E8 in the intact protein and peptide E6 with peptide E7. This view is further substantiated when one considers the amino acid composition and partial sequence of peptides E3 and E4 and the full sequence of peptides E1 and E2. Clearly, peptides E6 and E1 and E2 are all derived from the same sequence in the intact protein. Further, peptide E7 is the same as peptides E3 and E4 except that in the last two one probably has an oxidized *N*-terminal tryptophan residue. Peptides E5 and E4, on the other hand, are peptides

Table 5. *Amino acid sequence of cysteic acid peptides from a chymotryptic and peptic digest of elastase*

The symbols are defined in Table 2.

Peptide	Sequence	Other details
E5	His-$CySO_3H$-Leu	
E6	$CySO_3H$-Val-Asp-Arg	Acidic at pH6·5
E7	Ala-His-Thr-$CySO_3H$-Gly-Gly-Thr-Leu	
E8	Ala-Asx-Asx-Ser-($CySO_3H$,Pro,Tyr)	

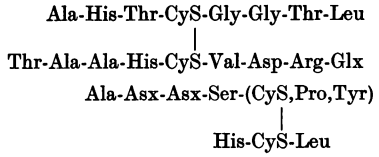

Fig. 6. Amino acid sequences around two of the disulphide bonds of elastase.

derived from another cystine-bridged structure in elastase in which only one histidine residue is closely adjacent to the disulphide bond. It is possible therefore to write the structure shown in Fig. 6 for the sequences around two of the disulphide bonds of elastase.

DISCUSSION

From the experimental evidence described, it can be concluded that there is present in each of the four enzymes, chymotrypsin A and B, trypsin and elastase, a common structure involving two histidine residues brought close to each other in the primary structure by a cystine bridge. The amino acid sequences of these structures are presented in Fig. 7. Identical results have been reported for trypsin by Walsh *et al.* (1964). It is significant that considerable variation relative to the sequence about His-40 of chymotrypsin A is permissible. Thus in elastase the aromatic residues at positions 39 and 41 of both trypsin and chymotrypsin are replaced by alanine and threonine. Thus the aromaticity of these residues is apparently not essential to the esterase activity of these enzymes although it may be of some importance in their differing specificities. However, it is significant that the positions of the two histidine residues relative to the disulphide bond are a constant feature. Further, the sequence following CyS-42 and preceding CyS-58 is remarkably similar for all four enzymes. In trypsin and chymotrypsin A, this near identity extends to the complete sequence between these two half-cystine residues in what may be called a 'histidine loop' (Fig. 8) (Hartley, 1964*a*; Walsh & Neurath, 1964). Whether the complete sequence enclosed by this loop in chymotrypsin B and elastase will show such close similarities awaits further work.

Our studies with the reagent TPCK confirm previous reports that it reacts specifically with His-57 of chymotrypsin (Ong, Shaw & Schoellmann, 1964, 1965; Smillie & Hartley, 1964*a*; Pospíšilová, Meloun & Šorm, 1964). Presumably TPCK is concentrated at the active site by virtue of interaction between the phenylalanyl moiety and a hydrophobic substrate-binding site. It is reasonable to conclude therefore that His-57 is within a few ångstroms of the substrate carbonyl group at the active site of chymotrypsin and therefore supplies the imidazole ring that has so frequently been implicated as a catalytically functional component of this enzyme. The discovery that a second histidine residue is brought close to His-57 by the 42–58 disulphide bridge, and that this structural feature occurs with remarkable sequence homology in three other pancreatic proteinases, has led to the design of hypothetical mechanisms in which two histidine residues participate in the catalysis. Hartley (1964*b*) has suggested that the two imidazole rings might form a planar tautomeric system whereby hydrogen-bonding of one ring to the serine hydroxyl group encourages attack of the substrate carbonyl group on the serine oxygen. Electrons would then be transferred via the two imidazole rings to a water molecule involved in the hydrolysis of the acyl-enzyme. Bender & Kézdy (1964) have suggested a rather similar system in which the two imidazole rings interact by virtue of stacking in a π complex. It is probable that only a knowledge of the precise spatial distribution of the relevant groups in the active centre will provide a final answer to these questions.

Other aspects of the remarkable homology between the pancreatic proteinases have been discussed by Walsh & Neurath (1964) and Hartley *et al.* (1965) and point both to divergent evolution

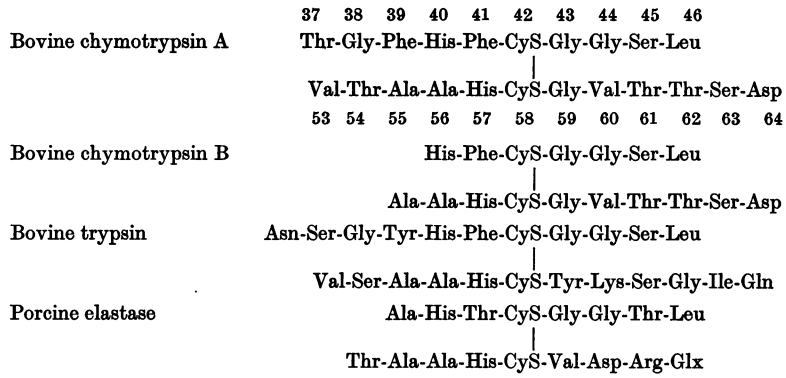

Fig. 7. Sequences of di-histidine cystine peptides of bovine chymotrypsin A and B, bovine trypsin and porcine elastase.

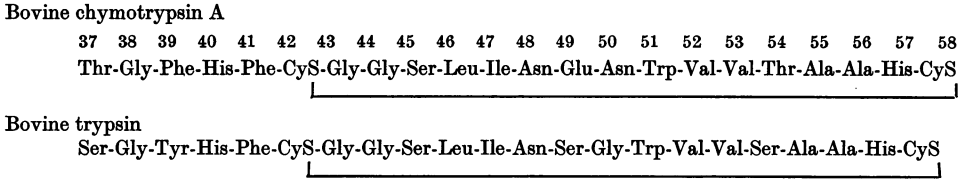

Fig. 8. Comparison of the 'histidine loops' of chymotrypsin A and trypsin.

from a common ancestral gene and to similar tertiary structures in their catalytic sites. Kinetic evidence also supports the concept of similar binding sites in chymotrypsin and trypsin. Inagami & Sturtevant (1960) demonstrated that trypsin catalyses the cleavage of certain aromatic substrates and that this activity is an intrinsic property of trypsin itself. Further, the recent findings of Inagami & Murachi (1963), of Inagami (1964) and of Mares-Guia & Shaw (1965) indicate that the specificity site of trypsin is composed of both an anionic site and a hydrophobic slit or crevice to which the carbon side chain of substrates or inhibitors is bound. Bender, Killheffer & Kézdy (1964*a*) have also shown that the rates of deacylation of a number of non-ionic acyl-chymotrypsin A and acyl-trypsin compounds are essentially identical over a 10^5-fold range of rate constants. This kinetic identity must reflect a similar interaction of the various acyl groups with very similar specificity sites of both enzymes. Bender, Killheffer & Kézdy (1964*b*) have been impressed with the structural homology within the histidine loops of these enzymes and suggest that the hydrophobic residues between CyS-42 and CyS-58 might contribute a major binding site whereas the sequence from CyS-191 to CyS-201 could serve as a subsidiary binding site. However, trypsin is well known to have a high affinity for lysyl and arginyl substrates and is therefore likely to possess a carboxyl group in its binding site. This feature should be absent from chymotrypsin. As candidates for the hydrophobic binding sites of trypsin and chymotrypsin we therefore also commend the areas of residues 95–114, 162–170 and 181–194, where homologies of sequence are accompanied by an acidic residue in trypsin that is absent from chymotrypsin.

The authors are indebted to Dr J. Brown and other members of the Laboratory of Molecular Biology for many helpful discussions. This work was supported in part by the National Research Council of Canada, the Life Insurance Medical Research Fund and the National Institutes of Health, U.S. Public Health Service (AM-06287).

REFERENCES

Ambler, R. P. (1963). *Biochem. J.* **89**, 349.
Bender, M. L. & Kézdy, F. J. (1964). *J. Amer. chem. Soc.* **86**, 3704.
Bender, M. L., Killheffer, J. V., jun. & Kézdy, F. J. (1964*a*). *J. Amer. chem. Soc.* **86**, 5330.
Bender, M. L., Killheffer, J. V., jun. & Kézdy, F. J. (1964*b*). *J. Amer. chem. Soc.* **86**, 5331.
Brown, J. R. & Hartley, B. S. (1963). *Biochem. J.* **89**, 59P.
Brown, J. R. & Hartley, B. S. (1966). *Biochem. J.* **101**, 214.
Dent, C. E. (1947). *Biochem. J.* **41**, 240.
Dixon, G. H., Kauffman, D. L. & Neurath, H. (1958). *J. biol. Chem.* **233**, 1373.

Enenkel, A. G. & Smillie, L. B. (1963*a*). *Biochemistry*, **2**, 1445.
Enenkel, A. G. & Smillie, L. B. (1963*b*). *Biochemistry*, **2**, 1449.
Gray, W. R. & Hartley, B. S. (1963*a*). *Biochem. J.* **89**, 59P.
Gray, W. R. & Hartley, B. S. (1963*b*). *Biochem. J.* **89**, 379.
Hartley, B. S. (1960). *Annu. Rev. Biochem.* **29**, 45.
Hartley, B. S. (1961). *Proc. 5th int. Congr. Biochem., Moscow*, vol. 4, p. 104. Oxford: Pergamon Press Ltd.
Hartley, B. S. (1964*a*). *Nature, Lond.*, **201**, 1284.
Hartley, B. S. (1964*b*). *Abstr. 1st Meet. Fed. Europ. biochem. Soc., Lond.*, p. 47. London: Academic Press (Inc.) Ltd.
Hartley, B. S., Brown, J. R., Kauffman, D. L. & Smillie, L. B. (1965). *Nature, Lond.*, **207**, 1157.
Heilmann, J., Barollier, J. & Watzke, E. (1957). *Hoppe-Seyl. Z.* **309**, 219.
Hunt, J. A. & Ottesen, M. (1956). *Biochim. biophys. Acta*, **48**, 411.
Inagami, T. & Murachi, T. (1963). *J. biol. Chem.* **238**, PC1905.
Inagami, T. (1964). *J. biol. Chem.* **239**, 787.
Inagami, T. & Sturtevant, J. M. (1960). *J. biol. Chem.* **235**, 1019.
Keil, B., Prusík, Z. & Šorm, F. (1963). *Biochim. biophys. Acta*, **78**, 559.
Koshland, D. E., Strumeyer, D. H. & Ray, W. J., jun. (1962). *Brookhaven Symp. Biol.* **15**, 101.
Lewis, U. J., Williams, D. E. & Brink, N. G. (1956). *J. biol. Chem.* **222**, 705.
Mares-Guia, M. & Shaw, E. (1965). *J. biol. Chem.* **240**, 1579.
Michl, H. (1951). *Mh. Chem.* **82**, 489.
Naughton, M. A. & Sanger, F. (1961). *Biochem. J.* **78**, 156.
Naughton, M. A., Sanger, F., Hartley, B. S. & Shaw, D. C. (1960). *Biochem. J.* **77**, 149.
Ong, E. B., Shaw, V. & Schoellmann, G. (1964). *J. Amer. chem. Soc.* **86**, 1271.
Ong, E. B., Shaw, E. & Schoellmann, G. (1965). *J. biol. Chem.* **240**, 694.
Pospíšilová, D., Meloun, B. & Šorm, F. (1964). *Abstr. 1st Meet. Fed. Europ. biochem. Soc. Lond.* p. 27, London: Academic Press (Inc.) Ltd.
Ryle, A. P., Sanger, F., Smith, L. F. & Kitai, R. (1955). *Biochem. J.* **60**, 541.
Schoellmann, G. & Shaw, E. (1963). *Biochemistry*, **2**, 252.
Smillie, L. B. & Hartley, B. S. (1964*a*). *Abstr. 1st Meet. Fed. Europ. biochem. Soc., Lond.*, p. 26. London: Academic Press (Inc.) Ltd.
Smillie, L. B. & Hartley, B. S. (1964*b*). *J. molec. Biol.* **10**, 183.
Smillie, L. B. & Hartley, B. S. (1965). *J. molec. Biol.* **12**, 933.
Spackman, D. H., Stein, W. H. & Moore, S. (1958). *Analyt. Chem.* **30**, 1190.
Stevenson, K. J. & Smillie, L. B. (1965). *J. molec. Biol.* **12**, 937.
Walsh, K. A., Kauffman, D. L., Sampath Kumar, K. S. V. & Neurath, H. (1964). *Proc. nat. Acad. Sci., Wash.*, **51**, 301.
Walsh, K. A. & Neurath, H. (1964). *Proc. nat. Acad. Sci., Wash.*, **52**, 884.
Yamasaki, M., Brown, J. R., Cox, D. J., Greenshields, R. N., Wade, R. D. & Neurath, H. (1963). *Biochemistry*, **2**, 859.