Biochem. J. (1973) **131**, 643-675
Printed in Great Britain

Evidence for the Amino Acid Sequence of Porcine Pancreatic Elastase

By DAVID M. SHOTTON* and BRIAN S. HARTLEY
Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

(*Received* 7 *August* 1972)

The preparation and purification of tryptic peptides from aminoethylated Dip-elastase and [^{14}C]carboxymethylated Dip-elastase, and of peptic peptides from native elastase is described. A summary of the results of chemical studies used to elucidate the amino acid sequence of these peptides is presented. Full details are given in a supplementary paper that has been deposited as Supplementary Publication SUP 50016 at the National Lending Library for Science and Technology, Boston Spa, Yorks. LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in *Biochem. J.* (1973), **131**, 1–20. These results, together with those from previously published papers, are used to establish the complete amino acid sequence of elastase, which is a single polypeptide chain of 240 residues, molecular weight 25900, containing four disulphide bridges.

Hartley *et al.* (1959) and Naughton *et al.* (1960) first investigated the amino acid sequence of porcine pancreatic elastase (EC 3.4.4.7) and proved that the enzyme was homologous with bovine trypsin and chymotrypsin-A, possessing the -Asp-Ser-Gly- active-centre sequence characteristic of this class of serine proteinases (Hartley, 1960). Smillie & Hartley (1966) and Brown *et al.* (1967) extended this work by determining the amino acid sequences around the active-centre serine and histidine residues, serine-195 and histidine-57, the four disulphide bridges and the amino terminus (Fig. 1), making extensive use of the cystine diagonal technique of Brown & Hartley (1966). [To aid comparison of the amino acid sequence of elastase with those of other pancreatic serine proteinases, we have used the chymotrypsinogen-A numbering scheme to describe the position of amino acid residues in the elastase polypeptide chain; see Fig. 21 and Shotton & Hartley (1970).]

We have completed the determination of the primary structure of porcine pancreatic elastase, and have published a description and discussion of the amino acid sequence homologies which exist between this enzyme and bovine trypsin, chymotrypsin-A and chymotrypsin-B (Shotton & Hartley, 1970). This sequence information has been further used to inter-

* Present address: Department of Botany, University of California, Berkeley, Calif. 94720, U.S.A.

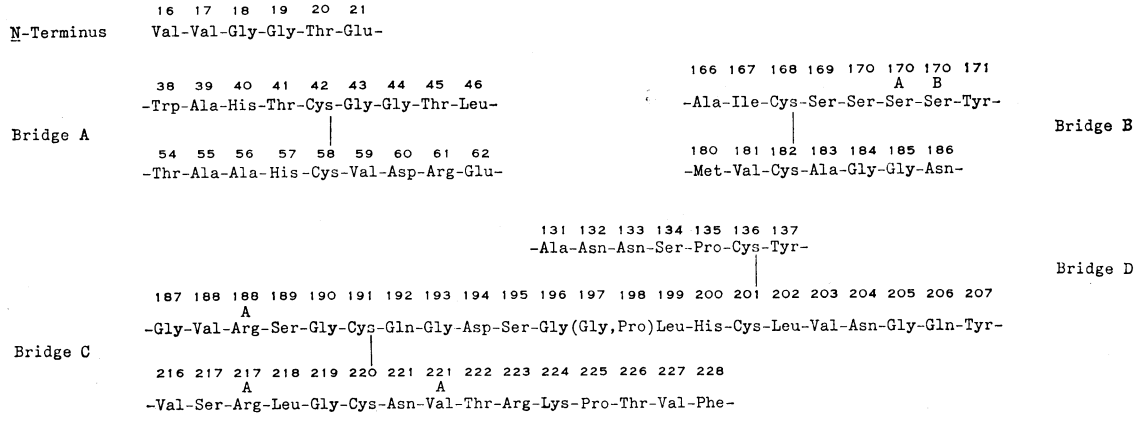


Fig. 1. *N-Terminus and disulphide bridges of elastase* (*Brown et al.*, 1967)

Residue numbers have been added, with the chymotrypsinogen-A numbering scheme, 'insertions' in the elastase sequence relative to that of chymotrypsinogen-A being numbered 170A, 170B etc. A serine residue has been deleted from the published sequence of Brown *et al.* (1967) at position 170C, a glycine residue has been added at position 205 and residue 186 has been corrected from aspartic acid to asparagine, on the basis of evidence described in this paper.

pret a 0.35nm electron-density map of the enzyme obtained by X-ray-diffraction studies, establishing that the three-dimensional structure of elastase is very similar to that of α-chymotrypsin (Shotton & Watson, 1970*a*,*b*; Watson *et al.*, 1970).

We give here the results of various chemical experiments on elastase and peptides derived from it, which form the evidence used to establish its complete amino acid sequence.

Experimental

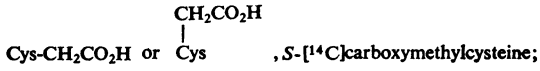
General strategy

The now classical approach of studying the sequences of tryptic peptides obtained from chemically modified derivatives of the enzyme, and obtaining the necessary sequence-overlap information from a study of small arginine- and lysine-containing peptides derived from other digests, was used to determine unambiguously the complete amino acid sequence of elastase.

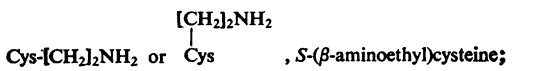
The tryptic peptides were obtained from two modified forms of Dip-elastase,* which had been prepared by inhibition of native elastase with di-isopropyl fluorophosphate (Dip-F) to prevent subsequent autolysis (Naughton & Sanger, 1961). The first derivative studied was aminoethylated Dip-elastase, in which the four cystine bridges of Dip-elastase had been reduced and then alkylated with ethyleneimine to yield aminoethylcysteine residues, providing eight additional sites of tryptic cleavage in areas of known sequence (Brown *et al.*, 1967; Fig. 1). It was hoped that this modification would decrease the amount of insoluble 'core' remaining after tryptic digestion, and give relatively small tryptic peptides with easily determined sequences. Peptides accounting for almost all the elastase molecule were isolated from a tryptic digest of this aminoethylated Dip-elastase (digest X), and most of the sequence studies were performed on them. However, because of the failure to recover some of the more hydrophobic peptides during column chromatography, and because of low yields and impurities due to incomplete tryptic cleavage at some aminoethylcysteine residues and unwanted secondary cleavages elsewhere (Cole *et al.*, 1967), it was necessary, in addition, to study the tryptic peptides from another derivative of Dip-elastase. This second derivative was [^{14}C]carboxymethylated Dip-elastase, prepared by alkylation of the reduced cystine bridges with iodo[$^{14}C_2$]acetic acid. Cyanogen bromide treatment of this derivative yielded three fragments, which were then subjected to tryptic digestion. From this digest (digest Y) all but the *C*-terminal peptide were purified, and where necessary were subjected to sequence study. The insoluble *C*-terminal peptide was isolated separately from the purified *C*-terminal cyanogen bromide fragment of this [^{14}C]carboxymethylated Dip-elastase.

The arginine-containing overlap peptides were obtained from a peptic digest of native elastase (digest P), and the required lysine overlaps were established by a study of [^{14}C]acetyl-lysine-containing peptides isolated from a combined tryptic plus chymotryptic digest (digest L) of aminoethylated Dip-elastase, which had been fully acetylated before digestion by reaction with [$^{14}C_2$]acetic anhydride.

* Abbreviations: The IUPAC–IUB Commission on Biochemical Nomenclature recommendations of symbols for amino acid derivatives and peptides [*Biochem. J.* (1972) **126**, 773–780] have been used in this and in the Supplementary Publication. Unusual or recently changed abbreviations used are:



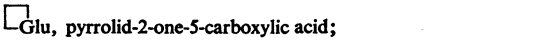
, *S*-[^{14}C]carboxymethylcysteine;



, *S*-(β-aminoethyl)cysteine;

Dip, di-isopropylphosphoryl-;

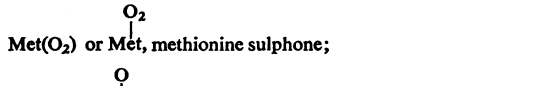
Dns or dansyl, 5-dimethylaminonaphthalene-1-sulphonyl-;



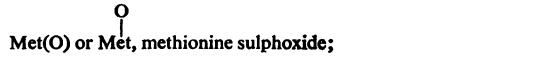
pyrrolid-2-one-5-carboxylic acid;

Hse, homoserine;

homoserine lactone;



methionine sulphone;



methionine sulphoxide;

Val-[CH_2]$_2NH_2$, *N*-(β-aminoethyl)valine.

Materials

Enzymes. Trypsin (2×crystallized, salt-free, batch TRSF 6188), α-chymotrypsin (3×crystallized, batch CD1 6108-9), pepsin (2×crystallized, batch PM C92), papain (2×crystallized, batch PAP 7KA) and Dip-F-treated carboxypeptidase-A (recrystallized, batch CPA-DFP 6139) were obtained from Worthington Biochemical Corp., Freehold, N.J., U.S.A.; Trypsin 1-300 was obtained from Nutritional Biochemicals Corp., Cleveland, Ohio, U.S.A.; thermolysin, from Daiwa Kasei K.K., Osaka, Japan, was a gift from Dr. R. P. Ambler; elastase was prepared from Trypsin 1-300 by the methods described by Shotton (1970). Aminopeptidase-M (batch 1/50932) was obtained from Röhm und Haas G.m.b.H., Darmstadt, West Germany.

Materials for chromatography and electrophoresis. Sephadex and Blue Dextran-2000 marker were obtained from Pharmacia (G.B.) Ltd., London W.5, U.K.; Zeo-Karb 225 (X2, SRC 7) ion-exchange resin was from The Permutit Co. Ltd., London W.4, U.K.; Whatman chromatography papers were from H. Reeve Angel and Co. Ltd., London E.C.4, U.K.; chromatographic and electrophoretic marker dyes were from G. T. Gurr Ltd., London, U.K.; polyamide layers (ε-polycaprolactam bonded to solvent-resistant polyester sheets) for the separation of Dns-amino acids were from the Cheng-Chin Trading Co. Ltd., Taipei, Taiwan.

Reagents. Analytical grade reagents were used if available; pyridine, β-mercaptoethanol, ethyleneimine trifluoroacetic acid, butyl acetate and phenylisothiocyanate were redistilled before use, and the last-named was then stored at −20°C under N_2; Dns-chloride and Dns-amino acids were obtained from BDH Chemicals Ltd., Poole, Dorset, U.K. *S*-(β-Aminoethyl)cysteine hydrochloride was a gift from Dr. R. D. Cole; iodo[$^{14}C_2$]acetic acid (15.5mCi/mmol) was obtained from The Radiochemical Centre, Amersham, Bucks., U.K. and was diluted before use with a solution of carrier iodoacetic acid, from which free iodine had been extracted by washing with chloroform, to a final specific radioactivity of 0.4mCi/mmol.

Buffers. The molarity of buffers described in this paper refers to the concentration of the anionic component of the buffer, with the exception of Tris–HCl and pyridine acetate buffers, where the molarities given refer to the concentrations of Tris and pyridine respectively.

Methods

Preparation of Dip-elastase. Dip-elastase was prepared by stirring a solution of elastase (20mg/ml) in 0.2M-Tris–HCl buffer, pH7.6, with a 15-fold molar excess of Dip-F in a sealed vessel at 2°C for 3h. The white precipitate of Dip-elastase was dissolved by adjusting the pH to 3.5 with 1M-HCl, and the solution was dialysed exhaustively against 1mM-acetic acid and freeze-dried (Naughton *et al.*, 1960). Assay of the residual elastolytic activity (Shotton, 1970) showed inhibition to be better than 99.8%.

Denaturation and reduction of Dip-elastase. Dip-elastase (500mg) was dissolved in 50ml of ice-cold 8M-urea–0.1M-acetic acid–0.1mM-EDTA that had been previously deoxygenated by bubbling N_2 through it for 30min. After the protein had been allowed to denature at 2°C for 15min at low pH, to prevent autolysis, solid Tris base was added to a concentration of 0.8M and the pH was adjusted to 8.6. The solution was deoxygenated by bubbling N_2 through it for a further 15min, a drop of octan-2-ol being used as an antifoaming agent, and then β-mercaptoethanol (40-fold molar excess over the half-cystine concentration) was added. The vessel was flushed with N_2 and sealed, and the solution stirred at room temperature for 4h. This preparation of denatured and reduced Dip-elastase was then either aminoethylated or carboxymethylated, as described below.

Preparation of aminoethylated Dip-elastase. To the solution containing 500mg of denatured and reduced Dip-elastase, ethylenimine (fivefold molar excess over all thiol groups) was added. The vessel was flushed with N_2 and sealed, and the reaction mixture stirred at room temperature. After 2h the amount of free thiol remaining [assayed by mixing a drop of the reaction mixture with a drop each of 10% (w/v) sodium nitroprusside solution and aq. 1M-NH_3, and observing the intensity of the purple colour formed] was negligible, and the reaction was stopped by adjusting the pH to 4.0 with HCl (Raftery & Cole, 1966). The aminoethylated Dip-elastase was precipitated by dilution with 20vol. of ethanol. The precipitate was washed three times in ethanol, redissolved in water and freeze-dried. Amino acid analysis showed no detectable cystine and 7.7mol of aminoethylcysteine/mol of Dip-elastase (96% conversion).

Preparation of [^{14}C]carboxymethylated Dip-elastase. The solution containing 500mg of denatured and reduced Dip-elastase was adjusted to pH3.5 with acetic acid and dialysed to equilibrium under N_2 against several changes of thoroughly deoxygenated 8M-urea–1mM-β-mercaptoethanol–0.8M-Tris–acetate, pH3.5, to remove the excess of β-mercaptoethanol. To this protein solution were added a deoxygenated solution of iodo[$^{14}C_2$]acetic acid (specific radioactivity 0.4mCi/mmol) in 8M-urea, sufficient to give a sixfold molar excess of iodoacetic acid over all thiol groups, and enough solid Tris base to titrate the mixture to pH8.6. The vessel was flushed with N_2 and sealed, and the reaction mixture stirred for 3h at room temperature, after which time the reaction was stopped by adding a sixfold molar excess of β-mercaptoethanol over iodoacetic acid. After the mixture had been stirred for a further hour, it was adjusted to pH3.5 with acetic acid, and the resulting [^{14}C]carboxymethylated Dip-elastase was purified by dialysis for 8h each time against two changes of 2 litres of 8M-urea titrated to pH3.5 with acetic acid, and for 24h each time against four changes of 10 litres of 1mM-acetic acid, before being freeze-dried (Milstein, 1966*b*). The specific radioactivity of the resulting protein indicated the presence of 8.02mol of [^{14}C]carboxymethylcysteine/mol of protein; the uncorrected recovery of carboxymethylcysteine on the amino acid analyser after 20h of hydrolysis was 7.0mol/mol (88% recovery).

Peptic digestion of native elastase, digest P. Elastase (200mg) was dissolved in 0.5ml of 98% formic acid and then diluted with 9.5ml of water. After 4h, 4mg

of pepsin was added and the digestion mixture incubated for 13h at 37°C. The clear solution was then fractionated on a column (bed volume 600ml) of Sephadex G-25 (fine grade) in 20mM-acetic acid, at a flow rate of 30ml/h, 4.5ml fractions being collected.

Tryptic digestion of aminoethylated Dip-elastase, digest X. Aminoethylated Dip-elastase is insoluble above pH6.5 in dilute buffer. Trial digests showed that the following conditions gave an optimum amount of material digested while minimizing the number of unwanted secondary cleavages. Aminoethylated Dip-elastase (400mg) was dissolved in 160ml of 10mM-HCl, and 0.8mg of trypsin, freshly treated with diphenylcarbamoyl chloride to inhibit chymotryptic contaminants (Erlanger *et al.*, 1966), was added. This digestion mixture was incubated at room temperature under N_2 at three progressively increasing pH values, in the hope that sufficient peptide bonds would be hydrolysed at the lower pH to increase solubility at the higher pH values, first for 1h at pH6.5, at which pH the aminoethylated Dip-elastase was fully soluble, next for 1h at pH7.5, during which time a slight precipitate formed, and finally at pH8.0, at which pH a substantial precipitate formed. Digestion was followed by titration with 1M-NaOH in a Radiometer TTT1c pH-stat. After 4.5h at pH8.0, when the rate of digestion became slow, the insoluble material was removed by centrifugation and the soluble digestion products of digest X, accounting for 70% of the original material absorbing at 280nm, were made 0.1M in pyridine and titrated to pH3.2 with acetic acid.

Fractionation of the tryptic peptides from digest X. The soluble digestion products of digest X were adsorbed on the top of a jacketed column (bed volume 400ml) of Zeo-Karb 225 (X2, SRC 7) sulphonated polystyrene resin, pre-equilibrated in the pyridinium form with 0.1M-pyridine acetate buffer, pH3.2. The products were fractionated at 40°C with 3.1 litres of a pyridine acetate gradient devised so that the rate of increase of pyridinium ion concentration and pH were as linear as possible, the former rising from 0.1 to 1.0M and the latter from pH3.2 to 5.1. [This effect was obtained by having two level buffer vessels connected by a siphon: the larger, which was stirred and from which buffer was pumped to the column top, contained 2.3 litres of the starting buffer, 0.1M-pyridine acetate, pH3.2, and the smaller, which was half the cross-sectional area of the first, containing 1.15 litres of 2.1M-pyridine acetate, pH6.0, 90% of the total gradient volume being used (Bock & Ling, 1954).] This gradient was followed by a step of 1.4 litre of 2.0M-pyridine acetate, pH5.0 (1.4M-pyridinium), at 40°C, and the most hydrophobic peptide was removed by a second step of 5M-pyridine acetate, pH6.0 (0.9M-pyridinium), at room temperature (see Fig. 3). The column was run at 60ml/h and 10ml fractions were collected.

Cyanogen bromide cleavage of [^{14}C]carboxymethylated Dip-elastase. [^{14}C]Carboxymethylated Dip-elastase (400mg) was dissolved in 16ml of 98% formic acid, and then the solution was diluted with water to 70% formic acid and 400mg of cyanogen bromide was added, giving a 30-fold molar excess of cyanogen bromide over methionine residues. The reaction mixture was incubated for 20h at room temperature, and then diluted with water to 600ml. It was concentrated by rotary evaporation at 30°C and diluted again twice to remove most of the formic acid, and finally freeze-dried (Gross & Witkop, 1962).

Fractionation of the cyanogen bromide fragments of [^{14}C]carboxymethylated Dip-elastase. The freeze-dried cyanogen bromide fragments (400mg) were dissolved in 4ml of 98% formic acid and diluted with 36ml of water to give a viscous solution. This was fractionated on a column (bed volume 1300ml) of Sephadex G-50 (fine grade) in 0.02M-acetic acid, run at 25ml/h, 10ml fractions being collected.

Tryptic digestion of [^{14}C]carboxymethylated Dip-elastase, digest Y. The substrate for this digest was the aggregating breakthrough material from the Sephadex G-50 fractionation of the cyanogen bromide fragments of [^{14}C]carboxymethylated Dip-elastase (see the Results section below). It consisted of all the large middle cyanogen bromide fragment, 88% of the *N*-terminal fragment and 79% of the *C*-terminal fragment, and was treated as equivalent to 350mg of unfragmented [^{14}C]carboxymethylated Dip-elastase. To this material, dissolved in 240ml of 0.02M-acetic acid, 4mg of trypsin, freshly treated with diphenylcarbamoyl chloride, was added and the mixture titrated to pH8.0 with solid NH_4HCO_3, a white precipitate forming at pH4.5. After incubation for 20h at room temperature, digestion was stopped by the addition of 2μl of Dip-F. The 'core' of undigested protein and insoluble peptides was removed by centrifugation, and the soluble peptides from digest Y, accounting for 55% of the material absorbing at 280nm, were concentrated to 20ml by rotary evaporation at 30°C and then fractionated on a column (bed volume 1300ml) of Sephadex G-50 (fine grade) in 0.5% (w/v) NH_4HCO_3, at a flow rate of 25ml/h, 10ml fractions being collected,

Purification and peptic digestion of the C-terminal tryptic peptide. The *C*-terminal tryptic peptide, residues 231–245, was not present among the soluble peptides purified from digest X and from digest Y. It was isolated as an insoluble tryptic 'core' by a preparative tryptic digest of the purified *C*-terminal cyanogen bromide fragment of [^{14}C]carboxymethylated Dip-elastase (residues 181–245), prepared as described above. Freeze-dried *C*-terminal fragment (1μmol) was suspended in 5ml of 0.5% (w/v) NH_4HCO_3, to which was added 0.25mg of trypsin, freshly treated with diphenylcarbamoyl chloride, giving an enzyme/substrate ratio of 1:100. After

digestion for 6h at 37°C the insoluble material was separated by centrifugation. 'Fingerprints' of the soluble material revealed only the presence of the tryptic peptides from residues 181–230. The insoluble material was digested again with twice the concentration of diphenylcarbamoyl chloride-treated trypsin for a further 18h at 37°C, and then freed of all soluble peptides by washing with water and 8M-urea, followed by extensive dialysis against water. It was evaporated to dryness *in vacuo* and dissolved in 0.2ml of 98% formic acid, giving a clear solution. This was diluted to 4ml with water to give a turbid solution in 5% formic acid, to which 50μl of pepsin (10mg/ml) was added, giving an enzyme/substrate molar ratio of 1:50. The mixture was incubated at 37°C for 20h, then dried *in vacuo* to remove the formic acid.

Determination of peptide positions in column eluent fractions. The eluents of the column fractionations described above were monitored for u.v.-absorbing material by using an LKB Uvicord II recording u.v. absorptiometer, and where appropriate for radioactivity by using a Nuclear–Chicago flow-through scintillation counter. Suitable portions were removed from every third or every fourth eluent fraction and evaporated to dryness; residues were redissolved in 20μl of water and applied as contiguous 1cm bands along the origins of sheets of Whatman no. 1 chromatography paper, and fractionated by paper electrophoresis at pH6.5, and pH2.0 for the neutral peptides, to give peptide 'fingerprints', from which the location of each peptide in the eluent fractions was clearly seen (see Figs. 2, 3 and 5). Pooled fractions were concentrated by rotary evaporation at 30°C, and then applied to Whatman 3MM chromatography paper for preparative peptide purification by high-voltage paper electrophoresis. With some pooled Zeo-Karb 225 eluent fractions from digest X, rotary evaporation of the pyridine acetate buffer left a brown oily liquid, despite dilution and re-evaporation to remove the pyridine and acetic acid. This did not interfere with the subsequent electrophoretic purification.

Quantitation of radioactive material. Labelled proteins were first solubilized by gentle heating in a small volume of 1M-'Hydroxide of Hyamine 10-X' (Packard). Radioactive quantitation was performed with a Packard Unilux liquid-scintillation counter, with the scintillation fluid toluene–2-methoxyethanol (3:1, v/v) containing 2,5-diphenyloxazole (7g/l) and 1,4-bis-(5-phenyloxazol-2-yl)benzene (0.25g/l). [^{14}C]-Toluene was used as an internal standard to correct for quenching. Counting rates were at least 30 times background.

High-voltage paper electrophoresis. Peptide solutions were applied as wide bands to Whatman chromatography paper, dried in a current of cool air and sharpened to the origin by careful application of buffer before electrophoresis. Internal fluorescent markers (Brown & Hartley, 1966), and dyes and amino acid mixtures as side markers (Milstein, 1966*a*) were used. Electrophoresis was performed at pH6.5, 3.5 or 2.0 under the conditions and loadings described by Ambler (1963), Whatman 3MM or no. 1 paper being used as appropriate. Where necessary, bands were cut out, stitched to a fresh sheet and subjected to chromatography or to electrophoresis at a different pH until purification was achieved.

Paper chromatography. Descending paper chromatography in butan-1-ol–acetic acid–water (3:1:1, by vol.) (Partridge, 1948), or in butan-1-ol–acetic acid–water–pyridine (BAWP) (15:3:10:12, by vol.) (Waley & Watson, 1953), was used to prepare 'fingerprints' of peptide mixtures fractionated in the first dimension by electrophoresis. Because of the difficulties discussed by Ambler (1963), and the greater speed and resolution of paper-electrophoretic methods, paper chromatography was employed for preparative peptide fractionations only when these could not be conveniently achieved by electrophoresis. In these cases the BAWP system of Waley & Watson (1953) was used, papers being developed once for 12–15h.

Detection of peptides on paper. After each electrophoretic or chromatographic run the paper was dried in a current of warm air and then examined under u.v. light to detect the positions of the fluorescent internal markers and the fluorescent bands that accompany tryptophan-containing peptides. The positions of radioactive peptides were determined by radioautography, the paper being stored in contact with Ilford Industrial-G X-ray film for 1–3 days.

The whole sheet or, in preparative runs, guide strips cut from the edge of the sheet, was then stained with cadmium–ninhydrin reagent (Heilmann *et al.*, 1957) and heated for about 3min in a current of hot air until the initial colours of the peptide spots developed. The papers were then stored overnight in polyethylene bags at room temperature, and any changes in the spot colour noted, the initial and final colours giving some indication of the *N*-terminal residue of the peptide stained (usually glycine or threonine, persistent yellow; serine, asparagine, cysteic acid or carboxymethylcysteine, yellow changing to orange–red; valine or isoleucine, weak red becoming strong only after several hours; others, orange or red). Peptides staining weakly with cadmium–ninhydrin reagent, often large acidic peptides, were detected by the chlorine stain of Reindel & Hoppé (1954). The specific colour tests for peptides containing tryptophan (Harley-Mason & Archer, 1958), tyrosine (Jepson & Smith, 1953), histidine (Dalgliesh, 1952), arginine (Irreverre, 1965) and *N*-terminal proline (Acher *et al.*, 1950) were also used where appropriate, those for tryptophan and tyrosine and the chlorine stain working well on

papers first stained with cadmium–ninhydrin reagent and then washed in acetone.

Electrophoretic mobilities. The electrophoretic mobilities of peptides at pH6.5 (abbreviated *m*) and at pH2.0 (abbreviated *n*) were measured from the neutral bands rather than the origin to allow for electroendosmotic flow, mobilities at pH6.5 being calculated relative to free aspartic acid ($m = 1.0$) (those of basic peptides being indicated by a negative sign), and mobilities at pH2.0 being calculated relative to free serine ($n = 1.0$). These mobilities were used to determine the correct allocation of amide groups in peptides containing asparagine and glutamine (Offord, 1966).

Elution. Purified peptides were eluted from paper strips into small test tubes with 20mM-acetic acid (for basic and neutral peptides) or aq. 20mM-NH_3 (for acidic peptides) and stored as frozen solutions.

Amino acid analysis. Protein or peptide samples for analysis were hydrolysed in evacuated sealed glass tubes in 0.2ml of constant-boiling HCl (6.1M), or of concentrated AristaR grade HCl (BDH Chemicals Ltd.)–water (1:1, v/v), to which 1% (w/v) phenol had been added to protect tyrosine residues from chlorination (Chaplin *et al.*, 1965), at 105°C for 12–100h, but usually for 18h.

Before quantitative analysis, samples (about 5nmol) of the hydrolysed peptides were usually analysed semi-quantitatively by paper electrophoresis on Whatman no. 1 paper at pH2.0 for 20min at 120V/cm, alongside known amino acid mixtures, followed by staining with collidine–ninhydrin reagent (Levy & Chung, 1953). Positive identification of all amino acids except serine, leucine and isoleucine, which are not resolved, is possible from their distinctive staining colours and mobilities. Amino acid derivatives can also be identified (homoserine lactone, $n = 2.1$, yellow; homoserine, $n = 1.0$, blue; *S*-(β-aminoethyl)cysteine, $n = 1.6$, blue; *N*-(β-aminoethyl)valine, $n = 1.5$, grey; carboxymethylcysteine, $n = 0.6$, grey–green), and where tryptophan is present a fluorescent non-staining hydrolysis product can usually be observed before staining, running with proline ($n = 0.8$).

Quantitative analyses were performed on a Beckman–Spinco model 120B amino acid analyser by the method of Spackman *et al.* (1958), incorporating the accelerated flow system of Spackman (1963) and high-sensitivity colorimeters (Evans Electroselenium Ltd., Halstead, Essex, U.K.). Sample loadings were 5–50nmol, and a 20cm basic column was used for analysis of peptides containing aminoethylcysteine (Raftery & Cole, 1966).

Enzymic digestion of peptides. Large purified peptides were often subjected to secondary digests, producing smaller fragments that were then purified by paper electrophoresis and subjected to sequence analysis. Digestion with trypsin, α-chymotrypsin, aminopeptidase-M or carboxypeptidase-A [solubilized by the method of Fraenkel-Conrat *et al.* (1956) and used for *C*-terminal residue identification] was performed in 0.5% (w/v) NH_4HCO_3 buffer, with pepsin in 5% (v/v) formic acid and with papain in pH6.5 electrophoresis buffer diluted with an equal volume of 1% (v/v) β-mercaptoethanol in water, for 3–6h at 37°C or 12–24h at room temperature. Digestion with thermolysin was performed in 5mM-$CaCl_2$–50mM-ammonium acetate buffer, pH8.5, at 60°C for 90min. About 50–500nmol of peptide, with an enzyme/peptide molar ratio of 1:50 or 1:100, was usually used in such digests.

Partial acid hydrolysis. This was performed in concentrated AristaR grade HCl (BDH Chemicals Ltd.) in evacuated sealed glass tubes at 37°C for 72h.

Sequence methods. N-Terminal amino acid residues of peptides were identified by the 'dansyl' method of Gray & Hartley (1963*a*), as described by Gray (1967*a*) and Hartley (1970), a positive identification being made only when the unknown Dns-amino acid appeared as a clear strongly fluorescent spot that could be unequivocally identified from its chromatographic position relative to known reference Dns-amino acids. Dns-peptides suspected to have *N*-terminal Dns-proline were subjected to acid hydrolysis for only 6h, and those thought to have *N*-terminal Dns-tryptophan were digested with chymotrypsin rather than hydrolysed by acid. A hydrolysis product of ε-Dns-aminoethylcysteine, which co-chromatographed with Dns-taurine, and Dns-dipeptides resulting from incomplete hydrolysis of valyl and isoleucyl peptide bonds, run in characteristic positions on the polyamide layers and were helpful for sequence confirmation.

Sequential analysis of peptides was performed by the 'dansyl'–Edman method of Gray & Hartley (1963*b*), as detailed by Gray (1967*b*) and Hartley (1970). The locations of amino groups of asparagine and glutamine residues were determined, where necessary, by measuring the electrophoretic mobility of 5nmol of the free residual peptide left after each round of Edman degradation.

Presentation of Results

Peptide nomenclature

Peptides isolated from the preparative digests of elastase and its derivatives, digests P, X and Y, and used in subsequent sequence studies are described by a number, according to their elution position from the column employed for their fractionation and to their electrophoretic mobility (see Figs. 2, 3 and 5), prefixed by a letter to indicate from which digest they were obtained (thus P1–P13, X1–X43, Y1–Y29). Peptides from the peptic digest of native elastase, digest P, which were not required for sequence or overlap evidence, are not numbered or described.

The insoluble *C*-terminal tryptic peptide of elastase, isolated from the purified *C*-terminal cyanogen bromide fragment of [$^{14}C_2$]carboxymethylated Dip-elastase as described above, is designated Y30. Three lysine overlap-peptides, isolated from a combined tryptic plus chymotryptic digest of [^{14}C]acetylated aminoethylated Dip-elastase (digest L) as described below, are designed L1, L2 and L3.

The sequence studies on P1–P13 and L1, L2 and L3 are described separately below. As many peptides were common to both digest X and digest Y, the studies on peptides X1–X43 and Y1–Y30 are described together, by considering their contributions to the sequence determination of the 15 'theoretical' tryptic peptides one would expect to be produced by perfect tryptic digestion at every arginine and lysine residue in the elastase molecule. These 'theoretical' tryptic peptides are numbered T1–T15, according to their position in the complete elastase sequence, numbering consecutively from the *N*-terminal end.

Peptides produced by secondary chemical or enzymic degradation of the peptides isolated from the primary digests are identified by an additional letter indicating the type of degradation employed and a further number for each of the products, with the code: A, partial acid hydrolysis; C, chymotrypsin; M, thermolysin; N, papain; P, pepsin; T, trypsin; R1, R2, R3 etc., residual peptides after successive rounds of Edman degradation, the number indicating the number of amino acid residues removed from the *N*-terminal end of the original peptide. Thus, for example X14N2R1 is the residual peptide produced by one round of Edman degradation of one of the papain digestion fragments, X14N2, of the peptide X14, a tryptic peptide from aminoethylated Dip-elastase.

Presentation of the sequence data: abbreviations and conventions

The methods used to purify primary digestion products subsequent to their initial column fractionations, and to purify secondary digestion products, are simply abbreviated 6.5, 3.5, 2.0, BAWP and G-25, where 6.5, 3.5 and 2.0 indicate the pH at which preparative high-voltage paper electrophoresis was conducted; BAWP indicates paper chromatography in butan-1-ol–acetic acid–water–pyridine (15:3:10:12, by vol.) and G-25 indicates gel filtration on Sephadex G-25 (fine grade) in a dilute volatile buffer. The order of these abbreviations indicates the order in which the various purification steps were performed.

Although the various peptide-staining tests described under 'Methods' were extensively used to detect and characterize peptides, they are not reported unless they provide essential sequence information not obtained by subsequent studies on the eluted peptides.

Quantitative amino acid analysis results are expressed as mol of each amino acid/mol of peptide, uncorrected for destruction of serine and threonine and for incomplete cleavage at valyl and isoleucyl bonds, trace contaminants (less than 0.2mol/mol of peptide) being ignored. Where necessary, the results of hydrolyses of varying lengths of time are reported. Semi-quantitative analyses with collidine–ninhydrin are scored +++ (very strong), ++ (strong), + (moderate) or trace. Positive results of the Ehrlich test for tryptophan are indicated thus: Trp (+).

Yields of primary peptides are given to the nearest 0.1 μmol, with no correction for losses, known or unknown, in their multistage preparation. Yields of secondary products are expressed as approximate percentage recoveries of the amount of peptide subjected to secondary digestion, or are scored +++, ++, + or trace.

Amino acid compositions that have been quantitatively established are indicated by continuous lines beneath the residues of the peptide in question, whereas those that have been determined only qualitatively are indicated by broken lines.

N-Terminal, *C*-terminal and 'dansyl'–Edman sequential analysis results are presented in an abbreviated form wherever possible, making use of the following symbols: $\rightarrow$ represents the unequivocal identification of the Dns-amino acid at the appropriate step of the Edman degradation; $\hookrightarrow$ indicates strong indirect evidence for the correct identification of the indicated *N*-terminal amino acid, in the absence of 'dansyl' information, from the cadmium–ninhydrin-staining colour, or from the slightly anionic mobilities at pH6.5 ($m = 0.05$–0.15) of 'neutral' peptides staining yellow or orange with cadmium–ninhydrin, whose *N*-terminal residue is asparagine, or more rarely serine or threonine (Offord, 1966; Ambler & Brown, 1967), or from the complete lack of reactivity of the *N*-terminal amino acid with cadmium–ninhydrin, Dns-chloride or phenylisothiocyanate, where it seems probable that an *N*-terminal glutamine residue has cyclized to form a pyrrolidone carboxylic acid residue; $\rightarrow$ represents the release of the indicated amino acid by aminopeptidase-M digestion; $\leftarrow$ represents the release of the indicated amino acid by carboxypeptidase-A digestion; $\hookleftarrow$ indicates strong indirect evidence for the correct identification of the *C*-terminal residue from the known specificity of trypsin, chymotrypsin or cyanogen bromide used to cleave the peptide bond *C*-terminal to this residue.

In Figs. 6–20 the results of experiments that establish the amino acid sequences of all the tryptic peptides of elastase are summarized. These figures, and similar sequences appearing in the text of the Supplementary Publication, share a common format.

The top line of each figure shows the amino acid sequence as it occurs in native elastase. Each peptide derived from this sequence is indicated by the peptide name and horizontal lines appearing vertically below the amino acid residues making up that peptide, with the use of the conventions described above. Any change introduced into the native sequence by chemical modification of an amino acid residue is indicated by the appropriate abbreviation for the modified residue appearing vertically below the original residue, beneath any unmodified peptides. All peptides appearing immediately below this abbreviation contain the modified, not the original, amino acid. For example, in Fig. 18 native elastase has the sequence Leu-Gly-Cys-Asn-Val-Thr-Arg, peptide X15 has the sequence

$$\begin{array}{c}[CH_2]_2NH_2\\ |\\ \text{Leu-Gly-Cys,}\end{array}$$

peptide X21 has the sequence Asn-Val-Thr-Arg, and peptide Y20 has the sequence

$$\begin{array}{c}CH_2CO_2H\\ |\\ \text{Leu-Gly-Cys-Asn-Val-Thr-Arg.}\end{array}$$

Results

Amino acid analysis of elastase

Identical samples of pure recrystallized elastase were hydrolysed for various lengths of time and analysed as described under 'Methods', a delayed long-column buffer change being used to separate valine and cystine, and appropriate internal standards being employed to correct for transfer errors. The results of these analyses are shown in Table 1, together with the analyses of other workers and the amino acid composition of elastase determined from the sequence studies (see below) and X-ray-crystallographic structure determination (Shotton & Watson, 1970*a*,*b*). The observed agreement between these analysis results and the amino acid composition of elastase is good, except that the values for serine and valine, always difficult to quantitate exactly, are a little in error. The values for half-cystine are predictably low and are included merely for completeness, accurate quantitation of half-cystine as cysteic acid by Brown *et al.* (1967) showing eight residues/molecule. The crystallographic studies show conclusively the presence of only seven tryptophan residues/molecule. The recalculated values of Gertler & Hofmann (1967), obtained from good chromatographically purified elastase, are also in good agreement (except that those of glutamic acid, proline and valine are slightly high), whereas the earlier results of Brown *et al.* (1967) reflect the slight impurity of their preparation (Shotton, 1970).

Sequence studies on the peptic peptides of native elastase from digest P

The electrophoretic 'fingerprints' of the peptic peptides of native elastase, produced by digest P and fractionated on Sephadex G-50 as described under 'Methods', are shown in Fig. 2. For a peptic digest it was very clean, giving relatively few peptides, which were present in good yields and were easily purified. Those peptides on which subsequent sequence studies were made are numbered P1–P12. P13 is another peptic peptide from native elastase, which was isolated not from digest P but from a similar digest during the studies of Brown *et al.* (1967) on the sequences of *N*-terminal and the disulphide-bridge peptides of elastase. The purification, and the results of amino acid analyses and *N*-terminal determinations on peptides P1–P13, are described in Table 2; Table 3 shows the results of sequence studies performed on them, and their final positions in the elastase sequence. In addition to these studies, peptide P7 was subjected to a thermolysin digest, and the fragments were studied to help establish the sequence of the tryptic peptide T2, and peptide P1 was digested with papain in parallel with peptide X14 to help establish the sequence of peptide T6. The details of these studies are described below together with those of the tryptic peptides.

Sequence studies on the lysine overlap-peptides from digest L

During other studies (Kaplan *et al.*, 1969, 1971), Dr. K. J. Stevenson, in our laboratory, performed a combined tryptic plus chymotryptic digest on a preparation of aminoethylated Dip-elastase in which the three lysine residues and the *N*-terminus had been radioisotopically labelled by acetylation with [^{14}C]-acetic anhydride. This digest (digest L) was fractionated on a column of Zeo-Karb 225 in a pyridine acetate gradient, similar to the one described for the fractionation of digest X under 'Methods'. Three of the radioactive peptides isolated, designated L1, L2 and L3, proved useful to these studies in establishing the required overlap evidence at two of the lysine positions in elastase, that for the third lysine having been established by Brown *et al.* (1967). The sequence studies performed on these peptides are described in Table 4.

Fractionation of the tryptic peptides from digest X

The tryptic peptides of aminoethylated Dip-elastase produced by digest X were fractionated on a Zeo-Karb 225 ion-exchange column, as described under 'Methods'. Fig. 3 shows the electrophoretic 'fingerprints' of the eluent fractions from this column. This fractionation resolved the complicated

Table 1. *Amino acid analysis of elastase*

Identical portions of recrystallized elastase were hydrolysed for varying lengths of time and analysed as described under 'Methods'. After correcting the values for serine and threonine for destruction, by extrapolation to zero hydrolysis time, and those for valine and isoleucine for incomplete hydrolysis, by extrapolation to infinite hydrolysis time, the observed recoveries were converted into mol of each amino acid/mol of elastase by using a factor obtained by dividing the sum of the mean values of the recoveries for all amino acids except cysteine and tryptophan by the number of these residues per molecule of elastase (namely 225). (An identical factor was obtained when the values for serine, theonine, proline, valine and isoleucine were also omitted.) All values have been rounded off to one decimal place. Beside the mean analytical results, the actual amino acid composition of elastase determined from the sequence and crystallographic studies (see Fig. 21) is given. For comparison, the published analytical results of Brown *et al.* (1967) and Gertler & Hofman (1967) are also included, the former having been multiplied by 1.039, and the latter by 1.062, to give the best fit with the amino acid composition.

		Amino acid composition (mol/mol of elastase)								
		Analytical results						Crystallographic and sequence results	Brown *et al.* (1967)	Gertler & Hofmann (1967)
Amino acid	Time of hydrolysis (h) ...	12	24	40	75	100	Mean±S.D.			
Lys		3.4	3.3	2.8	2.9	2.7	3.0±0.3	3	2.9±0.3	3.1
His		5.9	6.4	5.8	5.9	6.0	6.0±0.2	6	5.0±0.3	6.2
Arg		11.5	12.4	11.9	11.1	11.8	11.7±0.4	12	10.8±0.6	12.1
Asp		23.8	23.5	23.3	23.2	23.2	23.4±0.2	24	23.4±0.4	24.5
Thr		18.4	18.1	18.0	16.6	16.6	18.8	19	18.9	18.9
Ser		23.0	21.5	21.8	19.3	18.1	23.2	22	22.8	22.2
Glu		19.1	19.3	18.8	18.9	18.9	19.0±0.2	19	19.6±0.5	20.0
Pro		7.4	7.5	7.7	6.5	6.7	7.2±0.5	7	7.2±0.6	7.5
Gly		26.0	25.8	25.4	25.4	25.6	25.6±0.2	25	26.8±1.1	25.2
Ala		16.4	16.9	17.3	16.4	16.8	16.8±0.3	17	17.6±0.3	17.0
Cys		6.3	6.7	6.7	5.2	5.2	6.0±0.7	8	7.9±0.2	6.5
Val		20.2	23.4	24.2	24.9	25.2	26.0	27	27.0	27.7
Met		1.6	1.6	1.7	1.7	1.7	1.7±0.1	2	1.8±0.2	1.9
Ile		7.4	8.9	9.1	9.6	9.5	9.9	10	9.9	9.9
Leu		17.9	18.1	17.7	17.9	18.0	17.9±0.1	18	17.8±0.4	18.1
Tyr		11.2	11.2	11.5	11.4	11.4	11.3±0.1	11	10.6±0.5	10.6
Phe		3.1	3.5	3.1	3.4	3.6	3.3±0.2	3	3.4±0.2	3.1
Trp								7	7.7	7.3±0.5

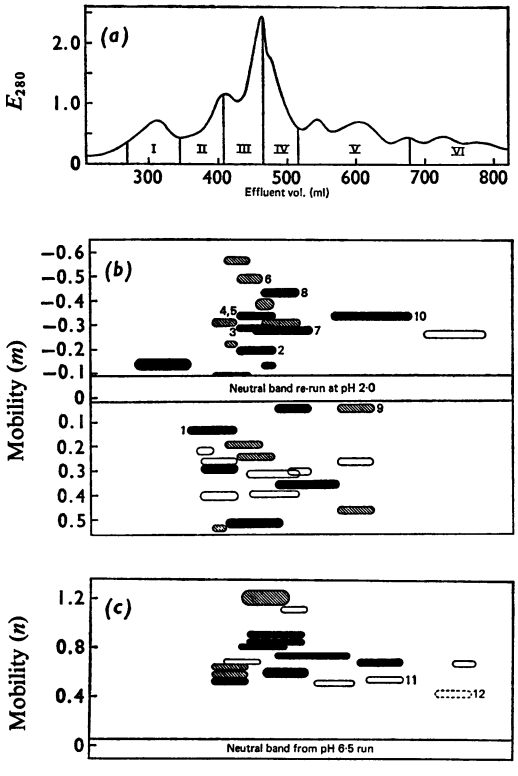


Fig. 2. *Peptic digest of elastase (digest P)*

(*a*) Elution profile on Sephadex G-50 (fine grade, bed volume 600 ml) in 20 mM-acetic acid. Roman numerals indicate pooled fractions. (*b*) Paper electrophoresis at pH 6.5 of samples from this column. The shading gives an indication of the relative intensities of bands after staining with cadmium–ninhydrin. Their elution volumes are indicated on the abscissa of (*a*). Arabic numerals indicate peptides described in Tables 2 and 3. (*c*) Paper electrophoresis at pH 2.0 of the neutral band from (*b*). This was cut out, stitched to another sheet of paper and subjected to electrophoresis. Ninhydrin-positive bands are shown as in (*b*). Peptide 12 was ninhydrin-negative but gave a purple colour with the Ehrlich stain for tryptophan.

mixture of peptides, each of which was eluted from the column in a relatively small volume. Both of the buffer steps after the elution gradient were useful in removing fairly large hydrophobic basic peptides from the resin. The eluent fractions were pooled for further purification into the 16 cuts shown in Fig. 3. Subsequent studies on the purified peptides obtained from this digest revealed that although tryptic digestion at aminoethylcysteine residues had been fairly efficient, it had been accompanied by several secondary cleavages of a non-tryptic nature, which produced several smaller peptides in low yield. These complicated the paper purification steps. Peptides accounting for all of the elastase sequence except residues 147–168, 208–215 and 231–245 were isolated from this digest and subjected to sequence studies as described below.

Fractionation of the cyanogen bromide fragments of [^{14}C]carboxymethylated Dip-elastase

The three polypeptide fragments produced by cyanogen bromide treatment of [^{14}C]carboxymethylated Dip-elastase were fractionated on Sephadex G-50 as described under 'Methods'. The elution profile from this fractionation is shown in Fig. 4. The large breakthrough peak contains all the middle cyanogen bromide fragment of 131 residues (54–180) in aggregation with much of the *N*-terminal and *C*-terminal fragments. No attempt was made to dissociate and refractionate this mixture, as the second and third peaks contained respectively 21% of the *C*-terminal fragment of 68 residues (181–245) and 12% of the *N*-terminal fragment of 41 residues (16–53), each in a high state of purity and in sufficient yield for further studies. Instead, the aggregating breakthrough peak was used directly as the substrate for a tryptic digestion, digest Y, which was then itself fractionated on Sephadex G-50. The purified *N*-terminal and *C*-terminal fragments were used as described below to confirm the accuracy of the local overlaps by which the tryptic peptides were ordered to establish the complete sequence of elastase, and the *C*-terminal fragment was further used as a source of the insoluble *C*-terminal tryptic peptide, as described under 'Methods'.

Fractionation of tryptic peptides from digest Y

The elution profile and electrophoretic 'fingerprints' of the tryptic peptides from digest Y, fractionated on Sephadex G-50 as described under 'Methods', are shown in Fig. 5. Although the peptides were eluted in quite large volumes, a moderately successful fractionation was obtained by pooling the eluted material into five cuts for subsequent purification. As the digest contained only clean tryptic peptides in good yield, their subsequent electrophoretic purification, which was aided by the ^{14}C label on the carboxymethylcysteine-containing peptides, was relatively straightforward. Peptides accounting for all of the elastase molecule except the *C*-terminal 15 residues (231–245) were purified from this digest, and where necessary were subjected to sequence analysis as described below.

Table 2. *Purification and composition of peptic peptides from native elastase*

The analytical data for the peptic peptides P1–P13 from digest P (see Fig. 2), obtained as described under 'Methods', are presented with the nomenclature and conventions detailed in the Presentation of Results section.

Peptide	Purification procedure	Yield (μmol)	Mobility *m*	Mobility *n*	Amino acid analysis (mol/mol of peptide)	*N*-Terminus (Dns-)	Comments
P1	6.5, 3.5	1.7	0.27	—	His (1.0), Arg (1.0), Asp (4.2), Thr (0.9), Glu (4.0), Gly (2.1), Val (2.5), Leu (0.8)	Arg	72h hydrolysis
P2	6.5	1.9	−0.18	—	His (0.8), Thr (1.0), Ser (1.0), Gly (1.1), Ala (1.0), Val (1.7), Phe (0.8)	Ala	
P3	6.5, BAWP	0.6	−0.30	—	Arg (1.1), Asp (0.9), Thr (1.7), Glu (1.0), Gly (1.1), Leu (0.9)	Thr	
P4	6.5, BAWP	0.8	−0.35	—	Arg (0.9), Thr (0.8), Ser (1.1), Ala (1.0), Val (0.9)	Thr	
P5	6.5, BAWP	0.1	−0.35	—	Lys (0.3), Arg (0.7), Thr (0.7), Ser (1.4), Glu (0.4), Pro (++), Gly (1.4), Ala (1.0), Val (0.8), Leu (0.7)	Gly	This low-yield peptide was impure. Despite this it gave unambiguous 'dansyl'–Edman results
P6	6.5	0.5	−0.54	—	Lys (1.1), Arg (1.2), Thr (1.8), Pro (++), Val (1.8), Phe (1.0)	Val	
P7	6.5, 2.0	0.6	−0.30	—	Arg (0.9) Asp (1.1), Ser (2.3), Glu (2.0), Pro (1.1), Ala (1.1); Trp (+)	Ala	
P8	6.5	1.0	−0.49	—	Arg (1.0), Leu (2.0)	Leu	
P9	6.5	1.4	0.13	—	Asp (1.1), Glu (1.0), Tyr (1.0)		Orange with cadmium–ninhydrin, and slightly anionic, therefore *N*-terminal Asn-
P10	6.5	1.1	−0.33	—	Arg (0.7), Asp (1.0), Glu (1.1), Ile (0.9); Trp (+)	Ile	
P11	6.5, 2.0, 3.5	0.4	0.00	0.54	Asp (2.0), Ser (1.1), Ile (2.1); Trp (+)	Ile	
P12	6.5, 2.0	0.2	0.00	0.46	Arg (++), Ser (3.0), Glu (1.0), Gly (1.0), Tyr (++); Trp (+)	Negative	No colour with cadmium–ninhydrin and unreative, therefore *N*-terminal $\overline{\text{Glu}}$-
P13	6.5, 2.0, BAWP		0.00	—	Thr (0.9), Glu (3.0), Ala (1.0), Leu (1.0)	Ala	

Table 3. *Sequence studies on peptic peptides from native elastase*

The symbols used to summarize these sequence results obtained from peptides P1–P13 are described in the Presentation of Results section. Peptides P1 and P7 were further analysed as described in the text.

Peptide	Sequence	Position in final sequence
P1	Arg-Val-Val-Val-Gly(Glu,His,Asn,Leu,Asn,Gln,Asn,Asn,Gly,Thr,Glu,Gln)	65A–81
P2	Ala-Val-His-Gly-Val-Thr-Ser-Phe	208–215
P3	Thr-Arg-Thr-Asn-Gly-Gln-Leu	144–151
P4	Thr-Arg-Val-Ser-Ala	229–233
P5	Gly-Val-Leu-Pro-Arg-Ala	121–126
P6	Val(Thr,Arg,Lys,Pro,Thr,Val,Phe)	221A–228
P7	Ala-Gln-Arg-Asn(Ser,Trp,Pro,Ser)Gln	22–30
P8	Leu-Arg-Leu	106–108
P9	Asn(Ser,Tyr)	115–117
P10	Ile-Arg-Gln-Asn-Trp	47–51
P11	Ile(Ser,Trp,Ile,Asn,Asn)	235–240
P12	Glu(Tyr,Arg,Ser,Gly,Ser,Ser,Trp)	34–38
P13	Ala-Gln-Thr-Leu-Gln-Gln	152–157

Results of peptic digestion of the insoluble C-terminal tryptic peptide

The *C*-terminal tryptic peptide of elastase, residues 231–245, is insoluble in water, salt solutions and 8M-urea. It was purified from the *C*-terminal cyanogen bromide fragment of [^{14}C]carboxymethylated Dip-elastase and subjected to peptic digestion as described under 'Methods'. Trial 'fingerprints' of the peptic digest revealed the presence of several soluble peptides in good yield, including two which gave a positive result to the Ehrlich test for tryptophan. This residue was known to be present in the *C*-terminal tryptic peptide, by comparison of the total amino acid analysis of elastase with the compositions of the peptides comprising residues 16–230 of the molecule, but was absent from the rest of the *C*-terminal cyanogen bromide fragment (residues 181–230). The absence from the 'fingerprints' of any radioactive peptides, arising from the four [^{14}C]carboxymethylcysteine residues elsewhere in the *C*-terminal cyanogen bromide fragment, indicated the high purity of the insoluble *C*-terminal tryptic peptide prepared by this procedure. The soluble peptides from this peptic digest were purified by paper electrophoresis and used

Table 4. *Lysine overlap peptides from [^{14}C]acetylated aminoethylated Dip-elastase*

The analytical data and sequence results for peptides L1–L3, obtained as described under 'Methods', are presented with the nomenclature, symbols and conventions detailed in the Presentation of Results section.

Peptide	Purification procedure	Mobility *m*	Mobility *n*	Amino acid analysis (mol/mol of peptide)	Sequence studies	Position in final sequence
L1	6.5 —	−0.19		Lys (1.0), His (1.0), Glu (1.0), Pro (1.2), Gly (1.1), Val (2.5) after 18h, 3.5 after 72h), Ile (0.5 after 18h, 1.2 after 72h), Tyr (1.0)	[^{14}C]acetyl | Val-Gly-Val-Gln-Lys-Ile-Val-Val-His-Pro-Tyr	82–93
L2	6.5, 2.0	0.00	0.62	Lys (0.9), Asp (1.0), Thr (1.0), Ser (1.2), Gly (1.1), Val (0.9)	[^{14}C]acetyl | Gly-Ser-Thr-Val-Lys-Asn	173–178
L3	6.5, 2.0	0.00	0.49	Lys (0.9), Asp (1.0), Thr (0.9), Ser (1.0), Gly (1.0), Val (1.0); Trp (+)	[^{14}C]acetyl | Trp-Gly-Ser-Thr-Val-Lys-Asn	172–178

to establish the complete sequence of the *C*-terminal tryptic peptide, as described under 'Peptide T15' below.

Sequence determination of the tryptic peptides of elastase (summary)

For convenience and brevity, the evidence for the amino acid sequences of the 15 'theoretical' tryptic peptides of elastase, T1–T15, deduced from studies of the tryptic peptides X1–X43 and Y1–Y30 and of the peptic peptides P1 and P7, is presented here in tabulated summary form, the conventions described in the Presentation of Results section above being used. The sequence evidence, some 50 pages of detailed typescript describing fully the amino acid compositions, mobilities and sequence analyses of the many secondary digestion products obtained from the primary peptides, is given in Supplementary Publication SUP 50016.

The purification, mobilities, amino acid analyses and *N*-terminal determinations of the peptides obtained from digest X are given in Table 5, and those of the peptides obtained from digest Y are given in Table 6. The results of the sequence analyses and secondary digests performed on these peptides are described below, in relation to the 15 'theoretical' tryptic peptides, T1–T15, of which they are part and whose sequences they help to determine. Some of the peptides are minor variants, which were not used in the sequence determination and are described merely for completeness.

Peptide T1 (residues 16–24). Peptides X12 and Y11 were identical. Studies on X12 extended the *N*-terminal sequence Val-Val-Gly-Gly-Thr-Glu- determined by Brown *et al.* (1967) (see Fig. 1) and showed the sequence of peptide T1 to be as shown in Fig. 6.

Peptide X27 was a variety of X12, in which the α-amino group had become aminoethylated.

Peptide T2 (residues 25–36). The sequence of T2 was established by studies on peptide X38 and the peptic peptide P7, and is shown in Fig. 7.

Peptide Y28 was identical with peptide X38. It was present in soluble form in the eluent fractions of the Sephadex G-50 column used to fractionate digest Y (see Fig. 5) and was isolated in purest form and greatest yield as well-shaped crystals, which formed when these eluent fractions were stored for 4 weeks at 0°C. The crystals were insoluble in water, aq. 0.2M-NH_3, pyridine and aq. 70% (v/v) pyridine, but dissolved easily in 0.2M-acetic acid. Unlike that of peptide X38, the α-amino group of the *N*-terminal asparagine residue of peptide Y28 did not become blocked during isolation, and yielded Dns-aspartic acid cleanly upon reaction with Dns-chloride and hydrolysis.

Peptides Y23 and Y24 were minor fragments of T2 isolated together in almost equimolar amounts. It was concluded that the sequence of peptide Y23 was

Asn-Ser-Trp-Pro-Ser-Gln-Ile

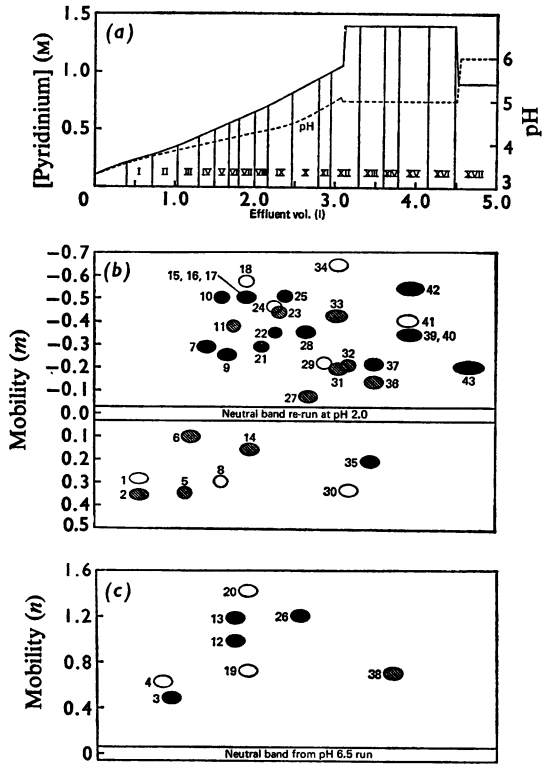


Fig. 3. *Tryptic digest of aminoethyl Dip-elastase (digest X)*

(*a*) Profile of the gradient on Zeo-Karb 225 resin, described under 'Methods'. ——, Pyridinium ion concentration; ----, pH. Roman numerals indicate pooled fractions. (*b*) Paper electrophoresis at pH6.5 of samples from this column. (*c*) Paper electrophoresis at pH2.0 of the neutral band from (*b*). Arabic numerals indicate peptides described in Table 5. The details are as given in Fig. 2.

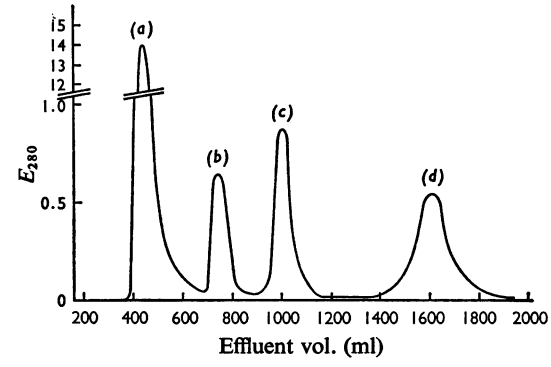


Fig. 4. *Cyanogen bromide fragments from* [^{14}C]*carboxymethyl Dip-elastase*

Elution profile on Sephadex G-50 (fine grade, bed volume 1300ml) in 20mM-acetic acid. (*a*) Aggregated fragments; (*b*) *C*-terminal fragment; (*c*) *N*-terminal fragment; (*d*) salt.

and that of peptide Y24 was

Asn-Ser-Trp-Pro-Ser-Gln-Ile-Ser-Leu.

Peptide T3 (*residues* 36*A*–48). Peptide T3 contains a half-cystine residue. Consequently, two peptides of digest X, X39 and X22, were derived from it, being overlapped by peptide Y16. Peptide X29 was a minor fragment of X39, and Y14 was a variety of Y16 in which the α-amino group has become blocked. From a study of these peptides the sequence of peptide T3 was completely determined (Fig. 8).

This confirms the sequence Trp-Ala-His-Thr-Cys-Gly-Gly-Thr-Leu determined by Smillie & Hartley (1966) (see Fig. 1).

Peptide T4 (*residues* 49–61). The sequence analysis of peptide T4 was complicated by two factors. The first was the tendency for the *N*-terminal glutamine residue to form pyrrolidone carboxylic acid during purification, which prevented conventional Edman degradation, and the second was the susceptibility of the particular aminoethylcysteine residue present in this peptide to modification, probably by oxidation, which rendered it resistant to tryptic attack. Many of the peptides from digest X used to determine the sequence of this region were minor products of non-tryptic cleavages. From the studies on peptides X2, X20, X32, X34, X36, X40 and X41 and their secondary digestion products, the sequence of peptide T4 was completely determined (Fig. 9).

This work confirmed the sequence Thr-Ala-Ala-His-Cys-Val-Asp-Arg-Glx determined by Smillie & Hartley (1966) (see Fig. 1).

Peptide Y21 was the expected *N*-terminal cyanogen bromide fragment of peptide T4 from digest Y, in which the *N*-terminal glutamine had been converted into pyrrolidone carboxylic acid, having the sequence

Glu-Asn-Trp-Val-Hse,

and peptide Y13 was the corresponding *C*-terminal cyanogen bromide fragment, having the sequence

CH_2CO_2H
|
Thr-Ala-Ala-His-Cys-Val-Asp-Arg.

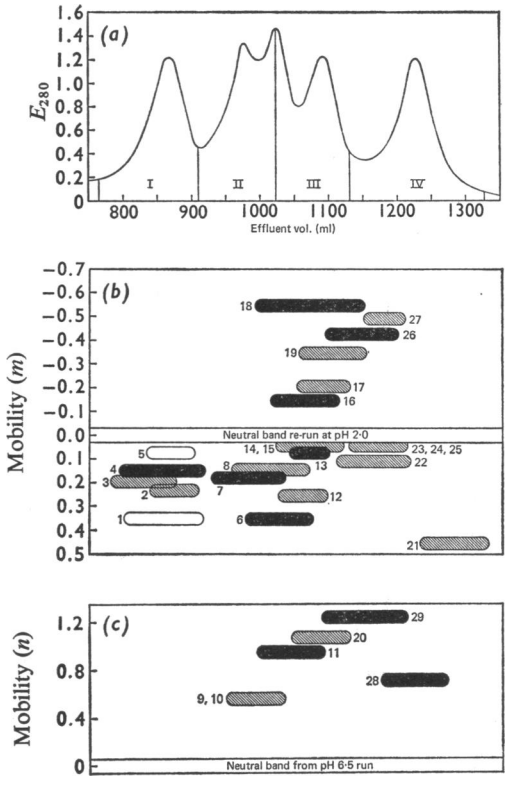
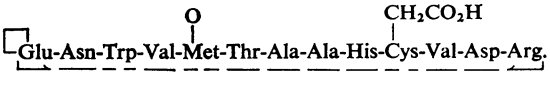


Fig. 5. *Tryptic digest of cyanogen bromide fragments from [^{14}C]carboxymethyl Dip-elastase (digest Y)*

(*a*) Elution profile on Sephadex G-50 (fine grade, bed volume 1300ml) in 0.5% (w/v) NH_4HCO_3. Roman numerals indicate pooled fractions. (*b*) Paper electrophoresis at pH6.5 of samples from this column. (*c*) Paper electrophoresis at pH2.0 of neutral band from (*b*). Arabic numerals indicate peptides described in Table 6. The details are as given in Fig. 2.

Peptide Y12 was an overlap peptide produced by incomplete cyanogen bromide cleavage at methionine-53, possibly because of prior partial oxidation to methionine sulphoxide, having the sequence



Peptides Y25 and Y15 were the parent peptides of Y21 and Y12 respectively, in which the *N*-terminal residues were uncyclized glutamine.

Peptide T5 (*residues* 62–65*A*). Peptides X26 and Y29 were identical. The sequence of peptide X26 was determined completely, giving the sequence of T5 as shown in Fig. 10.

Peptide T6 (*residues* 66–87). Peptides X14 and Y4 were identical. Peptide X14 was studied together with the peptic peptide P1 to determine the sequence of peptide T6, which is shown in Fig. 11. The allocation of amide groups in peptides P1N1 and P1N2 was made on the basis of the electrophoretic mobilities of their residual peptides after successive rounds of Edman degradation, as described and illustrated in the Supplementary Publication.

Peptide X19 was identical with peptide X14, but with an electrophoretic mobility that was only slightly acidic at pH6.5, which showed it to have only two acidic residues. The presence of three acidic residues in peptides X14, Y4 and P1 was therefore the result of deamidation of one amide residue during purification. Although neither glutamic acid in these peptides could easily have been derived from a glutamine during the conditions used in preparation and purification, the other three -Asp-Gly- sequences of elastase peptides (see peptides T10, T11 and T12) were all derived from -Asn-Gly- sequences by deamidation, the deamidated varieties of peptides containing this sequence invariably occurring in greater yield than the amidated variety. This phenomenon has also been noted in peptides of *Pseudomonas* cytochrome *c*-551 (Ambler, 1963). It was therefore concluded that the sequence of peptide T6 was as shown in Fig. 11, and that the extra acidic residue in peptides X14, Y4 and P1 was formed by the accidental deamidation of the labile -Asn-Gly- sequence during purification.

Peptide T7 (*residues* 88–107). Peptides X35 and Y7 were identical. Both preparations of this peptide were used in determining the sequence of peptide T7, as were the peptides X30, X33 and Y26 derived from them by secondary non-tryptic cleavages, as shown in Fig. 12. The allocation of amide groups in peptides Y7P5C1, Y7P6C1, Y7P5C2 and Y7P6C2 was made on the basis of the electrophoretic mobilities of their residual peptides after successive rounds of Edman degradation, as described and illustrated in the Supplementary Publication.

Peptide T8 (*residues* 108–125). This peptide was never isolated intact. The *N*-terminal half was represented by peptides X3 and Y9, which were identical; the sequence of peptide X3 was determined. Similarly the *C*-terminal half was represented by the identical peptides X31 and Y17; the sequence of peptide X31 was determined, and the two halves of peptide T8 were overlapped by the peptic peptide P9 (see Table 3). The sequence of peptide T8 is shown in Fig. 13.

Peptide T9 (*residues* 126–145). As peptide T9 contains a half-cystine residue, two tryptic peptides in digest X, X7 and X43, were obtained from it. Peptide X4 was a minor fragment of peptide X7, and

Table 5. *Purification and composition of tryptic peptides from digest X*

The analytical data for the tryptic peptides X1–X43 from digest X (see Fig. 3), obtained as described under 'Methods', are presented with the nomenclature and conventions detailed in the Presentation of Results section.

Peptide	Purification procedure	Yield (μmol)	Mobility *m*	Mobility *n*	Amino acid analysis (mol/mol of peptide)	*N*-Terminus (Dns-)	'Theoretical' tryptic peptide and position in final sequence	Comments
X1	6.5	0.2	0.29	—	Cys-$[CH_2]_2NH_2$ (0.5), Asp (1.0), Ser (0.8), Glu (1.1), Pro (0.9), Gly (2.9), Val (0.8), Leu (0.7)	Negative	Impurity*	No colour with cadmium–ninhydrin. N-Terminal —Glu
X2	6.5	0.5	0.33	—	Met(O_2) (0.2), Asp (1.0), Thr (1.0), Glu (1.0), Ala (1.0), Val (0.8), Met (0.6); Trp (+)	Negative	T4 49–55	No colour with cadmium–ninhydrin. *N*-Terminal —Glu
X3	2.0	1.4	0.00	0.50	Asp (1.0), Thr (1.0), Ser (1.1), Glu (1.1), Ala (1.1), Val (1.0), Leu (1.9)	Leu	T8 108–115	Identical with Y9
X4	2.0	0.2	0.00	0.58	Asp (1.1), Thr (0.9), Gly (1.2), Ala (2.1), Ile (0.8), Leu (1.1)	Ala	T9 126–132	Minor fragment of X7
X5	6.5	0.3	0.35	—	Asp (1.0), Glu (1.3), Gly (1.2), Val (0.9), Leu (0.9), Tyr (0.7)	Leu	T12 202–207	
X6	6.5	1.0	0.09	—	Cys-$[CH_2]_2NH_2$ (0.7), His (0.9), Asp (1.1), Ser (0.7), Glu (1.0), Pro (1.0), Gly (3.0), Leu (1.0)	Negative	T12 192–201	Active-centre Dip-peptide with *N*-terminal —Glu
X7	6.5	1.9	−0.29	—	Cys-$[CH_2]_2NH_2$ (0.9), Asp (2.0), Thr (0.9), Ser (1.2), Pro (1.1), Gly (1.0), Ala (2.2), Ile (1.0), Leu (1.0)	Ala	T9 126–136	
X8	6.5	0.3	0.30	—	Arg (1.0), Asp (1.0), Glu (1.0), Gly (2.0), Ala (1.0), Val (0.5 after 18h; 1.5 after 72h)	Val plus Val-Val	Impurity*	
X9	6.5, 3.5	0.7	−0.25	—	Cys-$[CH_2]_2NH_2$ (0.7), Asp (0.9), Ser (1.2), Val (1.0), Met (1.0)	Negative	T11 178–182	*N*-Terminal Asn had become blocked during isolation
X10	6.5, 2.0	1.5	−0.50	1.60	Cys-$[CH_2]_2NH_2$ (0.8), Ser (0.8), Gly (1.1)	Ser	T12 189–191	
X11	6.5, BAWP, 3.5	0.2	−0.38	—	Arg (1.0), Asp (1.3), Gly (3.0), Ala (1.0), Val (1.0)	Ala	T11 183–188A	
X12	6.5, 2.0	5.8	0.00	0.96	Arg (1.0), Thr (1.0), Glu (2.0), Gly (2.0), Ala (1.1), Val (0.7 after 18h; 1.4 after 72h)	Val plus Val-Val	T1 16–24	Identical with Y11

X13	6.5, 2.0, 3.5	3.4	0.00	1.19	Arg (0.7), Asp (1.0), Gly (3.0), Ala (1.0), Val (0.7)	Ala	T11	183–188A	Variety of X11, deamidated at -Asn-Gly-
X14	6.5, 2.0	0.6	0.15	0.77	Lys (0.9), His (1.1), Asp (3.8), Thr (1.0), Glu (5.0), Gly (3.0), Val (3.5 after 18h), Leu (1.0), Tyr (1.0)	Val plus Val-Val	T6	66–87	Identical with Y4. Variety of X19, deamidated at -Asn-Gly-
X15	6.5, BAWP, 3.5	5.0	−0.49	—	Cys-$[CH_2]_2NH_2$ (0.9), Gly (1.0), Leu (0.9)	Leu	T13	218–220	
X16	6.5, BAWP, 3.5	0.3	−0.49	—	Lys (1.0), Ser (0.9), Val (1.0)	Val	Impurity*		
X17	6.5, BAWP, 3.5	0.1	−0.49	—	Lys (0.9), Glu (1.0), Ala (2.2)	Ala	Impurity*		
X18	6.5, 3.5	0.1	−0.58	—	Cys-$[CH_2]_2NH_2$ (0.8), Val (1.1)	Val	T11	181–182	Minor fragment of X9. Bright purple with cadmium–ninhydrin
X19	6.5, 2.0	0.1	0.04	0.70	Lys (0.9), His (0.7), Asp (3.7), Thr (1.1), Glu (5.0), Gly (3.4), Val (3.7 after 18h), Leu (1.1), Tyr (1.2)	Val plus Val-Val	T6	66–87	
X20	6.5, 3.5	0.2	0.00	1.40	Arg (1.0), Asp (1.0), Val (1.0)	Val	T4	59–61	
X21	6.5	5.0	−0.29	—	Arg (1.1), Asp (1.0), Thr (1.0), Val (1.0)	Asp	T13	221–223	
X22	6.5	1.3	−0.34	—	Arg (1.2), Thr (1.0), Gly (1.9), Ile (1.0), Leu (1.1)	Gly	T3	43–48	
X23	6.5, 3.5	0.2	−0.44	—	Arg (1.1), Thr (1.2), Gly (1.0), Leu (0.9)	Gly	T9	142–145	Minor fragment of X43
X24	6.5, 3.5	0.1	−0.45	—	Arg (1.1), Ser (1.0), Val (1.0)	Val	T12	216–217A	
X25	6.5	1.2	−0.49	—	Arg (1.0), Thr (0.9)	Thr	T14	229–230	Identical with Y27. Fragment of X42
X26	3.5	5.5	0.00	1.20	Arg (0.9), Thr (0.9), Glu (1.0), Leu (1.1), Phe (1.0)	Glu	T5	62–65A	Identical with Y29
X27	6.5, 2.0	0.4	−0.05	1.30	Arg (0.8), Val-$[CH_2]_2NH_2$ (0.7 after 72h), Thr (0.9), Glu (1.9), Gly (2.0), Ala (1.1), Val (1.0 after 72h)	Val-$[CH_2]_2NH_2$	T1	16–24	α-Aminoethylated *N*-terminal peptide
X28	6.5	3.1	−0.34	—	Lys (1.0), Thr (1.0), Pro (1.0), Val (1.0), Phe (1.0)	Lys	T14	224–230	Identical with Y19. Fragment of X42
X29	6.5	0.2	−0.21	—	His (1.0), Ser (2.6 after 18h), Gly (1.0), Ala (1.0); Trp (+)	Ser	T3	36A–40	Minor fragment of X39

Table 5—*continued*

Peptide	Purification procedure	Yield (μmol)	Mobility *m*	Mobility *n*	Amino acid analysis (mol/mol of peptide)	*N*-Terminus (Dns-)	'Theoretical' tryptic peptide and position in final sequence		Comments
X30	6.5	0.3	0.32	—	His (0.8), Asp (4.0), Thr (1.0), Pro (1.0), Gly (1.2), Ala (3.0), Val (2.1 after 18h), Ile(1.2 after 18h), Tyr (0.6); Trp (+)	Ile and Ile-Val	T7	88–104	Minor fragment of X35
X31	6.5, 3.5	1.4	–0.19	—	Arg (1.0), Ser (1.0), Glu (1.0), Pro (0.9), Gly (1.1), Val (1.8 after 72h), Leu (1.8 after 72h), Tyr (0.9)	Ser	T8	116–125	Identical with Y17
X32	6.5, 3.5	0.5	–0.20	—	Cys-$[CH_2]_2NH_2$ (trace), taurine (++), Arg (0.9), Asp (1.0), Val (1.0)	Negative	T4	58–61	Fragment of X36. *N*-Terminal Cys-$[CH_2]_2NH_2$ modified and resistant to trypsin
X33	6.5, 2.0	0.5	–0.42	1.48	Arg (1.1), Leu (1.9)	Leu	T7	105–107	Identical with Y26. Minor fragment of X35
X34	6.5, 3.5	0.1	–0.63	—	Cys-$[CH_2]_2NH_2$ (0.6), His (1.0), Ala (1.1)	Ala	T4	56–58	Fragment of X36
X35	6.5, 2.0	0.8	0.19	0.65	His (0.9), Arg (1.0), Asp (3.7), Thr (1.0), Pro (1.0), Gly (1.2), Ala (3.0), Val (2.1 after 18h), Ile (1.2 after 18h), Leu (2.1), Tyr (1.9); Trp (+)	Ile and Ile-Val	T7	88–107	Identical with Y7
X36	2.0, 6.5	0.5	–0.13	0.89	Cys-$[CH_2]_2NH_2$ (0.7), His (1.0), Arg (1.1), Asp (2.0), Thr (1.0), Glu (1.2), Ala (1.9), Val (1.8 after 18h), Met (0.6); Trp (+)	Negative	T4	49–61	*N*-Terminal —└Glu. Cys-$[CH_2]_2NH_2$ modified and resistant to trypsin
X37	6.5, BAWP, G-25	1.1	–0.21	—	Lys (0.9), Thr (1.0 after 12h; 0.9 after 24h), Ser (5.0 after 12h; 4.8 after 24h), Gly (1.2), Val (1.0), Tyr (1.0); Trp (+)	Ser	T10	169–177	
X38	2.0, 3.5, 2.0, G-25	0.7	Insoluble	0.67	Arg (0.9), Asp (1.0), Ser (2.9 after 12h; 2.6 after 24h), Glu (2.0), Pro (1.0), Ile (1.0), Leu (1.1), Tyr (1.0); Trp (+)	Negative	T2	25–36	Identical with Y28. *N*-Terminal Asn had become blocked during isolation

X39	6.5, 3.5, G-25	1.8	−0.32	—	His (1.0), Cys-$[CH_2]_2NH_2$ (0.8), Thr (1.0), Ser (2.7 after 12h; 2.5 after 24h), Gly (1.0), Ala (1.0); Trp (+)	Ser	T3	36A–42	
X40	6.5, 3.5	0.3	−0.32	—	Cys-$[CH_2]_2NH_2$ (0.2), His (0.8), Arg (0.9), Asp (1.0), Ala (0.9), Val (1.0)	Ala	T4	56–61	Fragment of X36. Cys-$[CH_2]_2NH_2$ modified and resistant to trypsin
X41	6.5, 3.5	0.2	−0.38	—	Cys-$[CH_2]_2NH_2$ (trace), His (0.9), Arg (1.5), Asp (1.0), Ala (1.7), Val (1.2)	Ala	T4	55–61	Fragment of X36. Cys-$[CH_2]_2NH_2$ modified and resistant to trypsin
X42	6.5, 3.5	3.5	−0.54	—	Lys (1.0), Arg (1.0), Thr (2.0), Pro (+), Val (1.0), Phe (0.9)	Lys	T14	224–230	Identical with Y18
X43	6.5, 2.0	1.0	−0.20	0.80	Arg (1·0), Thr (1.9), Gly (2.0), Ile (0.9), Leu (1.1), Tyr (0.8); Trp (+)	Tyr	T9	137–145	

* For description see Supplementary Publication SUP 50016.

peptide X23 a minor fragment of peptide X43. Studies on these peptides and the uncleaved peptide Y10 showed the sequence of peptide T9 to be as shown in Fig. 14. Peptide Y8 was identical with peptide Y10, except that it was acidic, indicating that one of its two asparagine residues had become accidentally deamidated during preparation. These studies confirmed the sequence Ala-Asn-Asn-Ser-Pro-Cys-Tyr determined by Brown *et al.* (1967) (see Fig. 1).

*Peptide T*10 (*residues* 147–177). This peptide contains a half-cystine residue, and consequently was expected to give rise to two tryptic peptides in digest X. However, the larger *N*-terminal peptide of 22 residues was not isolated from the digestion products. Careful amino acid analyses and Edman degradations on the smaller *C*-terminal peptide, peptide X37, and its chymotryptic digestion products, described in Supplementary Publication SUP 50016, proved its sequence to be Ser-Ser-Ser-Ser-Tyr-Trp-Gly-Ser-(Thr,Val)Lys (see Fig. 15). As peptide X37 was a major tryptic peptide from digest X, which was absent from digest Y, it follows that it was preceded by an aminoethylcysteine residue before tryptic digestion, giving an overlap sequence

$$\begin{array}{l}[CH_2]_2NH_2\\ \quad|\\ \text{-Cys-Ser-Ser-Ser-Ser-Tyr-Trp-Gly-Ser(Thr,Val)Lys}\end{array}$$

This corrects the sequence Ala-Ile-Cys-Ser-Ser-Ser-Ser-Ser-Tyr reported by Brown *et al.* (1967), where the sequence of five serine residues was erroneously suspected, in the absence of direct sequence evidence, on the basis of a single analysis of an overnight hydrolysate of the peptide, which was reported as being slightly contaminated.

Peptide T10 was represented in digest Y by peptide Y2, which was subjected to sequence analysis as shown in Fig. 15. This peptide and peptide Y2M1 had mobilities that indicated an *N*-terminal sequence Thr-Asp-Gly-Gln-, but as the peptic peptide P3, Thr-Arg-Thr-Asn-Gly-Gln-Leu (Table 3), was basic and not neutral at pH6.5, having no acidic residues, it was concluded that this was another example (see also peptides T6, T11 and T12) of the deamidation of a labile -Asn-Gly- sequence, which had occurred during the purification of peptide Y2, the true *N*-terminal sequence of peptide T10 therefore being Thr-Asn-Gly-Gln-. From these studies on peptides X37 and Y2, together with the known sequences of the peptic peptides P3 and P13 (see Table 3), the sequence of T10 was completely determined, and is shown in Fig. 15.

*Peptide T*11 (*residues* 178–188*A*). Two peptides were expected in digest X from this region, which contains a half-cystine residue. Peptide X9 was the expected *N*-terminal peptide, which on isolation was found to have a blocked α-amino group, and peptide

Table 6. *Purification and composition of tryptic peptides from digest Y*

The analytical data for the tryptic peptides Y1–Y29 from digest Y (see Fig. 5), obtained as described under 'Methods', are presented with the nomenclature and conventions detailed in the Presentation of Results section.

Peptide	Purification procedure	Yield (μmol)	Mobility *m*	Mobility *n*	Amino acid analysis (mol/mol of peptide)	*N*-Terminus (Dns-)	'Theoretical' tryptic peptide and position in final sequence	Comments
Y1	6.5, 3.5	0.2	0.35	—	His (1.7), Arg (1.1), Cys-CH_2CO_2H (1.4), Asp (2.4), Thr (1.0), Ser (2.9 after 24h), Glu (2.3), Pro (1.0), Gly (6.0), Ala (1.2), Val (3.7), Leu (1.8), Tyr (0.9), Phe (0.9)	—	T12 189–217A	Minor monoisopropyl derivative of Y5, deamidated at -Asn-Gly-
Y2	6.5, 3.5	0.7	0.22	—	Lys (0.8), Cys-CH_2CO_2H (0.7), Asp (2.1), Thr (3.5 after 18h), Ser (4.9 after 18h), Glu (3.9), Pro (1.1), Gly (2.1), Ala (3.0), Val (1.8), Ile (0.8), Leu (2.9), Tyr (2.7); Trp (+)	Thr	T10 147–177	
Y3	6.5, 3.5	0.8	0.20	—	His (1.9), Arg (1.0), Cys-CH_2CO_2H (1.6), Asp (2.2), Thr (1.1), Ser (3.3 after 18h), Glu (2.1), Pro (1.2), Gly (5.8), Ala (1.1), Val (3.6 after 18h), Leu (1.9), Tyr (0.9), Phe (1.0)	Ser	T12 189–217A	Minor variety of Y5, deamidated at -Asn-Gly-
Y4	6.5, 3.5	0.5	0.15	—	Lys (0.9), His (1.0), Asp (3.9), Thr (1.0), Glu (5.1), Gly (3.4), Val (3.6 after 18h), Leu (1.1), Tyr (1.0)	Val and Val-Val	T6 66–87	Identical with X14
Y5	6.5	0.2	0.09	—	His (2.2), Arg (1.2), Cys-CH_2CO_2H (1.0), Asp (2.0), Thr (1.2), Ser (3.2 after 18h), Glu (2.2), Pro (0.9), Gly (5.0), Ala (1.2), Val (3.4 after 18h), Leu (2.0), Tyr (1.2), Phe (0.8)	Ser	T12 189–217A	
Y6	6.5	1.9	0.34	—	Arg (1.0), Cys-CH_2CO_2H (0.6), Asp (1.1), Gly (3.2), Ala (1.3), Val (1.8)	Val	T11 181–188A	Cyanogen bromide fragment, deamidated at -Asn-Gly-
Y7	6.5, 3.5	0.6	0.19	—	His (0.9), Arg (0.9), Asp (3.7), Thr (1.0), Pro (0.9), Gly (1.0), Ala (3.1), Val (2.2 after 18h), Ile (1.3 after 18h), Leu (2.0), Tyr (1.8); Trp (+)	Ile and Ile-Val	T7 88–107	Identical with X35

Y8	6.5, 3.5	0.3	0.15	—	Arg (++), Cys-CH_2CO_2H (0.6), Asp (1.8), Thr (2.6), Ser (1.0), Pro (0.9), Gly (3.0), Ala (2.1), Ile (1.8), Leu (2.0), Tyr (1.0); Trp (+)	Ala	T9	126–145	Variety of Y10, deamidated at -Asn-Asn-
Y9	6.5, 3.5, 3.5	0.3	0.00	0.50	Asp (1.0), Thr (1.0), Ser (1.1), Glu (0.8) ,Gly (0.6), Ala (1.0), Val (0.7), Leu (1.7)	—	T8	108–115	Identical with X3. Contaminated with glycine. Not repurified or studied further
Y10	6.5, 3.5	0.1	0.00	0.54	Arg (++), Cys-CH_2CO_2H (0.8), Asp (2.0), Thr (2.7), Ser (1.0), Pro (0.9), Gly (3.0), Ala (2.2), Ile (1.7), Leu (2.0), Tyr (1.0); Trp (+)	Ala	T9	126–145	
Y11	6.5, 3.5	0.7	0.00	0.90	Arg (1.0), Thr (1.0), Glu (2.0), Gly (2.1), Ala (1.1), Val (0.7 after 18h)	Val plus Val-Val	T1	16–24	Identical with X12
Y12	6.5	0.4	0.25	—	His (1.0), Arg (1.0), Cys-CH_2CO_2H and Met(O) (0.6), Asp (1.9), Thr (1.0), Glu (1.3), Ala (2.1), Val (1.9); Trp (+)	Negative	T4	49–61	Uncleaved by cyanogen bromide, probably due to oxidation of Met to Met(O). No colour with cadmium–ninhydrin. *N*-Terminal —Glu
Y13	6.5	0.5	0.09	—	His (0.9), Arg (0.9), Cys-CH_2CO_2H (0.6), Asp (1.1), Thr (1.0), Ala (2.2), Val (0.9)	—	T4	54–61	Cyanogen bromide fragment. Yellow with cadmium–ninhydrin, therefore *N*-terminal Thr
Y14	6.5	0.2	0.05	—	His (1.3), Arg (1.3), Cys-CH_2CO_2H and Met(O) (0.9), Asp (1.0), Thr (2.3), Ser (2.9), Glu (0.7), Gly (3.0), Ala (2.2), Val (0.9), Ile (0.8), Leu (1.1); Trp (+)	Glu only	T3	36A–48	Y14 and Y15 were minor peptides eluted and hydrolysed together as a 2:1 mixture. Y14 was identical with Y16, except that its α-amino group was blocked. Y15 was identical with Y12, except that it had *N*-terminal Gln
Y15	6.5	0.1	0.05	—			T4	49–61	

Table 6—*continued*

Peptide	Purification procedure	Yield (μmol)	Mobility *m*	Mobility *n*	Amino acid analysis (mol/mol of peptide)	*N*-Terminus (Dns-)	'Theoretical' tryptic peptide and position in final sequence	Comments
Y16	6.5, 3.5	0.9	–0.14	—	His (0.8), Arg (0.7), Cys-CH_2CO_2H (0.6), Thr (1.7), Ser (2.6 after 18h), Gly (2.9), Ala (1.1), Ile (0.8), Leu (1.0); Trp (+)	Ser	T3 36A–48	
Y17	6.5	0.3	–0.19	—	Arg (++), Ser (1.0), Glu (1.0), Pro (1.0), Gly (1.3), Val (1.8 after 24h), Leu (1.7 after 24h), Tyr (0.9)	Ser	T8 116–125	Identical with X31
Y18	6.5	—	–0.54	—	—	—	T14 224–230	Identical with X42. Not eluted for separate study
Y19	6.5	—	–0.34	—	—	—	T14 224–228	Identical with X28. Not eluted for separate study
Y20	6.5, 2.0	0.5	0.00	1.04	Arg (1.1), Cys-CH_2CO_2H (0.6), Asp (0.9), Thr (1.1), Gly (1.4), Val (1.0), Leu (1.0)	Leu	T13 218–223	
Y21	6.5	0.5	0.45	—	Asp (1.0), Hse (0.4), Glu (1.2), Val (1.0); Trp (+)	Negative	T4 49–53	Cyanogen bromide fragment. Variety of Y25. No colour with cadmium–ninhydrin. *N*-Terminal —□Glu
Y22	6.5, 3.5	0.5	0.12	—	Asp (1.0), Ser (1.1), Hse (0.8)	—	T11 178–180	Cyanogen bromide fragment. Yellow to orange with cadmium–ninhydrin and slightly anionic, therefore *N*-terminal Asn
Y23	6.5, 2.0	0.2	0.06	0.54	Asp (2.0), Ser (4.9), Glu (2.1), Pro (1.8), Ile (1.8), Leu (1.0); Trp (+)	Asp only	T2 25–31	Y23 and Y24 were minor peptides eluted and hydrolysed together as a 1:1 mixture
Y24	6.5, 2.0	0.2	0.06	0.54			T2 25–33	

Y25	6.5, 3.5	0.1	0.04	—	Asp (1.0), Ser (0.7), Hse (+), Glu (1.0), Gly (0.4), Ala (0.5), Val (0.8); Trp (+)	—	T4	49–53	Very impure cyanogen bromide fragment, Gln-Asn-Trp-Val-Hse. Sequence in this region already determined, therefore not studied further
Y26	6.5	—	–0.42	—	Arg (++), Leu (+++)	—	T7	105–107	Identical with X33, therefore not quantitated or studied further
Y27	6.5	—	–0.49	—	—	—	T14	229–230	Identical with X25. Not eluted for separate study
Y28	crystals; 2.0	1.0	Insoluble	0.67	Arg (1.0), Asp (1.0), Ser (2.8 after 18h), Glu (2.0), Pro (1.0), Ile (1.0), Leu (1.0), Tyr (1.0); Trp (+)	Asp	T2	25–36	Identical with X38
Y29	6.5, 2.0	—	0.00	1.20	Arg (+++), Thr (+++), Glu (+++), Leu (+++), Phe (+++)	Glu	T5	62–65A	Identical with X26, therefore not quantitated or studied further

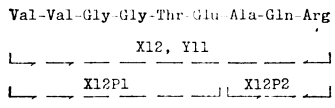


Fig. 6. *Sequence of peptide T1*

In this and ensuing figures (Figs. 7–20) the nomenclature, symbols and conventions used are those described in the Presentation of Results section.

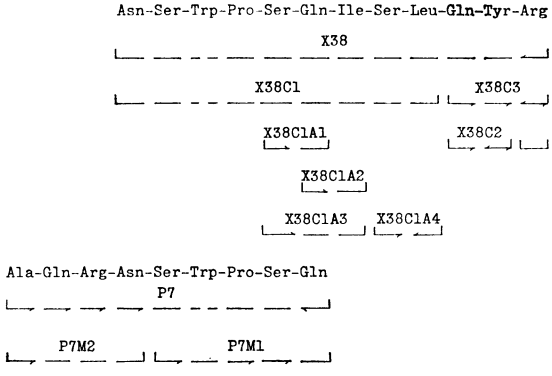


Fig. 7. *Sequence of peptide T2*

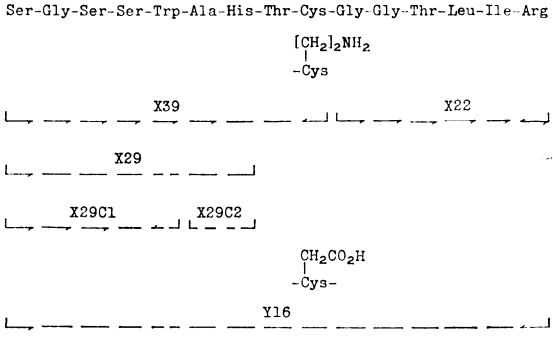


Fig. 8. *Sequence of peptide T3*

X18 was a minor fragment of this peptide produced by a chymotryptic-like cleavage at the methionine residue. Peptide X11 was the expected basic *C*-terminal peptide, but was isolated in much lower yield than the neutral version of the same sequence, peptide X13, which had been produced by the accidental deamination of another -Asn-Gly- sequence. Because of the presence of a methionine residue, two peptides were also expected from digest Y. These were peptides Y22 and Y6, the latter, like

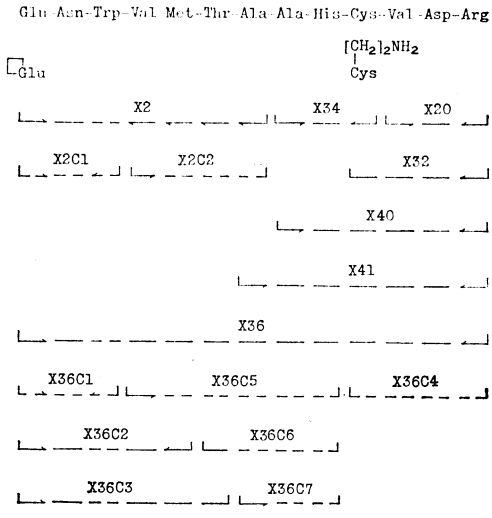


Fig. 9. *Sequence of peptide T4*

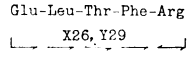


Fig. 10. *Sequence of peptide T5*

peptide X13, being the deamidated form of the sequence. Studies on these peptides completely determined the sequence of peptide T11 (Fig. 16).

This confirmed the deamidated sequence Met-Val-Cys-Ala-Gly-Gly-Asp reported by Brown *et al.* (1967) and overlapped it with their peptide starting Gly-Val-Arg- (see Fig. 1).

Peptide T12 (residues 189–217A). This peptide contains two half-cystine residues, the active centre Dip-serine residue and the fourth labile -Asn-Gly- sequence of the elastase molecule (see also peptides T6, T10 and T11). Consequently several peptides from each digest were derived from this region, and were studied as shown in Fig. 17. Peptide X6 was used to resolve the -(Gly,Pro)- ambiguity in the sequence of the large peptic active-centre peptide Gly-Val-Arg-Ser-Gly-Cys-Gln-Gly-Asp-Ser-Gly(Gly,Pro)Leu-His-Cys-Leu-Val-Asn-Gln-Tyr reported by Brown *et al.* (1967), but as its composition agreed with the corresponding region of the peptic peptide, this region of the sequence was not otherwise reinvestigated. Studies on peptide X5 (see Fig. 17) showed it to have the sequence Leu-Val-Asx(Gly,Glx-Tyr). This corrects the results of

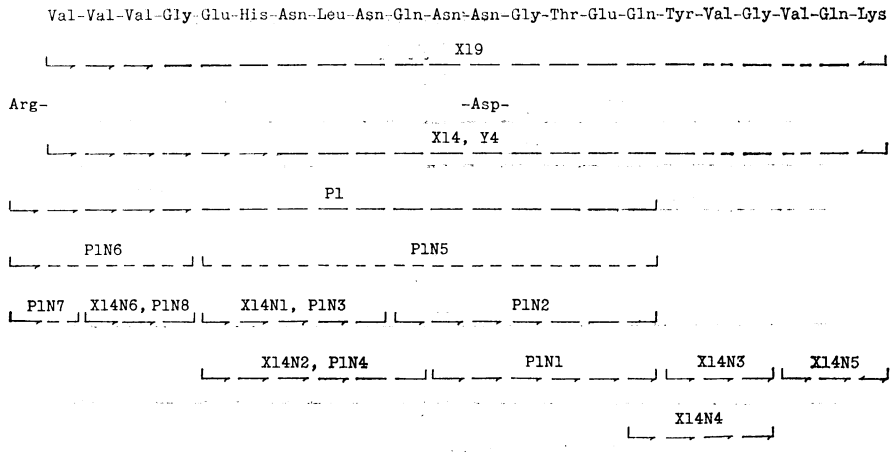


Fig. 11. *Sequence of peptide T6*

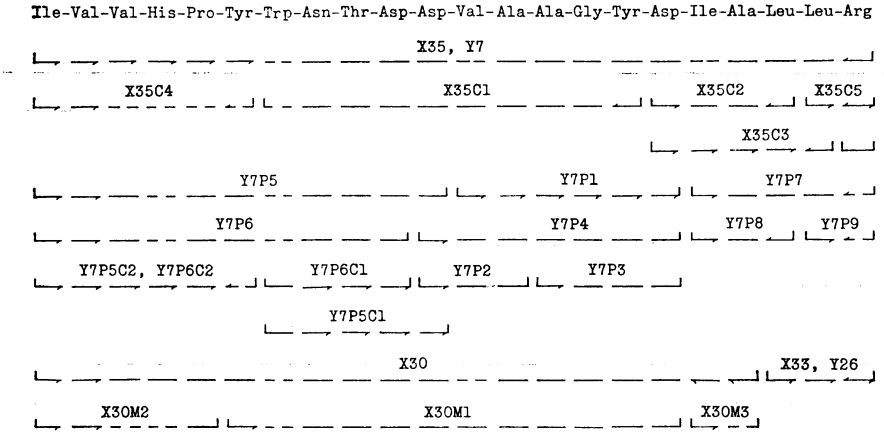


Fig. 12. *Sequence of peptide T7*

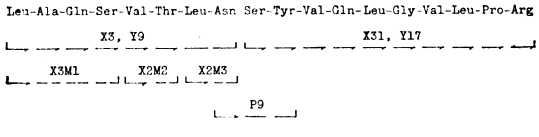


Fig. 13. *Sequence of peptide T8*

Brown *et al.* (1967), who reported the *C*-terminal sequence of their peptic active-centre peptide to be -Cys-Leu-Val-Asn-Gln-Tyr, lacking the glycine residue found in peptide X5. They found good evidence from a carboxypeptidase digest for a *C*-terminal -Gln-Tyr sequence, but their assignment of the -Val-Asn- sequence adjacent to this glutamine was made without any direct sequence evidence, the omission of the glycine residue being based upon the results of a single analysis of an overnight hydrolysate of the complete 22-residue peptic peptide, which contained five other glycine residues. Their evidence for the correct allocation of the amide groups in this sequence was also inconclusive. Evidence that the correct sequence of peptide X5 is Leu-Val-Asp-Gly-Gln-Tyr was given by peptide Y5. This peptide and

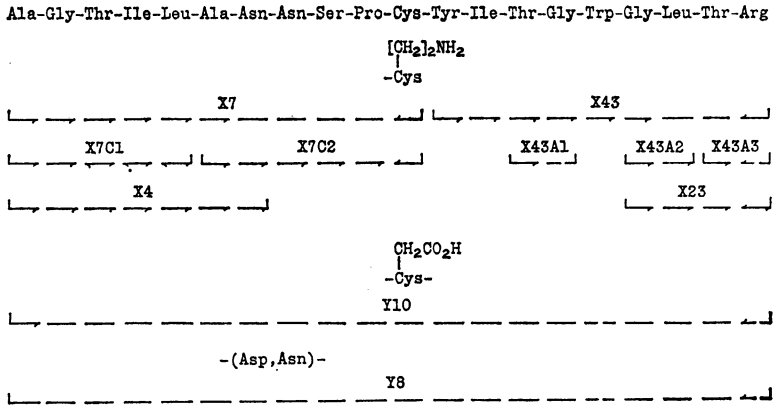


Fig. 14. *Sequence of peptide T9*

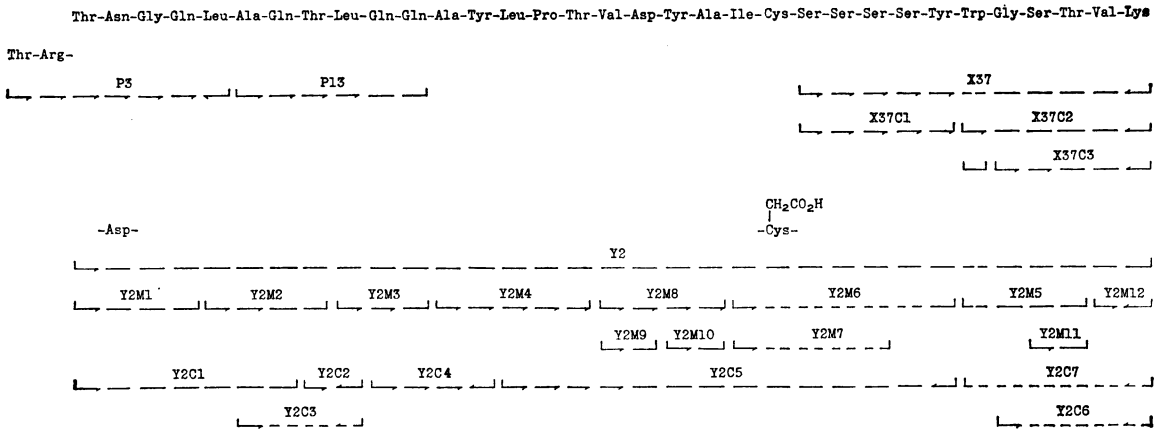


Fig. 15. *Sequence of peptide T10*

peptide Y3 were identical, except that peptide Y5 had one less negative charge and was present in smaller yield. It was suspected that the accidental deamidation of an -Asn-Gly- sequence had resulted in peptide Y3 being the major product. Peptide Y5 was therefore subjected to a chymotryptic digestion, which yielded, among the other expected fragments, peptides Y5C1 and Y5C2. The only difference between these (see Fig. 17) was that the one in lowest yield, peptide Y5C2, had a mobility indicating the presence of two amide groups, as in the parent peptide, whereas peptide Y2C1 had an extra negative charge, suggesting that even during the relatively mild treatment of chymotryptic digestion at pH8.0 followed by electrophoresis at pH6.5, conditions under which a glutamine residue would not be deamidated, the majority of this -Asn-Gly- sequence had deamidated to -Asp-Gly-. These studies also provided the additional evidence for the *C*-terminal position of the tyrosine residue required to determine completely the sequence of peptide X5. As the sequence Ala-Val-His-Gly-Val-Thr-Ser-Phe, not found among the products of digest X, was already known from peptide P2, the sequence of the chymotryptic fragment of peptide Y5 identical with P2, which was present in good yield in this chymotryptic digest, was not determined. Another chymotryptic peptide from this digest, peptide Y5C3, was shown

unequivocally to contain glutamine rather than glutamic acid, by analysis of the products of its digestion by aminopeptidase-M.

From these studies on peptides X5, X6, X10, X24, Y3, Y5 and P2, and from the corrected results of Brown *et al.* (1967) (see above and Fig. 1), the sequence of peptide T12 was completely determined, and is shown in Fig. 17.

Peptide Y1 was a minor peptide identical with Y3 except that it possessed one extra negative charge at pH6.5, probably produced by the partial breakdown of the di-isopropylphosphoryl group attached to the active-centre serine residue to monoisopropylphosphoryl-serine.

*Peptide T*13 (*residues* 218–223). Peptides X15, X21 and Y20 were the expected peptides from this region. Their sequences, and the sequence of peptide T13 derived from them, are shown in Fig. 18.

*Peptide T*14 (*residues* 224–230). Peptides X42 and Y18 were identical and represented the complete tryptic peptide, whereas the identical pairs of peptides X28 and Y19, and X25 and Y27, were the results of accidental hydrolysis in the two digests at the chymotrypsin-sensitive -Phe-Thr- bond. Sequence analysis of the peptides from digest X proved the sequence of peptide T14 shown in Fig. 19. These studies on peptides T13 and T14 confirm the sequence of residues 216–228 determined by Brown *et al.* (1967) (see Fig. 21).

*Peptide T*15 (*residues* 231–245). This tryptic peptide

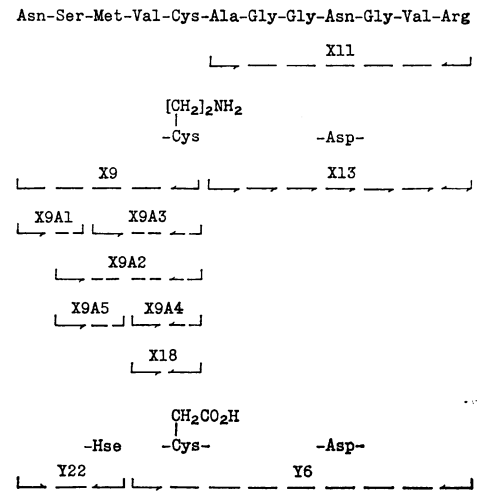


Fig. 16. *Sequence of peptide T*11

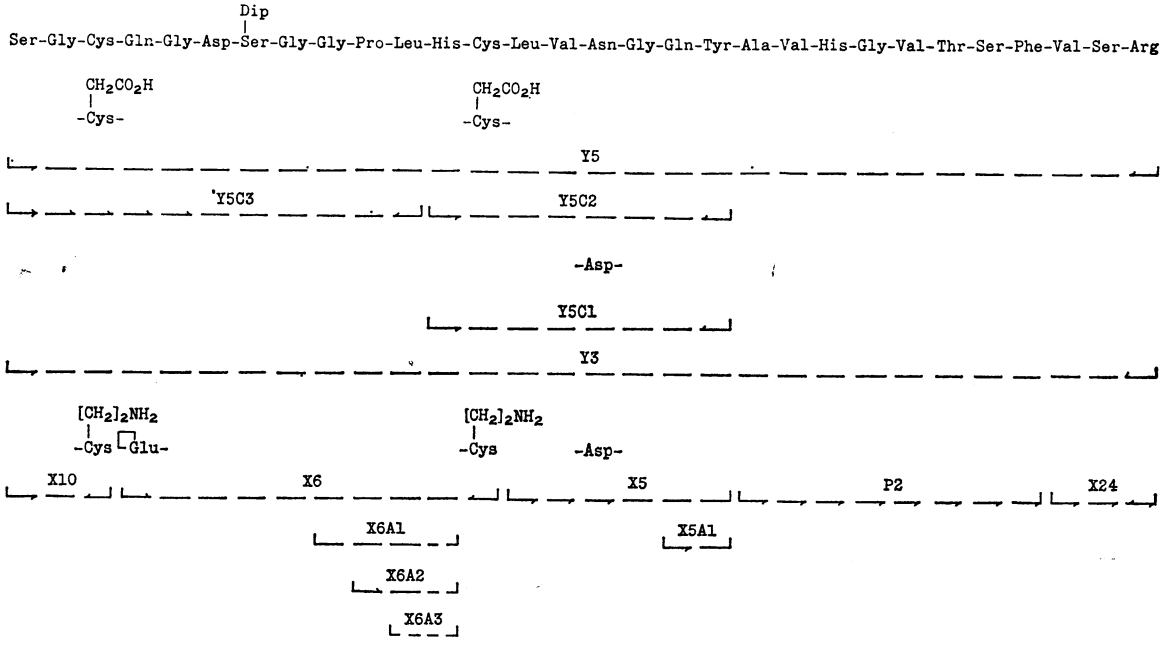


Fig. 17. *Sequence of peptide T*12

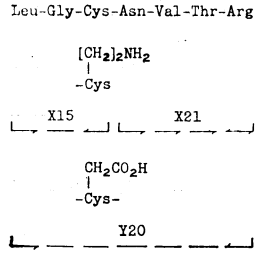


Fig. 18. *Sequence of peptide T13*

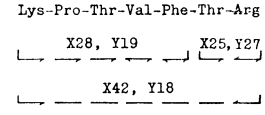


Fig. 19. *Sequence of peptide T14*

was insoluble and thus not present among the soluble products of digests X and Y. Instead it was prepared in pure form as an insoluble 'tryptic core' peptide by thorough tryptic digestion of the *C*-terminal cyanogen bromide fragment (residues 181–245) of [^{14}C]carboxymethylated Dip-elastase; it was named peptide Y30 and was subjected to peptic digestion as described under 'Methods'. Because of the low specificity of pepsin, the peptic digestion products of peptide Y30 were a set of overlapping peptides, whose compositions, mobilities and sequence analyses (fully described in the Supplementary Publication) enabled the complete sequence of peptide T15 to be unambiguously determined (Fig. 20).

Discussion

Complete amino acid sequence of elastase

The sequence information described above permits the unambiguous determination of the complete amino acid sequence of elastase. The corrected published sequences of Smillie & Hartley (1966) and Brown *et al.* (1967) (Fig. 1), the peptic peptides shown in Table 3 and the lysine-containing peptides shown in Table 4 all contribute to provide the necessary overlaps for the correct ordering of the 15 'theoretical' tryptic peptides, which make up the elastase molecule, proving this to be a single polypeptide chain of 240 residues. The complete sequence, with these overlaps indicated, is shown in Fig. 21. Each overlap, except those between peptides T5 and T6, T7 and T8, T9 and T10, and T14 and T15, is unique for both tryptic peptides involved. Peptides

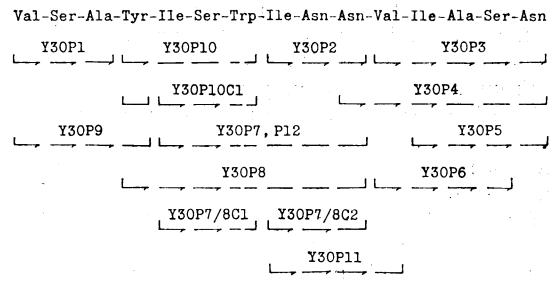
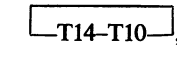


Fig. 20. *Sequence of peptide T15*

T8 and T13 both possess an *N*-terminal leucine residue, but since T13 is unambiguously overlapped with T12, peptide P8, Leu-Arg-Leu, must overlap peptide T7, the only peptide to end -Leu-Arg, with peptide T8. Peptide T5 is the only one from among peptides T5, T9 and T14 not to end -Thr-Arg. It must therefore be overlapped with peptide T6 by peptide P1, leaving peptides T9 and T14 to be overlapped with peptides T10 and T15 by peptides P3 and P4, which both commence Thr-Arg-. The only way in which both these overlaps can be made to give a single linear chain, rather than the nonsense of a shorter linear chain T1–T9-T15 plus a circular peptide



is for peptide T9 to be overlapped with peptide T10 by peptide P3, and peptide T14 to be overlapped with peptide T15 by peptide P4, thus putting T15, the only tryptic peptide to lack a basic residue, in its expected *C*-terminal position.

Confirmation of the sequence

Confirmation of the accuracy of the elastase sequence presented in Fig. 21 is given by evidence from four independent sources.

(*i*) *Amino acid composition of elastase.* The comparison between the expected amino acid composition of elastase calculated by summing the compositions of peptides T1–T15, and the experimentally determined composition obtained from the amino acid analyses of the whole protein, is shown in Table 7. The excellent agreement between the observed and the calculated results indicates that no region of the sequence has been omitted from, and no extra residues mistakenly included in, the elastase sequence shown in Fig. 21.

(*ii*) *Cyanogen bromide fragments.* The small and the medium-sized cyanogen bromide fragments of [^{14}C]-carboxymethylated Dip-elastase were purified as described under 'Methods'. The results of amino

Table 7. *Compositions of the elastase tryptic peptides and of elastase*

The exact amino acid compositions of the 15 'theoretical' tryptic peptides of elastase, T1–T15 (whose completely determined sequences are shown in Figs. 6–20), are tabulated and summed to give the exact amino acid composition of the elastase molecule. This is compared with the best amino acid analysis results obtained from elastase hydrolysates. The quoted analytical results are from this present work, except that of cysteine, which was determined as cysteic acid by Brown *et al.* (1967), and that of tryptophan, which was measured colorimetrically by Gertler & Hofmann (1967) (see Table 1).

	Composition of 'theoretical' tryptic peptides															Composition	Amino acid
Amino acid	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	of elastase	analysis
Lys						1				1				1		3	3.0
His			1	1		1	1					2				6	6.0
Arg	1	1	1	1	1		1	1	1		1	1	1	1		12	11.7
Asp				1			3			1		1				6	} 23.4
Asn		1		1		4	1	1	2	1	2	1	1		3	18	
Thr	1		2	1	1	1	1	1	3	4		1	1	2		19	18.8
Ser		3	3					2	1	5	1	4			3	22	23.2
Glu	1				1	2										4	} 19.0
Gln	1	2		1		3		2		4		2				15	
Pro		1					1	1	1	1		1		1		7	7.2
Gly	2		3			3	1	1	3	2	3	6	1			25	25.6
Ala	1		1	2			3	1	2	3	1	1			2	17	16.8
Cys			1	1					1	1	1	2	1			8	7.9
Val	2			2		5	3	3		2	2	4	1	1	2	27	26.0
Met				1							1					2	1.7
Ile		1	1				2		2	1					3	10	9.9
Leu		1	1		1	1	2	4	2	3		2	1			18	17.9
Tyr		1				1	2	1	1	3		1			1	11	11.3
Phe					1							1		1		3	3.3
Trp		1	1	1			1	1	1						1	7	7.3
Totals	9	12	15	13	5	22	22	18	20	33	12	30	7	7	15	240	240.0

acid analyses and *N*-terminal determinations on these fragments are shown in Table 8 and agree excellently with the compositions and *N*-termini predicted from the sequence studies on elastase described above, proving that the small fragment was derived from the *N*-terminus of elastase (residues 16–53) and that the medium-sized fragment was from the *C*-terminus of the molecule (residues 181–245). Further, comparative 'fingerprinting' experiments on tryptic digests of [^{14}C]carboxymethylated Dip-elastase, unfractionated cyanogen bromide-treated [^{14}C]carboxymethylated Dip-elastase and the two purified cyanogen bromide fragments, with mobility measurements, radioautography and selective peptide stains used to characterize the peptides present in the 'fingerprints', revealed that the small cyanogen bromide fragment yielded only the expected tryptic peptides from residues 16–53 (peptides Y11, Y16, Y21, Y25, Y28 and the homoserine lactone derivatives of Y21 and Y25), and the medium-sized cyanogen bromide fragment yielded only the expected tryptic peptides from residues 181–245 [peptides Y3, Y5, Y6, Y18, Y19, Y20, Y27 and a tryptophan-containing peptide that did not move from the origin during electrophoresis at pH6.5, but which moved with an R_F value of 0.42 in the butan-1-ol–acetic acid–water–pyridine (15:3:10:12, by vol.) system used to develop the second dimension of the 'fingerprints', namely Y30, the 'insoluble' *C*-terminal peptide]. By difference, the other major tryptic peptides of digest Y belonged to the large cyanogen bromide fragment, residues 54–180, which was not

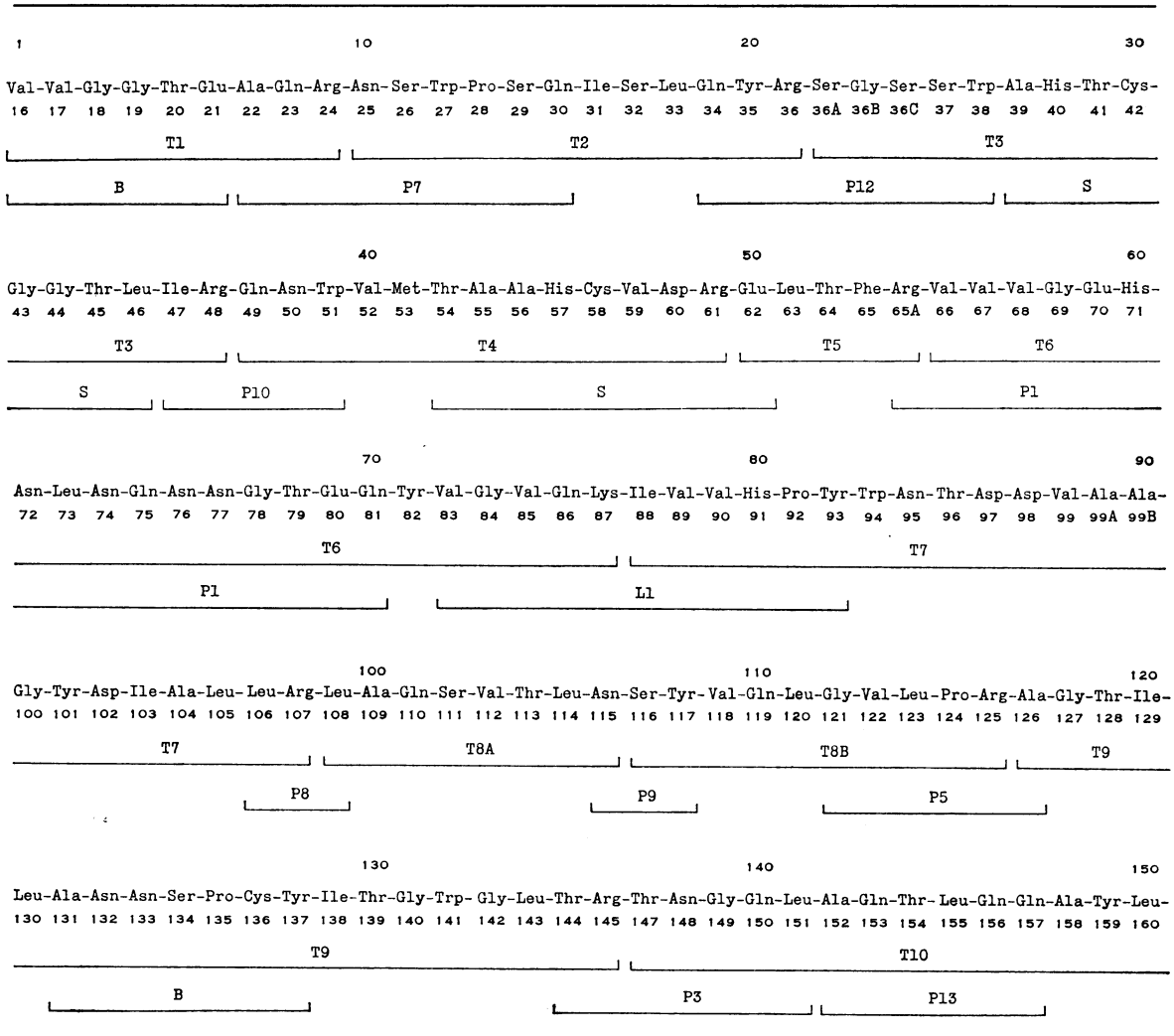


Fig. 21

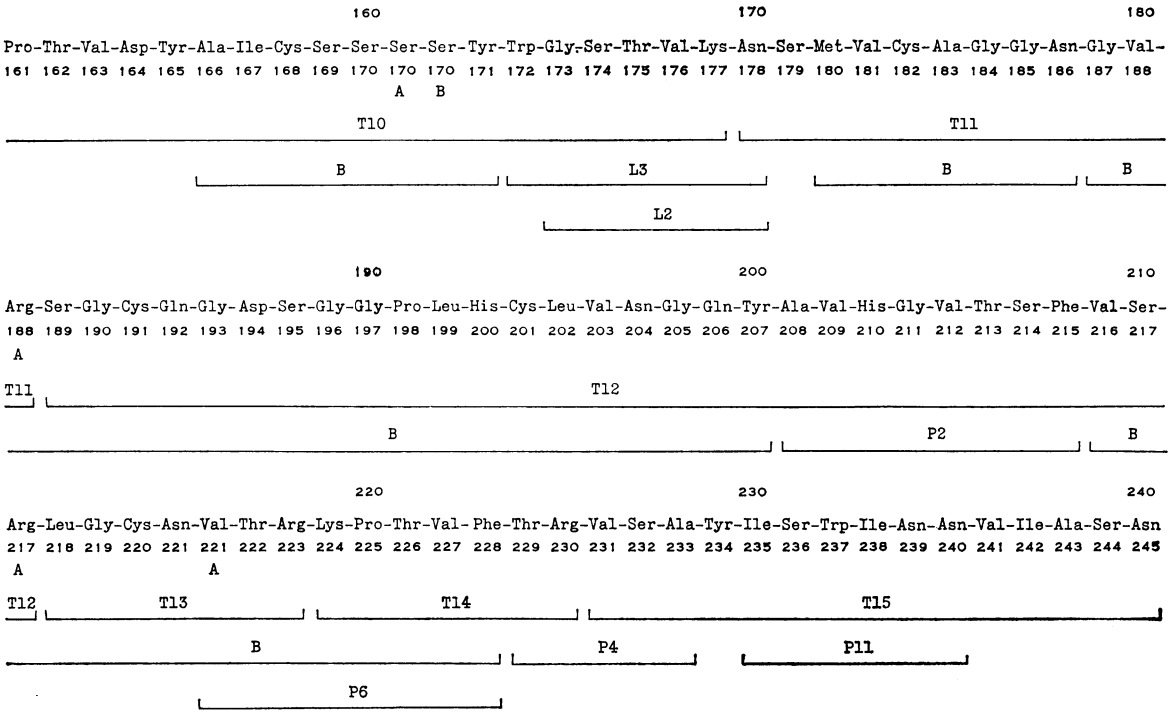


Fig. 21. *Amino acid sequence of porcine pancreatic elastase*

The various overlapping peptides used to determine the primary structure of porcine pancreatic elastase are indicated by labelled lines delineating their positions in the elastase polypeptide chain. Details of how the sequences of these peptides overlap uniquely to reveal the amino acid sequence of the enzyme are given in the text. Key to peptides: T, 'theoretical' tryptic peptides of elastase, whose sequences are shown in Figs. 6–20; P, peptic peptides of elastase from digest P, described in Table 3; L, lysine-containing chymotryptic plus tryptic peptides of elastase from digest L, described in Table 4; S, published sequences of Smillie & Hartley (1966), shown in Fig. 1; B, published sequences of Brown *et al.* (1967) (corrected), shown in Fig. 1. Disulphide bridges: 42–58, 136–201, 168–182 and 191–220. The chymotrypsinogen-A numbering is given below the elastase sequence. Insertions are indicated 36A, 36B etc. (Shotton & Hartley, 1970).

isolated in a pure form. These results reinforce those obtained by comparing the calculated and observed amino acid composition of the whole protein, and confirm the accuracy of the composition, determined from the sequence studies, of the insoluble *C*-terminal tryptic peptide, upon which an amino acid analysis was not performed. In addition, they provide a general check upon the accuracy of the local overlaps.

(*iii*) *Homologies of the elastase sequence with other serine proteinases.* Throughout the elastase sequence, extensive homologies exist with the sequences of bovine trypsin, chymotrypsin-A and chymotrypsin-B, which we have discussed in detail elsewhere (Shotton & Hartley, 1970). These homologies are not fortuitous, but are closely related to the function and three-dimensional structure of the enzymes. They would be lessened by any re-ordering of the elastase sequence, and provide strong confirmation of the peptide overlaps and of much of the detailed sequence.

(*iv*) *Three-dimensional electron density of tosyl elastase.* The crystallographic studies on elastase described by Shotton & Watson (1970*a*,*b*) and Watson *et al.* (1970) have led to the calculation of a map of three-dimensional electron density of tosyl elastase, which contains detailed information about the order of large, small, polar, non-polar and aromatic amino acid residues along the single polypeptide chain, many of which can be identified from their electron density alone. This map has provided the most stringent check on the accuracy of the elastase sequence and overlaps, and has, for instance, eliminated any possibility of repeating -Trp-Trp- sequences, which could possibly have been mistaken for a single

Table 8. *Amino acid compositions and N-terminal residues of the cyanogen bromide fragments of [^{14}C]carboxymethylated Dip-elastase*

In the upper part of the table, the amino acid-analysis results from hydrolysates of the purified *N*-terminal and *C*-terminal cyanogen bromide fragments of elastase, obtained as described under 'Methods,' are tabulated for comparison with the amino acid compositions expected for these fragments from the sequence of elastase shown in Fig. 21. Below, a similar comparison is made between the expected and observed Dns-amino acids and Dns-peptides produced by acid hydrolysis of the cyanogen bromide fragments after reaction with Dns-Cl.

	N-Terminal fragment (residues 16–53)		*C*-Terminal fragment (residues 181–245)	
Amino acid	Analytical results	Expected composition	Analytical results	Expected composition
Lys	0.0	0	1.0	1
His	0.9	1	1.9	2
Hse	+	1	0.0	0
Arg	2.7	3	3.8	4
Cys-CH_2CO_2H	0.8	1	3.6	4
Asp	2.3	2	6.7	7
Thr	2.6*	3	3.8*	4
Ser	5.3*	6	6.5*	7
Glu	4.4	5	2.5	2
Pro	1.0	1	1.8	2
Gly	5.0	5	9.4	10
Ala	2.0	2	4.3	4
Val	2.0*	3	8.4*	10
Met	0.0	0	0.0	0
Ile	1.5*	2	2.6*	3
Leu	2.1	2	3.0	3
Tyr	1.2	1	2.0	2
Phe	0.0	0	2.0	2
Trp	+	3	+	1
	Dansylation result	Expected derivatives	Dansylation result	Expected derivatives
Present:	Dns-Val (+)	Dns-Val	Dns-Val (+)	Dns-Val
	Dns-Val-Val (+)	Dns-Val-Val	ε-Dns-Lys (trace)	ε-Dns-Lys
	O-Dns-Tyr (+)	*O*-Dns-Tyr	*O*-Dns-Tyr (+)	*O*-Dns-Tyr
Absent:	ε-Dns-Lys		Dns-Val-Val	

* Uncorrected result after 18h hydrolysis.

tryptophan residue in the chemical studies. By using the sequence data shown in Fig. 21, a model of the elastase molecule has been built that satisfactorily accounts for all the electron density in the map, there being no point at which the crystallographic and sequence evidence conflict.

The data in this paper thus establish the amino acid sequence of porcine pancreatic elastase to be that which is shown in Fig. 21. For descriptions and discussion of the sequence homologies and similarities of three-dimensional structure between this enzyme and other serine proteinases, reference should be made to the papers mentioned earlier in this section, and to Shotton (1971) and Hartley & Shotton (1971).

We thank Mr. R. Jakes for excellent technical assistance. D. M. S. acknowledges with gratitude the financial support of a Medical Research Council Scholarship and a Beit Memorial Medical Research Fellowship received during the course of this work.

References

Acher, R., Fromageot, C. & Jutisz, M. (1950) *Biochim. Biophys. Acta* **5**, 81–88

Ambler, R. P. (1963) *Biochem. J.* **89**, 349–378

Ambler, R. P. & Brown, L. H. (1967) *Biochem. J.* **104**, 784–825

Bock, R. M. & Ling, N.-S. (1954) *Anal. Chem.* **26**, 1543–1546

Brown, J. R. & Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228

Brown, J. R., Kauffman, D. L. & Hartley, B. S. (1967) *Biochem. J.* **103**, 497–507
Chaplin, H., Cohen, S. & Press, E. M. (1965) *Biochem. J.* **95**, 256–261
Cole, R. D., Raftery, M. A. & Plapp, B. V. (1967) *J. Biol. Chem.* **242**, 265–270
Dalgliesh, C. E. (1952) *Biochem. J.* **52**, 3–14
Erlanger, B. F., Cooper, A. G. & Cohen, W. (1966) *Biochemistry* **5**, 190–196
Fraenkel-Conrat, H., Harris, J. I. & Levy, A. L. (1956) *Methods Biochem. Anal.* **2**, 359–425
Gertler, A. & Hofmann, T. (1967) *J. Biol. Chem.* **242**, 2522–2527
Gray, W. R. (1967*a*) *Methods Enzymol.* **11**, 139–150
Gray, W. R. (1967*b*) *Methods Enzymol.* **11**, 469–475
Gray, W. R. & Hartley, B. S. (1963*a*) *Biochem. J.* **89**, 59P
Gray, W. R. & Hartley, B. S. (1963*b*) *Biochem. J.* **89**, 379–380
Gross, E. & Witkop, B. (1962) *J. Biol. Chem.* **237**, 1856–1860
Harley-Mason, J. & Archer, A. A. P. G. (1958) *Biochem. J.* **69**, 60P
Hartley, B. S. (1960) *Annu. Rev. Biochem.* **29**, 45–72
Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822
Hartley, B. S. & Shotton, D. M. (1971) *Enzymes*, 3rd edn., **3**, 323–373
Hartley, B. S., Naughton, M. A. & Sanger, F. (1959) *Biochim. Biophys. Acta* **34**, 243–244
Heilmann, J., Barollier, J. & Watzke, E. (1957) *Hoppe-Seyler's Z. Physiol. Chem.* **309**, 219–220
Irreverre, F. (1965) *Biochim. Biophys. Acta* **111**, 551–552
Jepson, J. B. & Smith, I. (1953) *Nature (London)* **172**, 1100–1101
Kaplan, H., Stevenson, K. J. & Hartley, B. S. (1969) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **28**, 533
Kaplan, H., Stevenson, K. J. & Hartley, B. S. (1971) *Biochem. J.* **124**, 289–298
Levy, L. A. & Chung, D. (1953) *Anal. Chem.* **25**, 396–399
Milstein, C. (1966*a*) *Biochem. J.* **101**, 338–351
Milstein, C. (1966*b*) *Biochem. J.* **101**, 352–368
Naughton, M. A. & Sanger, F. (1961) *Biochem. J.* **78**, 156–163
Naughton, M. A., Sanger, F., Hartley, B. S. & Shaw, D. C. (1960) *Biochem. J.* **77**, 149–163
Offord, R. E. (1966) *Nature (London)* **211**, 591–593
Partridge, S. M. (1948) *Biochem. J.* **42**, 238–248
Raftery, M. A. & Cole, R. D. (1966) *J. Biol. Chem.* **241**, 3457–3461
Reindel, F. & Hoppé, W. (1954) *Chem. Ber.* **87**, 1103–1107
Shotton, D. M. (1970) *Methods Enzymol.* **19**, 113–140
Shotton, D. M. (1971) *Proc. Int. Res. Conf. Proteinase Inhibitors, 1st*, 1970 47–55
Shotton, D. M. & Hartley, B. S. (1970) *Nature (London)* **225**, 802–806
Shotton, D. M. & Watson, H. C. (1970*a*) *Phil. Trans. Roy. Soc. London, Ser. B* **257**, 111–118
Shotton, D. M. & Watson, H. C. (1970*b*) *Nature (London)* **225**, 811–816
Smillie, L. B. & Hartley, B. S. (1966) *Biochem. J.* **101**, 232–241
Spackman, D. H. (1963) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **22**, 244
Spackman, D. H., Stein, W. H. & Moore, S. (1958) *Anal. Chem.* **30**, 1190–1206
Waley, S. G. & Watson, J. (1953) *Biochem. J.* **55**, 328–337
Watson, H. C., Shotton, D. M., Cox, J. M. & Muirhead, H. (1970) *Nature (London)* **225**, 806–811