The Amino Acid Sequence of *Pseudomonas fluorescens* Azurin

By R. P. AMBLER* and LARAMIE H. BROWN†

*Department of Chemistry, University of California at San Diego, La Jolla, Calif., U.S.A.
and Medical Research Council Laboratory of Molecular Biology, Cambridge*

The amino acid sequence of *Pseudomonas fluorescens* azurin has been determined. The protein consists of a single peptide chain of 128 residues. There is one intra-chain disulphide bridge. The sequence was determined by isolation of the soluble tryptic peptides, and by exhaustive examination of the products of chymotryptic and peptic digestion. The sequence has been confirmed by the purification and analysis of the seven fragments obtained by cyanogen bromide treatment of the protein.

Several naturally occurring proteins are known that contain stoicheiometric amounts of copper. Such proteins include several phenol oxidases (EC 1.10.3.–), the serum protein caeruloplasmin, plastocyanin from photosynthetic plants and the bacterial azurins. All these proteins have been prepared in a highly purified state, and all are then found to have a very pronounced blue colour. Although early reports (Keilin & Mann, 1940) suggested that in some cases the colour might not depend on the presence of copper, more recent evidence (e.g. Nakamura, 1958) suggests that the colour is due to copper atoms being bound in a specific manner to a polypeptide chain.

The blue colour is very intense, with absorption at the maximum of a broad peak centred at about 600 mμ about 80 times as great as that given by the same concentration of copper as cuprammonium ion. The colour disappears reversibly if a reducing agent such as sodium dithionite is added, or irreversibly if the protein is denatured. With some proteins the copper can be removed (e.g. by dialysis against cyanide; Yamanaka, Kijimoto & Okunuki, 1963), whereupon the colour disappears, to reappear when copper, and only copper (as Cu^{2+} ion), is added back to the system.

The status of the copper in caeruloplasmin and in azurin from two different species of bacteria has been investigated by physical methods (Broman, Malmström, Aasa & Vänngård, 1963; Mason, 1963). Both the azurins and the plasma protein had the same very characteristic electron-spin-resonance spectrum, despite the greater complexity of the caeruloplasmin molecule, which possessed several copper atoms, of both valency states, in each sedimentable unit.

Verhoeven & Takeda (1956) reported that *Pseudomonas aeruginosa* contained a blue protein. It was purified by Horio (1958a), and its physical and chemical properties were investigated (Horio, 1958b; Coval, Horio & Kamen, 1961). Sutherland & Wilkinson (1963) showed that similar proteins are widely distributed in the genera *Pseudomonas*, *Bordetella* and *Alcaligenes*, and proposed the name 'azurin' for this class of protein. It has been shown (R. P. Ambler, unpublished work) that the azurins from all three genera have many similarities in their primary structure.

The azurins are the simplest of all the copper proteins so far discovered and are of low molecular weight (about 14000), containing only 1 copper atom/mol., and (unlike laccase or caeruloplasmin) containing no carbohydrate. The proteins can readily be prepared in quantity (e.g. Ambler, 1963a) and so it was decided to investigate the chemical structure of an azurin as an approach to finding out how copper is bound to polypeptide in copper-containing blue proteins.

A preliminary account of this work has been published (Ambler & Brown, 1964).

MATERIALS

Chymotrypsin, carboxypeptidase A (di-isopropyl phosphorofluoridate-treated), carboxypeptidase B (pig), soya-bean trypsin inhibitor and leucine aminopeptidase were obtained from Worthington Biochemical Corp., Freehold, N.J., U.S.A. The trypsin used was from the same source; a single batch (TRSF 813/814), known to be low in chymo-

* Present address: Department of Molecular Biology, University of Edinburgh.

† Present address: 199 Yarkerdale Drive, Rochester, N.Y., U.S.A.

tryptic-like activity, was used throughout. Pronase (Pronase-P, B grade) was obtained from California Corp. for Biochemical Research, Los Angeles, Calif., U.S.A. Subtilopeptidase B ('bacterial trypsin') was obtained from Novo Terapeutisk Laboratorium, Copenhagen, Denmark. Bovine carboxypeptidase B was a gift from Dr E. Wintersberger.

CM-cellulose (grade CM-70) and DEAE-cellulose (grade DE-50) were from Whatman. 1-Dimethylaminonaphthalene-5-sulphonyl chloride was obtained from British Drug Houses Ltd., Poole, Dorset.

Synthetic dipeptides used were obtained from Koch-Light Laboratories Ltd., Colnbrook, Bucks., and Yeda, P.O. Box 26, Rehovoth, Israel.

Acetone-dried cells of *Pseudomonas fluorescens* were prepared from strain P6009/1 as described by Ambler (1963a).

METHODS

Preparation of azurin. The method used was very similar to that described by Ambler (1963a). Up to step 5 (concentration by adsorption on to and elution from CM-cellulose) the procedure was exactly as described. The protein solution was then equilibrated with 0·01M-tris–HCl buffer, pH 8·7, by gel filtration through Sephadex G-25. At this stage the yield from 300g. of acetone-dried organisms was 200mg. of protein, which was in 50ml. of buffer. The protein solution was added to a column (10cm. × 2cm. diam.) of DEAE-cellulose in 0·01M-tris–HCl buffer, pH 8·7. The column was eluted with 0·05M-tris–HCl buffer, pH 8·7. The blue eluate was then made 75% saturated with $(NH_4)_2SO_4$ (at 20° and pH 6·5), and the very small amount of red–brown precipitate that formed was removed by centrifugation at 10000g for 10min. The blue solution (containing about 2mg. of protein/ml.) was then saturated with $(NH_4)_2SO_4$; a blue precipitate formed very slowly, but after 16hr. at 20° it could be collected by centrifugation to leave an almost colourless supernatant. The precipitate was dissolved in water and $(NH_4)_2SO_4$ was removed by gel filtration through Sephadex G-25 in 0·05M-ammonium acetate buffer, pH 3·9. The protein was kept as a 2·5% (w/v) solution in this volatile buffer. At 4° the protein solution was stable.

Removal of copper. The protein solution (in 0·05M-ammonium acetate buffer, pH 3·9) was cooled in an ice bath, and an equal volume of 10% (w/v) trichloroacetic acid added. A white precipitate formed slowly, and the blue colour faded, disappearing within about 10min. After centrifugation at 1000g for 5min., the supernatant was discarded and the white precipitate was dissolved in 0·1N-NH₃ solution. The protein was again precipitated at 0° by the addition of 10% (w/v) trichloroacetic acid, and the precipitate obtained by centrifugation was washed four times with cold acetone.

Performic acid oxidation. The method of Hirs (1956) was used. Oxidation was for 4hr. at −15°, and the reaction mixture was diluted with 25vol. of ice-cold water before freeze-drying.

Starch-gel electrophoresis. Horizontal gels were prepared by the method of Smithies (1959). The gel buffers used were: pH 6·0, 0·02M-sodium cacodylate adjusted with HCl; pH 8·5, 0·02M-boric acid adjusted with NaOH. Both buffers were made up to contain 1mM-$K_3Fe(CN)_6$. The protein was visible before staining as a blue band, but the

band was enhanced and fixed by staining with Naphthalene Black.

Amino acid analysis. The method of Spackman, Stein & Moore (1958) was used, with Beckman–Spinco model MS amino acid analysers. Some were fitted with the accelerated modification of Spackman (1963) and a high-sensitivity attachment (Evans Electroselenium Ltd., Halstead, Essex). Sample loadings were 0·75mg. of protein/column or 0·03–0·10μmole of peptide/column.

Samples for hydrolysis were evaporated to dryness in 16mm. × 125mm. Pyrex test tubes and then dissolved in 0·5ml. of freshly diluted 6N-HCl. Analytical-grade acid that had been stored in the dark was used. and bottles were discarded a month after opening. A constriction was formed on the tube and the contents were frozen in acetone–solid CO_2. The tubes were evacuated to a pressure of 0·1–0·05mm. Hg while the acid thawed, and after 5min. evacuation the tubes were sealed. The tubes were heated at 105±0·5° for 12–120hr. in an oven fitted with forced circulation and a temperature recorder (Laboratory Thermal Equipment Ltd., Oldham, Lancs.). After hydrolysis the acid was removed at room temperature in a desiccator evacuated to about 0·1mm. Hg pressure. As soon as the acid had evaporated, the samples were removed from the vacuum and stored stoppered at 4°.

Unless otherwise stated, peptide samples were hydrolysed for 24hr. To compensate for destruction during hydrolysis, recoveries of serine were increased by 6% and those of threonine by 3%.

Tryptophan determination. The tryptophan content of the whole protein was determined by the Harrison & Hoffman (1961) modification of the method of Spies & Chambers (1948), the protein being digested with pepsin before analysis. Tryptophan in peptides was determined spectrophotometrically (after removal of contaminating ultraviolet-absorbing material by gel filtration), by amino acid analysis after exopeptidase digestion and by the Spies & Chambers (1948) colorimetric method.

Amide content. The amide content of the whole protein was determined from the difference in NH₃ content between an unhydrolysed sample of protein and a portion that had been hydrolysed for 1hr. at 100° in 2N-HCl. The NH₃ content was measured by the automatic amino acid analyser.

Digestion of protein and peptides with proteolytic enzymes. The methods described by Ambler (1963b) were used. For primary digests 8–12μmoles of protein were used.

Partial acid hydrolysis. The peptide (0·1–0·5μmole) was evaporated to dryness in a 12mm. × 75mm. Pyrex tube, dissolved in 0·3ml. of 6N-HCl, evacuated (0·5–0·05mm. Hg) and sealed. The tube was then heated in a boiling-water bath for 10–45min. The time of hydrolysis was chosen after taking account of the amino acid composition of the peptide to be treated.

High-voltage paper electrophoresis. The conditions used for the separation of peptides and amino acids were as described by Ambler (1963b). DNS*-amino acids were

* Abbreviations: DNS, 1-dimethylaminonaphthalene-5-sulphonyl; PTC, phenyl isothiocyanate; LAP, leucine aminopeptidase; CP, carboxypeptidase (A); CPB, carboxypeptidase B; Glx and Asx, residues where the evidence given does not differentiate between Glu or Gln and between Asp or Asn respectively; Hoser, homoserine; MetSO₂, methionine sulphone; CySO₃H, cysteic acid.

separated by using a flat-plate apparatus (The Locarte Co., London, S.W.3), Whatman 3MM paper and pyridine–acetic acid–water (3:6:500, by vol.) buffer, pH 4·38 (Gray & Hartley, 1963a,b; Gray, 1967); electrophoresis was continued for 2–3 hr. at 80 v/cm. External coloured markers (Milstein, 1966) were used to check the lengths of runs.

The methods used for location and identification of peptides on paper were as described by Ambler (1963b); peptides containing N-terminal proline were detected with isatin (Acher, Fromageot & Jutisz, 1950).

Paper chromatography. The conditions used for separation of peptides were as described by Ambler (1963b). For some difficult separations, multiple development of the paper was used. For reasons discussed by Ambler (1963b), chromatography was never used as the last purification step for a peptide.

Separation of peptides by gel filtration. Up to 20 μmoles (280 mg.) of peptides were separated on columns (120 cm. × 1·5 cm. diam.) of Sephadex G-25 (fine bead form), normally equilibrated with 0·1 N-NH₃ solution. The columns flowed under gravity at flow rates of about 15 ml./hr. and 3·5 ml. fractions were collected. Peptides were detected by high-voltage paper electrophoresis of portions of each fraction (e.g. see Fig. 3).

Cyanogen bromide cleavage (Gross & Witkop, 1961). Unoxidized apoprotein (150 mg.) was dissolved in 6 ml. of 50% (v/v) formic acid and treated with an equal weight of CNBr for 20 hr. at 20°. At the end of this time all the reaction products were still in solution, and samples on analysis showed that less than 5% of the methionine remained unconverted. The reaction mixture was diluted with 10 vol. of water and freeze-dried. The residue was dissolved in 2 ml. of 50% (v/v) formic acid, and fractionated by gel filtration through a column (120 cm. × 1·5 cm.) of Sephadex G-25 equilibrated with 50% (v/v) formic acid. Most of the peptide material was little retarded, and this first fraction was freeze-dried (after dilution) and oxidized with performic acid. The peptides in the later fractions were further purified by paper electrophoresis. The oxidized material was fractionated by gel filtration through a column (120 cm. × 1·5 cm.) of Sephadex G-50 equilibrated with 50% (v/v) formic acid, followed by paper electrophoresis. If the reaction mixture was oxidized before gel filtration, the Br⁻ ions present formed bromine, which completely destroyed all tyrosine and histidine.

Most peptides were present in two electrophoretically distinct but interconvertible forms. The more acidic form, containing C-terminal homoserine, could be converted into the homoserine lactone form by treating with anhydrous trifluoroacetic acid at 20° for 1 hr. The lactone form could be slowly converted into the homoserine form by treatment with 2 N-NH₃ solution at 37°, but conversion was not complete even after 16 hr.

Before amino acid analysis, acid hydrolysates of the fractions obtained by CNBr cleavage were treated with pyridine–acetate buffer, pH 6·5, at 105° for 1 hr., a procedure that converted all homoserine lactone into homoserine. After removal of the volatile buffer, the sample was dissolved in sodium citrate buffer, pH 2·2 (Spackman et al. 1958), and at once applied to the ion-exchange column of the amino acid analyser. Recovery of homoserine, assuming the same colour yield with ninhydrin as serine, but making no allowance for destruction during hydrolysis, was about 90%. The separation of homoserine and glutamic acid was im-

proved by adjusting the pH of the first eluting buffer from 3·28 to 3·20 (Ambler, 1965).

N-Terminal sequence of azurin. The N-terminal amino acid was determined by the DNP method. To investigate the N-terminal sequence, the DNP-protein was prepared from performic acid-oxidized azurin by the second method of Ambler (1963a) and then, while still in solution, digested at pH 8·5 with subtilopeptidase B. The resultant yellow acidic peptides were purified by paper electrophoresis and chromatography, and analysed for amino acid composition (Ambler, 1963a).

C-Terminal sequence of azurin. The methods used are discussed in detail by Ambler (1967). Equal 0·1–0·2 μmole portions of performic acid-oxidized azurin were treated with solutions of CP, CPB or a mixture of the two enzymes (10–100 μg.) in 0·2 M-N-ethylmorpholine–acetate buffer, pH 8·5, at 37° for 0·5–12 hr.; the digests were evaporated to dryness, dissolved in 0·2 M-sodium citrate buffer, pH 2·2, and the free amino acids analysed directly, without removal of the protein, on the automatic amino acid analyser. The concentration of azurin was determined by amino acid analysis of one portion of the oxidized protein after total acid hydrolysis.

N-Terminal groups of peptides. The DNS method (Gray & Hartley, 1963a,b; Gray, 1967) was used both for the determination of peptide N-termini and for the identification of the new N-terminus exposed in a peptide after the removal of an amino acid residue by the PTC method. No attempt was made to use the method at the lower limit of its sensitivity. The peptide (0·005–0·05 μmole) was evaporated to dryness in a 28 mm. × 3·5 mm. Durham tube, 0·01 ml. of 1% (w/v) NaHCO₃ was added and the mixture was dried *in vacuo*. Then 0·01 ml. of water and 0·01 ml. of 0·5% (w/v) DNS chloride in acetone were added, and the tube was covered and left at room temperature for 4 hr., after which time the mixture was no longer yellow (owing to the reaction or hydrolysis of all the DNS chloride). The liquid was removed under vacuum, 0·05 ml. of 6 N-HCl added and the tube sealed. The DNS-peptide was hydrolysed at 105°, normally for 12 hr., but for longer times (up to 120 hr.) if initial experiments showed the presence of a stable N-terminal peptide. The hydrolysates were dried *in vacuo* and the DNS-amino acids identified by high-voltage paper electrophoresis at pH 4·38.

Some DNS-amino acids could not be identified unequivocally by electrophoresis at pH 4·38 alone. Ambiguous fluorescent spots were cut out and sewn on to another sheet of paper for identification by another system (Milstein, 1966). The identity of DNS-Val, DNS-Phe, DNS-Leu, DNS-Ile and DNS-Pro were confirmed by paper chromatography in light petroleum (b.p. 40–60°)–acetic acid–water (10:1:1, by vol.) (Boulton & Bush, 1964). In this system the acid-stable DNS-dipeptides (DNS-Val-Val, DNS-Ile-Val, DNS-Val-Ile and DNS-Ile-Ile) have much lower R_F values than the DNS-amino acids that are released by prolonged hydrolysis. The identities of DNS-Ala, DNS-Gly and DNS-Ser were confirmed by high-voltage paper electrophoresis at pH 2·0, and of DNS-Arg, α-DNS-His and α- and ε-DNS-Lys by high-voltage paper electrophoresis at pH 9·0. The apparatus of Michl (1951) was used; at pH 2·0 and 9·0 the DNS-amino acids concerned were not lost from the electrophoresis paper into the organic solvent used for cooling.

Peptides were treated with PTC under the conditions described by Konigsberg & Hill (1962). No attempt was

made to isolate phenylthiohydantoins. After completion of the two stages of the reaction the residual peptide was purified by high-voltage paper electrophoresis before the determination of its N-terminal amino acid and amino acid composition.

The conditions for use of LAP were as described by Ambler (1963b).

C-*Terminal groups of peptides.* CP and CPB were used as described by Ambler (1963b, 1967).

RESULTS

Nomenclature and abbreviations

Peptide nomenclature. Peptides are identified by a capital letter (indicating the method used for the primary degradation of the protein) followed by an arabic numeral (indicating the relative mobility of the peptide on electrophoresis at pH 6·5, the most basic peptide having the lowest number). The series of numbers are not always complete. Suffixes 'a', 'b' etc. are used for peptides with similar electrophoretic mobilities at pH 6·5. When the primary peptide has been further degraded, the secondary peptides formed are identified by the same system, with a second capital letter to indicate the second method of degradation and a number for each of the products.

The letters used are as follows. For enzymic digests: T, trypsin; C, chymotrypsin; P, pepsin; S, subtilisin (subtilopeptidase B); Q, Pronase; L, LAP; K, CP; B, CPB. For chemical degradation: X, cleavage with cyanogen bromide; A, partial hydrolysis with acid; F (with a number indicating the number of steps), the residual peptide after PTC degradation.

Abbreviations used in the Tables, Figures and text. Most of the results from which the sequences of peptides were deduced are presented in Tables; these results are then summarized in Figures, in which the sequence of a particular region of polypeptide chain is derived.

The electrophoretic mobility at pH 6·5 (m) was measured relative to lysine for basic peptides ($+$ sign) and to aspartic acid for acidic peptides ($-$ sign), the true origin being taken as the position to which the monoaminomonocarboxylic acids moved. Reproducibility is about $\pm 5\%$, except for histidine-containing peptides, which have more variable mobilities, as the pK of imidazole groups is close to 6·5. The few other peptide types with pK values near 6·5 are considered in the Discussion section. The electrophoretic mobilities at pH 3·5 and 2·0 (not expressed numerically) are referred to as m' and m''. Unless otherwise stated, chromatographic mobility (R_F) refers to the system butan-1-ol–acetic acid–water (3:1:1, by vol.).

Yields of peptides are given in two forms. In the Tables summarizing the properties of the primary peptides from digests of azurin, the yields of the final pure peptides are expressed as percentages (μmole/μmole) of the protein digested, with no corrections at all for losses, known or unknown, in the multistage purification. In the other Tables relative yields are given, to provide a rough indication of the amounts of peptide relative to each other produced in the secondary digest. Relative yields were estimated visually by comparison of the colour intensities given by total acid hydrolysates of samples of the peptides; the hydrolysates were separated by electrophoresis at pH 2·0 and stained with ninhydrin.

The purification procedure is abbreviated as follows. Gel filtration: G, on Sephadex G-25; G′, on Sephadex G-50. Paper electrophoresis: 6, at pH 6·5; 3, at pH 3·5; 2, at pH 2·0; 9, at pH 9. Paper chromatography: B (with superscript number), butan-1-ol–acetic acid–water (3:1:1, by vol.); the superscript shows the number of times the paper was developed and dried before staining. The symbols are arranged in the order the purification methods were used in preparing the peptides.

The amino acid composition, after total acid hydrolysis, was measured either with an automatic amino acid analyser (in which case the results are expressed numerically, the calculation being made on the basis that the average of the amounts of the amino acids present in other than trace amounts was integral), or by paper electrophoresis at pH 2·0, followed by staining with ninhydrin. In the latter case the intensities of the spots were estimated visually, and expressed as the number of '+' signs. Impurities present in less than 5% amounts are not shown.

In the Figures, peptides represented by a solid line ($\longleftrightarrow$) were analysed for amino acid composition quantitatively, whereas those presented by a broken line ($\leftarrow\!\!-\!\!-\!\!\rightarrow$) were only analysed qualitatively. Residues removed completely, partially or in trace amounts by endopeptidases or PTC degradation are marked by arrows ($\leftarrow$, $-\rightarrow$, $\cdots\rightarrow$) and the appropriate single-letter abbreviation for the method involved (L, B, K, F). N-Terminal residues identified by the DNS method are marked by a second line under the left of the line representing the peptide (e.g. $\underleftarrow{\quad}\rightarrow$) or by underlining the residue.

Properties of azurin

The spectrum of the final azurin preparation was similar to that reported by Horio, Sekuzu, Higashi & Okunuki (1961), with the same very sharp subsidiary maximum at 292 mμ, as well as the principal maxima at 280 and 625 mμ. The best material had spectral purity (E^{ox}_{625}/E_{280}) 0·60, but the spectral purities of preparations judged by the

other criteria given below to be homogenous have been variable, ranging from 0·48 to 0·60. The ultraviolet spectra of all the preparations have been the same, the $E_{1cm}^{1\%}$ value at 280mμ being 6·0. The ultraviolet spectra of solutions of the apoprotein were indistinguishable from those of the native protein, but, on denaturation (by heat, urea or extremes of pH) or proteolysis, the spectrum altered greatly, and the peak at 292mμ disappeared.

Most preparations showed a trace of a Soret band at 415mμ. The amount of cytochrome present was calculated to be always less than 0·002mole/mole, showing that all but a trace of proteins very closely associated with the azurin throughout the purification had been removed.

The protein migrated as a single band in starch-gel electrophoresis at pH 6·0 and pH 8·5. The gels contained 1mM-potassium ferricyanide to maintain the azurin completely in the oxidized form; without this agent the protein bands were always double.

Material of spectral purity (E_{625}^{ox}/E_{280}) 0·5 had a copper content (determined by the method of

Table 1. Amino acid composition of azurin

Experimental details are given in the text.

Amount of amino acid recovered (μmoles)

	Hydrolysis time (hr.)					Ox. (A)*	Ox. (B)*	Blank*	Amide*	Best value†	Best value/ 0·0585†	From sequence
	12	24 (A)	24 (B)	48	96							
Gly	0·654	0·643	0·648	0·655	0·655	0·660	0·657	—	—	0·651	11·1	11
Ala	0·416	0·409	0·415	0·420	0·417	0·416	0·414	—	—	0·415	7·1	7
Val	0·440	0·527	0·537	0·578	0·607	0·530	0·517	—	—	0·622	10·6	10
Leu	0·563	0·582	0·591	0·592	0·588	0·583	0·588	—	—	0·582	9·9	10
Ile	0·182	0·208	0·211	0·226	0·231	0·217	0·211	—	—	0·240	4·1	4
Ser	0·569	0·548	0·538	0·532	0·414	0·542	0·539	—	—	0·582	9·9	10
Thr	0·564	0·583	0·575	0·550	0·521	0·580	0·574	—	—	0·598	10·2	10
Asp	1·079	1·090	1·083	1·090	1·083	1·094	1·111	—	—	1·085	18·6	18
Glu	0·618	0·627	0·623	0·623	0·622	0·634	0·599	—	—	0·623	10·7	10
Phe	0·343	0·348	0·356	0·353	0·357	0·346	0·350	—	—	0·351	6·0	6
Tyr	0·116	0·116	0·121	0·121	0·118	0·105	0·105	—	—	0·118	2·0	2
Trp	—	—	—	—	—	—	—	—	—	—	0·8‡	1
CySO₃H	—	—	—	—	—	0·169	0·163	—	—	0·166	2·8	3
CyS	0·125	§	§	§	§	—	—	—	—	—	—	—
Met	0·350	0·347	0·350	0·346	0·354	0·356‖	0·354‖	—	—	0·349	6·0	6
Pro	0·244	0·240	0·244	0·236	0·233	0·234	0·217	—	—	0·239	4·1	4
Lys	0·655	0·651	0·657	0·650	0·647	0·633	0·636	—	—	0·652	11·2	11
His	0·236	0·238	0·237	0·236	0·234	0·228	0·224	—	—	0·236	1·0	4
Arg	0·057	0·056	0·053	0·057	0·055	0·059	0·060	—	—	0·056	1·0	1
NH₃	0·780	0·834	0·846	0·907	1·047	0·917	0·802	0·044	0·752	0·706	12·1	13
Correction factor¶	(1·00)	0·99	1·00	1·00	1·04	0·95	0·83¶	(1)	(1)			
Total	—	—	—	—	—	—	—	—	—	—	—	128

* Ox.: duplicate samples oxidized with performic acid and then hydrolysed for 24hr. The recoveries were not used for the calculation of 'best values' other than that of cysteic acid. Blank: unhydrolysed sample analysed for NH₃ content. Amide: sample analysed for NH₃ content after hydrolysis with 2N-HCl for 3hr. at 100°.

† Best value: average of 12hr., 24hr., 48hr. and 96hr. recoveries for amino acids other than serine, threonine, valine and isoleucine. The 'best values' for serine and threonine were the recoveries extrapolated to zero time of hydrolysis, and for valine and isoleucine extrapolated to infinite time of hydrolysis. 0·0585: this divisor was obtained by dividing the sums of the 'best values' of phenylalanine, tyrosine, methionine, histidine and arginine by 19.

‡ Determined colorimetrically; see the text.

§ Recoveries not calculated at longer times of hydrolysis owing to asymmetry of peak.

‖ As methionine sulphone.

¶ To minimize the effect of pipetting inaccuracies, a correction was applied to the recoveries before inclusion in the Table. The factor used was the ratio of the sums of the recoveries of all amino acids except serine, threonine, valine, isoleucine and cysteic acid to the sum of these amino acids in the 12hr. hydrolysis sample. The very low correction factor in the 'Ox. (B)' sample was due to the loss of an unknown amount of this sample (estimated at the time to be 'about 10%') in a transfer operation between oxidation and hydrolysis.

Gahler, 1954) of 0·37%, corresponding to 1 atom of copper/mol. of mol.wt. 16000. The carbohydrate content (Devor, 1950) was less than 0·2%, and the nitrogen content (by the micro-Kjeldahl method) 15·8%. Amino acid analysis showed that at least 98% of the dry weight of the protein could be accounted for in the amino acids formed after hydrolysis.

The native protein was completely resistant to proteolysis by trypsin, chymotrypsin and subtilopeptidase B. The apoprotein was readily digested, but, if any copper were present, the apoprotein very rapidly reverted to the resistant blue native form. No other metal has been found that will recombine with the apoprotein to form a coloured complex, but cobalt appears to protect the protein from proteolysis in 1 atom/mol. amounts. The effect of pH on the recombination of apoprotein with copper has been studied spectrophotometrically; recombination was found to be complete and very rapid in the pH range 3·1–9·0, but to occur to only a small extent at pH 2·6 or 10·0.

The azurin can easily be crystallized from 90%-saturated ammonium sulphate buffered with ammonium phosphate at pH 6–6·5, but the crystals obtained have always been very small and unsuitable for X-ray-diffraction studies.

Amino acid analysis of azurin. Amino acid analyses of samples of azurin hydrolysed for different times have been performed on several different preparations of protein. One such set has already been published (Ambler, 1963a). Another set, obtained by using the more recent methods of hydrolysis and analysis, is shown in Table 1.

The tryptophan content, determined after the protein had first been digested with pepsin, was found to be 1·2%, corresponding to 1 residue/unit of mol.wt. 17000. If the determination was carried out on undigested protein, the tryptophan content was consistently found by this method to be less than 0·3%.

The amino acid and amide content shown in Table 1 are in complete agreement with the amino acid sequence of the protein deduced from overlapping peptides (Fig. 15.)

N-Terminal sequence of azurin. Preliminary experiments (Ambler, 1960) with the DNP method and the 'paper-strip PTC' method (Fraenkel-Conrat, Harris & Levy, 1955) had shown that the N-terminal residue of azurin was alanine and that the penultimate residue was either glutamic acid or glutamine.

At an early stage in the sequence investigation, evidence suggested that a very acidic peptide (peptide C13; Table 15) included the N-terminus of the protein molecule. If this hypothesis was true, α-DNP-peptides derived by enzymic digestion of the DNP-protein would be very acidic and easily purifiable (Ambler, 1963a). This proved to be the

Table 2. *Dinitrophenyl-peptides isolated after subtilopeptidase B digestion of dinitrophenylated azurin*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Yield (%)	Electrophoretic mobility (m)	Amino acid composition										
				Gly	Ala	Val	Ile	Ser	Asp	Glu	CySO₃H	α-DNP-Ala	O-DNP-Tyr	ε-DNP-Lys
8S1a	6B6	2	−0·40	−	±	±	−	+	±	+	+	−	−	+
8S1b	6B6	8	−0·49	−	−	−	−	+	±	±	+	−	−	+
8S1c	6B6	25	−0·63	+	−	−	−	−	+	+	+	−	−	+
8S1d	6B6	15	−0·36	+	−	±	−	±	±	+	−	−	+	+
8S2	6B6	15	−0·93	−	<0·02	0·97	0·90	0·99	1·05	2·05	1·04	0·89	−	−
8S3a	6B6	24	−1·05	−	<0·02	−	−	0·94	−	1·02	1·04	0·89	−	−
8S3b	6B6	5	−1·12	−	−	−	−	−	−	++	++	++	−	−

case, and the peptides shown in Table 2 were isolated from a subtilopeptidase B digest of soluble DNP-protein prepared from performic acid-oxidized azurin. This evidence shows that the N-terminal sequence of the protein is:

$$\text{Ala-Glx-Cys-Ser-(Val,Ile,Asx,Glx)} \ldots$$
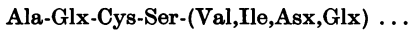
The sequence of three amino acids from the N-terminus was confirmed by the DNS–PTC method (Gray & Hartley, 1963a; Gray, 1967).

C-terminal sequence of azurin. Performic acid-oxidized azurin was treated with bovine CPB and with a mixture of bovine CPB and CP; the whole reaction mixture was then analysed for free amino acids. The results are shown in Table 3, and are consistent with the C-terminal sequence of the protein being:

$$\ldots \text{(Leu,Thr)-(Leu,Lys)}$$
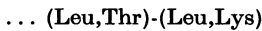
and with the protein having mol.wt. 14000.

Peptides from tryptic digests of azurin

The best yields of all the tryptic peptides except one (peptide T4a; see below) were obtained from digests of oxidized protein. In most digests trypsin was used in the ratio 1:40 (by weight), and digestion was for 1 hr. at 37°. Digestion for longer times increased the yield of peptide T9a, but decreased that of peptide T8a, and new peptides characteristic of chymotryptic cleavage (e.g. peptide T7b; Table 4) were produced.

Although all the digest was soluble at pH 8.5, it did not separate well on paper when subjected to electrophoresis at pH 6.5. There was much material that either remained at the origin or moved as a smear towards the cathode, and that affected the movement of the rest of the digest. This 'core'

Table 3. Amino acids released by carboxypeptidases A and B from azurin

Equal portions of the protein were taken for enzymic digestion and acid hydrolysis, and the amount of protein was calculated on the assumption that each protein molecule contained 10 residues each of leucine, serine, threonine and lysine (Table 1). Experimental details are given in the text. Amounts of amino acids are moles/mole of protein.

Treatment	Leu	Ser	Thr	Lys
CPB*	0.94	0.03	0.18	0.86
CP+CPB†	1.66	0.06	0.98	0.96
Acid hydrolysis‡	9.7	9.8	9.8	10.6

* 50 μg. of CPB/0.17 μmole of protein for 5 hr. at 37°.

† 50 μg. of CPB and 25 μg. of CP/0.17 μmole of protein for 5 hr. at 37°.

‡ 6 N-HCl for 24 hr. at 105°; amounts of serine and threonine were corrected for destruction during hydrolysis.

Table 4. Properties of peptides formed by tryptic digestion of azurin

Experimental details and definition of symbols are given in the text. The minor peptides in the lower part of the Table are described further in Table 14.

Peptide	Electrophoretic mobility (m)	Purification procedure	Approx. yield (%)	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	$CySO_3H$	$MetSO_2$	Pro	Lys	His	Arg	No. of residues	N-Terminus (DNS)
T1	+0.70	G6B3	30	—	—	0.57*	0.99	0.57*	—	—	—	—	—	—	—	—	—	1.01	—	—	2	Leu
T2	+0.68	G6B3	30	—	1.00	—	—	—	—	1.00	—	—	—	—	—	—	—	1.04	0.95	—	6	Val
T3	+0.39	G6B3	30	0.95	—	—	2.03	—	1.05	2.05	—	—	—	—	—	0.10	—	0.99	—	—	6	Gly
T4a†	+0.28	G6B3	10	1.05	—	1.04	2.00	1.00	—	0.96	2.08	0.93	0.90	—	—	—	2.13	0.94	0.93	—	14	Glx
T6a	+0.05	G6B3	20	2.03	—	—	0.94	—	0.99	—	2.08	1.06	—	—	—	—	—	0.99	—	—	7	Leu
T8a	0	G6B3	30	—	—	—	1.02	—	0.97	—	—	1.06	—	—	0.99	—	0.98	1.01	—	—	3	Ser
T7a	-0.10	G6B3	30	1.99	0.99	—	1.05	—	0.96	0.99	0.09	3.04	3.04	1.01	—	2.04	0.78	0.99	0.95	—	19	Glx
T9a	-0.24	G6B3	30	—	—	—	—	—	0.97	—	3.14	—	—	0.93	1.02	—	—	0.99	—	—	9	Asx
T9b	-0.23	G6B3	10	—	—	1.97	—	—	2.03	1.03	2.08	—	0.94	—	—	—	—	0.94	—	0.98	9	Asx
T10	Stuck	G26	<5	1.11	1.76	1.72	0.07	1.74	1.09	1.78	5.35	3.96	0.91	—	0.89	1.02	—	0.91	0.06	—	(23)	Ala
T4b	+0.45	G6B3	<5	+1.07	+	+0.70*	+	—	+	—	+	—	—	—	—	+	—	+	—	—	(5)	Ser
T4c	+0.42	G6B3	<5	0.29	0.97	1.05	0.07	1.31	0.06	0.97	1.00	0.20	—	—	—	+0.65*	—	—	—	—	5	Asx
T6b	+0.05	G6B3B4	<5	—	—	1.05	0.07	1.31	0.16	1.05	1.29	—	—	—	—	0.65*	—	1.08	0.93	—	6	Asx
T7b	0	G6B3	<5	—	—	+	0.92	—	+	1.02	+	—	—	—	—	—	—	+	—	—	(4)	Asx
T8b	-0.14	G6B3B3	10	2.00	—	2.06	0.90	0.90	3.11	1.02	2.10	1.12	0.96	—	—	—	—	1.82	—	—	16	Leu

* Low yield due to partial hydrolysis of stable peptide bond.

† Isolated from digest of unoxidized protein.

could be separated from the other peptides by gel filtration through Sephadex G-25 in 0·1 N-ammonia solution, and comprised about 40% (by weight) of the digest. The remaining peptides were not much fractionated by the gel filtration, but could afterwards be separated satisfactorily by paper electrophoresis and chromatography (Table 4).

The 'core' has not been satisfactorily frac-tionated. N-Terminal analysis (both DNS and DNP methods) showed it to contain only two N-terminal residues (alanine and aspartic acid or asparagine). Paper electrophoresis at pH 2·0 sep-arated several components, but much material remained at the origin. This material still remained at the origin when subjected to electrophoresis at pH 6·5, and after elution (with 2 N-ammonia

Table 5. *Amino acids released from peptide T2 by various treatments*

Experimental details are given in the text.

Treatment	Amino acids released					
	Ala	Val	Ile	Thr	Lys	His
Acid hydrolysis for 24 hr.	1·00	0·57	0·57	1·00	1·04	0·95
Acid hydrolysis for 96 hr.	0·97	0·77	0·79	0·88	1·03	0·81
One cycle of PTC degradation, then acid hydrolysis for 24 hr.	1·03	(0·01)	0·94	1·03	1·03	0·98
Two cycles of PTC degradation, then acid hydrolysis for 24 hr.	0·89	—	(0·04)	0·98	1·00	1·11
CP (I)*	0·08	—	—	0·39	0·68	0·32
CP (II)†	0·91	0·44	0·43	0·98	1·00	0·90
LAP‡	0·89	1·00	1·00	0·86	0·86	0·72

* 10 μg. of CP/0·2 μmole of peptide for 5 hr. at 37°.
† 50 μg. of CP/0·2 μmole of peptide for 5 hr. at 37°.
‡ 25 μg. of LAP/0·2 μmole of peptide for 5 hr. at 37°.

Table 6. *Peptides formed by subtilopeptidase B digestion of peptide T2*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)
T2S1	639	+0·67	12	Thr (+ +), Lys (+ +)	Thr
T2S2	639	+0·65	6	Ala (+ +), His (+ +)	Ala
T2S4	63	+0·47	4	Ala (+), Val (+), Ile (+), His (+)	Val*
T2S6a	632	0	4	Val (+), Ile (+)	Val*
T2S6b	632	0	2	Ala (+), Val (+), Ile (+)	Val*

* Normal hydrolysis (12 hr. for 105°) formed mainly a DNS compound with the same electrophoretic and chro-matographic properties as DNS-peptide T2S6a, but different from those of DNS-peptide T2S6b. After longer hydrolysis (96 hr. at 105°) DNS-Val predominated.

Table 7. *Peptides formed by chymotryptic digestion of peptide T3*

Experimental details and definition of the symbols are given in the text.

Peptide	Purifica-tion procedure	Electro-phoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
T3C1	63	+0·60	10	Leu (1·04), Thr (0·95), Lys (1·01)	Thr	CPB: Lys (+ + +) released; residual peptide had m 0, (Leu,Thr)
T3C2	63	0	10	Gly (0·94), Leu (1·03), Thr (1·03), [Asp (0·06)]	Gly	CP: Leu (+ + +) released; residual peptide had different m′

solution) was found to have alanine as the only N-terminal residue (DNS method), and to have the amino acid composition shown (as peptide T10) in Table 4. Several of the bands that moved when subjected to electrophoresis (at pH 6·5 or 2·0) were fluorescent, suggesting the presence of oxidized tryptophan, and all these contained aspartic acid or asparagine as the N-terminal residue, although amino acid analysis showed them to be too impure for further investigation. The amino acids present

and the N-terminal aspartic acid or asparagine were consistent with their being derived from the region of the polypeptide chain (residues 47–70; Fig. 15) not found in any of the other tryptic peptides.

In the rest of this section the evidence for the sequence of each of the major tryptic peptides is presented.

Peptide T1. The evidence in Table 4 showed that the sequence was:

Leu-Lys

Table 8. *Peptides formed by chymotryptic digestion of peptide T4a*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
T4aC1	36	+0·45	8	Gly (1·01), Leu (1·01), Ser (0·90), Asp (1·00), Pro (2·14), Lys (0·98), His (0·96)	Ser	Equivalent to peptide C3aT1 (Table 20)
T4aC2	36	+0·46	2	Gly (1·0), Leu (1·6), Ser (1·1), Asp (1·0), Pro (1·9), Lys (+ +), His (+ +), [Glu (0·07)]	Leu	Equivalent to peptide P3aT1 (Table 31)
T4aC3a	32	0	10	Val (1·00), Leu (1·03), Thr (0·92), Asp (1·04), [Ser (0·07)]	Thr	Equivalent to peptide C7d CP: Leu (+ +) and Asn (±) released; main residual peptide had composition (Val,-Thr,Asx)
T4aC3b	32	0	3	Val (+ +), Thr (+ +), Asp, (+ +), [Leu (±)]	Thr	Equivalent to peptide C7m
T4aC4	32	0	10	Glu (1·05), Phe (0·95), [Gly (0·08), Asp (0·06), Val (0·06)]	Glx	As m is 0, Gln (not Glu) must be present

Table 9. *Peptides formed by leucine aminopeptidase and carboxypeptidase B treatment of peptide T6a*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
T6aL1	3	0	1	Gly (+), Glu (+), Lys (+)		CPB: Lys (+) released; residual peptide had m −0·61
T6aL2	3	0	12	Gly (+ +), Ser (+ +), Glu (+ +), Lys (+ +)		CPB: Lys (+ + +) released; residual peptides described below
T6aL2aB	6	−0·43	6	Gly (1·94), Ser (0·99), Glu (1·07), [Asp (0·07)]	Gly	PTC degradation: one cycle: peptide had m −0·53
T6aL2bB	6	−0·53	6	Gly (1·03), Ser (0·94), Glu (1·03)	Ser	PTC degradation: one cycle: peptide had m −0·62, (Gly,-Glu), N-terminal Gly (DNS); two cycles: free Glu

Table 10. *Peptides formed by chymotryptic, peptic and Pronase digestion of peptide T8a*

Experimental details and definition of the symbols are given in the text. The following amino acids (with relative yields) were released during Pronase digestion: Ala (16), Leu (20), Phe (20), Tyr (12), MetSO₂ (16), His (some).

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
T8aC1	63	+0·59	10	MetSO₂ (0·91), Lys (1·06)	MetSO₂	
T8aC2a	632	0	3	Ala (1·04), Leu (1·04), Ser (0·93), [Gly (0·07), Asp (0·07), Glu (0·07)]	Ser	Equivalent to peptide C7g
T8aC3a	6B3	−0·10	3	Gly (++), Thr (+), Phe (+), Pro (+), CySO₃H (+), His (+)	CySO₃H	Equivalent to peptide C7e (same R_F)
T8aC3b	6B3	−0·10	4	Gly (1·07), Ala (0·99), Leu (1·04), Ser (1·04), Thr (1·01), Phe (0·96), CySO₃H (0·99), Pro (0·90), His (0·73)	CySO₃H	Equivalent to peptide C7c (same R_F and m')
T8aC3d	6B3	−0·10	2	Phe (2·01), MetSO₂ (0·99)	MetSO₂	Equivalent to peptide C7h CP: completely degraded
T8aC4	63	−0·69	12	Gly (0·99), Glu (3·01), Tyr (0·80)	Glx	Equivalent to peptide C9aF1 (Table 26)
T8aP3	63	+0·24	8	Gly (1·13), Ala (1·10), Leu (0·93), Ser (1·06), Thr (1·06), Phe (1·70), MetSO₂ (1·03), CySO₃H (1·03), Pro (1·23), Lys (0·93), His (0·80)		CPB: Lys (++) and traces of MetSO₂, Ala and Ser/Leu released LAP: Phe (+) released
T8aP5	63	−0·58	12	Gly (1·02), Glu (3·00), Phe (1·05), Tyr (0·95), MetSO₂ (0·95)	Glx	CP: Phe (++), Tyr (++), MetSO₂ (++) and Gln (+) released; two equal residual peptides had m −0·89 and −0·98
T8aQ1	63	+0·61	12	MetSO₂ (++), Lys (++)	MetSO₂	
T8aQ3	63	0	6	Ala (++), Ser (++)	Ser	
T8aQ5	63	−0·46	6	Gly (1·07), Thr (1·02), Phe (1·07), CySO₃H (0·92), Pro (0·92), [Asp (0·19)]	CySO₃H	
T8aQ7	63	−0·83	16	Gly (1·07), Glu (2·93), [Asp (0·08)]	Glx	PTC degradation: one cycle: peptide had m −0·50, (Gly,Glu); two cycles: peptide had m −0·61, (Glu) (not free Glu)

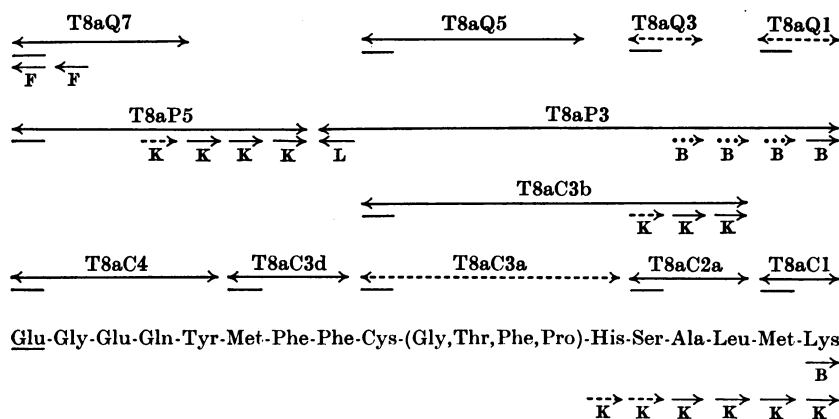Glu-Gly-Glu-Gln-Tyr-Met-Phe-Phe-Cys-(Gly,Thr,Phe,Pro)-His-Ser-Ala-Leu-Met-Lys

Fig. 1. Amino acid sequence of peptide T8a (Table 10). Definition of the symbols is given in the text.

The electrophoretic mobilities (m and m') were the same as for synthetic Lys-Leu, but the R_F was markedly different.

Peptide T2. This peptide did not yield valine or isoleucine in stoicheiometric amounts even after 96hr. hydrolysis at 105° (Table 5). *N*-Terminal analyses (DNS method) yielded (after the normal 12hr. hydrolysis) a fluorescent compound with electrophoretic mobility (at pH4·38) the same as DNS-Phe, but with a lower R_F (about equal to that of DNS-Ala) on chromatography (in light petroleum–acetic acid–water; 10:9:1, by vol.), together with about a tenth of the amount of DNS-Val. With a much longer time of hydrolysis (96hr.) about equal amounts of the two fluorescent compounds were formed. These results suggested that the *N*-terminal sequence of the peptide was Val-Ile ... This was confirmed by PTC degradation followed by amino acid analysis (Table 5). CPB released Lys (+ + +) from the peptide, leaving a less basic residual peptide (m +0·46). Both LAP and CP degraded the peptide almost completely (Table 5). Subtilisin digestion produced the peptides shown in Table 6. The sequence of the peptide was deduced to be:

<div align="center">Val-Ile-Ala-His-Thr-Lys</div>

Peptide T3. CPB released Lys (+ + +) and Leu

Table 11. Peptides formed by subtilopeptidase B digestion of peptide T9a, and by Pronase digestion and partial acid hydrolysis of peptide T9aS3

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	*N*-Terminus (DNS)	Other details
T9aS1	63	0	12	Ser (1·02), Asp (1·00), Arg (0·96)	Asx	CPB: Arg (+ +) released; residual peptide had m −0·76 PTC degradation: one cycle: peptide had m +0·73, (Ser,Arg) Value of m shows Asp, not Asn
T9aS3	63	−0·30	12	Leu (0·95), Asp (2·00), Tyr (0·75), Pro (1·05), Lys (1·00)	Asx	CP: no action Pronase: see below Partial acid hydrolysis (6N-HCl for 45min. at 100°): see below Values of m of peptide and of Pronase peptides show both residues Asp, not Asn
T9aS3Q1a	3	+0·11	10	Asp (+), Pro (+), Lys (+		PTC degradation: one cycle: peptide had m −0·65, (Asp,Pro), and gave blue colour with isatin reagent on paper; two cycles: free Asp
T9aS3Q2	36	−0·60	10	Asp (+ +), Tyr (+ +)	Asx	
T9aS3A1	6	+0·76	2	Pro (+), Lys (+)	Lys	
T9aS3A2	6	+0·68	3	Leu (+), Lys (+)	Leu	
T9aS3A3	6	+0·59	1	Leu (+), Pro (+), Lys (+)	Leu	
T9aS3A4	6	+0·51	1	Leu (+), Tyr (+), Pro (+), Lys (+)	Tyr	
T9aS3A5a	63	0	2	Asp (+), Pro (+), Lys (+)	Lys	
T9aS3A5b	63	0	1	Leu (+), Asp (+), Pro (+), Lys (+)	Leu	
T9aS3A5c	63	0	4	Leu (+), Tyr (+)	Tyr	
T9aS3A6	6	−0·49	2	Leu (+), Asp (+), Tyr (+)	Asx	
T9aS3A7	6	−0·58	2	Asp (+), Tyr (+)	Asx	
T9aS3A8	6	−0·71	5	Asp (+ +), Pro (+ +)	Pro	Peptide gave blue colour with isatin reagent on paper

A5b

A4

A6   A5a

T9aS3

A5c   A1

A7   A2   A8

T9aS3Q2   T9aS3Q1a
                F   F

T9aS3   T9aS1
                        F

Asp-Tyr-Leu-Lys-Pro-Asp-Asp-Ser-Arg
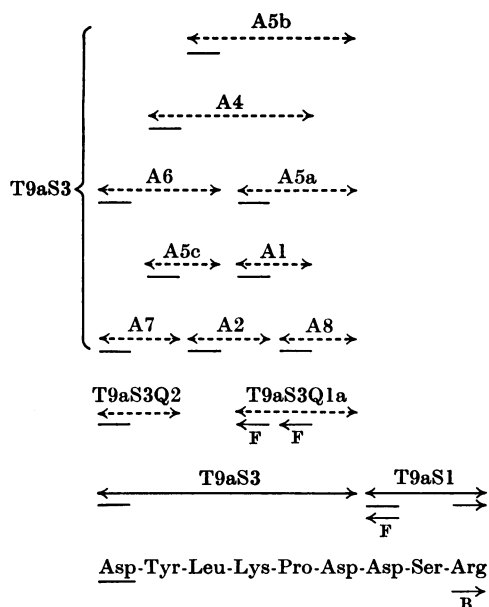   B

Fig. 2. Amino acid sequence of peptide T9a (Table 11). Definition of the symbols is given in the text.

acid residue much more easily in the digest of the oxidized protein.

No histidine-containing peptide with the electrophoretic mobility expected for the pyrrolidone-carboxylic acid derivative of peptide T4a was detected on peptide 'maps', but it is considered likely that the modified peptide would be insoluble.

Peptide T6a. CPB released Lys (+ + +), to leave an acidic (m -0.28) residual peptide. LAP released Leu, Ile, Gly and Ser (in the proportions $1:1:0.5:0.1$), and left the residual peptides shown in Table 9. These peptides were then treated with CPB before further degradation by the PTC method. The sequence of the peptide was deduced to be:

Leu-Ile-Gly-Ser-Gly-Glu-Lys

Peptide T7a. CPB released Lys (+ + +), to leave a very acidic residual peptide. The sequence is:

Ser-Cys-Lys

Peptide T8a. CPB released Lys (+ + + +), with only slight traces of neutral amino acids, and left an acidic (m -0.30) residual peptide. CP released Lys (+ + +), MetSO$_2$ (+ + +), Ala (+ + +), Leu (+ + +), Ser (+) and His (+); the predominant residual peptide had m -0.40. Table 10 shows the properties of peptides formed from peptide T8a by digestion with Pronase, chymotrypsin and pepsin, and Fig. 1 shows the deduction of the sequence.

Peptide T9a. CPB released Arg (+ + +), to leave a more acidic (m -0.43) residual peptide. Neither chymotrypsin nor pepsin had any effect on peptide T9a, but subtilisin split it into two peptides (Table 11). The larger peptide was then further degraded by partial acid hydrolysis and gave the products that are also shown in Table 11. The sequence of the peptide is deduced in Fig. 2.

Peptide T9b. CPB released Lys (+ + +) and traces of Ser. The sequence was deduced from the degradation of the two peptides formed by chymotryptic digestion (Table 12). Hydrolysis with pepsin produced the same peptides. The sequence was deduced to be:

Asp-Ser-Val-Thr-Phe-Asp-Val-Ser-Lys

(+ + +), leaving a single residual peptide. CP released Lys (+ + +), Leu (+ + +) and Thr (+ +), and left two residual peptides, both with electrophoretic mobilities (m') different from that of the residual peptide from CPB digestion. Chymotrypsin digested the peptide to give the products shown in Table 7, from which the sequence:

Gly-Thr-Leu-Thr-Leu-Lys

was deduced. The results of digestion with CP and CPB show that it is likely that this peptide forms the C-terminus of the whole protein.

Peptide T4a. CPB released Lys (+ +), leaving a much less basic residual peptide. Chymotrypsin digested the peptide to give the products shown in Table 8. The sequence of the peptide was deduced to be:

Gln-Phe-Thr-(Val,Asn)-Leu-Ser-(Gly,Leu,Asx,Pro,Pro,His)-Lys

Peptide T4a was obtained in good yields from tryptic digests of unoxidized protein, but was only present in very small amounts in digests of oxidized protein. The peptide does not contain any residues that are normally affected by performic acid, and the reason for the anomaly is not known. The most likely cause is that for some reason the *N*-terminal glutamine is converted into a pyrrolidonecarboxylic

Peptide T10. CPB released Lys (+ +). The properties of peptides formed by the action of chymotrypsin are shown in Table 13. This peptide was purified in only very small amounts, but the low yield is a reflection of the difficulty of the purification. Insufficient peptide was obtained to enable extensive sequence experiments to be carried out,

Table 12. *Peptides formed by chymotryptic digestion of peptide T9b*

Experimental details and definition of the symbols are given in the text.

Peptide	Purifica-tion procedure	Electro-phoretic mobility (*m*)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
T9bC1	63	0	15	Val (1·01), Ser (1·03), Asp (1·01), Lys (0·93), [Glu (0·06)]	Asx	CP: Lys (+ + +) and Ser (+) released; residual peptides had *m* − 0·59 and − 0·69. Pronase: Lys (+ + +) released; peptide T9bC1Q2 (+ + +), *m* − 0·59, (Val,Ser,Asx), and peptide T9bC1Q3 (+), *m* − 0·69, (Val,Asx), formed. Values of *m* show Asp, not Asn
T9bC2	63	−0·41	15	Val (1·03), Ser (0·99), Thr (1·00), Asp (1·04), Phe (0·96)	Asx	CP*: (*a*) Phe (+) and Thr (+) released; residual peptides had *m* − 0·50, (Val,Ser,Thr,Asx) (+ +) and *m* − 0·58, (Val,Ser,Asx) (+). (*b*) Phe (+ +), Thr (+ +) and Val (+ +) released; residual peptide had *m* − 0·77. Pronase: Thr (+ +) and Phe (+ +) released; peptide T9bC2Q2 (+ +), *m* − 0·58, (Val,Ser,Asx), and peptide T9bC2Q3 (+ + +), *m* − 0·77, (Ser,Asx), formed. Values of *m* show Asp, not Asn

* (*a*) 2·5 μg. of CP/0·08 μmole of peptide for 30 min. at 20°; (*b*) 10 μg. of CP/0·015 μmole of peptide for 5 hr. at 37°.

Table 13. *Peptides formed by chymotryptic digestion of peptide T10*

Experimental details and definition of the symbols are given in the text.

Peptide	Purifica-tion procedure	Electro-phoretic mobility (*m*)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
T10C1b	63	0	4	Ala (1·00), Val (0·89), Ile (1·03), Thr (1·86), Asp (3·14), Lys (0·80), [Gly (0·15), Ser (0·12), Glu (0·08)]	Asx	
T10C1c	63	0	4	Thr (+), Asp (+)	Asx	Probably equivalent to peptide C7b
T10C2	63	−0·23	4	Ala (+), Val (+), Leu/Ile (+), Thr (+ +), Asp (+ +), Lys (+)	Asx	Probably deamidated peptide T10C1b; cf. peptides C6d and C9d
T10C3	63	−0·70	15	Gly (1·00), Ala (0·92), Val (1·01), Ile (0·91), Ser (0·99), Asp (3·19), Glu (4·05), Phe (0·91), MetSO₂ (1·03), CySO₃H (1·03), [Thr (0·06)]	Ala	Equivalent to peptide C13 (same results when treated with subtilopeptidase B or carboxypeptidase)
T10C4	63	−0·80	6	Amino acid composition not qualitatively different from that of peptide T10C3	Ala	Equivalent to peptide C14 (same results when treated with subtilopeptidase B or carboxypeptidase); probably deamidated peptide T10C3

Table 14. *Minor peptides formed by tryptic digestion of azurin*

Definition of the symbols is given in the text.

| Peptide	Sequence (deduced from Fig. 15)	Remarks
T4b	Ser-Ala-Leu-Met-Lys	Formed by tryptic hydrolysis of a histidinyl peptide bond
T4c	Asn-Val-Met-Gly-His	Formed by tryptic hydrolysis of a histidinyl peptide bond
T6b	Ala-Ile-Thr-Val-Asp-Lys	Formed by splitting of -Thr-Asn↓-Ala- bond; peptide very difficult to separate from peptide T6a, and best preparation still contained 20% contamination
T7b	Asp-Val-Ser-Lys	Formed by splitting of -Thr-Phe↓-Asp- bond
T8b	Leu-Ile-Gly-Ser-Gly-Glu-Lys-Asp-Ser-Val-Thr-Phe-Asp-Val-Ser-Lys	Sum of peptides T6a and T9b; tryptic hydrolysis of -Glu-Lys-Asp- bond slow
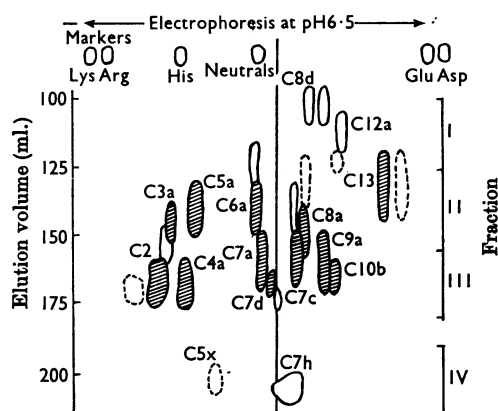
Fig. 3. Separation of peptides formed by chymotryptic digestion of *Pseudomonas* azurin (12 μmoles) by gel filtration followed by paper electrophoresis. Gel filtration (in 0·1 N-NH₃ solution) was through Sephadex G-25 (120cm. × 1·5 cm. diam. column). Then 0·075ml. of each 3·4ml. fraction was subjected to electrophoresis at pH 6·5 (60v/cm. for 1hr.). ε-DNP-lysine was eluted after 210ml. Definition of the symbols is given in the text.

but the evidence in Table 13 shows that the peptide contains the part of the polypeptide chain found also in peptides C13 and C6a (Figs. 4 and 5).

Minor tryptic peptides. The yields, mobilities and amino acid compositions of the tryptic peptides that were present only in minor amounts are shown at the bottom of Table 4. Table 14 gives further details of these peptides.

Peptides from chymotryptic digests of azurin

Azurin contains a high proportion of amino acids that form bonds that are hydrolysable by chymotrypsin (Table 1), and consequently the mixture of peptides produced by this enzyme was very complex. Most digestions were of oxidized protein, but digests of unoxidized material have been used to

obtain information about the sequence around the tryptophan residue.

The peptides were fractionated by gel filtration on Sephadex G-25 (equilibrated with 0·1 N-ammonia solution), followed by paper electrophoresis and chromatography. Fig. 3 shows a peptide 'map' obtained by electrophoresis of a portion of each fraction from a gel-filtration separation. Table 15 summarizes the methods of purification, approximate yields, electrophoretic mobilities and amino acid compositions of the isolated chymotryptic peptides.

In the rest of this section the evidence for the sequence of the principal chymotryptic peptides is presented. As two or more peptides were generally found that contained each region of sequence, families of peptides derived from the same region of polypeptide chain are dealt with together. The peptides are considered in the order in which they were eventually deduced to occur in the whole protein (Fig. 15).

Peptide C13. CP released Gln (1·11), MetSO₂ (0·84) and Phe (1·00), to leave a more acidic (*m* − 0·87) residual peptide. The peptides produced by pepsin and subtilisin digestion are shown in Tables 16 and 17, the results of mild acid hydrolysis of one of the subtilisin peptides are summarized in Table 18 and the sequence of peptide C13 is deduced in Fig. 4.

Peptide C14a had the same amino acid composition and terminal groups as peptide C13, but the residual peptide from CP digestion was more acidic. When digested with subtilisin, the peptides shown in Table 17 were formed. The evidence given shows that peptide C14a differs from peptide C13 either by the loss of the amide group on residue 10 or by loss of that on residue 12 (Fig. 15). Chymotryptic digests also contained small amounts of a peptide C14b, with a very similar amino acid composition to those of peptides C13 and C14a, the same *m* as peptide C14a, but a significantly less acidic *m'*. The mobilities were consistent with peptide C14b having lost an amide group from glutamine (either from residue

Table 15. *Properties of peptides formed by chymotryptic digestion of azurin*

Experimental details and definition of the symbols are given in the text. The peptides are arranged in groups in order of decreasing yield.

Peptide	Electrophoretic mobility (m)	Approx. yield (%)	Purification procedure	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	CySO3H	MetSO2	Pro	Lys	His	Arg	No. of residues	N-Terminus (DNS)
C2b	+0.59	10	G6B*B*B36	1.06	—	—	1.04	—	—	0.98	—	—	—	—	—	—	—	—	0.98	—	—	3	Thr
C3a	+0.50	20	G6B*B3	—	—	1.00	1.03	—	0.90	—	2.17	—	—	—	—	—	0.99	1.89	0.99	0.97	—	11	Ser
C4a	+0.43	20	G6B*B3	1.02	—	—	1.00	—	—	1.00	—	—	—	—	—	—	0.84	—	0.97	0.99	—	5	MetSO2
C4b*	+0.36	30	G6G	1.01	—	—	—	—	—	—	0.77	—	—	—	0.87	—	—	—	—	0.99	0.98	4	Gly
C5a	+0.40	10	G6B*B3	—	1.06	0.53†	0.98	0.49†	1.02	—	2.13	1.03	—	—	—	0.99	—	0.95	0.98	0.95	0.98	11	Leu
C6a	+0.40	10	G6B*B3	1.06	1.00	1.06	1.06	1.02	1.00	0.97	1.00	—	1.03	—	—	—	—	—	1.94	—	—	11	Ala
C7a	0	30	G6B*B3	—	—	1.03	1.02	—	0.97	1.00	0.98	—	0.97	—	—	0.95	—	—	0.99	0.98	—	5	Asx
C7c	−0.12	20	G6B*B3	1.04	1.01	—	1.02	—	1.00	0.99	—	0.97	1.03	—	—	—	—	0.99	—	0.98	—	9	CySO3H
C7d	0	20	G6B*B3	1.01	1.01	0.94	1.03	—	1.00	0.99	—	—	—	—	—	—	—	—	0.99	—	—	4	Thr
C7h	−0.10	—	G6B3	—	—	0.94	—	—	—	—	1.04	1.95	—	—	—	—	1.05	—	—	—	—	3	MetSO2
C8a	−0.21	30	G6B*B3	1.99	—	0.99	—	0.93	1.95	0.99	1.03	1.06	1.95	—	—	—	1.05	—	0.99	—	—	11	Ile
C8d	−0.25	10	G6B*B3	3.02	3.92	1.83†	1.95	0.62†	2.82	2.14	6.28	1.08	1.05	0.90	—	—	2.09	0.96	1.95	0.94	1.02	33	Ser
C9a	−0.29	30	G6B*B3	1.04	—	—	—	—	—	—	4.15	3.06	—	0.96	—	—	2.05	—	0.92	—	—	6	Lys
C12a	−0.41	10	C6B*B3	3.00	3.04	1.17†	1.01	—	1.85	1.98	—	—	—	—	—	—	—	—	0.99	—	—	22	Ser
C13	−0.69	30	G63	0.98	0.92	0.97	1.01	0.96	0.93	1.98	3.22	4.09	0.94	0.91	—	1.01	0.98	—	—	—	—	15	Ala
C1a	+0.70	5	G6B*B3	0.12	—	—	1.04	—	—	—	—	—	—	—	—	—	—	—	0.96	—	—	2	Lys
C2c	+0.59	5	G6B*B3B36	—	—	—	1.00	—	—	1.00	—	—	—	—	—	—	—	—	1.00	—	—	3	Thr
C2f	+0.56	5	G6B*B3B36	0.93	—	1.01	1.06	—	0.07	1.10	—	—	—	—	—	1.05	—	—	0.90	—	—	4	Lys
C6d	0	5	G6B*B3	—	1.05	—	—	0.98	1.02	2.00	3.05	1.05	0.98	—	—	—	—	—	1.89	—	—	14	Asx
C7b	0	5	G6B*B3	0.08	0.06	—	—	—	—	0.99	2.00	0.06	—	—	—	1.02	—	—	—	—	—	3	Asx
C7e	−0.08	5	G6B*B3	1.07	1.07	—	—	—	0.10	1.05	1.05	—	0.94	—	—	—	—	0.97	—	0.94	—	6	CySO3H
C7m	0	5	G6B*B3	0.08	0.08	—	—	—	0.06	0.93	1.03	—	—	—	—	—	—	—	—	—	—	3	Thr
C8c	−0.16	5	G6B*B3	1.02	1.02	1.03	1.00	—	0.98	0.98	2.04	3.09	—	0.86	—	—	—	—	1.89	—	—	11	Asx
C10b	−0.37	5	G6B*B3	—	2.07	0.98	—	—	—	—	—	—	—	0.96	—	—	—	—	0.99	—	—	4	Asx
C11a‡	−0.40	5	G6B3	2.98	2.86	1.10†	0.98	—	0.97	0.98	1.00	3.09	—	—	—	—	0.98	—	—	—	—	6	Ser
C12b	−0.38	5	G6B*B3	1.05	—	—	—	—	1.80	2.00	2.19	1.10	—	—	—	—	2.08	—	—	—	—	18	Ser
C14a	−0.79	5	G63	—	0.98	1.02	—	0.97	0.95	1.04	2.97	4.06	1.02	—	—	1.02	0.99	—	—	—	—	15	Ala
C1b	+0.70	<5	G6B*B3	—	1.00	1.07	1.07	—	—	1.04	1.05	1.05	—	—	—	—	—	1.94	1.00	0.89	—	5	Ala
C2a	+0.56	<5	G6B*B3	0.98	—	—	0.96	—	0.81	—	1.10	1.10	—	—	—	—	—	—	1.03	0.98	—	8	Ser
C5x	+0.30	<5	G63	0.94	—	—	0.94	—	0.90	—	1.10	1.07	—	—	—	—	—	—	—	1.04	—	6	Gly
C6b||	+0.03	<5	G6B*B3	—	—	0.96	0.94	0.96	0.90	0.90	2.09	—	—	—	§	—	—	1.08	1.01	0.93	0.99	9	Leu
C6c	+0.04	<5	G6B*B3	—	1.12	0.38†	1.01	0.44†	1.12	0.14†	4.10	4.10	0.99	0.85	—	—	—	0.93	1.78	1.02	1.01	15	Asx
C7f	−0.10	<5	G6B*B3	1.02	1.02	—	0.99	—	1.04	1.01	—	0.99	0.99	—	—	0.99	1.02	0.88	—	—	—	10	CySO3H
C7g	0	<5	G6B*B32	1.02	1.02	—	1.06	—	0.92	—	—	—	—	—	—	—	—	—	—	—	—	3	Ser
C7j	0	<5	G6B*B3B32	++	—	++	+++	—	—	—	++	—	—	++	—	—	++	—	—	—	(2)	Val
C7k	0	<5	G6B*B3	—	—	++	—	—	—	—	0.10	—	—	—	—	—	1.14	—	—	—	(3)	Asx
C7n	0	<5	G6B*B3	—	1.02	—	1.02	—	—	—	0.10	—	—	—	—	—	1.0	—	—	—	4	Ser
C9b	−0.30	<5	G6B*B3	1.4	2.0	1.8	1.13	0.63†	0.9	1.8	6.23	1.1	—	—	—	—	2.12	0.92	1.85	—	11	Ser
C9c	−0.32	<5	G6B3	3.06	2.97	1.71†	0.3	0.92	2.84	2.00	2.90	1.07	0.95	0.65	—	—	—	—	1.95	—	0.93	31	¶
C9d	−0.19	<5	G6B*B3	0.11	1.10	0.98	1.98	—	1.05	1.95	+++	1.15	—	—	—	1.02	—	—	++	—	—	14	Asx
C9e	−0.30	<5	G6B33	++	—	++	1.02	—	0.96	—	1.02	—	—	++	—	—	—	—	—	—	—	(6)	¶
C10a	−0.40	<5	G6B*B3	2.03	0.99	0.74†	1.02	—	1.98	2.06	1.02	1.06	—	—	—	—	0.98	—	—	—	—	7	Asx
C12c	−0.40	<5	G6B*B3	2.10	1.80	—	0.17	—	2.06	—	2.06	1.06	—	—	—	—	1.93	—	—	—	—	16	Ser

* Peptide obtained from digest of unoxidized protein.
† Low yield due to partial hydrolysis of stable peptide bond.
‡ Peptide obtained from digest of unoxidized protein, but oxidized before final purification.
§ Peptide contained oxidized tryptophan.
|| Peptide hydrolysed for 72 hr.
¶ N-Terminal analysis not performed.

Table 16. *Peptides formed by peptic digestion of peptide C13*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C13P1a	63	0	2	Glu (2·11), Phe (0·93), MetSO₂ (0·96), [Gly (0·22), Ser(0·20), Asp (0·10), Leu (0·10), Ile (0·10), Ala (0·07), Val (0·07)]	Glx	LAP: complete degradation; Gln (+ + + +), Phe (+ +) and MetSO₂ (+ +) released, but no Glu released. Value of m also indicates that both residues are Gln, not Glu
C13P1b	6	−0·19	5	Glu (1·03), Phe (0·95), MetSO₂ (1·01)	MetSO₂	Anomalous m due to N-terminal MetSO₂
C13P1c	63	0	1	Gly (1·04), Ile (0·97), Glu (0·99), [Ser (0·10), Thr (0·09)]	Ile	Value of m shows Gln, not Glu
C13P2	6	−0·27	10	Gly (1·03), Ile (0·91), Asp (2·32*), Glu (3·12), Phe (0·94), MetSO₂ (1·24*)	Ile	Equivalent to peptide P12b (same m and m')
C13P3	636	−0·36	2	Gly (1·02), Ile (0·92) Asp (2·06), Glu (2·00)	Ile	Contains one more Glx residue than peptide C13P4b; lower m value shows this is Gln, not Glu
C13P4a	63	−0·38	1	Asp (+ +), Glu (+ +), Phe (+), MetSO₂ (+ +)	Asx	
C13P4b	636	−0·40	2	Gly (1·01), Ile (0·94), Asp (2·03), Glu (1·02)	Ile	
C13P5	6	−0·73	4	Amino acid composition not qualitatively different from that of peptide C13	Ala	Undigested peptide
C13P6	6	−0·98	15	Ala (0·96), Val (0·98), Ser (0·96), Asp (1·00), Glu (1·04), CySO₃H (1·04)	Ala	Equivalent to peptide P18a (same m and m')

* Aspartic acid and methionine sulphone not sufficiently resolved by automatic amino acid analyser for satisfactory quantitative determination.

12 or 14), and peptide C14a having lost an amide group from asparagine (i.e. from residue 10).

The N-terminal sequence of peptide C13 corresponds exactly to that of the whole protein (Table 2).

Peptide C7b. CP released Asn (+ +), leaving a residual peptide with a different m'. LAP released Asn (+ +) and Thr (+), but no Asp. The value of m (− 0·04) was consistent with the sequence:

Asn-Thr-Asn

Peptide C6a. CP released Gln (+ +) and Phe (+ +). With trypsin, the peptides shown in Table 19 were formed. The sequence deduced is shown in Fig. 5.

Peptides C6d and C9d. Both these peptides had N-terminal aspartic acid or asparagine (DNS method), and amino acid compositions corresponding to the sum of those of peptides C7b and C6a. The electrophoretic mobility (m) of peptide C9d

was consistent with it having lost one of the three amide groups.

Peptide C7d. CP released Leu (+ + + +) and Asn (+ +), to leave two residual peptides, with m 0 and different values of m', and compositions (Val,Thr) and (Val,Thr,Asx). The sequence was therefore:

Thr-Val-Asn-Leu

Small amounts of peptide C7m, Thr-Val-Asn, were also found in chymotryptic digests.

Peptide C3a. CP released Val (+ + +), Asn (+ +) and MetSO₂ (+ +), to leave a residual peptide with m + 0·59. Trypsin split the peptide into two (Table 20), but the larger of these peptides (C3aT1) was not further degraded by pepsin or subtilisin, and Pronase only removed the C-terminal lysine. Peptide C3aT1 was therefore degraded by partial acid hydrolysis, and the peptides shown in Table 21 were

Table 17. *Peptides formed by subtilopeptidase B digestion of peptides C13 and C14a*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C13S1a	63	0	3	Glu (+), Phe (+)	Glx	Value of m shows Gln, not Glu
C13S1b	63	-0.13	3	Glu (1·01), Phe (1·04), MetSO$_2$ (0·96)	MetSO$_2$	Anomalous m due to N-terminal MetSO$_2$ CP: Phe (+ +) released; residual peptide had m -0.10, (Glx,MetSO$_2$)
C13S2	63	-0.32	6	Gly (0·93), Asp (2·10), Glu (2·11), Phe (0·87), MetSO$_2$ (0·99)	Gly	CP: Gln (+ +), Phe (+ +) and MetSO$_2$ (+ +) released; two residual peptides, with properties of peptides C13S3 and C13S4
C13S3	63	-0.43	3	Gly (+ +), Asp (+ + + +), Glu (+ +), MetSO$_2$ (+ +)	Gly	CP: MetSO$_2$ (+ +) released; residual peptide had properties of peptide C13S4
C13S4	63	-0.48	3	Gly (0·94), Asp (2·03), Glu (1·03)	Gly	CP: no action PTC degradation: one cycle: peptide had m -0.54, N-terminal Asx (DNS); two cycles: peptide had m -0.64, N-terminal Asx (DNS); three cycles: free Gln Mobilities show sequence to be Gly-Asn-Asp-Gln
C13S5	63	-0.82	12	Ala (0·94), Val (0·97), Ile (0·96), Ser (0·97), Asp (1·04), Glu (2·07), CySO$_3$H (1·08)	Ala	CP: Gln (1·0), Ile (1·0), Val (0·7) and Asp (+ + +) released; two residual peptides, both with m -0.93, but with different m', (Ala,Val,Ser,Asx,Glx,-CySO$_3$H) and (Ala,Ser,Glx,-CySO$_3$H) LAP: Ala (+ +) and Glu (+ +) released; residual peptide had m -0.72 Exopeptidases show that Glu is near the N-terminus and Asp and Gln near the C-terminus Partial acid hydrolysis: see Table 18
C14aS2	63	-0.64	12	Gly (0·93), Asp (2·10), Glu (2·00), Phe (1·00), MetSO$_2$ (1·10)	Gly	CP: Gln (+ +), MetSO$_2$ (+ +) and Phe (+ +) released; two residual peptides had m -0.82 (+) and m -0.95 (+ +) Same amino acid composition and terminal residues as peptide C13S2; more acidic value of m suggests loss of amide group
C14aS3	63	-0.83	12	Ala (0·94), Val (1·00), Ile (0·97), Ser (0·99), Glu (1·97), Asp (1·06), CySO$_3$H (1·06), [Gly (0·06)]	Ala	Equivalent to peptide C13S5

Table 18. *Peptides formed by partial acid hydrolysis of peptide C13S5*

Experimental details and definition of the symbols are given in the text. The peptide was hydrolysed with 6 N-HCl for 30 min. at 100°.

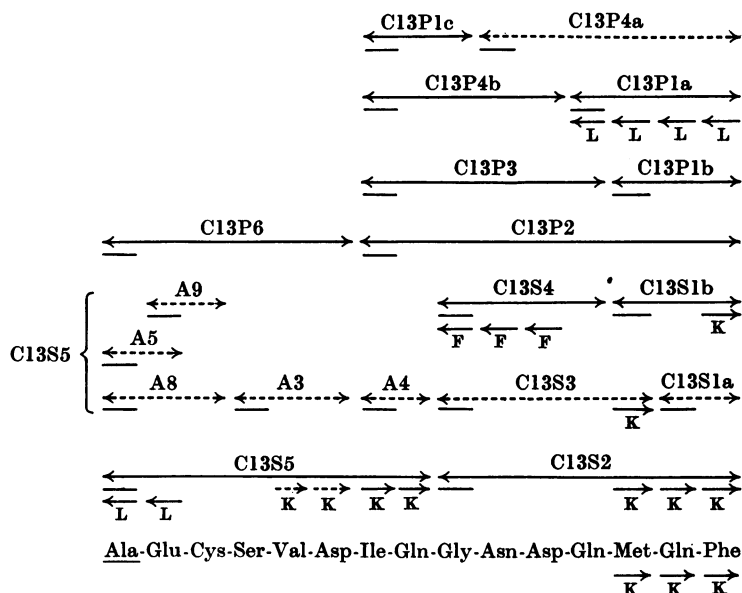| Peptide	Purifica-tion procedure	Electro-phoretic mobility (*m*)	Relative yield	Amino acid composition	N-Terminus (DNS)
C13S5A1	63	0	1	Val (+), Ser (+)	Ser
C13S5A3	63	−0·45	5	Val (+), Ser (+), Asp (+)	Ser
C13S5A4	63	−0·51	10	Ile (++), Glu (++)	Ile
C13S5A5	63	−0·58	10	Ala (++), Glu (++)	Ala
C13S5A8	63	−1·05	6	Ala (++), Glu (++), CySO₃H (++)	Ala
C13S5A9	63	−1·20	3	Glu (++), CySO₃H (++)	Glx

Fig. 4. Amino acid sequence of peptide C13 (Tables 16, 17 and 18). Definition of the symbols is given in the text.

isolated and characterized. The sequence of peptide C3a is deduced in Fig. 6.

Peptides C2a and C7k, present in small amounts in chymotryptic digests, were the same as peptides C3aT1 and C3aT2 in amino acid composition and electrophoretic mobility, indicating that there was some tryptic activity in the chymotrypsin used.

Peptide C4b. CP released Trp (++++) and Asn (++), and left peptides (Gly,Asx,His), *m* +0·60, and (Gly,His), *m* +0·79. The sequence was therefore:

<p style="text-align:center">Gly-His-Asn-Trp</p>

Peptide C4b was only present in digests of unoxidized protein. In digests of oxidized protein, the predominant fluorescent peptide, containing oxidized tryptophan, was peptide C5x. There were also present smaller amounts of a more basic fluorescent peptide with probable composition (Gly,Asx,His,'oxidized Trp'). Peptide C5x had N-terminal glycine (DNS method), and CP released Val (++) and Leu (++).

In digests of unoxidized protein the region corresponding to the C-terminus of peptide C5x was present as peptide C7j [*m* 0, composition (Val,Leu), N-terminal valine (DNS method); same *m'* as synthetic Val-Leu], but the amount that could be isolated was much less than that of peptide C4b.

Peptides C8d, C10b, C12a and C12b. The largest peptides in the chymotryptic digest that were least

26

Table 19. *Peptides formed by tryptic digestion of peptide C6a and by Pronase digestion of peptide C6aT1*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C6aT1	32	+0·04	10	Ala (1·04), Val (0·96), Ile (1·04), Thr (0·88), Asp (1·12), Lys (0·96), [Ser (0·10)]	Ala	CPB: Lys (+ +) released; residual peptide had m −0·36 LAP: Ala (+ +), Ile (+ +) and Thr (+ +) released Pronase: see below Equivalent to peptide T6a
C6aT2	32	0	10	Ser (+ +), CySO₃H (+ +), Lys (+ +)	Ser	CPB: Lys (+ +) released Equivalent to peptide T7a
C6aT3	32	0	10	Glu (0·97), Phe (1·03)	Glx	Value of m shows Gln, not Glu
C6aT1Q1a	63	0	4	Val (+ +), Asp (+ +), Lys (+ +)	Val	
C6aT1Q1b	63	0	2	Asp (+ +), Lys (+ +)	Asx	
C6aT1Q1c	63	0	2	Ala (+ +), Ile (+ +), Thr (+ +)	Ala	
C6aT1Q2	6	−0·34	1	Ala (+), Val (+), Ile (+), Thr (+), Asp (+)	Ala	
C6aT1Q3	6	−0·50	2	Val (+), Thr (+), Asp (+)	Thr	
C6aT1Q4	6	−0·64	4	Val (+ +), Asp (+ +)	Val	

C6aT1Q1b
←------→

C6aT1Q4
←------→

C6aT1Q3
←-------→

C6aT1Q2
←-----------------→

C6aT1Q1c   C6aT1Q1a
←---------→ ←---------→

C6aT1 ──────────────→ ← C6aT2 → ← C6aT3 →
←─ ←─ ←─   B ←------→   B ─ ─
 L   L   L

Ala-Ile-Thr-Val-Asp-Lys-Ser-Cys-Lys-Gln-Phe
                              →   →
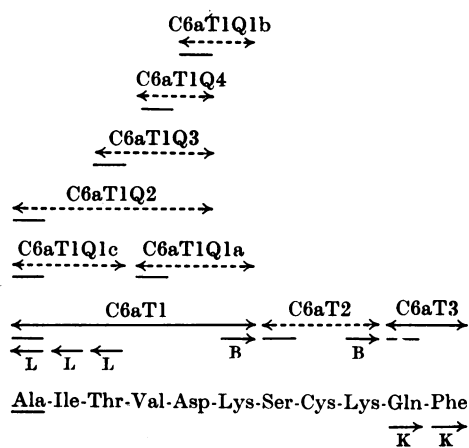   K   K

Fig. 5. Amino acid sequence of peptide C6a (Table 19). Definition of the symbols is given in the text.

retarded by gel filtration were peptides C8d, C12a and C12b (Fig. 3). All had N-terminal serine, and their amino acid compositions were consistent with their all being derived from the same region of polypeptide chain.

None of these peptides gave a simple or clean peptide pattern when digested with chymotrypsin

or pepsin. Most of the evidence for sequence was derived from subtilisin digestion of peptide C12b (Table 22). The experiments with peptides C8d and C12a were designed to show that they included the sequence of peptide C12b.

The N-terminal sequence of peptide C8d was examined by PTC degradation, the N-terminal residue exposed at each step being identified by the DNS method. The sequence was shown to be:

Ser-Thr-Ala-Ala-Asx-MetSO₂ . . .

CP released Ala (+ + + +), Val (+ + + +), Ile (+ + + +), His (+ + + +) and Arg (+ +), leaving residual peptides with m −0·31 and −0·43.

From peptide C12a, CP (10 μg./0·02 μmole for 30 min. at 37°) released Tyr (+ +) and Asp (+), leaving two residual peptides (neither containing tyrosine), with m −0·46 and −0·34. More intense treatment (20 μg./0·02 μmole for 5 hr. at 37°) released Tyr (+ +) and Asp (+ +), leaving only one residual peptide (m +0·34). The action of subtilisin on peptide C12a was to produce all the peptides produced from peptide C12b (Table 22), in approximately the same relative yields, but with in addition a peptide of composition [Asp (+ +), Tyr (+), Lys (+)] and the same mobilities (m and m') as peptide C10b. When digested with chymotrypsin, peptide C12a formed the peptides shown in Table 23.

The amino acids released from peptide C12b by

Table 20. *Peptides formed by tryptic digestion of peptide C3a*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C3aT1	63	+0·59	10	Gly (0·98), Leu (0·96), Ser (0·81), Asp (1·05), Pro (1·94), Lys (1·03), His (0·98)	Ser	CPB: Lys (+ +) released: residual peptide had m +0·32 Value of m shows that peptide contains Asn, not Asp Trypsin, subtilopeptidase B, pepsin: no action Pronase: a little Lys released; otherwise no action Partial acid hydrolysis: see Table 21 Equivalent to peptides T4aC1, P3aT1, C2a and S3a
C3aT2	63	0	10	Val (0·98), Asp (1·04), MetSO$_2$ (0·98)	Asx	CP: MetSO$_2$ (+ +) released Value of m shows Asn, not Asp

Table 21. *Peptides formed by partial acid hydrolysis of peptide C3aT1 or S3a*

Experimental details and definition of the symbols are given in the text. The peptides were hydrolysed with 6 N-HCl for 15 min. at 100°.

| Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C3aT1A1	63	+0·77	6	Pro (+ +), Lys (+ +)	Pro	Peptide gave blue colour with isatin reagent on paper
C3aT1A2	63	+0·65	12	Ser (+ +), His (+ +)	Ser	
C3aT1A3	63	+0·59	12	Leu (+ +), Pro (+ +), Lys (+ +)	Leu	
C3aT1A4	63	+0·40	15	Gly (1·00), Ser (0·82), Pro (1·00), His (0·80)	Ser	PTC degradation: one cycle: peptide had m +0·34, (Gly,Pro,His); two cycles: peptide had m 0, (Gly, Pro), gave blue colour with isatin reagent on paper, and had same m' as peptide C3aT1A7b
C3aT1A6	63	+0·05	6	Leu (+), Asp (+), Pro (+), Lys (+)	Asx	Possibly $\alpha\beta$-aspartyl form of peptide C3aT1A7a
C3aT1A7a	632	0	12	Leu (1·04), Asp (1·01), Pro (0·98), Lys (0·96)	Asx	CPB: Lys (+ +) released; residual peptide had m −0·55 (the same as peptide C3aT1A8)
C3aT1A7b	632	0	6	Gly (+ +), Pro (+ +)	Pro	Peptide gave blue colour with isatin reagent on paper
C3aT1A8a	63	−0·55	3	Leu (+ +), Pro (+ +), Asp (+ +)	Asx	
C3aT1A8b	63	−0·55	1	Leu (+ +), Asp (+ +)	Asx	Peptide gave yellow colour with ninhydrin on paper Possibly $\alpha\beta$-aspartyl form of peptide C3aT1A9
C3aT1A9	63	−0·66	6	Leu (+ +), Asp (+ +)	Asx	
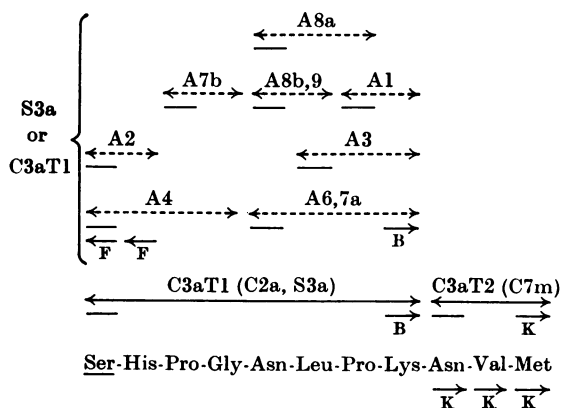
Fig. 6. Amino acid sequence of peptides C3a and S3a (Tables 20 and 21). Definition of the symbols is given in the text.

CP were determined quantitatively. Mild treatment (20 μg./0·05 μmole for 5 hr. at 37°) released Leu (1·00), Gly (0·94), Ala (0·92), Ser (0·80) and MetSO$_2$ (0·08). The residual peptides were not well separated, having mobilities (m) in the range − 0·41 to − 0·45.

From peptide C10b, CP released Tyr (+ +), leaving a more acidic residual peptide (m − 0·42). Two cycles of PTC degradation produced a peptide (Asx,Tyr) with m − 0·58 and N-terminal aspartic acid or asparagine (DNS method).

Besides peptides C12a, C12b, C10b and C8d, several other peptides were isolated, in very variable amounts, that can be accounted for in this region. The evidence for their location (Fig. 7) was derived from amino acid compositions and from knowledge of which bonds in the sequence are slightly susceptible to chymotrypsin.

Peptides C8d, C12a and C12b were absent from digests of unoxidized protein. Instead, peptide C11a was isolated in good yield (Fig. 7). No trace of a peptide extending between peptides C11a and C10b was found. As this would have had N-terminal glutamine, it is likely that it became converted into a pyrrolidonecarboxylic acid peptide during purification and was thus lost.

Peptide C5a. CP released Ala (+ +), Val (+ +), Ile (+ +), His (+ +) and Arg (±), the same amino acids as were released from peptide C8d. The amino acid composition was equal to the difference between those of peptides C8d and C12a. Trypsin split peptide C5a into two peptides (Table 24); the larger of these was further degraded with subtilisin (Table 24).

Peptides C6b and C6c, present in small amounts in chymotryptic digests, have amino acid composi-

tions and terminal groups consistent with their being derived from this same region of polypeptide chain. Fig. 8 shows the derivation of the sequence.

Peptide C2c. CP released Leu (+ + +), to leave a more basic (m + 0·74) residual peptide. The sequence was therefore:

Thr-Lys-Leu

Peptide C8a. CP released Val (+ + +), Thr (+ + +) and Phe (+ + +), and left a more acidic residual peptide (m − 0·27). LAP released Ile (+ + +), Gly (+ +) and Ser (+). Trypsin split peptide C8a into two peptides (Table 25). From these results the sequence was deduced to be:

Ile-(Gly$_2$,Ser,Glu)-Lys-Asp-Ser-Val-Thr-Phe

Peptide C7a. CP released Lys (+ +), Leu (+ +) and Ser (+), leaving residual peptides (Val,Ser,Asx) (m − 0·60) and (Val,Asx) (m − 0·70). Subtilisin split the peptide into two peptides, C7aS1 [(Leu,-Lys), m + 0·70, N-terminal lysine (DNS method)] and C7aS2 [(Val,Ser,Asx), m − 0·60, N-terminal aspartic acid or asparagine (DNS method)]. Peptide C7aS1 had the same R_F as synthetic Lys-Leu. The sequence of peptide C7a was therefore:

Asp-Val-Ser-Lys-Leu

the N-terminal residue being taken to be aspartic acid rather than asparagine because of the acidic mobility of peptides C7aS2 and the residual peptides from CP digestion.

Peptide C9a. CP released Tyr (+ +) and traces of Gln, leaving a strong residual peptide with m − 0·37 and a trace of one with m − 0·41. More intense treatment with CP released more Gln (+ +), and caused the more acidic residual peptide to predominate. The peptide was subjected to PTC degradation; at each step the peptide was purified by electrophoresis (at pH 6·5), and samples were tested for amino acid composition and N-terminal residue (Table 26). LAP released Lys (+ + + +), Glu (+ + + +), Gln (+ +), Tyr (+ +) and Gly (+ +), leaving residual peptides with the same mobilities as peptides C9aF1, C9aF2 and C9aF3 (Table 26). The sequence of the peptide was deduced to be:

Lys-Glu-Gly-Glu-Gln-Tyr

Peptide C8c. This peptide, present in varying amounts, had the amino acid composition of the sum of peptides C7a and C9a, N-terminal aspartic acid or asparagine (DNS method) and C-terminal tyrosine (CP). Chymotrypsin split it into peptides equivalent to peptides C7a and C9a, and trypsin split it into peptides corresponding to peptides T1, T7b (Table 14) or T9bC1 (Table 12) and T8aC4 (Table 10) or C9aF1 (Table 26).

Table 22. *Peptides formed by subtilopeptidase B digestion of peptide C12b*

Experimental details and definition of the symbols are given in the text. The equivalents of all the peptides in the Table except peptides C12bS2e, C12bS6 and C12bS8 were identified in the products of digestion of peptide P13c with subtilopeptidase B.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C12bS1a	632	0	2	Ala (+++)		Value of m' the same as that of synthetic Ala-Ala
C12bS1b	632	0	3	Gly (++), Leu (++)	Gly	
C12bS1c	632	0	5	Ala (+++), Ser (+++)	Ala	PTC degradation: one cycle: free Ser
C12bS1d	632	0	3	Gly (++), MetSO$_2$ (++)	Gly	PTC degradation: one cycle: free MetSO$_2$
C12bS2a	63	−0·07	4	Ser (++), Thr (++)	Ser	
C12bS2b	63	−0·16	6	Glu (+++), MetSO$_2$ (+++)	MetSO$_2$	Anomalous m due to N-terminal MetSO$_2$; value shows Gln, not Glu
C12bS2e	63	−0·07	2	Ala (+), Ser (+), Thr (+)	Ser	Value of m″ the same as that of peptide P8c and different from that of peptide P8e
C12bS3a	63	−0·38	8	Gly (1·03), Val (1·13), Thr (0·95), Asp (1·02), Val-Val (0·76/2*)	Gly	Equivalent to peptide S13b PTC degradation: one cycle: peptide had m −0·41, (Val,Thr,Asp), N-terminal Val† (DNS); two cycles: peptide had m −0·47, (Val,Thr,Asp), N-terminal Val‡ (DNS); three cycles: peptide had m −0·69, (Thr,Asp), N-terminal Thr (DNS) Values of m show Asp, not Asn
C12bS4	63	−0·49	4	Ala (+), Asp (+), Glu (+), MetSO$_2$ (+)	Ala	PTC degradation: one cycle: peptide had m −0·57, the same as that of peptide C12bS5 CP: Gln (++) released; residual peptide had m −0·57
C12bS5	63	−0·57	4	Asp (+), Glu (+), MetSO$_2$ (+)	Asx	Value of m shows at least one acidic residue to be present
C12bS6	63	−0·63	4	Ala (++), Asp (+)	Ala	PTC degradation: one cycle: peptide had m −0·75, the same as that of peptide C12bS8
C12bS8	63	−0·75	4	Ala (+), Asp (+)	Ala	Value of m shows Asp, not Asn

* 24 hr. hydrolysis; Val-Val identified and determined quantitatively on automatic amino acid analyser calibrated with synthetic peptide.

† Peptide formed an acid-stable DNS derivative; DNS-Val only identified after 96 hr. hydrolysis.

‡ No acid-stable DNS derivative formed; DNS-Val formed in good yield after 12 hr. hydrolysis.

Peptide C7h. CP degraded the peptide completely. The sequence was deduced to be:

Met-Phe-Phe

This peptide was missed in several of the early digests, as it has a very high R_F, was much retarded during gel filtration and was sparingly soluble.

Peptide C7c. CP released Ala (+++), Ser (++), Leu (+++) and His (+). Chymotrypsin split peptide C7c into peptides with the same mobilities

Table 23. *Peptides formed by further chymotryptic digestion of peptide C12a*

Experimental details and definition of the symbols are given in the text. Peptide (0·55 μmole) was digested with 200 μg. of chymotrypsin for 6 hr. at 37°.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C12aC1a	63	−0·27	2	Gly (1·01), Leu (1·05), Asp (2·02), Tyr (1·05), Lys (0·89), [Ser (0·15)]	Gly	Equivalent to peptide C9e
C12aC1b	63	−0·30	1	Gly (1·7)*, Ala (2·1), Val (1·1), Ser (1·0), Thr (1·8), Asp (1·1), Glu (1·1), MetSO₂ (1·0), [Leu (0·1)]	Ser	Equivalent to peptide C9b CP: Val (+ + + +) and Thr (+ +) released; main residual peptide had m −0·38
C12aC2a	63	−0·39	10	Asp (+ + + +), Tyr (+ +), Lys (+ + +)	Asx	Equivalent to peptide C10b
C12aC2b	63	−0·39	2	Gly (2·1), Ala (1·1), Leu (1·0), Ser (1·0), Asp (1·0), MetSO₂ (1·1)	Asx	Equivalent to peptide C10a
C12aC3	63	−0·50	<1	Gly (+), Ala (+), Ser (+), Asp (+), MetSO₂ (+)	Asx	

* From the sequence (Fig. 7) this peptide should contain only one residue of glycine; as the amino acid analysis was performed on a very small amount of peptide (0·010 μmole), not using a high-sensitivity attachment, the inaccuracy is understood.
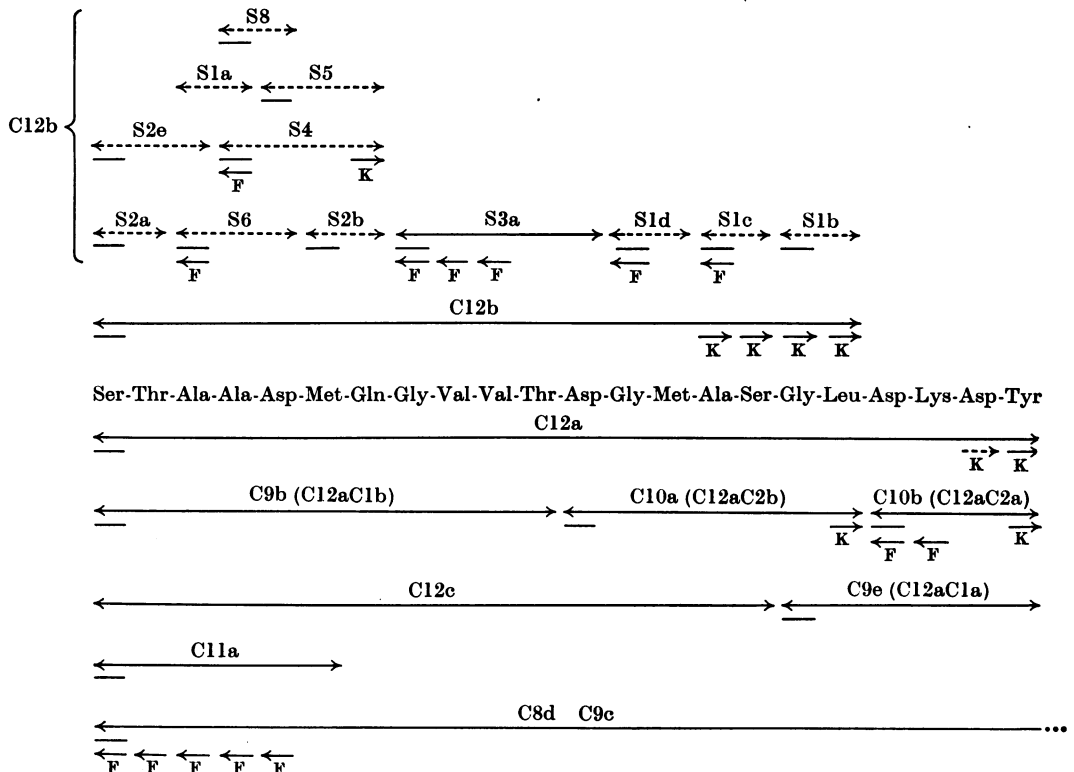

Fig. 7. Amino acid sequence of peptides C12a and C12b (Tables 22 and 23). Definition of the symbols is given in the text.

Table 24. *Peptides formed by tryptic digestion of peptide C5a and by subtilopeptidase B digestion of peptide C5aT2*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C5aT1	6	+0·53	12	Ala (1·04), Val (0·48), Ile (0·52), His (0·94)	Val*	CPB: His (+ +) released PTC degradation: one cycle: peptide had m +0·59, (Ala, Ile, His), N-terminal Ile (DNS); two cycles: peptide had m +0·71, (Ala, His), N-terminal Ala (DNS)
C5aT2	63	0	12	Leu (1·0), Ser (1·1), Asp (2·1), Pro (0·6), Lys (1·0), Arg (0·8), [Gly (0·1)]	Leu	CPB: Arg (+ +) released; residual peptide had m −0·31 Subtilopeptidase B: see below
C5aT2S1	3	0	8	Leu (0·83), Asp (1·21), Pro (0·97), Lys (0·97), [Gly (0·07), Ala (0·07), Ser (0·07)]	Leu	Partial acid hydrolysis (6 N-HCl for 45 min. at 100°): peptides with same mobilities as peptides T9aS3A1, T9aS3A2, T9aS3A3, T9aS3A5a and T9aS3A5b (Table 11) formed Value of m shows Asp, not Asn
C5aT2S2	3	0	8	Ser (1·14), Asp (0·95), Arg (0·90), [Ala (0·12), Gly (0·09), Leu (0·09), Val (0·08)]	Asx	CPB: Arg (+ +) released; residual peptide had m −0·75 Value of m shows Asp, not Asn

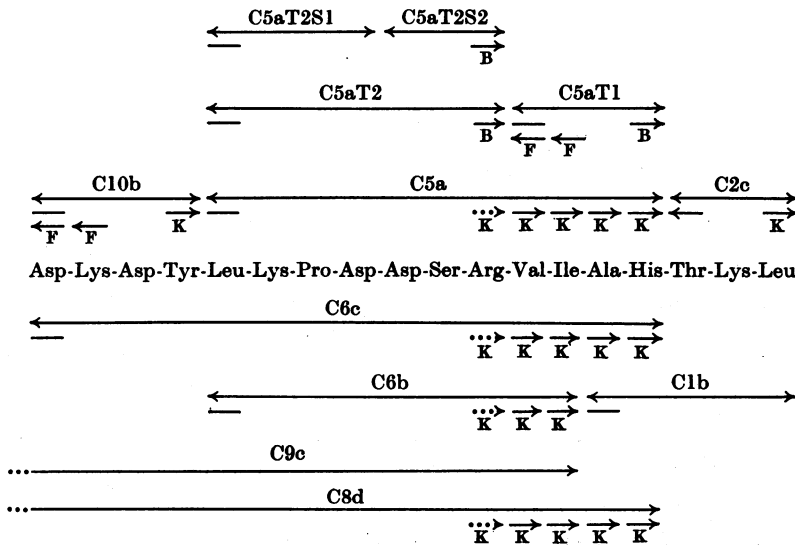* Acid-stable DNS derivative formed; DNS-Val detected only after 96 hr. hydrolysis.

Fig. 8. Amino acid sequence of peptide C5a (Table 24). Definition of the symbols is given in the text.

Table 25. *Peptides formed by tryptic digestion of peptide C8a*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C8aT1	63	+0·04	8	Gly (1·98), Ile (0·87), Ser (1·00), Glu (1·00), Lys (1·09)	Ile	
C8aT2	63	−0·42	8	Val (1·02), Ser (0·98), Thr (0·99), Asp (0·99), Phe (1·01)	Asx	Equivalent to peptide T9bC2 CP: same amino acids released and residual peptides formed as from peptide T9bC2 (Table 12) Value of m shows Asp, not Asn

Table 26. *Phenyl isothiocyanate degradation of peptide C9a*

Experimental details and definition of the symbols are given in the text. The predicted mobilities were derived from unpublished graphs of electrophoretic mobility (m) against molecular weight for all the peptides isolated during the determination of the amino acid sequence of *Pseudomonas* cytochrome-551 (Ambler, 1963b). Similar though more elaborate Tables have been published by Offord (1966).

Peptide	Purification procedure	Electrophoretic mobility (m) Predicted 1 amide	No amide	Observed	Gly	Glu	Tyr	Lys	N-Terminus (DNS)	Proposed sequence
C9a	G6B²3	−0·31	−0·58	−0·28	1·04	3·06	0·96	0·92	Lys	Lys-Glu-Gly-Glu-Gln-Tyr
C9aF1	6	−0·66	−0·99	−0·70	+ +	+ + + +	+	−	Glx	Glu-Gly-Glu-Gln-Tyr
C9aF2	63	−0·41	−0·77	−0·39	1·00	2·41	0·91	−	Gly	Gly-Glu-Gln-Tyr
C9aF3	6	−0·44	−0·84	−0·44	−	+ +	±	−	Glx	Glu-Gln-Tyr
C9aF4	3	0	−0·56	0	−	+	±	−	Glx	Gln-Tyr

(m and m'), N-terminal groups and amino acid compositions as peptides C7g and C7e (Table 15).

CP released Leu (+ +) from peptide C7g to leave a peptide with a different m'; hence peptide C7g has the sequence:

Ser-Ala-Leu

Peptide C7e. This was largely resistant to proteolytic enzymes, the only degradation being partial removal of the histidine by CP or Pronase, which formed a peptide with m −0·46. The sequence was therefore investigated by partial acid hydrolysis. Table 27 shows the properties of peptides formed from peptides C7c and C7e, and Fig. 9 the deduction of the sequence.

Peptide C7f. Digests of oxidized protein also contained small amounts of peptide C7f, which had the same amino acid composition as peptide C7c

except for the addition of a residue of methionine sulphone, and from which CP released Ala (+ +), Ser/Leu (+ + +) and MetSO₂ (+ +).

Peptide C4a. CP released Leu (+ + +) and Thr (+ +), to leave two basic (m +0·48 and +0·53) residual peptides. Trypsin formed two peptides, C4aT1 [(MetSO₂,Lys), m +0·67] and C4aT2, [(Gly,Leu,Thr), m 0]. CP released only Leu (+ +) from peptide C4aT2, to leave a peptide with different m'. The sequence was therefore:

Met-Lys-Gly-Thr-Leu

Small amounts of peptide C2f (Lys,Gly,Thr,Leu) were also present in digests of oxidized protein. Digests of unoxidized protein contained little or no peptide C4a, but much larger amounts of peptide C2f and also free methionine.

Peptide C2b. CP had no action, but CPB released

Table 27. *Peptides formed by partial acid hydrolysis of peptides C7c, C7e and S4a, and by phenyl isothiocyanate degradation of peptide S4a*

Experimental details and definition of the symbols are given in the text. The peptides were hydrolysed with 6N-HCl for 15 min. at 100°.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
C7cA9*	6	−0·52	6	Gly (+ +), Thr (+ +), Phe (+), CySO₃H (+ +), Pro (+)		
C7cA10*	6	−0·61	2	Thr (+ +), Phe (+), CySO₃H (+ +)		CP: Phe (+ +) released; residual peptide had *m* −0·80 (the same as peptide C7cA11)
C7cA11*	6	−0·80	1	Thr (+ +), CySO₃H (+ +)		
S4aA1*	36	+0·75	1	Gly (+), His (+)	Gly	
S4aA2†	36	+0·40	2	Ser (+ +), His (+)	His	
S4aA4*	32	0	2	Gly (+), Pro (+)	Pro	Peptide gave blue colour with isatin reagent on paper
S4aA5	32	0	10	Gly (+ +), Thr (+ +), Phe (+), Pro (+)	Thr	
S4aA6a	32	0	4	Thr (+), Phe (+)	Thr	
S4aA6b	32	0	2	Thr (+ +), Phe (+), Pro (+)	Thr	
S4aF1	6	+0·40	—	Gly (+ +), Ser (+ +), Phe (+), Pro (+), His (+)	Phe	
S4aF2	6	+0·48	—	Gly (+), Ser (+), Pro (+), His (+)	Pro	Peptide gave blue colour with isatin reagent on paper

* Peptides also present in hydrolysate of peptide C7e.
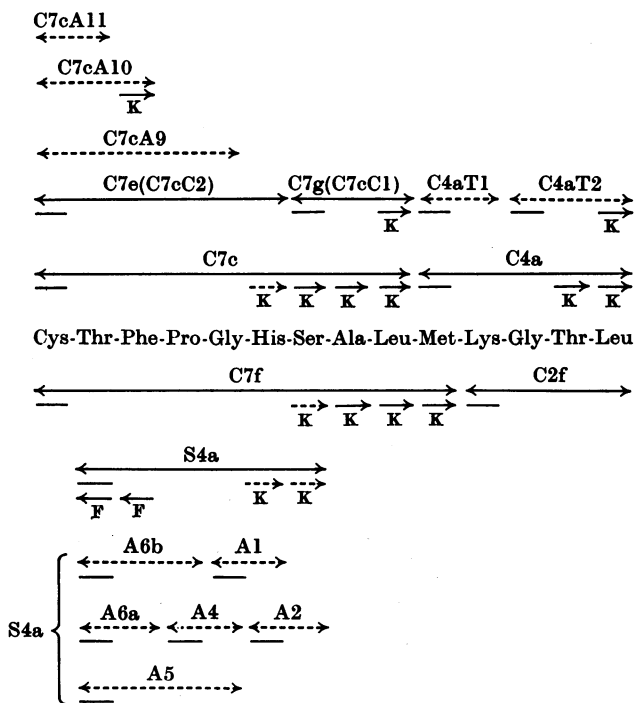† Peptides also present in hydrolysate of peptide C7c.

Fig. 9. Amino acid sequence of peptides C7c, C4a and S4a (Table 27). Definition of the symbols is given in the text.

Table 28. Properties of peptides formed by peptic digestion of azurin

Experimental details and definition of the symbols are given in the text. The peptides are arranged in groups in order of decreasing yield.

Peptide	Electrophoretic mobility (m)	Purification procedure	Approx. yield (%)	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	CySO$_3$H	MetSO$_2$	Pro	Lys	His	Arg	No. of residues	N-Terminus (DNS)
P2	+0·56	G6B3	20	—	—	—	—	—	—	1·00	—	—	—	—	—	—	—	—	0·97	—	—	3	Thr
P3a	+0·50	G6B3	10	2·11	—	0·80*	2·05	—	1·04	2·01	3·02	—	—	—	—	—	0·82*	2·21	1·04	2·00	—	15	Leu
P3b	+0·50	G6B3	10	0·98	1·00	—	2·03	—	1·07	2·06	—	—	—	—	—	1·00	0·95	0·93	2·02	0·94	—	8	MetSO$_2$
P6a	+0·18	G6B3	10	1·99	—	1·92	2·20	0·96	1·06	1·90	1·99	1·06	1·84	—	—	—	0·94	—	0·96	—	—	15	Phe
P7b	0	G6B3	20	0·06	—	0·95	1·05	—	—	—	1·99	1·06	—	—	0·86‡	1·03	—	—	2·14	—	—	13	Ile
P7g†	0	G93	10	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·97	—	—	—	3	‖
P8a	−0·10	G6B3	10	1·04	0·97	—	1·03	—	1·02	1·03	—	—	1·95	—	—	1·04	—	0·97	—	0·95	—	10	Phe
P10	−0·19	G6B3	20	1·08	—	1·06	0·94	—	1·01	—	0·97	3·00	1·00	0·94	—	—	1·01	—	1·98	—	—	13	Asx
P11a	−0·24	G6B3	20	1·91*	—	1·03	—	0·92*	2·05	1·01	1·03	1·05	1·03	—	—	—	1·01	—	0·96	—	—	11	Ile
P14a	−0·47	G6B3	10	—	—	—	0·96	—	—	—	2·04	—	—	—	—	—	—	—	0·96	—	—	3	Asx
P16	−0·60	G6B3	20	1·00	1·80	1·04	—	0·96	0·97	0·99	4·94	4·03	1·02	—	—	0·99	0·91	—	0·96	—	—	19	Ala
P1	+0·59	G6B3	5	—	1·00	0·62*	1·03	0·66*	0·10	1·03	—	0·07	—	—	—	—	—	—	1·02	0·90	0·77*	8	Arg
P4a	+0·40	G6B3	5	0·99	—	—	1·03	—	0·96	0·99	0·07	0·07	0·98	0·70	—	—	0·88	—	0·99	—	—	5	MetSO$_2$
P7d	0	G6B3	5	1·09	—	1·02	1·06	—	—	—	3·11	3·11	1·00	0·84	—	—	0·99	—	1·84	—	—	12	Val
P7e	0	G93	5	—	1·00	—	—	—	—	—	0·07	—	—	—	—	—	0·96	—	—	—	—	3	Tyr
P8c	−0·08	G6B32	5	1·00	—	—	—	—	0·93	1·00	—	—	—	—	—	—	—	—	—	—	—	3	Ser
P9	−0·11	G663	5	3·00	3·76	1·39*	2·98	0·52*	2·10	2·12	5·97	1·01	—	0·88	—	—	1·93	1·01	2·94	0·97	0·94	34	Ala
P13a	−0·39	G63B6	5	1·81*	0·06	1·01	0·07	0·76*	1·99	1·00	2·17	1·00	—	—	—	—	—	—	0·97	—	—	12	Ile
P13c	−0·42	G63	5	2·8	2·6	1·0*	0·8	0·2	1·0	1·0	2·3	1·2	0·88	—	—	—	1·8	—	—	—	—	15–16	Ala
P4c	+0·31	G6B3	<5	0·10	1·02	0·95	2·20	0·56*	1·06	0·89	1·16	1·03	0·96	0·84	—	1·23	—	0·91	1·89	—	0·76*	9	Lys
P6c	+0·12	G636	<5	—	—	0·51*	0·81	—	1·04	1·03	3·87	—	—	0·95	—	—	—	—	2·87	1·00	—	18	Asx
P6e	+0·03	G6B32	<5	0·06	—	1·02	0·92	—	0·89	—	2·00	—	—	—	—	—	—	0·80	1·02	—	0·89	9	¶
P7a	0	G6B3	<5	0·06	1·01	1·03	1·01	—	1·06	—	0·98	—	—	—	—	—	—	—	0·92	—	—	5	Asx
P7f	0	G6B3	<5	1·12	—	—	0·96	—	0·91	—	—	—	—	—	—	—	—	—	—	—	—	4	Ala
P7h	0	G3B3	<5	0·06	0·06	1·05	0·99	—	0·99	0·96	—	—	—	—	§	—	—	—	—	—	—	5	‖
P7j	0	G6B3	<5	1·02	—	1·02	1·03	—	1·01	—	—	—	—	—	—	—	0·06	—	1·87	—	—	9	Val
P9b	−0·09	G6B3	<5	1·04	1·04	—	—	—	0·86	0·96	1·99	3·06	—	—	—	—	—	—	—	—	—	4	Asx
P9e	−0·08	G6B3	<5	2·00	2·00	—	—	—	1·16	1·00	—	—	—	—	—	—	—	—	—	—	—	4	Ser
P11b	−0·22	G63B6	<5	1·26	—	1·02	—	—	0·96	0·14	0·98	2·70	0·18	0·63	—	—	0·79	—	1·72	—	0·76*	12	Asx
P12a	−0·27	G6B3	<5	—	—	—	0·92	0·79	—	—	2·11	0·10	—	0·64	—	—	—	—	0·96	—	0·89	7	Tyr
P12b	−0·27	G63B6	<5	1·02	—	1·03	—	—	—	—	1·98	2·89	0·97	—	—	—	0·79	0·78	—	—	—	9	Ile
P12c	−0·22	G63B6	<5	1·02	—	—	—	—	—	—	0·99	3·02	0·97	—	—	—	1·02	—	1·86	—	—	10	Asx
P12d	−0·29	G63B6	<5	+	+	+*	1·02	—	0·99	+	1·14	1·11	—	—	—	—	+	—	—	—	—	(10)	Val
P12e	−0·30	G63B6	<5	1·78	0·09	0·92	0·15	0·86	1·94	0·92	0·30	3·19	1·03	0·91	—	—	0·06	—	0·98	—	—	10	Ile
P12f	−0·24	G6B3	<5	1·06	—	—	—	—	—	—	2·04	3·19	—	—	—	—	1·02	—	0·99	—	—	8	¶
P12g	−0·35	G63B6	<5	1·00	1·00	—	0·96	—	1·02	—	1·06	1·01	—	—	—	—	0·96	—	0·96	—	—	8	MetSO$_2$
P13b	−0·41	G63B6	<5	0·97	1·90	—	—	—	1·00	—	0·99	—	—	—	—	—	1·05	—	—	—	—	6	Ala
P13d	−0·40	G6B3	<5	2·05	1·03	—	0·95	0·16	—	—	1·04	1·10	—	—	—	—	1·02	—	—	—	—	7	Asx
P13e	−0·45	G63B6	<5	+	1·93	—	—	—	0·93	0·94	1·01	—	—	—	—	—	1·02	—	—	—	—	(5)	¶
P18f	−0·35	G63B6	<5	1·05	—	0·71*	—	—	—	0·94	1·04	—	—	—	—	—	1·03	—	—	—	—	8	Ser
P13g	−0·38	G63B6	<5	1·01	—	—	—	—	—	0·96	1·01	—	—	—	—	—	+	—	—	0·95	—	6	Val
P13h	−0·37	G6B3	<5	1·05	—	—	—	—	—	—	0·17	3·20	—	—	—	—	—	—	—	—	—	5	¶
P14b	−0·42	G63B6	<5	+	+	+	—	—	+	+	+	+	—	—	—	—	++	—	—	—	—	(6)	Val
P14c	−0·42	G63B6	<5	+	+	*	—	—	+	+	+	+	—	—	—	—	++	—	—	—	—	(5)	Asx
P14d	−0·50	G63B6	<5	—	+++	—	—	—	—	—	+	+	—	—	—	—	+++	—	—	—	—	(4)	¶
P14e	−0·51	G03B6	<5	—	+	—	—	—	—	—	+	+	—	—	—	—	+++	—	—	—	—	(4)	Asx
P14g	−0·58	G63B6	<5	—	—	—	—	—	—	—	—	—	—	—	—	—	1·0	—	—	—	—	(3)	¶
P15c	−0·58	G63B6	<5	2·0	0·2	0·8*	—	—	0·2	1·0	1·9	1·2	—	—	—	—	—	—	—	—	—	9	Asx
P18a	−0·97	G63	<5	—	0·95	0·94	—	—	1·01	1·0	1·05	1·64	—	—	—	0·99	—	—	—	—	—	6	Ala

* Low yield due to partial hydrolysis of stable peptide bond.
† Peptide obtained from digest of unoxidized protein.
‡ Tryptophan determined colorimetrically.
§ Peptide contained oxidized tryptophan.
‖ No N-terminal residue detected by DNS method.
¶ N-Terminal analysis not performed.

Lys $(++)$ to leave a neutral $(m\ 0)$ residual peptide. The sequence was deduced to be:

Thr-Leu-Lys

The peptides C2b (Thr-Leu-Lys), C2c (Thr-Lys-Leu) and C2f (Lys-Gly-Thr-Leu) were extremely difficult to separate. Peptides C2b and C2c had identical electrophoretic mobilities at all pH values tried (2, 3·5, 6·5 and 9). Peptide C2f would separate at pH 6·5, but only if low loadings were used. Peptides C2b and C2c had very slightly different R_F values, but peptide C2f ran between them. The final separation method was two successive chromatographic steps, in each of which the paper was developed three times with the solvent, followed by electrophoresis at pH 6·5.

Peptides from peptic digest of azurin

As so much of the polypeptide chain could not conveniently be isolated from tryptic digests, it was necessary to use peptides produced by peptic and subtilopeptidase B digestion and by cyanogen bromide cleavage to establish satisfactorily the relative positions of all the chymotryptic peptides.

The peptic digests were fractionated by gel filtration followed by paper electrophoresis and chromatography (Table 28). All the peptides were soluble, but the mixture was very complex and included peptides that contained as many as 33 residues. Not all the peptides shown in Table 28 could be found in each digest, peptides from the P13c region (Fig. 10) being particularly erratic.

The evidence for the sequence of individual pepsin peptides is given in the order in which they were ultimately shown to occupy in the whole polypeptide chain.

Peptide P16. CP released Ala $(++)$, Thr $(++)$, Asn $(+++)$, Gln $(++)$, Phe $(++)$ and MetSO$_2$ $(++)$, leaving a residual peptide of amino acid composition (Gly, Ala, Val, Ile, Ser, Asp$_3$, Glu$_3$,-CySO$_3$H) and $m\ -0·65$. Chymotrypsin split peptide P16 into two peptides (Table 29); the larger one had all the properties of peptide C13 (Fig. 3), and the sequence of the small one, peptide P16C1, was deduced to be:

Asn-Thr-Asn-Ala

and formed the C-terminus of peptide P16, as shown by the release of the constituent amino acids when CP acted on peptide P16.

Peptic digests also contained peptides P12b, P18a and P8b. Peptides P12b and P18a had all the properties of peptides C13P2 and C13P6 (Table 16), the major products of peptic hydrolysis of peptide C13. Peptide P8b was identical with peptide P16C1.

Peptide P7b. CP released Val $(++)$, Thr $(++)$, Asn $(++)$, Gln $(++)$ and Phe $(++)$, to leave a neutral $(m\ 0)$ peptide with m' different from that of

---

Table 29. *Peptides formed by chymotryptic digestion of peptide P16*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
P16C1	63	−0·09	12	Ala (0·99), Thr (1·01), Asp (2·00), [Gly (0·06)]	Asx	CP: Ala $(++)$ and Asn $(+++)$ released; residual peptide (Thr,Asx) had same m' as peptide P16C1, but different m''. PTC degradation: one cycle: peptide (Ala,Thr,Asx), N-terminal Thr (DNS); two cycles: peptide (Ala-Asx), N-terminal Asx (DNS); three cycles: free Ala. Values of m show both residues Asn, not Asp
P16C2	63	−0·70	12	Gly (0·97), Ala (1·00), Val (0·95), Ile (0·86), Ser (0·99), Asp (3·21), Glu (4·08), Phe (0·90), MetSO$_2$ (0·95), CySO$_3$H (1·08)	Ala	Equivalent to peptide C13

Table 30. *Peptides formed by tryptic digestion of peptide P7b and by chymotryptic digestion of peptide P7bT2b*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
P7bT1	32	0	10	Val (1·04), Ile (0·99), Thr (0·97), Asp (1·02), Lys (0·96)	Ile	CPB: Lys (+ +) released; residual peptide had m −0·44 Pronase (only acidic peptides examined): peptide P7bT1Q2, m −0·53, (Val,Thr,Asx), N-terminal Thr (DNS), and peptide P7bT1Q3, m −0·66, (Val,Asx), N-terminal Val, (DNS), converted into free Asp by one cycle of PTC degradation Values of m show Asp, not Asn
P7bT2a	32	0	10	Ser (0·92), CySO₃H (1·12), Lys (0·96)	Ser	CPB: Lys (+ + +) released; residual peptide had m −0·80
P7bT2b	32	0	10	Val (0·90), Thr (1·02), Asp (1·09), Glu (1·03), Phe (0·95)	Glx	CP: Val (+ + +), Thr (+ + +), Asn (+ + +), Gln (+ + +) and Phe (+ + +) released; peptide completely degraded Chymotrypsin: see below. Values of m and CP results show Asn and Gln, not Asp and Glu
P7bT2bC1	B3	0	8	Val (+ +), Thr (+ +), Asp (+ +)	Thr	CP: Asn (+ +) released; residual peptide had different m′
P7bT2bC2	B3	0	8	Glu (+ +), Phe (+ +)	Glx	

peptide P7b. Table 30 shows the properties of the peptides formed by the action of trypsin on peptide P7b. The sequence was deduced to be:

Ile-Thr-Val-Asp-Lys-Ser-Cys-Lys-Gln-Phe-Thr-Val-Asn

Some digests also contained the peptide P4b, with properties that showed it was derived from the same region as was peptide P7b.

Peptide P3a. CP had very little action. Trypsin produced the peptides shown in Table 31. The sequence was deduced to be:

Leu-(Gly,Leu,Ser,Asx,Pro,Pro₂,His)-Lys-Asn-Val-Met-Gly-His-Asn

Peptide P7g. Small amounts of a neutral fluorescent peptide that on hydrolysis yielded valine, leucine and the fluorescent bands characteristic of oxidized tryptophan were isolated from digests of

oxidized protein. From digests of unoxidized protein, moderate amounts of peptide P7g were

isolated, together with much smaller amounts of peptide P7h (Table 28). The amino acid compositions of these peptides showed that the sequence around the single tryptophan residue was:

(Val,Leu,Trp)-(Ser,Thr)

No N-terminal residue was detected in either of these peptides by the DNS method.

P13c region. The region of polypeptide chain that

Table 31. *Peptides formed by tryptic digestion of peptide P3a and by subtilopeptidase B digestion of peptide P3aT2*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
P3aT1	63	+0·45	12	Gly (1·04), Leu (2·08), Ser (1·02), Asp (1·04), Pro (1·87), Lys (0·98), His (0·94)	Leu	CPB: Lys (+ + +) released; residual peptide had m +0·23 Equivalent to peptide T4aC2 (Table 8)
P3aT2	63	+0·30	12	Gly (1·12), Val (0·89), Asp (1·96), MetSO₂ (0·89), His (0·91)	Asx	CP: Asn (+ + + +), Gly (+ + +), His (+ +) and MetSO₂ (+) released; residual peptides had m +0·45 and +0·05 LAP: Asn (+ +), Val (+ + +), MetSO₂ (+ +), Gly (+) and His (+ +) released; main residual peptide had same m' as peptide P3aT2S2 (below) Value of m and exopeptidase results show both residues Asn, not Asp
P3aT2S1	36	+0·74	3	Gly (+ +), His (+ +)	Gly	
P3aT2S2	36	+0·55	4	Gly (+ +), Asp (+ +), His (+ +)	Gly	
P3aT2S3	32	0	9	Val (+ + +), Asp (+ + +), MetSO₂ (+ + +)	Asx	Equivalent to peptide C3aT2 (Table 20)

...Ser-Thr-Ala-Ala-Asp-Met-Gln-Gly-Val-Val-Thr-Asp-Gly-Met-Ala-Ser-Gly-Leu-Asp-Lys-Asp...
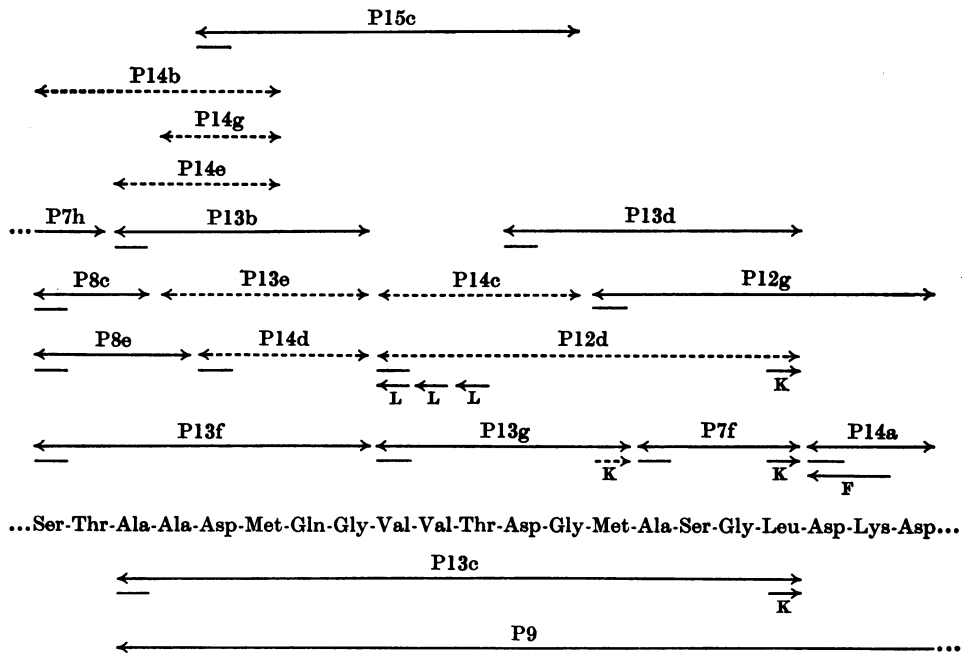
Fig. 10. Amino acid sequence of the P13c region. Definition of the symbols is given in the text.

yields peptide C12a on chymotryptic hydrolysis (Fig. 7) gave rise to an even more complex mixture of peptides in peptic digests. About 20 distinct peptides have been found in digests of oxidized or unoxidized protein (Fig. 10). The peptides were very difficult to separate, as most of them had electrophoretic mobilities in the range −0·3 to −0·5; also, the yield of any one peptide was low and the peptides obtained varied from digest to digest. The peptide most studied was peptide P13c (Fig. 10), for which it was unfortunately not found possible to get a satisfactory stoicheiometrical amino acid analysis. CP released Leu (+ +) from peptide P13c, and on subtilisin digestion it yielded

Table 32. *Amino acids released from peptide P1 by various treatments*

Experimental details are given in the text.

	Amino acids released							
Treatment	Ala	Val	Leu	Ile	Thr	Lys	His	Arg
Acid hydrolysis for 24 hr.	1·00	0·62	1·03	0·66	1·03	1·02	0·90	0·77
LAP*	0·16	0·46	0·07	0·43	0·05	0·08	0·07	0·51
CP (I)†	0·31	0·07	0·93	0·32	0·60	+++	++	±
CP (II)‡	0·85	0·69	1·00	0·88	0·81	+++	+++	+++

* 25 μg. of LAP/0·08 μmole of peptide for 5 hr. at 37°.
† 10 μg. of CP/0·07 μmole of peptide for 5 hr. at 37°.
‡ 60 μg. of CP/0·07 μmole of peptide for 5 hr. at 37°.

Table 33. *Peptides formed by subtilopeptidase B digestion of peptide P11a*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
P11aS1	32	+0·04	>10	Asp (+ +), Lys (+ +)	Lys	Value of m shows Asp, not Asn
P11aS5	32	0	4	Val (+ +), Ser (+ +), Thr (+ +)		CP: Thr (+) released; residual peptide had different m′
P11aS6b	32	0	8	Val (+ +), Thr (+ +), Phe (+ +)	Val or Phe	DNS-amino acid identified only by electrophoresis at pH 4·38 CP: Phe (+ +) released; residual peptide had different m′
P11aS7	32	0	15	Val (+ +), Ser (+ +), Thr (+ +), Phe (+ +)	Ser	CP: Phe (+ + +) and Thr (+) released; two residual peptides, the main one with same m′ as peptide P11aS5
P11aS8	36	−0·43	30	Gly (1·87), Ile (0·80), Ser (1·01), Glu (0·99)	Ile	CP: no action PTC degradation: one cycle: peptide had m −0·51, (Gly,Ser,Glx), N-terminal Gly (DNS); two cycles: peptide had m −0·60, (Gly,Ser,Glx), N-terminal Ser (DNS); three cycles: peptide had m −0·70, (Gly,Glx), N-terminal Gly (DNS); four cycles: free Glu Values of m show Glu, not Gln

all the peptides isolated from the digest of peptide C12b with this enzyme (Fig. 7) except peptides C12bS2a and C12bS2e.

Peptide P14a. Two cycles of PTC degradation converted the peptide into aspartic acid. The acidic mobility (*m*) despite the presence of a lysine residue indicated that no amide groups were present. The sequence was therefore:

Asp-Lys-Asp

Peptide P12a. CP had little action, but LAP released Tyr (+ +) and Leu (+ +). Although the peptide contains lysine, the mobility (*m*) was towards the anode and so no amide groups were present. The sequence was therefore:

Tyr-Leu-(Ser,Asp₂,Pro,Lys)

Peptide P1. The amino acids released by the action of CP and LAP were quantitated (Table 32) and from these results the sequence was deduced to be:

Arg-Val-Ile-Ala-(Leu,Thr,Lys,His)
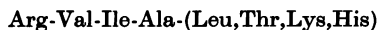
Peptide P9. The amino acid composition of peptide P9, which was found in digests of oxidized protein, was the sum of those of peptides P13c,

P14a, P12a and P1. Peptide P9 was the biggest peptic peptide isolated and corresponded to very nearly the same region of polypeptide chain as that of the biggest chymotryptic peptide, C8d.

Peptide P11a. The actions of CP and trypsin were the same as those on peptide C8a (Table 25), which had the same mobilities, *N*-terminal residue and amino acid composition. When digested with subtilisin, the peptides shown in Table 33 were formed, allowing the sequence shown in Fig. 11 to be deduced.

Peptide P10. CP released Phe (+ + +), Tyr (+ + +), MetSO₂ (+ + +) and Gln (±); the predominant residual peptide had *m* −0·22. The peptide was split by pepsin to give the products shown in Table 34, which also shows the products of tryptic digestion. From this information the sequence shown in Fig. 11 was deduced.

The region covered by peptides P11a and P10 also gave rise to several other peptides in peptic digests. Those isolated in sufficient amount were analysed for amino acid composition and end groups; Fig. 11 shows how they can be fitted into the sequence.

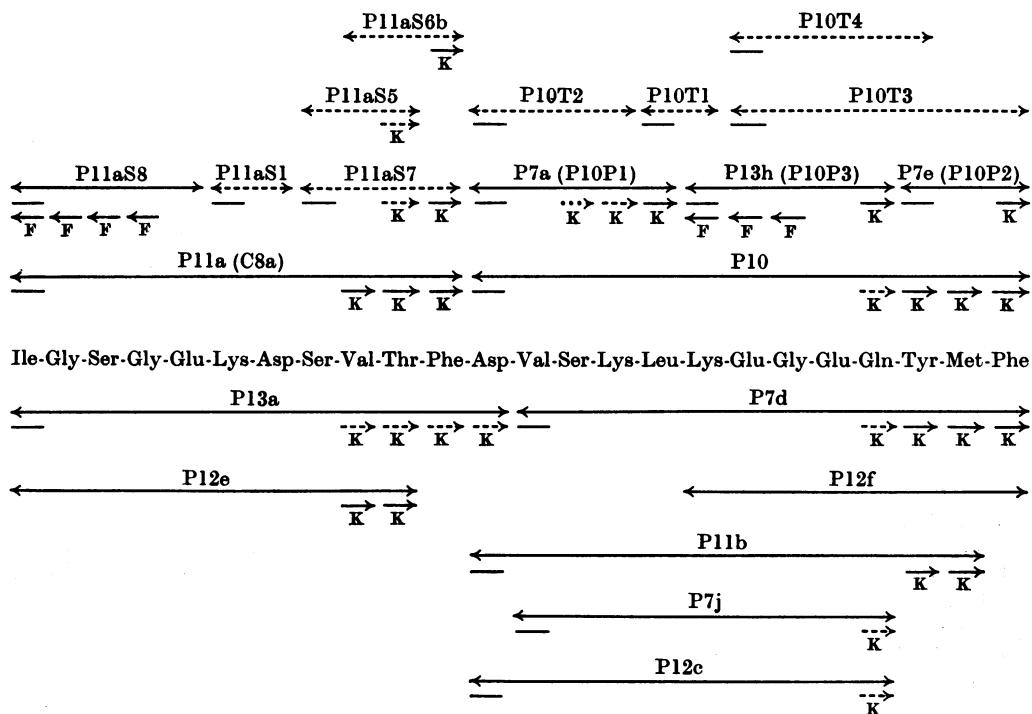*Peptide P8a.* CP released Ala (+ + +), Leu (+ + +), Ser (+) and His (+). After one cycle of

Fig. 11. Amino acid sequence of peptides P11a and P10 (Tables 33 and 34). Definition of the symbols is given in the text.

Table 34. *Peptides formed by tryptic digestion and by further peptic digestion of peptide P10*

Experimental details and definition of the symbols are given in the text.

Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
P10T1	63	+0·70	8	Leu (+ +), Lys (+ +)	Leu	Equivalent to peptide T1
P10T2	63	0	12	Val (+ + +), Ser (+ + +), Asp (+ +), Lys (+ + +)	Asx	Equivalent to peptide T9bC1 (Table 12)
P10T3	63	−0·60	10	Gly (+ +), Glu (+ + + + +), Phe (+ +), Tyr (+ +), MetSO$_2$ (+ +)	Glx	Equivalent to peptide T8aP5 (Table 10)
P10T4	63	−0·74	2	Gly (+ +), Glu (+ + + +), Tyr (+)	Glx	Formed by 'chymotryptic-like' activity in the trypsin Equivalent to peptide T8aC4 (Table 10)
P10P1	63	0	12	Val (1·00), Leu (1·01), Ser (1·03), Asp (0·89), Lys (1·01)	Asx	Equivalent to peptides C7a and P7a CP: Leu (+ + +), Lys (+ +) and Ser (+) released; three residual peptides had m 0 (different m'' from that of peptide P10P1), (Val,Ser, Asx,Lys) (+ +), m −0·61, (Val,Ser,Asx), (+ +), and m −0·69, (Val,Asx), (+ +) Values of m show Asp, not Asn
P10P2	63	0	12	Phe (0·97), Tyr (0·80), MetSO$_2$ (1·03)	Tyr	Equivalent to peptide P7e CP: Phe (+ +) released; residual peptide had different m'
P10P3	63	−0·34	12	Gly (1·03), Glu (3·15), Lys (0·81)	Lys	Equivalent to peptide P13h CP: some Gln (+) released; residual peptide had m −0·40, but much unchanged peptide left PTC degradation: one cycle: peptide had m −0·84, (Gly,Glx) N-terminal Glx (DNS); two cycles: peptide had m −0·52, (Gly,Glx), N-terminal Gly (DNS); three cycles: peptide had m −0·61 CP result and m values show sequence to be: Lys-Glu-Gly-Glu-Gln

PTC degradation the peptide had the same m', N-terminal residue and R_F as peptide C7c (Fig. 9). Chymotryptic digestion produced two peptides, one identical with peptide C7g, and the other with N-terminal phenylalanine and the same amino acid composition as peptide C7e, except that it contained an additional residue of phenylalanine. The sequence was therefore:

Phe-Cys-(Gly,Thr,Phe,Pro)-His-Ser-Ala-Leu

This peptide was scarcely present in digests of unoxidized protein, being replaced by one identical (after oxidation) with peptide C7c.

Peptide P3b. CPB released Lys (+ + +) and a neutral amino acid or peptide (+ + +), to leave a residual peptide of m +0·42. When peptide P3b was digested with a mixture of trypsin and chymotrypsin, the peptides shown in Table 35 were formed.

Peptic digests of azurin also contained varying amounts of peptide P4a (identical with peptide C4a; Fig. 9) and peptide P2 (identical with peptide

Table 35. *Peptides formed by mixed tryptic and chymotryptic digestion of peptide P3b*

Experimental details and definition of the symbols are given in the text.

Peptide	Purifica-tion procedure	Electro-phoretic mobility (*m*)	Relative yield	Amino acid composition	*N*-Terminus (DNS)	Other details
P3bCT1a	62	+0·54	10	MetSO₂ (+), Lys(+)	MetSO₂	
P3bCT1b	62	+0·54	10	Leu (+), Thr (+), Lys (+)	Thr	Equivalent to peptides P2, C2b and T3C1 (Table 7), having same R_F CPB: Lys (++) released
P3bCT2	63	0	10	Gly (++), Leu (++), Thr (++)	Gly	CP: Leu (+++) released; residual peptide had different *m'*

Table 36. *Peptides formed by subtilopeptidase B digestion of azurin*

Definition of the symbols is given in the text. Only tri- and higher peptides purified with a yield greater than about 5% are shown. In addition about 20 dipeptides were isolated with at least this recovery. The sequences were deduced from electrophoretic mobility (*m*), amino acid composition, *N*-terminal residue (DNS method) and the sequence of the complete protein (Fig. 15).

Peptide	Approx. yield (%)	Amino acid sequence	Peptide	Approx. yield (%)	Amino acid sequence
S3a	30	Ser-His-Pro-Gly-Asn-Leu-Pro-Lys	S10a	20	Gly-Glu-Lys-Asp-Ser-Val-Thr-Phe
S3b	10	Thr-Leu-Lys	S11a	10	Lys-Glu-Gly-Glu-Gln-Tyr
S3c	20	Gly-His-Asn	S12a	10	Gly-Glu-Lys-Asp
S4a	20	Thr-Phe-Pro-Gly-His-Ser	S12b	5	Gly-Asn-Asp-Gln-Met-Gln-Phe
S5	5	Asp-Ser-Arg-Val-Ile-Ala-His	S13a	20	Gly-Asn-Asp-Gln-Met
S7g	5	Asn-Val-Met	S13b	20	Gly-Val-Val-Thr-Asp
S7i	5	Thr-Val-Asn	S13d	5	Ala-Ile-Thr-Val-Asp
S7r	5	Asn-Thr-Asn	S13e	5	Val-Thr-Asp-Gly-Met
S7t	10	Leu-Ile-Gly-Ser	S14a	5	Gly-Asn-Asp-Gln
S7u	5	Trp-Val-Leu	S14b	20	Ala-Asp-Met-Gln
S8a	20	Asp-Lys-Asp-Tyr-Leu-Lys-Pro-Asp	S15a	10	Asp-Met-Gln
S9a	10	Leu-Lys-Glu-Gly-Glu-Gln-Tyr	S16a	30	Asp-Val-Ser

C2c and different from peptide C2b). Small amounts of a peptide (P6a) with *N*-terminal phenylalanine and amino acid composition equal to the sum of peptides P8a and P4a have also been isolated. The sequence of peptide P3b was deduced to be:

Met-Lys-Gly-Thr-Leu-Thr-Leu-Lys

Peptides from subtilisin digests of azurin

Subtilisin digestion (of unoxidized protein) gave very good yields of peptides from regions that did not give high-yield peptides in any other enzymic digest. Table 36 shows peptides containing three or more residues that have been isolated with a yield of more than 5%, and Table 37 gives the properties of peptides that have been further investigated.

Peptide S3a. This peptide had the same end groups, amino acid composition, mobilities and gave the same products on partial acid hydrolysis as peptide C3aT1 (Table 20), but could be prepared more easily in much higher yield from the subtilisin digest. It was used with C3aT1 to obtain the results shown in Table 21 and Fig. 6.

Peptide S7u. This was the only peptide in the subtilisin digest to contain tryptophan. It was degraded by LAP to give the amino acid analysis shown in Table 37. One cycle of PTC degradation converted it into a non-fluorescent peptide of composition (Val,Leu), which had the same *m'* as synthetic Val-Leu, and two cycles converted it into free leucine. The sequence was therefore:

Trp-Val-Leu

Peptide S13b. This had the same amino acid composition, *N*-terminal residue and mobilities as

Table 37. *Properties of some of the peptides formed by subtilopeptidase B digestion of azurin*

Experimental details and definition of the symbols are given in the text.

Peptide	Electrophoretic mobility (m)	Approx. yield (%)	Purification procedure	Amino acid composition														No. of residues	N-Terminus (DNS)
				Gly	Ala	Val	Leu	Ser	Thr	Asp	Glu	Phe	Trp	Pro	Lys	His		
S3a	+0·39	30	G6B3	1·01	—	—	1·00	0·94	—	0·99	—	—	—	2·06	1·02	0·98	8	Ser
S4a	+0·31	20	G6B3	1·07	—	—	—	1·02	0·98	—	—	1·01	—	0·91	—	1·01	6	Thr
S7u*	0	5	G3	—	—	1·02	0·98	—	—	—	—	—	0·63	—	—	—	3	§
S13b	−0·37	20	G6B3														5	Gly
(i)†				1·00	0·07	0·99‡	—	—	1·01	1·00	—	—	—	—	—	—		
(ii)*				1·00	0·07	1·93	—	—	1·00	1·06	0·06	—	—	—	—	—		

* Amino acid composition determined after complete degradation with LAP.
† Sample hydrolysed with 6N-HCl for 24 hr. at 37°.
‡ Low yield due to partial hydrolysis of stable peptide bond.
§ N-Terminal analysis not performed.

peptide C12bS3a (Table 22), and could readily be prepared in good yield. LAP degraded the peptide completely, and the results of analysis are shown together with those after acid hydrolysis (6N-hydrochloric acid for 24 hr. at 105°) in Table 37. The remaining properties of the peptide are given in combination with those of peptide C12bS3a, in Table 22.

Peptide S4a. CP released Ser (+) and His (+); no residual peptide was detected, but there was some unchanged material. On partial acid hydrolysis, under the same conditions as used for peptides C7c and C7e, the peptides shown in Table 27 were formed. The properties of the peptides formed after one and two stages of PTC degradation are also shown. From these results the sequence was deduced to be (Fig. 9):

Thr-Phe-Pro-Gly-His-Ser

Peptides from cyanogen bromide cleavage of azurin

Azurin contains six methionine residues/mol., and two of these are present in the region that gave the least satisfactory peptides after chymotryptic or peptic digestion (Figs. 7 and 10), and that was lost in the 'core' after tryptic digestion. Cyanogen bromide was therefore used to degrade the protein, and it was possible to isolate and characterize all the predicted peptides (Fig. 15 and Table 38).

The peptides were fractionated by gel filtration and paper electrophoresis (Figs. 12 and 13). All the peptides that contained homoserine (i.e. all except the C-terminal peptide X1) were present in two electrophoretically distinct but interconvertible forms. Although the proportion of the more acidic form (containing C-terminal homoserine rather than the lactone) increased during purification, there was no suggestion that interconversion actually took place on paper during electrophoresis. The only peptide bonds cloven were the six on the C-terminal side of the methionine residues.

'Peptide' XI/II (Table 38) could not be purified at all by electrophoresis, being insoluble except in high concentrations of formic acid.

The sequence of peptide X4/8a was further investigated, as it bridged a gap in the sequence that could not be satisfactorily investigated in enzymic digests. CP released Hoser (+ +), Asp (+ +), Ala (+) and Thr (+). The peptides produced by the action of chymotrypsin are shown in Table 39, and the sequence deduced is shown in Fig. 14.

Cystine bridge of azurin

The apoprotein (100 mg.) was digested with pepsin and the products were separated by gel

Table 38. Properties of peptides formed by cyanogen bromide treatment of azurin

Experimental details and definition of the symbols are given in the text. The values in parentheses are the numbers of residues predicted from the sequence of the protein shown in Fig. 15.

Peptide	Electrophoretic mobility (m): Lactone form	Hoser form	Yield (%)	Purification procedure	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	Trp	$CySO_3H$	Met	Hoser	Hoser lactone	Pro	Lys	His	Arg	No. of residues	N-Terminus (DNS)
Whole mixture	—	—	(100)	—	11·0 (11)	7·5 (7)	9·6 (10)	9·7 (10)	3·8 (4)	9·7 (10)	9·5 (10)	18·0 (18)	10·1 (10)	5·7 (6)	2·1 (2)	* (1)	2·8 (1)	0·3	†	2·5 (6)	4·1 (4)	10·1 (11)	3·4 (4)	1·0 (1)	(128)	(Ala)
X1	+0·66	—	47	G6	1·03 (1)			2·02 (2)			2·00 (2)											1·96 (2)			7	Lys
X4/8a	+0·14	−0·08	36	G6	0·80 (1)	2·06 (2)	0·96 (1)	0·97 (1)		1·01 (1)	1·00 (1)	2·00 (2)		0·07			0·03	0·89 (1)					0·79 (1)		11	Gly
X4/8b	+0·14	−0·08	22	GoxG'6	1·07 (1)	1·08 (2)	0·03	1·10 (1)		1·02 (1)	1·03 (1)	0·07 (2)		2·86		(1)	0·98	0·88 (1)			0·97 (1)		0·90 (1)		12	Phe
X7/8	0	−0·09	16	GoxG'66‡	4·04 (4)	2·11 (2)	2·81§ (3)	3·99 (4)	1·74§ (2)	4·66 (5)	2·02 (2)	6·07 (6)	4·10 (4)	1·15 (3)	1·32 (1)	—	0·08 (1)	0·42		0·45	1·04 (1)	5·73 (6)	1·01 (1)	1·01 (1)	45	Ala
X7/9	0	−0·35	57	G	2·02 (2)	0·04	1·12§ (1)	0·04		—	0·99 (2)	0·98 (1)	1·01 (4)	0·02 (1)			—	0·90 (1)			(1)	(2)			8	Glx
X11	−0·66	−0·82	40	GoxG'6	1·04 (2)	0·98 (1)	0·97 (1)		0·99 (1)	0·98 (1)	3·00 (1)	3·00 (3)	3·06 (1)				0·98 (1)	0·89 (1)							13	Ala
XI/II	Stuck	Stuck	32	GoxG'	1·32 (1)	1·24 (1)	2·85 (3)	2·14 (2)	1·08 (1)	2·14‡ (3)	2·85 (6)	6·00 (2)	2·30 (2)	1·86 (2)			0·93 (1)	†			1·78 (2)	2·87 (3)	0·98 (1)		31	‖

* Tryptophan present, but not determined quantitatively.

† Homoserine present, but not sufficiently separated from glutamic acid for quantitative determination.

‡ Peptide treated with 2 N-NH₃ solution (see the text) between electrophoretic purification steps, to convert most into the homoserine form.

§ Low yield owing to partial hydrolysis of stable peptide bond.
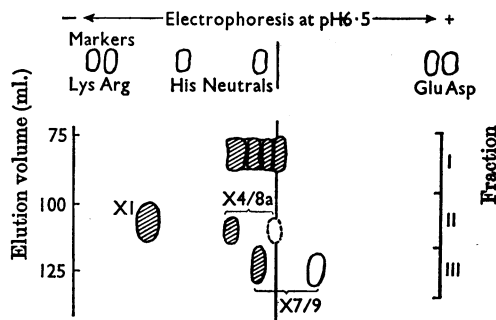
‖ N-Terminal analysis not performed.

Fig. 12. Separation of peptides formed by CNBr treatment of *Pseudomonas* azurin (12 μmoles) by gel filtration followed by paper electrophoresis. Gel filtration (in 50%, v/v, formic acid) was through Sephadex G-25 (120 cm. × 1·5 cm. diam. column). Then 0·04 ml. of each 2·5 ml. fraction was subjected to electrophoresis at pH 6·5 (60 v/cm. for 1 hr.). ε-DNP-lysine was eluted after 180 ml. Definition of the symbols is given in the text.
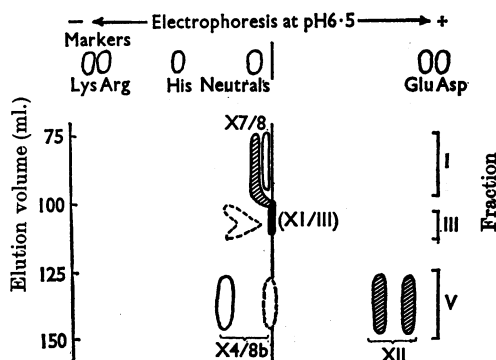

Fig. 13. Further fractionation of peptides formed by CNBr treatment of *Pseudomonas* azurin. Fraction I (Fig. 12) was freeze-dried, oxidized with performic acid and separated by gel filtration in 50% (v/v) formic acid through Sephadex G-50 (120 cm. × 1·5 cm. diam. column). Then 0·05 ml. of each 2·5 ml. fraction was subjected to paper electrophoresis at pH 6·5 (60 v/cm. for 1 hr.). ε-DNP-lysine was eluted after 185 ml. Definition of the symbols is given in the text.

filtration through Sephadex G-25 in 5% (v/v) formic acid, on a 120 cm. × 1·5 cm. diam. column. Samples (0·05 ml.) of each 4 ml. fraction were oxidized with performic acid, hydrolysed (6 N-hydrochloric acid for 24 hr. at 105°) and examined for cysteic acid content by paper electrophoresis. The cysteic acid was approximately equally divided between fraction I (eluted between 120 and 140 ml.) and fraction II (140–160 ml.), with very little in the later fractions. Cystine and cysteine peptides

present in unoxidized fractions I and II were isolated after detection by diagonal electrophoresis at pH 6·5 or 2·0 (Brown & Hartley, 1966; Milstein, 1966). The peptides isolated are described in Table 40. Fraction I contained peptide PIa, which could be purified by electrophoresis at pH 2·0 and which on performic acid oxidation formed peptides PIx and PIy; these peptides were easily purified by electrophoresis, and proved to be identical with peptides P7b and P18a (Table 28) isolated from peptic digests of oxidized protein. In fraction II the mobilities of two sets of peptides were altered by oxidation. A neutral fraction (m 0) formed peptide PIIv (equivalent to peptide P4b) and peptide PIIw (equivalent to peptide P18a), and a slightly basic fraction formed peptide PIIx (equivalent to peptide C7c) and (in much smaller amount) peptide PIIy (equivalent to peptide P8a).

These results indicate that the cysteine residues at positions 3 and 26 are joined in the apoprotein to form a disulphide bridge.

Attempts have been made to obtain a reaction of the cysteine residue at position 112 with bromoethylamine or iodoacetic acid. The native protein, apoprotein and apoprotein dissolved in 8 M-urea have all been treated with the alkylating agents, but in no case has reaction occurred, as judged by the absence of S-β-aminoethylcysteine or S-carboxymethylcysteine from acid hydrolysates of treated protein. When reduced protein was prepared by treatment with thioglycollic acid in 8 M-urea solution and the protein then treated with an excess of bromoethylamine, only 2·0 residues of S-β-aminoethylcysteine/mol. of protein were found after acid hydrolysis. After tryptic digestion, no peptide containing S-β-aminoethylcysteine and corresponding to residues 104–112 could be detected. As the molecular weight calculated from sedimentation and diffusion (13 000; Ambler, 1960) is very nearly the same as that from amino acid analysis (assuming one arginine residue/mol.), it would seem impossible for the cysteine residue at position 112 to be involved in an interchain disulphide bridge. As a result it appears likely that it exists as a free thiol group, but is prevented from reacting with alkylating reagents by the bulky side chains on the adjacent residues.

Deduction of the amino acid sequence of azurin

The amino acid sequence of azurin is shown in Fig. 15, together with the peptides needed to establish this overall sequence. The size of the overlap in each case is sufficient to ensure that there is no ambiguity. Except for peptide C1a (see below), all the evidence from yield and sequence of the peptides described in Tables 4, 15, 28, 36 and 38 is compatible with this scheme.

Table 39. *Peptides formed by chymotryptic digestion of peptide X4/8a*

Experimental details and definition of the symbols are given in the text.

| Peptide	Purification procedure	Electrophoretic mobility (m)	Relative yield	Amino acid composition	N-Terminus (DNS)	Other details
X4/8aC1	6	+0·40	8	Gly (+ + +), Asp (+ + +), His (+ + +), Trp present	Gly	Equivalent to peptide C4b CP: Trp (+ + +) and Asn (+ +) released; two residual peptides had m +0·78 and +0·59
X4/8aC2	6	+0·25	1	Gly (+ +), Val (+ +), Leu (+ +), Asp (+ +), His (+ +'), Trp present		
X4/8aC3a	63	0	1	Val (+), Leu (+)		Same m' and R_F as synthetic Val-Leu
X4/8aC4	6	−0·28	6	Ala (1·98), Val (0·96), Leu (1·01), Ser (0·99), Thr (0·99), Asp (1·09), Hoser (0·83), [Gly (0·18), Phe (0·08)]	Val	Value of m shows that peptide is in the homoserine form and contains Asp, not Asn
X4/8aC5	6	−0·39	2	Ala (+ + +), Ser (+), Thr (+), Asp (+), Hoser (+)	Ser	
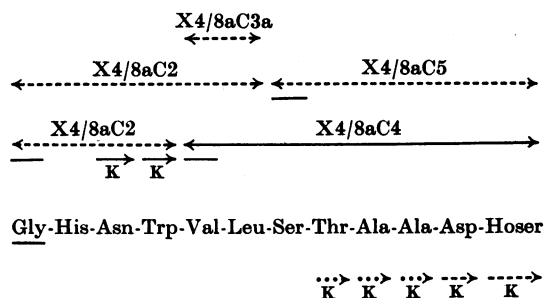
Fig. 14. Amino acid sequence of peptide X4/8a (Table 39). Definition of the symbols is given in the text.

DISCUSSION

Elucidation of the sequence. The general experimental approach to the elucidation of the sequence was very similar to that used in investigating *Pseudomonas* cytochrome-551 (Ambler, 1963b). The principal improvements in methodology have been the development of the DNS method (Gray & Hartley, 1963a,b) as a quick, easy and reliable (but not quantitative) method of determining the *N*-termini of peptides, and great improvements in speed (Spackman, 1963) and sensitivity of automatic amino acid analysis. It was again found that the most rapid and economic method of establishing sequences was to degrade peptides with endopeptidases of increasingly wide specificity, rather than to use chemical methods such as subtractive PTC degradation. Chemical methods were generally used only to elucidate regions of sequence (mostly around proline residues) where the purely enzymic approach failed.

Azurins have a high methionine content and have proved to be admirable proteins for investigation by the cyanogen bromide method (Gross & Witkop, 1961). It has been found possible to develop methods (Ambler, 1965) that avoid the trouble caused by the easy interconversion of the homoserine and homoserine lactone forms of the peptides.

The yields of tryptic, peptic and chymotryptic peptides were generally lower than those obtained from the smaller cytochrome molecule, but this is explained by the greater number of steps necessary to purify peptides from the more complex mixture. No attempts have been made to correct these yields for losses that certainly occur in purification, such as material lost in marker strips and material not eluted from paper, as the estimates would be unrealistic and deceptive.

Amide residues were recognized initially by the electrophoretic mobilities of peptides at pH 6·5 (m). The region of sequence containing the residue was degraded to as small a peptide as possible (containing only one acidic or amide residue) and the assignation made on the basis of whether the value of m was positive, negative or very near zero. With the hexapeptide C9a, which yielded three residues of glutamic acid on hydrolysis, the number and position of the amide group were determined by following the change in mobility of the peptide as residues were successively removed by PTC degradation. In Table 26 are shown the predicted and observed mobilities at each step, clearly

Table 40. Peptides formed by peptic digestion of azurin that had altered electrophoretic mobilities after performic acid oxidation

Experimental details and definition of the symbols are given in the text. Peptides PI and PIa were successive stages in the purification of the cystine peptide that formed peptides PIx and PIy when oxidized with performic acid. Samples were oxidized before acid hydrolysis. Peptides PIIv/PIIw and PIIx/PIIy were not obtained in a pure enough form for analysis before oxidation.

Peptide	Purification procedure	Electrophoretic mobility (m)	Gly	Ala	Val	Leu	Ile	Ser	Thr	Asp	Glu	Phe	Tyr	CySO₃H	MetSO₂	Pro	Lys	His	N-Terminus (DNS)	Equivalent peptides (Table 28)
														Amino acid composition						
PI	G	Stuck	0.53	0.99	2.60	0.34	0.82	1.56	1.95	3.49	1.80	0.79	—	2.00*	0.22	—	1.76	0.22	Ala, Ile	
PIa	G2	Stuck	0.10	0.97	2.96	0.05	0.87	2.23	1.78	3.20	2.15	0.87	—	2.05	—	—	1.88	—	Ala, Ile	
PIx	G2ox63	0	—	—	2.02	—	0.93	1.12	1.95	2.16	1.05	0.92	—	1.04	—	—	1.82	—	Ile	P7b
PIy	G2ox63	−0.97	—	0.97	0.99	—	—	0.99	—	0.98	1.05	—	—	1.02	—	—	—	—	Ala	P18a
PIIv	G6ox63	+0.31	0.08	—	0.89	—	—	1.10	—	1.08	1.10	0.98	—	1.05	—	—	1.81	—	Lys	P4b
PIIw	G6ox6	−0.97	—	++	++	—	—	++	0.98	++	++	—	—	++	—	—	—	—	Ala	P18a
PIIx	G6ox63	−0.10	1.05	1.05	—	1.02	—	1.07	1.03	++	++	0.99	—	1.00	—	—	0.91	0.89	CySO₃H	†
PIIy	G6ox63	−0.05	1.01	0.95	—	0.95	—	1.01	1.01	0.06	—	1.44	—	0.74	—	—	0.74	0.77	Phe	P8a

* Analysis calculated assuming 2·0 residues of cysteic acid.

† Same amino acid composition, electrophoretic mobility and N-terminal residue (DNS method) as peptide C7c (Table 15); equivalent to peptide P8a without N-terminal phenylalanine.

demonstrating the position of the one glutamine residue. The predicted mobilities were derived from plots of observed mobility versus the charge/mol. wt. ratio for all the peptides isolated during the elucidation of the sequence of Pseudomonas cytochrome-551 (Ambler, 1963b). These Tables are not given here, as Offord (1966) has published logarithmic plots of m versus mol.wt. that are more convenient and general in application than those used in the present investigation.

Offord (1966) describes some cases of anomalous electrophoretic mobilities. The most common cases at pH 6·5 are 'neutral' peptides (those with the same number of amino or guanidino and carboxyl groups), having N-terminal serine, threonine, asparagine or (most markedly) methionine sulphone, that have slight (m up to −0·16) anionic mobilities. Another type that has been noticed in the present investigation is neutral peptides that contain aspartic acid or glutamic acid adjacent to lysine (e.g. peptides T6c, T6b and P11aS1). In these cases the ionization of the ω-carboxyl group is probably partly suppressed, as the peptides have small (+0·05) cationic mobilities; the effect is greatly enhanced if the electrophoresis is at pH 6.

In all except two cases (the asparagine residues at positions 10 and 38) the presence of amide groups has been confirmed by the release of asparagine or glutamine by an exopeptidase.

In a few cases pairs of peptides were isolated after enzymic digestion in which the only difference between the major and minor form appeared to be the loss of an amide group in the latter. Examples were peptides C13 and C14a (residue 10), peptides C13 and C14b (residue 12 or 14) and peptides C6a and C9a (residue 16 or 18). Although this result could be due to heterogeneity in the starting protein (e.g. Ehrenstein, 1967), a much more likely explanation is that the difference is due to hydrolysis during isolation of the protein or peptide. The azurin does not appear to contain any very labile amide groups, such as the asparagine at position 50 in Pseudomonas cytochrome-551 (Ambler, 1963b), but such groups seem to be particularly characteristic of the sequence -Asn-Gly- (Haley, Corcoran, Dorer & Buchanan, 1966), not present in azurin.

No evidence of microheterogeneity of the protein, such as reported by Ehrenstein (1967), has been noticed in azurin. In the Tables of properties of primary peptides (Tables 4, 15, 28 and 38) contaminant amino acids are shown if present in the ratio 0·05mole/mole or more; most peptides contained less than this level of impurity. Where contamination occurred it was generally by amino acids such as glycine, alanine and serine that are present in chromatography paper in appreciable amounts, and was usually associated with peptides obtained only in low yield. In many cases an amino
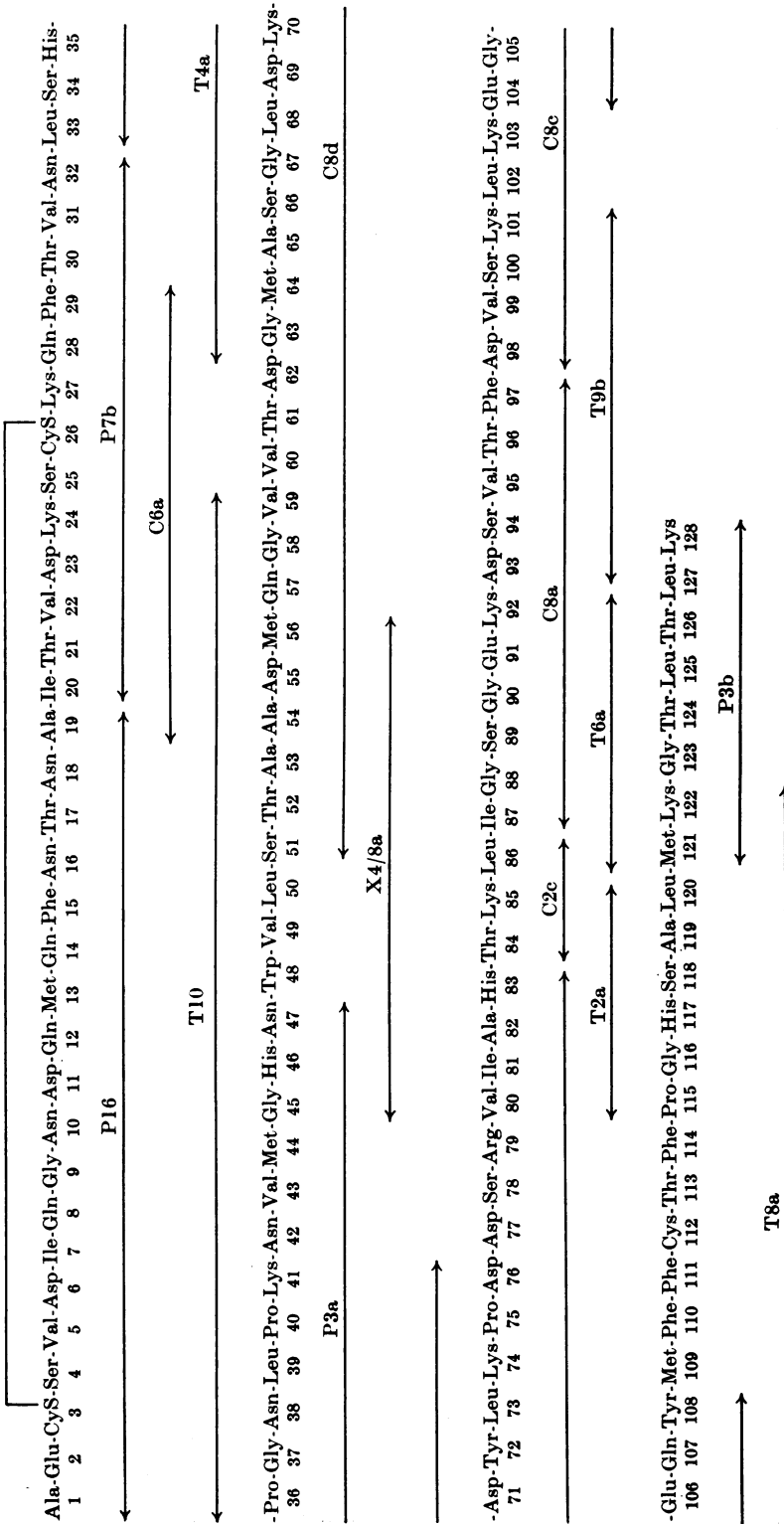
Ala-Glu-CyS-Ser-Val-Asp-Ile-Gln-Gly-Asn-Asp-Gln-Met-Gln-Phe-Asn-Thr-Asn-Ala-Ile-Thr-Val-Asp-Lys-Ser-CyS-Lys-Gln-Phe-Thr-Val-Asn-Leu-Ser-His-
1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   16   17   18   19   20   21   22   23   24   25   26   27   28   29   30   31   32   33   34   35

-Pro-Gly-Asn-Leu-Pro-Lys-Asn-Val-Met-Gly-His-Asn-Trp-Val-Leu-Ser-Thr-Ala-Ala-Asp-Met-Gln-Gly-Val-Val-Thr-Asp-Gly-Met-Ala-Ser-Gly-Leu-Asp-Lys-
36   37   38   39   40   41   42   43   44   45   46   47   48   49   50   51   52   53   54   55   56   57   58   59   60   61   62   63   64   65   66   67   68   69   70

-Asp-Tyr-Leu-Lys-Pro-Asp-Asp-Ser-Arg-Val-Ile-Ala-His-Thr-Lys-Leu-Ile-Gly-Ser-Gly-Glu-Lys-Asp-Ser-Val-Thr-Phe-Asp-Val-Ser-Lys-Leu-Lys-Glu-Gly-
71   72   73   74   75   76   77   78   79   80   81   82   83   84   85   86   87   88   89   90   91   92   93   94   95   96   97   98   99   100 101 102 103 104 105

-Glu-Gln-Tyr-Met-Phe-Phe-Cys-Thr-Phe-Pro-Gly-His-Ser-Ala-Leu-Met-Lys-Gly-Thr-Leu-Thr-Leu-Lys
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128

Fig. 15.   Amino acid sequence of *Pseudomonas fluorescens* azurin. Only sufficient peptides to establish the overall sequence are shown.   Definition of the symbols is given in the text.

acid alteration would have had sufficient effect on the properties of the peptide to have caused a minor form to separate from a major form during purification, but, as peptides present in small amounts (yields down to less than one-eighth of the maximum-yield peptides) were isolated and analysed, it is unlikely that there is appreciable microheterogeneity in the molecule.

The only peptides isolated during the investigation that could not be completely reconciled with the sequence shown in Fig. 15 were the dipeptides Lys-Leu and His-Leu. The former was isolated as peptide C1a from each chymotryptic digest, in yields of 5–10%. Although the sequence -Lys-Leu- occurs twice in the molecule (positions 85–86 and 101–102), there was no evidence that either sequence could be (or was) split out by chymotrypsin. The peptide His-Leu was isolated in about 2% yield from subtilopeptidase B digests of the protein. Peptides containing the four presumed histidine residues of azurin were isolated from the same digest in higher yields (Table 36), and the peptide could not be isolated from autolysates of subtilopeptidase B. No convincing explanation can yet be put forward to account for the occurrence of these dipeptides.

The most difficult region of the molecule to elucidate was between residues 42 and 69. In tryptic digests this region formed part of the 'core' and could not be purified by gel filtration or paper electrophoresis, whereas in peptic and chymotryptic digests this region gave rise to an extremely complex mixture of closely related peptides (Figs. 7 and 10). The whole sequence of this region could be elucidated from these peptides, except the proof of the peptide bond joining residues 50 and 51. This Leu-Ser bond was split rapidly by pepsin, chymotrypsin and subtilopeptidase B. A very minor peptic peptide (P7h) bridged the gap, but, to prove the sequence in this region, peptide X4/8a (Tables 38 and 39 and Fig. 14) was isolated after the protein had been cloven with cyanogen bromide.

In the early stages of the investigation it was believed that the protein did not contain any tryptophan residues, as the Spies & Chambers (1948) method did not detect any of this amino acid in the intact protein. The complex ultraviolet spectrum of the protein invalidated spectrophotometric determination, and the tryptophan residue was lost from tryptic digests in the core material. Later it was found that the Spies & Chambers (1948) method detected tryptophan if the protein was first digested with pepsin or chymotrypsin. It has been suggested (D. F. Teale, personal communication) that the 292 $m\mu$ absorption maximum in the ultraviolet spectrum of the native protein and the apoprotein is due to the tryptophan residue being constrained in a hydrophobic region by the tertiary structure of the molecule.

Specificity of proteases used. The trypsin used was a batch known to be very low in 'chymotryptic' activity, and which when used in the determination of the sequence of *Pseudomonas* cytochrome-551 had not produced any peptides as a result of non-tryptic cleavages. The minor tryptic peptides formed from azurin are shown in Table 14. Two peptides arose from slow tryptic cleavage of histidinyl bonds and one from incomplete hydrolysis at a lysine residue (position 92), lying between an aspartic acid and a glutamic acid residue. The Lys-Pro bond at residues 74–75 was not split at all by trypsin. The remaining two minor peptides were produced by hydrolysis of the bonds at positions 97–98 (Phe-Asp) and 18–19 (Asn-Ala). The former was rapidly hydrolysed by chymotrypsin, but the latter only slowly, chymotryptic attack being more rapid at most of the aromatic and leucine residues. It would therefore seem possible that minor tryptic splits of this type are not caused by contaminating chymotrypsin.

Chymotrypsin hydrolysed the expected aromatic and leucyl bonds. The His-Thr bond at position 83–84 was hydrolysed quite rapidly, and some cleavage was consistently observed at positions 18–19 (Asn-Ala), 32–33 (Asn-Leu), 81–82 (Ile-Ala) and 117–118 (His-Ser). No hydrolysis was detected at any of the other five asparaginyl bonds. In one digest appreciable hydrolysis occurred at positions 61–62 (Thr-Asp) and 66–67 (Ser-Gly). The expected peptides were isolated in good yield from most of the molecule, but the section of the sequence from residues 51–82 gave rise to a very complex mixture of peptides, with (for instance) seven distinct peptides isolated from one or other digest each containing the sequence of residues 51–56. This same region gave an even more complicated mixture of peptides in peptic hydrolysates, 13 bonds between positions 53 and 83 showing some susceptibility to the enzyme. About 20 different peptides derived from this region were isolated from different digests (Fig. 10), but despite this, peptide P9 (residues 54–86 inclusive) could be isolated in good yield (about 10%) from digests of oxidized protein. As with chymotrypsin, part of the complexity was due to the different susceptibilities to hydrolysis of bonds adjacent to methionine and to methionine sulphone.

Almost all the peptides formed by the digestion of the apoprotein with subtilopeptidase B were separated; Table 36 shows tri- and higher peptides isolated in yields greater than 5%. Some regions of the molecule that were difficult to isolate in good yield after digestion with the more specific proteases (like positions 69–76, peptide S8a, and 113–118, peptide S4a) could easily be prepared in good yields by this method. Out of the 127 peptide bonds in the azurin molecule, 52 appeared to be susceptible to

subtilopeptidase B (compared with 32 to pepsin and 26 to chymotrypsin). Subtilopeptidase B did not hydrolyse 14 of the bonds split by pepsin, but hydrolysed all but four of those sensitive to chymotrypsin. Eight bonds were susceptible to all three enzymes, but of these only four were split by all at rapid rates; these were the Phe-Asp bond at position 97–98 and the Leu-Ser, Leu-Met and Leu-Thr bonds at positions 50–51, 120–121 and 125–126. Subtilopeptidase B split bonds adjacent to each type of residue in the protein except proline and arginine; bonds on the amino side of glycine and alanine appeared to be very often susceptible, whereas those adjacent to valine and isoleucine seemed to be relatively insusceptible.

All samples of CP that were used were able to release lysine from suitable positions in peptides at rates comparable with the rate of release of neutral amino acids. There is slight evidence (Ambler, 1967) that this is an intrinsic property of the 'A' enzyme, and is not due to contamination with CPB.

Binding of copper by azurin. Evidence has been presented (Katoh & Takamiya, 1964) that, in the plastocyanins, thiol groups from two separate protein units provide two of the copper ligands. This does not seem likely to be the case in azurin, as the sedimentable unit is the monomer (Ambler, 1960), and the only likely thiol group (position 112) appears to be so sequestered by the large side chains on the adjacent residues that it cannot be alkylated even when the apoprotein is treated in 8M-urea solution. The evidence available from the study of the effect of pH on the recombination of copper with the apoprotein may suggest that no group with a pK between pH 4 and pH 8 (i.e. no imidazole side chain) is involved in the chelation.

Comparison of Pseudomonas *azurin with other copper proteins.* Sequence investigations (R. P. Ambler, unpublished work) have shown that the azurins of *Bordetella bronchiseptica, Alcaligenes denitrificans* and *Alcaligenes faecalis* (Sutherland & Wilkinson, 1963) and the blue protein from '*Pseudomonas denitrificans*' (Suzuki & Iwasaki, 1962) are homologous with *Pseudomonas fluorescens* azurin.

REFERENCES

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

Ambler, R. P. (1960). Ph.D. Thesis: University of Cambridge.

Ambler, R. P. (1963a). *Biochem. J.* 89, 341.

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

Ambler, R. P. (1965). *Biochem. J.* 96, 32 P.

Ambler, R. P. (1967). In *Methods in Enzymology,* vol. 7. Ed. by Hirs, C. H. W., Colowick, S. P. & Kaplan, N. O. New York: Academic Press Inc. (in the Press).

Ambler, R. P. & Brown, L. H. (1964). *J. molec. Biol.* 9, 825.

Boulton, A. A. & Bush, I. E. (1964). *Biochem. J.* 92, 11 P.

Broman, L., Malmström, B. G., Aasa, R. & Vänngård, T. (1963). *Biochim. biophys. Acta,* 75, 365.

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

Coval, M. L., Horio, T. & Kamen, M. D. (1961). *Biochim. biophys. Acta,* 51, 246.

Devor, A. W. (1950). *J. Amer. chem. Soc.* 72, 6347.

Ehrenstein, G. von (1967). *Cold Spr. Harb. Symp. quant. Biol.* 31. (in the Press).

Fraenkel-Conrat, H., Harris, J. I. & Levy, A. L. (1955). *Meth. biochem. Anal.* 2, 339.

Gahler, A. R. (1954). *Analyt. Chem.* 26, 577.

Gray, W. R. (1967). In *Methods in Enzymology,* vol. 7. Ed. by Hirs, C. H. W., Colowick, S. P. & Kaplan, N. O. New York: Academic Press Inc. (in the Press).

Gray, W. R. & Hartley, B. S. (1963a). *Biochem. J.* 89, 379.

Gray, W. R. & Hartley, B. S. (1963b). *Biochem. J.* 89, 59 P.

Gross, E. & Witkop, B. (1961). *J. Amer. chem. Soc.* 83, 1510.

Haley, E. E., Corcoran, B. J., Dorer, F. E. & Buchanan, D. L. (1966). *Biochemistry,* 5, 3229.

Harrison, P. M. & Hoffman, T. (1961). *Biochem. J.* 80, 38 P.

Hirs, C. H. W. (1956). *J. biol. Chem.* 219, 611.

Horio, T. (1958a). *J. Biochem., Tokyo,* 45, 195.

Horio, T. (1958b). *J. Biochem., Tokyo,* 45, 267.

Horio, T., Sekuzu, I., Higashi, T. & Okunuki, K. (1961). In *Haematin Enzymes.* Ed. by Falk, J. E., Lemberg, R. & Morton, R. K. Oxford: Pergamon Press Ltd.

Katoh, S. & Takamiya, A. (1964). *J. Biochem., Tokyo,* 55, 378.

Keilin, D. & Mann, T. (1940). *Nature, Lond.,* 145, 304.

Konigsberg, W. & Hill, R. J. (1962). *J. biol. Chem.* 237, 2547.

Mason, H. S. (1963). *Biochem. biophys. Res. Commun.* 10, 11.

Michl, H. (1951). *Mh. Chem.* 82, 489.

Milstein, C. (1966). *Biochem. J.* 101, 338.

Nakamura, T. (1958). *Biochim. biophys. Acta,* 30, 44.

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

Smithies, O. (1959). *Advanc. Protein Chem.* 14, 65.

Spackman, D. H. (1963). *Fed. Proc.* 22, 244.

Spackman, D. H., Stein, W. H. & Moore, S. (1958). *Analyt. Chem.* 30, 1190.

Spies, J. R. & Chambers, D. C. (1948). *Analyt. Chem.* 20, 30.

Sutherland, I. W. & Wilkinson, J. F. (1963). *J. gen. Microbiol.* 30, 105.

Suzuki, H. & Iwasaki, H. (1962). *J. Biochem., Tokyo,* 52, 193.

Verhoeven, W. & Takeda, Y. (1956). In *Organic Nitrogen Metabolism,* p. 159. Ed. by McElroy, W. D. & Glass, W. B. Baltimore: The Johns Hopkins Press.

Yamanaka, T., Kijimoto, S. & Okunuki, K. (1963). *J. Biochem., Tokyo,* 53, 256.