*Biochem. J.* (1973) **131**, 485–498
*Printed in Great Britain*



# The Amino Acid Sequences of Cytochromes *c*-551 from Three Species of *Pseudomonas*

By R. P. AMBLER and MARGARET WYNN
*Department of Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, U.K.*

(*Received* 9 *August* 1972)

The amino acid sequences of the cytochromes *c*-551 from three species of *Pseudomonas* have been determined. Each resembles the protein from *Pseudomonas* strain P6009 (now known to be *Pseudomonas aeruginosa*, not *Pseudomonas fluorescens*) in containing 82 amino acids in a single peptide chain, with a haem group covalently attached to cysteine residues 12 and 15. In all four sequences 43 residues are identical. Although by bacteriological criteria the organisms are closely related, the differences between pairs of sequences range from 22% to 39%. These values should be compared with the differences in the sequence of mitochondrial cytochrome *c* between mammals and amphibians (about 18%) or between mammals and insects (about 33%). Detailed evidence for the amino acid sequences of the proteins has been deposited as Supplementary Publication SUP 50015 at the National Lending Library for Science and Technology, Boston Spa, Yorks. LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in *Biochem. J.* (1973), **131**, 5.

Mitochondrial cytochrome *c* has been studied from a wider range of eukaryotic organisms than any other protein. Many complete amino acid sequences are known (see Dayhoff, 1969; Dickerson, 1972) and the tertiary structures of representative proteins have been determined (Dickerson *et al.*, 1971). The amino acid sequences of all the eukaryotic mitochondrial cytochromes *c* that have been examined are so similar that they must all be accepted as homologous. The amino acid sequences of several different bacterial cytochromes have been determined, and the structure of one of these, cytochrome $c_2$ from *Rhodospirillum rubrum* (Dus *et al.*, 1968), is sufficiently similar to that of mitochondrial cytochromes *c* for their homology to be accepted. The other known bacterial sequences are less similar, and their relationship to eukaryotic proteins is not obvious. A large number of different *c*-type cytochromes occur in prokaryotes (Kamen & Horio, 1970*a*,*b*) but their distributions are erratic and their functions generally obscure.

The cells of many species of *Pseudomonas* are rich in *c*-type cytochromes, and particularly large amounts are produced by denitrifying species. The major cytochrome from these organisms, with an α-band spectral absorption maximum at 551–552nm, was purified by Horio *et al.* (1960) from a strain of *Pseudomonas aeruginosa* and the amino acid sequence determined (Ambler, 1963*a*,*b*). The organism used for the latter work was called '*Pseudomonas fluorescens*' at that time, but it has now been positively identified as another strain of *Ps. aeruginosa* (R. Y. Stanier, personal communication). This cytochrome *c*-551 shares some structural features with mitochondrial cytochrome *c*, including the location of the haem attachment site near the *N*-terminus of the molecule, and the nature (methionine) and approximate location of the sixth iron ligand (Fanger *et al.*, 1967).

Much information is now becoming available about the amount of variation that exists between homologous proteins from eukaryotes, but very little is yet known about prokaryotic proteins or of the course of prokaryote evolution. In the present investigation cytochromes *c*-551 from three further species of *Pseudomonas* have been studied. The species chosen can now be clearly distinguished by bacteriological techniques, but are nevertheless so similar that it is reasonable to consider them to be closely related to each other. The species are *Pseudomonas stutzeri* (van Niel & Allen, 1952; Palleroni *et al.*, 1970), *Pseudomonas mendocina* (Palleroni *et al.*, 1970) and *Ps. fluorescens* biotype C (Stanier *et al.*, 1966). *Ps. stutzeri* is well known to produce exceptional amounts of *c*-type cytochromes, and two different ones have been purified and crystallized by Kodama & Shidara (1969). *Ps. mendocina* is a newly discovered species, similar to *Ps. stutzeri* in several respects. The *Ps. fluorescens* group is so similar to *Ps. aeruginosa* that until recently there was considerable difficulty in differentiating between the species (Rhodes, 1959, 1961), but clear criteria for separating the species are now recognized (Jessen, 1965; Stanier *et al.*, 1966). The *Ps. fluorescens* group has been divided into seven biotypes (A-G) (Stanier *et al.*, 1966). Biotype C is the only one from which we have managed to isolate cytochrome *c*-551.

In the present paper the elucidation of the amino

acid sequences of the cytochromes *c*-551 from a single representative strain of each of the three species is described, and the amount of difference revealed is discussed. The differences were so great that the elucidation of each structure had to be treated as a separate problem, and independent evidence has been obtained for each sequence. The sequences have been briefly reported elsewhere (Ambler, 1971*a*).

Satisfactory publication of amino acid-sequence studies is difficult. If the complete evidence is published, the resulting papers are excessively long (e.g. Ambler & Brown, 1967), but it is very necessary that authors should be required to assemble the full evidence for sequences that they propose, and to make it available for inspection, as experimental and interpretational mistakes are easy and prevalant. In this paper advantage has been taken of the data-deposition scheme organized by the National Lending Library for Science and Technology to make full numerical information accessible to anyone interested, and to include only sufficient detail in the published journal to allow readers to judge the experimental approach. The nature and format of the deposited information is described in the Results section. Standards for numerical values are defined, exceptions are noted, and attention is drawn to what are believed to be the weaker points in the evidence.

## Experimental

### *Materials*

*Organisms.* The strains used, *Pseudomonas fluorescens* biotype C no. 18 (Stanier *et al.*, 1966; A.T.C.C. 17400), *Ps. stutzeri* no. 221 (Stanier *et al.*, 1966; A.T.C.C. 17588) and *Ps. mendocina* CH110 (Palleroni *et al.*, 1970) were obtained from Dr. R. Stanier.

*Chromatographic material.* Whatman carboxymethylcellulose CM-23 and CM-52, and diethylaminoethylcellulose DE-52, were obtained from W. and R. Balston Ltd., Maidstone, Kent, U.K. Sephadex G-25, G-50 and G-75, fine and superfine grades, were obtained from Pharmacia (GB) Ltd., London W.5, U.K.

*Enzymes.* Trypsin (TRSF-6188) and chymotrypsin (three-times crystallized, CD16108-9) were obtained from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Soya-bean trypsin inhibitor and deoxyribonuclease (type II) were from Seravac Laboratories Ltd., Colnbrook, Bucks., U.K. Subtilisin B was obtained from Novo Terapeutisk Laboratorium, Copenhagen, Denmark. Thermolysin was obtained from Chugai Boyaki Co. Ltd., P.O. Box Higashi no. 106, Osaka, Japan. Carboxypeptidase A ('DFP-treated') was obtained from Sigma London Chemical Co., London S.W.6, U.K. Diphenylcarbamyl chloride was obtained from Koch–Light Laboratories Ltd., Colnbrook, Bucks., U.K.

### *Methods*

*Growth of organisms.* The bacteria were grown anaerobically in 35-litre batches on a nitrate medium as described by Ambler (1963*a*). *Ps. fluorescens* C-18 was grown at 25°C (taking about 36h), the other organisms at 37°C for 12–24h. The cells were collected by centrifugation (Alfa-Laval LAB102B-25 centrifuge) and stored frozen.

*Preparation of acetone-dried powder of cells.* The thawed cell-paste was blended with acetone precooled to −20°C, in an homogenizer fitted with an explosion-proof motor (Waring Products Co.; model EP-1, Baird and Tatlock Ltd., Chadwell Heath, Essex, U.K.), about 2.5 litres of acetone/100g of wet cells being used, and the liquid removed by filtration on a Büchner funnel. The pad was then washed with cold acetone and dried to constant weight in a continuously evacuated desiccator at room temperature.

*Spectra.* Absorption spectra were measured with a Unicam model SP.500 series II spectrophotometer, and positions of absorption bands in the visible region with a Hartridge reversion spectroscope. Cytochrome *c* concentrations were estimated as described by Ambler (1963*a*).

*Preparation of proteins.* The method originally developed for purifying cytochrome *c*-551 from one particular strain of *Pseudomonas* (Ambler, 1963*a*) can be used with slight modification to prepare cytochromes and azurin from many species of pseudomonads and related organisms. However, the ionic properties of the proteins vary with their source, and so the precise purification conditions that must be used vary from strain to strain. Since 1963 some modifications have been made to the original procedure, and these are described in (*a*)–(*h*) below, together with important differences in method for the species concerned. In this section the numbers of the steps are taken from the original description (Ambler, 1963*a*). The purification and properties of the cytochromes from the organisms discussed in this paper are summarized in Table 1.

(*a*) Step 1, extraction of acetone-dried powder of cells. The dried cells are nowadays ground into a fine powder without admixture of alumina.

(*b*) Step 2, treatment of extract with rivanol. This step is omitted.

(*c*) Step 3, acidification of extract before chromatography. The yield of proteins from some strains of organism is very small if the pH is adjusted to as low as 3.9 at this step. The proteins from *Ps. fluorescens* C-18 appear particularly sensitive: the best results are obtained if the pH is adjusted only to 4.2 although this makes adsorption on CM-cellulose more difficult.

(*d*) Step 4, adsorption of respiratory components on CM-cellulose. The CM-cellulose now used is Whatman CM-23, which can be regenerated and used again

indefinitely. The acid extract at step 3 is diluted to give the same conductivity as an ammonium acetate solution (0.05M with respect to acetate) of the same pH, and made 10$\mu$M with $K_3Fe(CN)_6$. The CM-cellulose is removed from the Büchner funnels and washed with pH3.9 buffer before being packed into a 2.5cm-diam. column for elution of the proteins with pH6.5 buffer. In this way the proteins from 100g of acetone-dried cells can be concentrated into 50–100ml.

(*e*) Step 5, chromatography on CM-cellulose. The CM-cellulose now used is Whatman CM-52. The size of the column (with height 2–3 × diam.) is chosen such that, when all the protein has been loaded, about half the bed is still colourless.

(*f*) Step 6, concentration and ammonium sulphate precipitation, and (*g*) step 7, chromatography on DEAE-cellulose. The order of steps 6 and 7 is now reversed. The proteins are concentrated on CM-cellulose (CM-52) as described but the columns are washed with water (2–5 column vol.) before elution with 25mM-Tris–HCl, pH8.5. The conductivity and pH of the eluate are then checked and adjusted if necessary before chromatography on DEAE-cellulose (Whatman DE-52), equilibrated with 25mM-Tris–HCl, pH8.5, and eluted with a linear NaCl gradient. The pooled fractions containing protein are adjusted to about pH5.5 by adding 5% (v/v) 4M-sodium acetate solution, pH5.5, and adjusted to the required saturation with $(NH_4)_2SO_4$ by addition of solid. After being left at room temperature for 30min the precipitate is collected by centrifugation, and dissolved in ammonium acetate buffer, pH5.1 (0.05M with respect to acetate).

(*h*) Step 8 (new) gel-filtration on Sephadex G-75. Columns of Sephadex G-75, fine or superfine grade, are used at room temperature, equilibrated with 0.05M-acetic acid adjusted to pH5.1 with aq. 2M-$NH_3$, and containing 2mM-sodium azide. The protein sample has normally just been precipitated with $(NH_4)_2SO_4$, and so the pellet is dissolved in the pH5.1 buffer for application. $\epsilon$-*N*-Dnp-lysine (0.25mg/ml) is added to the sample as a retention volume marker. Protein etc. is detected, and positions of peak maxima are accurately recorded, with a u.v.-absorption monitor (LKB Uvicord-II, measuring at 280nm). Two sizes of column are used: (*a*) 115cm × 2.5cm diam. Sephadex G-75 (fine grade), pumped at 80ml/h, 4ml fractions, 10–80mg of protein applied in 5ml; (*b*) 90cm × 1.5cm diam. Sephadex G-75 (superfine grade), pumped at 5ml/h, 1.8ml fractions, up to 20mg of protein applied in 2ml. The protein from step 8 is assessed for purity by starch-gel electrophoresis, spectrally and by amino acid analysis. It is in solution in a volatile buffer, and so can be used (after freeze-drying) for haem-removal and digestion.

*Starch-gel electrophoresis.* Horizontal gels were prepared by the method of Smithies (1959). The gel buffers used were: pH4.0, 0.02M-potassium hydrogen phthalate; pH8.5, 0.02M-boric acid adjusted with NaOH. Both buffers were made up to contain 1mM-$K_3Fe(CN)_6$. The proteins were visible before staining as brown bands but were enhanced by staining with Naphthalene Black. The bands were leached from the gel by prolonged washing.

*Amino acid analysis.* The method of Spackman (1963) was used, with a Beckman model 120-C amino acid analyser. Approximate sample loadings of 15nmol of protein/column and 20–50mnol of peptide/column were used. Samples for analysis were prepared and hydrolysed as described by Ambler & Brown (1967).

*Removal of haem.* The salt-free cytochrome (1.5$\mu$mol) was dissolved in 4ml of 8M-urea containing 0.1M-HCl in a 75mm × 16mm screw-cap tube, and 50mg of $HgCl_2$ added. The tube was shaken at 37°C for 20h, and then the protein was separated from haem and small molecules by gel-filtration through a 25cm × 1.5cm column of Sephadex G-25 (fine grade) equilibrated with aq. 0.1M-$NH_3$. The protein (generally visible as a pale-yellow band) was collected and freeze-dried. If not obvious it was located by u.v. absorptiometry. Much of the haem was left bound to the column, but this could be removed by passing through a portion (10ml) of 50% (v/v) pyridine. This method appears to be generally applicable for the removal of haem from *c*-type cytochromes. Trouble may be experienced with precipitation of the apoprotein during incubation or in the column during gel filtration. The former can be avoided by keeping the protein concentration low, and the latter by using an acid gel-filtration medium, e.g. 5% (v/v) formic acid, for basic cytochromes.

*Performic acid oxidation.* Two methods were used.

(i) The protein was treated with performic acid for 1h at 0°C (Hirs, 1956) before digestion and after removal of the haem, and the reagent removed by dilution with 25vol. of ice-cold water and freeze-drying.

(ii) Alternatively the protein was digested after removal of the haem, and selected peptides were oxidized on paper by performic acid vapour (Brown & Hartley, 1966), after high-voltage paper electrophoresis and before elution, at an intermediate stage in peptide purification. Such peptides were eluted with 1M-acetic acid (rather than aq. 0.1M-$NH_3$) to avoid contamination with ammonium formate.

*Digestion of proteins and peptides with proteolytic enzymes.* The methods described by Ambler (1963*b*) and (for thermolysin) Ambler & Meadway (1968) were used. Trypsin was treated before use with diphenylcarbamyl chloride, by the procedure of Erlanger & Cohen (1962).

*Peptide fractionation.* Initial fractionation of primary digests was by gel filtration through Sephadex G-25 (superfine grade) equilibrated with 5% (v/v) formic acid. A Whatman PC-1000 column was used,

giving bed dimensions 90cm × 1cm diam., pumped at 20ml/h, collecting 1.2ml fractions and applying the sample in 0.6ml. This size of column has proved satisfactory for digests of up to 50mg of protein provided that all the products of digestion are freely soluble in dilute formic acid. The eluate was passed through a u.v.-absorption monitor (LKB Uvicord II measuring at 280nm). The void volume was always apparent from the record. Portions of each fraction (corresponding to about 0.05μmol of the total protein digest) were separated in parallel by paper electrophoresis at pH6.5, and the resultant map was used to decide how to pool fractions for further purification by paper electrophoresis and paper chromatography, and to calculate $V/V_0$ values for peptides. Similar conditions were used for Sephadex G-50 fractionations.

Apparatus of the Michl (1951) type was used for high-voltage paper electrophoresis peptide fractionation. The buffers and conditions were as described by Ambler (1963*b*). Butan-2-ol–acetic acid–water–pyridine (15:3:12:10, by vol.) (Waley & Watson, 1953) was used for paper partition chromatography.

*Tryptophan peptides.* These were detected by their high u.v. absorption after gel filtration, and by specific colour reactions on paper after electrophoresis (Harley-Mason & Archer, 1958). Tryptophan peptides were appreciably more retarded on gel filtration than other peptides of comparable size. The effect was much more marked than for other aromatic peptides, and was sufficiently noticeable as to be diagnostic.

*C-Terminal groups of proteins and peptides.* Haem-free cytochromes were investigated by carboxypeptidase A hydrolysis (Ambler, 1972*a*). Both qualitative methods (d, Ambler, 1972*b*) and quantitative methods (a, Ambler, 1972*b*) were used for the investigation of peptides with carboxypeptidase A. In quantitative experiments, although the amino acid-analysis system used (Spackman, 1963) did not resolve serine, glutamine and asparagine, a predominance of asparagine could be recognized by the unusually high absorption at 440nm relative to 570nm of the coloured product formed by this amino acid with ninhydrin (cf. glutamic acid and lysine).

*N-Terminal groups of proteins and peptides and Dns–phenyl isothiocyanate degradation of peptides.* The Dns method (Gray, 1972*a*) was used for qualitative investigation of proteins and peptides.

Dns-amino acids were identified by paper electrophoresis at pH4.38 and 2, and by paper partition chromatography. The details of the conditions used were as described by Ambler & Brown (1967). Dns-homoserine lactone stays near ε-Dns-lysine at pH4.38 but separates clearly at pH2.

The conditions for phenyl isothiocyanate degradation were as described by Gray (1972*b*). In suitable cases the *C*-terminal amino acid remaining after the others had been removed was directly identified (Pink *et al.*, 1970) by high-voltage paper electrophoresis at pH2.

In some cases the electrophoretic mobility of residual peptides after particular steps of the phenyl isothiocyanate degradation were measured, so that the position of amide groups could be located.

*Cyanogen bromide cleavage.* The proteins were treated with cyanogen bromide before removal of haem, and the initial fractionation was done on Sephadex G-50 (superfine grade) as described above. The conditions for reaction and pre-analysis treatment of hydrolysates were as described by Ambler & Brown (1967). Peptides were treated under similar conditions, and the products fractionated by paper electrophoresis.

*Quantities of material used.* Cytochrome (1–3μmol) was used for each experiment involving haem removal, digestion and peptide fractionation and investigation. The yields of peptides (not normally accurately determined) were in the range 10–45% overall (with no corrections made for losses during purification). The approximate amounts of peptide used for sequence experiments were as follows: qualitative amino acid analysis (by high-voltage paper electrophoresis at pH2), 5–10nmol; Dns *N*-terminal analysis, 5–10nmol; 10nmol per step of Dns–phenyl isothiocyanate degradation; quantitative amino acid analysis, 20–100nmol; secondary digestion with quantitative analysis and Dns–phenyl isothiocyanate degradation of peptide products, 300nmol; carboxypeptidase A digest (qualitative) 10–20nmol, (quantitative) 20–50nmol.

## Results

The particular strain of each species used was chosen because of its vigorous growth on the standard medium. The cytochrome *c*-551 was obtained in a pure state from each of the three new organisms considered in this paper, by using methods of growth and isolation very similar to that originally devised for the protein from *Ps. aeruginosa* (Ambler, 1963*a*). The purification details and some of the properties of the proteins are given in Table 1. Each protein had different ionic properties (as judged by behaviour during ion-exchange chromatography and starch gel electrophoresis), but they were not distinguished spectroscopically.

Two distinct forms of cytochrome *c*-551 were isolated from *Ps. fluorescens* and *Ps. mendocina*, which differed in behaviour during ion-exchange chromatography (Table 1) but which could not be distinguished by amino acid analysis or peptide mapping. In both cases gel filtration (Table 1) showed that the minor form had a higher molecular weight, and was probably the dimer. *Ps. stutzeri* also yielded two different ionic forms of the cytochrome. When digested with

Table 1. *Purification and properties of Pseudomonas cytochromes c-551*

Details are given in the text. The proteins from several authentic strains of *Ps. aeruginosa* (including N.C.T.C. 10332, the neotype) are indistinguishable from that of strain P6009 (R. P. Ambler, unpublished observations). The electrophoretic mobility is that measured in starch gels relative to Xylene Cyanol FF.

| Organism ... | *Ps. aeruginosa* (Ambler, 1963*a*) | *Ps. fluorescens* (biotype C) | *Ps. stutzeri* | *Ps. mendocina* |
|---|---|---|---|---|
| Strain ... | P6009 | Stanier C-18 | Stanier 221 | CH110 |
| Cell yield (g dry wt./litre of culture) | 0.9–1.2 | 0.8 | 1.0–1.4 | 0.7–0.9 |
| Cytochrome yield ($\mu$mol/100 g of dry cells) (at step 4) | 10 | 2–5 | 35 | 10 |
| CM-cellulose elution pH (at step 5) | | | | |
| Cytochrome *c*-551 | 4.45 | I:4.45; II:4.8 | I:4.4; II:4.5 | I & II:4.45 |
| $c_4$ | 5.1 | ? | ? | 4.75 |
| $c_5$ | 5.1 | 4.6 | ? | 4.75 |
| Azurin | 4.65 | 4.7 | Not present | ? 4.4 (Trace) |
| DEAE-cellulose elution [NaCl] (at step 7) (mM) | | I: 5<br>II: 5 | I: 40<br>II: 50 | I: 15<br>II: 50 |
| Electrophoretic mobility pH4.0 | 0.5 | I & II: 0.7 | I: 0.2; II: 0.1 | I & II: 0.4 |
| pH8.5 | 0.7 | I & II: 0.3 | I & II: 1.0 | I & II: 0.6 |
| Gel filtration $V/V_0$ Sephadex G-75 | 2.2 | I: 2.2; II: 1.8 | | I: 2.2; II: 1.8 |
| Absorption spectrum extinction ratio | | | | |
| $E^{red.}/E_{280}$ | 1.2 | 1.2 | 1.2 | 1.2 |
| $\alpha$-band max. (nm) | 551 | 551 | 551 | 551 |
| Amino acids absent in pure protein | All present | Tyr, Arg | Arg | Tyr, Arg |
| *N*-Terminus (Dns) | Glx (Dnp) | I & II: Glx | I: none found<br>II: Glx | I & II: Ala |
| *C*-Terminus (carboxypeptidase A) | | See the text | | See the text |

trypsin, form II (Table 1) yielded two peptides (T310 and T313, Fig. 2) of identical amino acid composition but different electrophoretic mobility. In contrast, form I yielded only the more acidic peptide T313. *N*-Terminal glutamic acid could be detected in intact form II and in peptide T310, but no free $\alpha$-amino group could be found in form I or in peptide T313. These results suggest that the *N*-terminal amino acid in form II is glutamine, which is partly converted into pyrrolidone carboxylic acid both during isolation of the protein and during further degradative experiments such as removal of the haem group and fractionation of the peptides.

The criteria used to assess the purity of the protein preparations were amino acid analysis and homogeneity on starch-gel electrophoresis, the former being the more stringent. Each of the proteins completely lacks one or more amino acids (Table 1) and preparations with less than 0.2mol of these amino acids/mol of protein were judged pure by the other criterion. Partly purified preparations that contained 0.2–0.5mol/mol of these contaminant amino acids still gave acceptably clean peptide maps. The amino acid composition of these proteins is given in Table 2, values derived from both amino acid analysis and from the deduced sequence of the proteins being given. Each protein was homogeneous when examined for *N*-terminal residues by the Dns method (with the same reaction conditions as used for peptides).

The *C*-terminal sequences of the *Ps. fluorescens* and *Ps. mendocina* proteins were investigated by carboxypeptidase A digestion, in each case at three different intensities. The results were consistent with *C*-terminal sequences:

*Ps. fluorescens* C-18 -Ala(Leu,Thr,Leu,Ala,Gln,Trp,Val,Leu,Ser)(Leu,Lys)
*Ps. mendocina* CH110 -(Ala,Lys)(Thr,Leu,Ala,Glu)(Trp,Val)(Leu,Thr,Leu,Lys)

In the former case glutamine and serine (or asparagine) were not distinguishable, and in the latter case tryptophan quantitation was poor. Full details are given in the Supplementary Publication SUP 50015.

In addition to the cytochromes *c*-551 each of the organisms contained small amounts of other *c*-type

Table 2. *Amino acid compositions of Pseudomonas cytochromes c-551*

The values for *Pseudomonas aeruginosa* P6009 are taken from Ambler (1963*a*). For the other proteins the values in the sequence columns are derived from the sum of the peptides constituting the protein (Figs. 1, 2 and 3), and those in the analysis column from analysis of 24h-hydrolysis samples of purified proteins.

| | *Ps. aeruginosa* P6009 | *Ps. fluorescens* C-18 | | *Ps. stutzeri* 221 | | *Ps. mendocina* CH110 | |
|---|---|---|---|---|---|---|---|
| | Sequence | Sequence | Analysis | Sequence | Analysis | Sequence | Analysis |
| Gly | 7 | 7 | 7.3 | 8 | 8.1 | 9 | 9.0 |
| Ala | 13 | 13 | 12.4 | 13 | 12.7 | 13 | 12.9 |
| Val | 7 | 5 | 4.8 | 4 | 4.2 | 8 | 7.6 |
| Leu | 4 | 7 | 6.6 | 6 | 6.0 | 7 | 6.7 |
| Ile | 3 | 3 | 2.9 | 5 | 5.3 | 1 | 0.9 |
| Ser | 3 | 3 | 3.3 | 4 | 4.5 | 4 | 3.9 |
| Thr | 2 | 5 | 4.8 | 1 | 1.1 | 4 | 3.8 |
| Asp | 5 | 5 | 9.2 | 4 | 6.1 | 3 | 6.4* |
| Asn | 3 | 4 | — | 2 | — | 3 | — |
| Glu | 5 | 3 | 6.0 | 6 | 8.6 | 6 | 7.1 |
| Gln | 5 | 3 | — | 4 | — | 1 | — |
| Phe | 2 | 1 | 1.0 | 2 | 2.0 | 1 | 1.0 |
| Tyr | 1 | 0 | 0.05 | 1 | 1.0 | 0 | 0.05 |
| Trp | 2 | 2 | † | 2 | † | 2 | † |
| Cys | 2 | 2 | 2.0‡ | 2 | † | 2 | 1.4 |
| Met | 2 | 2 | 2.0 | 1 | 0.9 | 2 | 1.0* |
| Pro | 6 | 6 | 6.1 | 7 | 7.0 | 6 | 5.5 |
| Lys | 8 | 9 | 9.1 | 8 | 8.4 | 8 | 8.0 |
| His | 1 | 2 | 2.1 | 2 | 2.1 | 2 | 1.9 |
| Arg | 1 | 0 | 0.05 | 0 | 0.1 | 0 | 0.05 |
| Total | 82 | 82 | | 82 | | 82 | |

* In this analysis some of the methionine had been converted into methionine sulphoxides, which were only partially resolved from the aspartic acid.

† Not determined.

‡ Value from an analysis of a sample from which the haem had been removed and the protein then oxidized.

cytochromes. These have been further characterized, and their properties will be described in further papers. They have amino acid sequences completely different from cytochromes of the *c*-551 type.

*Determination of amino acid sequences*

*Experimental approach.* The amino acid sequences of the three proteins have been determined (Figs. 1–4). Although comparison of the sequences (Fig. 5) shows considerable similarity, sufficient data were obtained for each protein for each sequence to be deduced independently.

Preparations of each protein were treated to remove the haem, and then digested with trypsin or chymotrypsin. For each protein at least two separate tryptic and chymotryptic digests have been examined, completely independently grown and prepared material being used. Digests were fractionated by gel filtration followed by high-voltage paper electrophoresis. Normally electrophoresis at two different pH values was necessary and sufficient to purify a peptide, but in a few cases a paper-chromatography step was also necessary. Absolute peptide yields are known accurately only in a few cases: in the other cases they could not be calculated as the exact quantity of starting material for the digest was not measured. Where known, the yields were in the same range (10–50% overall) that has been found for peptides from other small proteins purified in similar ways (Ambler & Brown, 1967; Ambler, 1971*b*). The amino acid compositions and purities of peptides were determined by quantitative amino acid analysis. *N*-Terminal residues were determined by the Dns method, and partial or complete sequences by the Dns–phenyl isothiocyanate method (Gray, 1972*b*). In suitable cases the results of phenyl isothiocyanate degradation were checked by identifying the free *C*-terminal amino acid by paper electrophoresis, after all the other residues had been successively removed (Pink *et al.*, 1970). Larger peptides were degraded by further proteolysis, and some *C*-terminal sequences

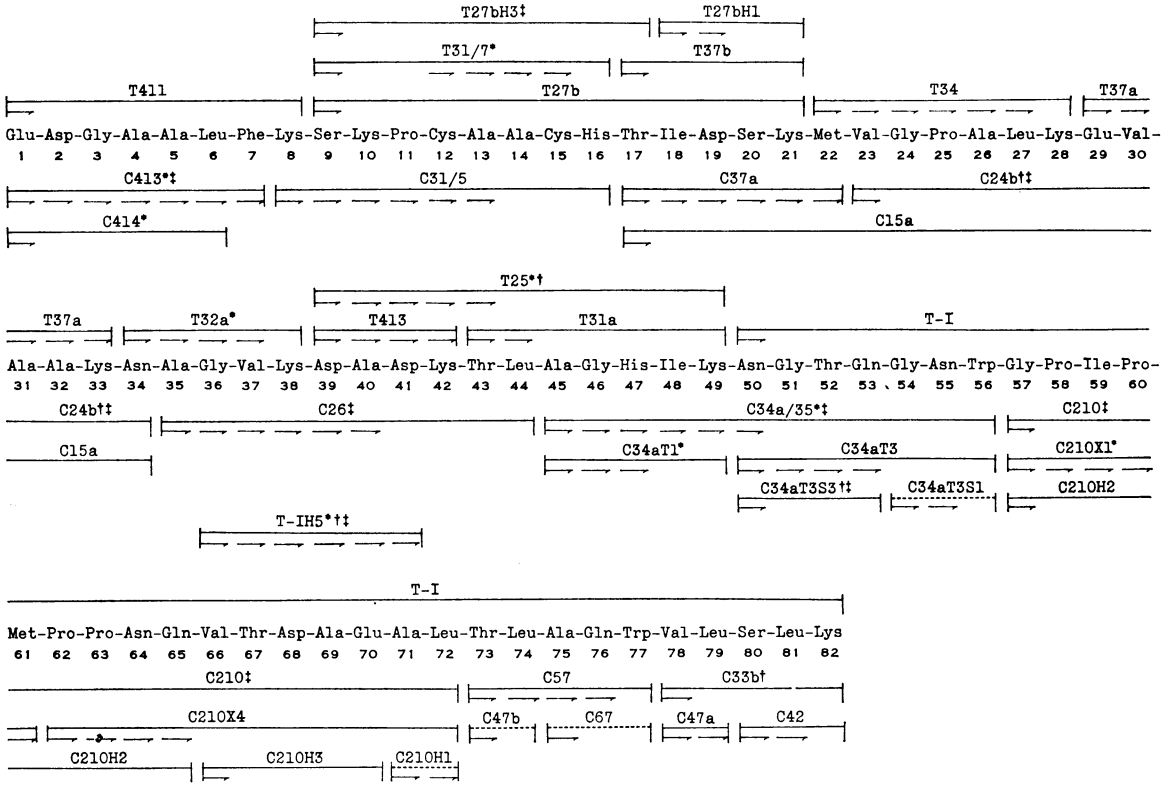


Fig. 1. *Amino acid sequence of Ps. fluorescens C-18 cytochrome c-551*

Peptides derived by tryptic and chymotryptic digestion of the protein are shown above and below the sequence respectively. Full lines indicate quantitative analyses (sub-standard if marked †). Broken lines indicate qualitative analyses. → Indicates end groups and subsequent residues revealed by phenyl isothiocyanate degradation (sub-standard in some respect if marked *). Peptides marked ‡ were examined by carboxypeptidase A digestion. An explanation of peptide nomenclature is given in the text.

examined by carboxypeptidase A digestion. Cyanogen bromide cleavage was used to characterize the *C*-terminal region of the *Ps. stutzeri* and *Ps. mendocina* proteins, and to degrade peptides formed by chymotryptic digestion of all three proteins.

Each protein has only two cysteine residues (Table 2), both involved in haem binding. In some experiments these residues were oxidized to cysteic acid after removal of the haem and before enzymic digestion. In later experiments the cysteine peptide was recognized at an intermediate stage in purification, and oxidized on paper by performic acid vapour (Brown & Hartley, 1966) before elution and final purification. The latter method was preferred, as tryptophan peptides could be preserved, and methionine peptides kept in a state suitable for cyanogen bromide cleavage. The state of the cysteine residues appeared to affect the susceptibility to trypsin of nearby peptide bonds.

Tryptophan was not estimated quantitatively either in the whole proteins or in isolated peptides. A direct colorimetric method (Spies & Chambers, 1948) had not proved satisfactory when applied to haem-containing samples (Ambler, 1963*a*) and more recent experiments have given poor recoveries when the same method has been attempted with cytochromes from which the haem has been removed (R. P. Ambler, unpublished work), presumably because of damage to the tryptophan residues during the prolonged acid treatment. The spectral ratios $E^{red.}_{551}/E_{280}$ for the three proteins (Table 1) were all very similar to that for *Ps. aeruginosa* cytochrome *c*-551 (Ambler, 1963*a*,*b*), for which quantitative determinations were attempted and which indicated two residues per mol,

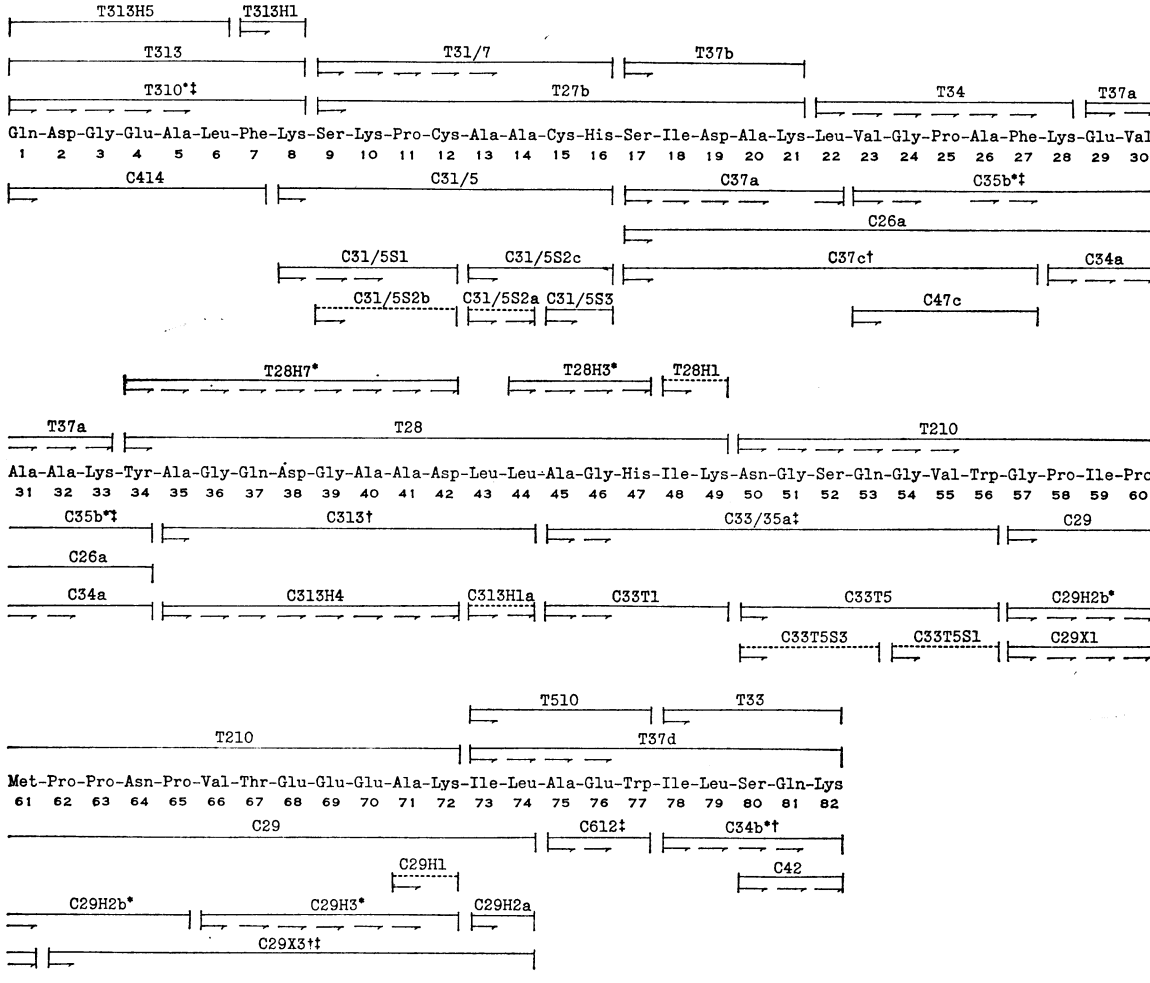


Fig. 2. *Amino acid sequence of Ps. stutzeri* 221 *cytochrome c*-551

Details are given in the legend to Fig. 1.

and in each protein two tryptophan residues were identified during the sequence determination. The possibility of the occurrence of a -Trp-Trp- sequence has been considered. Although there is little direct evidence against it, such a sequence would be expected to produce a more complex pattern of chymotryptic peptides than was found, and to alter the peptide patterns on gel filtration in a characteristic manner.

Amide residues were located from a consideration of the electrophoretic mobility at pH6.5 of simple peptides containing the residue under examination. In difficult cases the change in mobility of peptides as they were progressively degraded by the phenyl isothiocyanate method was determined (cf. Table 26, Ambler & Brown, 1967).

*Peptide nomenclature.* Peptides are identified by a capital letter (indicating the method used for the primary degradation of the protein) followed by two or three arabic numerals. The first indicates the gel-filtration fraction from which the peptide was isolated, and the second or second and third the relative mobility of the peptide on electrophoresis at pH6.5, the most basic peptide having the lowest number.

The mobilities of some peptides altered during purification, either because of deliberate action (e.g. oxidation with performic acid) or spontaneous modification. In these cases the name of the derived form contains an extra numeral representing the new mobility, separated by a solidus from the rest of the designation. Peptides, which for some reason could not be fractionated at any stage by pH6.5 electro-

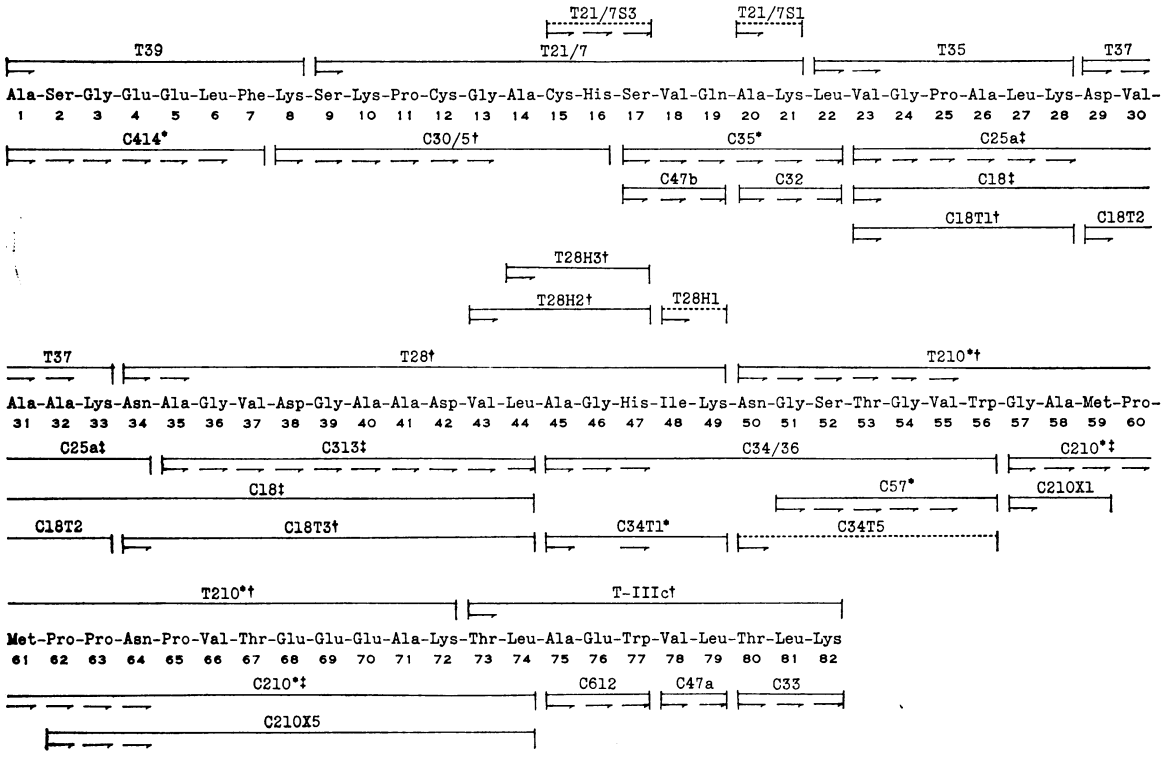


Fig. 3. *Amino acid sequence of Ps. mendocina CH*110 *cytochrome c*-551

Details are given in the legend to Fig. 1.

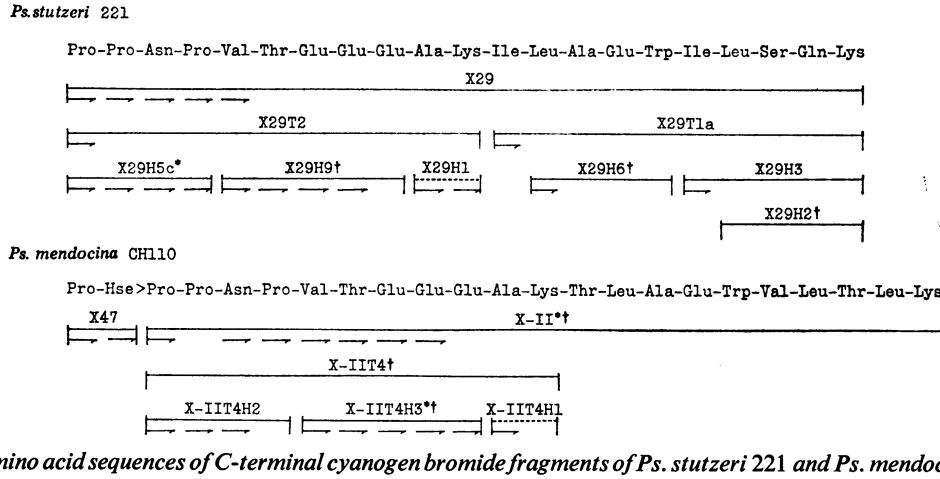


Fig. 4. *Amino acid sequences of C-terminal cyanogen bromide fragments of Ps. stutzeri* 221 *and Ps. mendocina CH*110 *cytochromes c*-551

Details are given in the legend to Fig. 1.

phoresis, have the gel-filtration fraction in which they occurred indicated by a roman numeral, with a suffix if necessary. Suffixes 'a', 'b' etc. are used for peptides with similar electrophoretic mobilities at pH6.5. When the primary peptide has been further degraded the secondary (and if necessary tertiary) peptides 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: T, trypsin; C, chymotrypsin; H, thermolysin; S, subtilisin B; X, cyanogen bromide.

*Evidence for proposed sequences.* The evidence for the amino acid sequences proposed for the three proteins is summarized in Figs. 1–4. These show all the peptides isolated from each digest, together with peptides formed by secondary digestion. Symbols show how much of the sequence of each peptide has been determined by the Dns–phenyl isothiocyanate method, and cases for which this evidence is not considered wholly satisfactory are indicated (*). Also indicated (†) are peptides for which quantitative amino acid data are in some way unsatisfactory (see below for criteria), and (‡) peptides examined by carboxypeptidase A digestion.

Peptides for which the Dns–phenyl isothiocyanate data were unsatisfactory comprised cases where more than one Dns-amino acid was present in significant amounts at one or more steps in the degradation, and cases where for some reason or other (accident or shortage of material) a Dns-amino acid was missed or not positively identified at one or more steps. The former type of trouble often consisted of contamination by the Dns-amino acid from the previous step. Such 'ragged' degradation frequently started at peptide bonds on the *N*-terminal side of glycine residues, which are known to be cloven more slowly than other bonds (Konigsberg & Hill, 1962). The criteria for satisfactory amino acid analyses were: (1) no impurity amino acid present in an amount as great as 0.2mol/mol. (2) The relative amounts of amino acids present were calculated on the basis that the average amount was integral. No values should fall outside the limits 0.8–1.2, 1.8–2.2, 2.7–3.3 or 3.7–4.3. However, values as low as 0.6 (=1) were considered acceptable for tyrosine and homoserine.

More than half the peptides analysed would have satisfied criteria twice as strict as these. The majority of unsatisfactory peptides failed because of non-acceptable (0.2–0.3) amounts of contaminating amino acids (normally glycine and serine, the common 'wet thumb-print' contaminants; Hamilton, 1965), small transgressions of the stoicheiometry conditions, or absence through instrument failure of quantitative values for one or more of the amino acids known qualitatively to be present. In a very few cases the basic amino acids were only analysed qualitatively.

The five cases of more serious failure are dealt with in detail below. The analysis results are calculated as mol of amino acid/mol of peptide, and impurities above the 0.1 level are shown.

*Ps. fluorescens* C-18 peptide T-IH5 (residues 66-71); analysis Ala 2.5, Val 0.9, Thr 0.6, Asp 0.9, Glu 1.1. The sequence of this peptide was deduced to be Val-Thr-Asp-Ala-Glu-Ala. It was probably contaminated with Ala-Glu-Ala (residues 69–71) derived by a minor thermolysin cleavage, and which would have a very similar electrophoretic mobility.

*Ps. stutzeri* 221 peptide C313 (residues 35–44); analysis Gly 2.0, Ala 2.9, Leu 1.5, Asp 2.0, Glu 1.2, (Ser 0.14). The material, isolated from a chymotryptic digest, was presumed to be a mixture of the -Leu peptide ending at residue 43 and the -Leu-Leu peptide ending at residue 44. These peptides would have closely similar electrophoretic properties, but, had it been essential, could probably have been separated by paper chromatography.

*Ps. stutzeri* 221 peptide C29X3 (residues 62–74); analysis Ala 1.1, Val 1.0, Leu 1.0, Ile 1.0, Thr 0.8, Asp 1.0, Glu 3.2, Pro 2.9, Lys 0.9, (Gly 0.23, Ser 0.4). The peptide was isolated from the products of cyanogen bromide cleavage of a chymotryptic peptide. The analysis represents an extreme example of contamination with glycine and serine, probably from dust or a human source. Contamination was not present in a sample from the same batch treated with carboxypeptidase A. The absolute level of contamination was high, as about 100nmol of peptide was hydrolysed.

*Ps. mendocina* CH110 peptide C30/5 (residues 8–16); analysis Gly 1.0, Ala 1.1, Ser 1.0, cysteic acid 2.6, (Pro lost), Lys 1.9, His 1.0. In this analysis of a chymotryptic peptide an instrument failure made the proline peak uncalculable. There is no good explanation for the high cysteic acid value. There was ample material (60nmol of peptide hydrolysed) and the peak appeared normal (breakthrough artifacts caused by buffer switching sometimes cause trouble at this stage in analyses).

*Ps. mendocina* CH110 peptide X-IIT4H3 (residues 66–70); analysis Val 0.8, Thr 0.9, Glu 3.1 (Gly 0.18, Ala 0.14, Asp 0.14). This peptide was derived from a thermolysin digest of a peptide formed by tryptic digestion of a cyanogen bromide fraction. The analysis figures satisfy the criteria set up above, but could equally well have been calculated to give Val 1.0, Thr 1.1, Glu 3.9. Such ambiguity is likely to occur only for peptides in which more than half of the residues are of a single type. Analytical data for other peptides of the same protein encompassing the same region, and sequence experiments on the peptide itself, provide convincing evidence that there should be three residues of glutamic acid.

*Supplementary information*

Detailed evidence for the amino acid sequence of the proteins has been deposited with the National Lending Library for Science and Technology, Boston Spa, Yorkshire LS23 7BQ, U.K., for storage on microfiche (Supplementary Publication SUP 50015). The information comprises:

(1) The details of the hydrolysis of *Ps. fluorescens* C-18 and *Ps. mendocina* CH110 cytochromes by

carboxypeptidase A, and an interpretation of the results.

(2) Tables showing the properties of all peptides shown in Figs. 1–4. The purification steps used for each peptide are given, and where applicable the values of $V/V_0$ [for gel filtration through Sephadex G-25 (superfine grade) in 5% (v/v) formic acid] and electrophoretic mobility at pH6.5. Relative overall yields for peptides from primary digests are given, as are amino acid analysis values (including impurities present in amounts greater than 0.1 mol/mol) and the *N*-terminal amino acid.

(3) Further tables show the individual sequence evidence for each peptide, including details of sub-standard results with the Dns–phenyl isothiocyanate method and details of the carboxypeptidase A digestions.

(4) Tables for each protein give the evidence for the presence or absence of amide groups on each aspartic acid or glutamic acid residue.

## Discussion

### *Accuracy and reliability of sequence determination*

The accuracy and reliability of a sequence determination is a function of the number of experiments performed and of the care with which they are done. Increasing the accuracy and reliability increases the time and expense of each determination, and so decreases the total amount of information obtained. A balance must be struck between an extravagantly pedantic approach to each determination, and work that is so unreliable as to be misleading. Each worker decides the point of balance for himself (consciously or unconsciously) but the reality of this compromise should be acknowledged, and clearly defined in publication.

Errors in sequence results fall into a number of categories: (i) technical errors during sequence determination [(*a*) wrong block sequence (incorrect ordering of fragments), (*b*) large deletions (missed peptides), (*c*) insertions or deletions of single amino acids, (*d*) transpositions of two amino acids, (*e*) wrong amide locations, (*f*) undetected non-standard amino acids]; (ii) typographical errors in publication; (iii) total mistakes.

Errors of type (iii) can arise from working with an impure protein or peptide, or by working with the wrong peptide (e.g. by mislabelling) or with the wrong protein (e.g. owing to misidentification of the source organism or to contamination of a microbial culture). Because the work described in the present paper was particularly susceptible to mistakes of this sort, as a precaution each digest of each protein was carried out on material grown and purified independently. Although only single cultures of each strain were obtained from the listed sources, the possibility of a total mistake at this stage has been lessened by examining proteins from other strains of each of the four organisms.

Sufficient evidence is presented for the sequence of each protein to be deduced independently, but of course this is not how the deduction was actually done. Despite the large differences in sequence (Fig. 5) the resemblances are strong, with exactly matching residues spread throughout the molecules. Such similarity was obvious from an early stage in each investigation, and knowledge of it prevented true independent deduction. The direct evidence for the block order of *Ps. fluorescens* peptides C57 and C33b (residues 73–82, Fig. 1) and for *Ps. mendocina* peptides C612, C47a and C33 (residues 75–82, Fig. 3) is weaker than for the rest of the molecules, being derived only from quantitative carboxypeptidase A digestion results. Stronger evidence could easily be obtained (e.g. by thermolysin digestion of intact proteins) but the effort was not considered justifiable. The cumulative evidence from the four homologous sequences makes it extremely unlikely that an error of type (*ia*) has occurred in any of the sequences.

Errors of type (*ib*) are unlikely in view of the adequate agreement between amino acid analysis figures for the whole proteins (Table 2) and the compositions deduced from the sequences. No unacceptable peptides were obtained in any digests.

Errors of type (*ic*) are likely if results from grossly sub-standard amino acid analyses are used: the specification for analyses for the present work is sufficiently rigorous to make errors unlikely. Errors can also arise from acid-resistant -Val-Val- or -Ile-Ile-sequences, but would have been detected in the present work by the recognition of the resistant peptides during Dns–phenyl isothiocyanate degradation of peptides. Sequence determination by the Dns–phenyl isothiocyanate method is susceptible to errors of type (*id*). To guard against them, we always tried to identify Dns-amino acids directly, independently and unequivocally. Separations of routine quality by high-voltage paper electrophoresis at pH4.38 (Gray, 1972*a*) can achieve this, but some separations (e.g. Dns-Gly/Ser/Ala) are very sensitive to slight changes in pH, and others (e.g. Dns-Val/Leu, Dns-Pro, Dns-Phe/Ile) depend to some extent on subjective considerations of spot shape and colour. In these cases to confirm identification the fluorescent spots were sewn on to another sheet of paper for positive identification after a further separation. The few cases where this procedure failed or was not used are noted in the Supplementary Publication. Nevertheless mislabelling of samples during Dns–phenyl isothiocyanate method of sequence determination can easily occur, and errors due to this would not always have been detected by redundant or duplicate experiments.

The evidence for amide ascriptions for each protein (error type *ie*) is given in detail in the Supplementary

Publication. The evidence is generally based on a qualitative assessment of the mobility of a small peptide containing a possible net charge of zero or one. Quantitative values of mobility are given to permit evaluation of the evidence.

No evidence, direct or circumstantial, has been obtained that would suggest the presence of non-standard amino acids or extra prosthetic groups (error type *if*) but rigorous experiments to exclude the possibility have not been done. Many eukaryotic cytochromes *c* contain ε-trimethyl-lysine (DeLange *et al.*, 1969). This has not been looked for specifically in the cytochromes *c*-551, but trypsin readily hydrolyses the peptide bonds adjacent to all but two of the putative lysine residues (the exceptions are the -Lys-Pro- at position 10-11 and the *C*-terminal residues), and ε-Dns lysine has been detected in hydrolysates of Dns-peptides which include each lysine residue in each protein. If errors of type (ii) creep in, they should be detectable by comparison of Figs. 1–4 and Fig. 5.

*Technical problems in sequence elucidation*

As ample quantities of each protein could be prepared, the determination of the sequences produced few major problems. Experience has shown that it is difficult to purify haem-containing peptides satisfactorily, so the haem group was always removed before enzymic digestion of the proteins. The peptide containing the haem-binding sequence (presumably now with a mercury atom bound to each cysteine residue) could be partially purified by gel filtration and paper electrophoresis, and then completely purified after oxidation with performic acid.

Few peptides were encountered with solubility properties that caused difficulty in purification. The *C*-terminal tryptic fragment from the *Ps. mendocina* protein (TIIIc, Fig. 3), although only a decapeptide, did not move from the origin when subjected to paper electrophoresis at pH6.5 or 3.5, but could be separated at pH2. This insolubility was surprising, particularly as no difficulty was encountered with the homologous and very similar peptide from *Ps. stutzeri* (T37d, Fig. 2). The 33-residue tryptic *C*-terminal fragment from the *Ps. fluorescens* C protein could not be purified by paper methods, but was readily purified by gel filtration.

All four proteins contain the sequence -Asn-Gly- at positions 50–51 and, as previously noticed (Ambler, 1963*b*), showed the now well-recognized (Konigsberg, 1972; Shotton & Hartley, 1970) property of very ready deamidation. Trouble from β-peptide bonds, which would block phenyl isothiocyanate degradation, was avoided by investigation of this region of sequence with the tryptic peptides in which residue 50 was *N*-terminal. If these peptides were purified as rapidly as possible, the labile residue could be removed by phenyl isothiocyanate degradation before excessive deamidation had occurred.

In both the *Ps. mendocina* and the *Ps. fluorescens* C protein chymotrypsin hydrolysed the bond on the *C*-terminal side of asparagine-34 rapidly. Only a small proportion of asparaginyl peptide bonds appear to be susceptible to this enzyme, and it has been difficult to explain such specificity. Interestingly, in the other cytochromes *c*-551, position 34 is occupied by tyrosine or phenylalanine. An analogous phenomenon has been observed in mitochondrial cytochromes *c* from higher plants, which contain asparagine at position 33. The adjacent peptide bond is rapidly split by chymotrypsin (Thompson *et al.*, 1971), as is the homologous bond in the cytochrome of some lower vertebrates where this site is occupied by an aromatic amino acid. These observations suggest that factors related to the tertiary structure of the substrate may contribute to the specificity of the protease.

No attempts were made to use micro methods for sequence determination. The peptide-purification methods used were optimum for quantities in the range 2–4 μmol of protein digest, and yielded ample quantities of most peptides for sequence determination. Other workers have managed to obtain sequences of mitochondrial cytochromes *c* with total quantities of material as small as 1 μmol (Thompson *et al.*, 1971), but when quantities as small as this are used there can be no margin for accident or error and very great care is necessary to avoid contamination with amino acids and peptides from dust or sweat (Hamilton, 1965).

*Comparison of sequences*

The four *Pseudomonas* cytochrome *c*-551 sequences are clearly homologous with one another. No internal deletions or insertions need be postulated to maximize the matching of the sequences, and the polypeptide chains are the same length (Fig. 5). The conserved residues are distributed fairly evenly through the sequence. The *N*-terminal sequence is variable, but the proline-rich region around the iron-ligand methionine is maintained. Nevertheless the amount of difference between each sequence is considerable, as shown by a difference matrix (Fig. 6). These values should be compared with the differences in the sequence of mitochondrial cytochrome *c* between mammals and amphibians (about 18%) or between mammals and insects (about 33%) (Dayhoff, 1969). The large differences observed between proteins from relatively closely related bacteria suggest that differences in corresponding proteins from more diverse micro-organisms may be so great that indications of a common evolutionary origin will not be obvious.

Many different *c*-type cytochromes are known, and several have been purified and partially characterized (Kamen & Horio, 1970*a*; Ambler, 1971*a*). In bacteria,

```
                    10                  20                  30                  40
(1) E D P E V L F K N K G C V A C H A I D T K M V G P A Y K D V A A K F A G Q A G A E A
(2) E D G A A     S   P   A       T I D S   M         L   E         N     V K D   D K
(3) Q D G E A     S   P   A       S I D A   L         F   E         Y     Q D G   A D
(4) A S G E E     S   P   G       S V Q A   L         L   D         N     V D G   A D

                50                  60                  70                  80
(1) E L A Q R I K N G S Q G V W G P I P M P P N A V S D D E A Q T L A K W V L S Q K
(2) T     G H       T Q   N     P I           Q   T D A     L T     Q   V   S L
(3) L     G H       S Q   V     P I           P   T E E     K I     E   I   S Q
(4) V     G H       S T   V     A M           P   T E E     K T     E   V   T L
```

Fig. 5. *Comparison of the amino acid sequences of cytochromes c-551 from different species of Pseudomonas*

(1) *Ps. aeruginosa* P6009 (Ambler, 1963*b*); (2) *Ps. fluorescens* biotype C no. 18; (3) *Ps. stutzeri* no. 221; (4) *Ps. mendocina* CH110. Gaps in the lower lines indicate where the sequence is common to all four proteins. The one-letter notation used is that recommended by the IUPAC–IUB Commission on Biochemical Nomenclature [*Biochem. J.* (1969) **113**, 1–4].

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| (1) *Ps. aeruginosa* P6009 | 0 | 32 | 32 | 39 |
| (2) *Ps. fluorescens* C-18 | — | 0 | 28 | 33 |
| (3) *Ps. stutzeri* 221 | — | — | 0 | 22 |
| (4) *Ps. mendocina* CH110 | — | — | — | 0 |

Fig. 6. *Difference matrix for amino acid sequences of Pseudomonas cytochromes c-551*

The values shown are differences per 100 residues.

proteins with one, two, three (cytochrome $c_7$; Ambler, 1971*b*) and four (cytochrome $c_3$; Ambler *et al.*, 1971) covalently attached haem groups per polypeptide chain are known. Other proteins (the cytochromes *c'*) exist in which, although the haem is attached to the protein at a conventional cysteine/histidine cluster (Cusanovich *et al.*, 1970), the single haem is attached close to the *C*-terminus of the polypeptide chain (Ambler, 1971*a*; T. Meyer & R. P. Ambler, unpublished work). The extent of variation between such extreme bacterial cytochromes *c* is so great that a polyphyletic origin would seem the simplest explanation. If the likelihood of separate origins for such diverse molecules as the cytochromes $c_3$ and *c'* is conceded, then the possibility of multiple origins for 'standard' cytochromes *c* (molecules with a single haem group covalently attached near the *N*-terminus of an 80–120 residue polypeptide chain, with a methionine side chain as the sixth iron ligand) must be seriously considered.

Several authors have discussed the relationship of the sequence of *Ps. aeruginosa* cytochrome *c*-551 to the mitochondrial proteins (Sackin & Sneath, 1965; Cantor & Jukes, 1966; Dus *et al.*, 1968; Needleman & Blair, 1969; Sackin, 1969; Dickerson, 1971; McLachlan, 1971). Most of these authors consider that the similarities detected are sufficiently extensive to indicate that the proteins are homologous, i.e. have a common evolutionary origin. These arguments are not cumulative, as a different scheme of matching is proposed in each paper. The additional information provided by the three new sequences does not greatly strengthen any of the proposed matchings, and we believe that the available evidence is insufficient to prove homology. By their nature, methods in which the sequences of amino acids or nucleotides are compared cannot disprove theories of common evolutionary origin. On the other hand comparisons of tertiary structure might provide such disproof, and it is to be hoped that such information will be available soon.

This work was supported by the Medical Research Council (Research Group on Bacterial Enzyme Variation).

## References

Ambler, R. P. (1963*a*) *Biochem. J.* **89**, 341–349
Ambler, R. P. (1963*b*) *Biochem. J.* **89**, 349–378
Ambler, R. P. (1971*a*) in *Recent Developments in the Chemical Study of Protein Structures* (Previero, A., Pechere, J.-F. & Coletti-Previero, M.-A., eds.), pp. 289–305, Inserm, Paris
Ambler, R. P. (1971*b*) *FEBS Lett.* **18**, 351–353
Ambler, R. P. (1972*a*) *Methods Enzymol.* **25**, 143–154
Ambler, R. P. (1972*b*) *Methods Enzymol.* **25**, 262–272
Ambler, R. P. & Brown, L. H. (1967) *Biochem. J.* **104**, 784–825
Ambler, R. P. & Meadway, R. J. (1968) *Biochem. J.* **108**, 893–895
Ambler, R. P., Bruschi, M. & Le Gall, J. (1971) *FEBS Lett.* **18**, 347–350

Brown, J. R. & Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228
Cantor, C. R. & Jukes, T. H. (1966) *Proc. Nat. Acad. Sci. U.S.* **56**, 177–184
Cusanovich, M. A., Tedro, S. & Kamen, M. D. (1970) *Arch. Biochem. Biophys.* **141**, 557–570
Dayhoff, M. O. (1969) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Silver Spring, Md.
DeLange, R. J., Glazer, A. N. & Smith, E. L. (1969) *J. Biol. Chem.* **244**, 1385–1388
Dickerson, R. E. (1971) *J. Mol. Biol.* **57**, 1–15
Dickerson, R. E. (1972) *Sci. Amer.* **226**, 58–72
Dickerson, R. E., Takano, T., Eisenberg, D., Kollai, O. B., Samson, L., Cooper, A. & Margoliash, E. (1971) *J. Biol. Chem.* **246**, 1511–1535
Dus, K., Sletten, K. & Kamen, M. D. (1968) *J. Biol. Chem.* **243**, 5507–5518
Erlanger, B. F. & Cohen, W. (1962) *J. Amer. Chem. Soc.* **85**, 348–349
Fanger, M. W., Hettinger, T. P. & Harbury, H. A. (1967) *Biochemistry* **6**, 713–720
Gray, W. R. (1972*a*) *Methods Enzymol.* **25**, 121–138
Gray, W. R. (1972*b*) *Methods Enzymol.* **25**, 333–344
Hamilton, P. B. (1965) *Nature (London)* **205**, 284–285
Harley-Mason, J. & Archer, A. A. P. G. (1958) *Biochem. J.* **69**, 60P
Hirs, C. H. W. (1956) *J. Biol. Chem.* **219**, 611–621
Horio, T., Higashi, T., Sasagawa, M., Kusai, K., Nakai, M. & Okunuki, K. (1960) *Biochem. J.* **77**, 194–201
Jessen, O. (1965) *Pseudomonas aeruginosa and other Green Fluorescent Pseudomonads: a Taxonomic Study*, Munksgaard, Copenhagen
Kamen, M. D. & Horio, T. (1970*a*) *Annu. Rev. Biochem.* **39**, 673–700
Kamen, M. D. & Horio, T. (1970*b*) *Annu. Rev. Microbiol.* **24**, 399–428
Kodama, T. & Shidara, S. (1969) *J. Biochem. (Tokyo)* **65**, 351–360
Konigsberg, W. (1972) *Methods Enzymol.* **25**, 326–332
Konigsberg, W. & Hill, R. J. (1962) *J. Biol. Chem.* **237**, 2547–2561
McLachlan, A. D. (1971) *J. Mol. Biol.* **61**, 409–424
Michl, H. (1951) *Monatsh. Chem.* **82**, 489–493
Needleman, S. B. & Blair, T. T. (1969) *Proc. Nat. Acad. Sci. U.S.* **63**, 1227–1233
Palleroni, N. J., Doudoroff, M., Stanier, R. Y., Solanes, R. E. & Mandel, M. (1970) *J. Gen. Microbiol.* **60**, 215–231
Pink, J. R. L., Buttery, S. H., DeVries, G. M. & Milstein, C. (1970) *Biochem. J.* **101**, 33–47
Rhodes, M. E. (1959) *J. Gen. Microbiol.* **21**, 221–263
Rhodes, M. E. (1961) *J. Gen. Microbiol.* **25**, 331–345
Sackin, M. J. (1969) in *Numerical Taxonomy* (Cole, A. J., ed.), pp. 241–256, Academic Press, London
Sackin, M. J. & Sneath, P. H. A. (1965) *Biochem. J.* **96**, 70P
Shotton, D. M. & Hartley, B. S. (1970) *Nature (London)* **225**, 802–806
Smithies, O. (1959) *Advan. Protein Chem.* **14**, 65–113
Spackman, D. H. (1963) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **22**, 244
Spies, J. R. & Chambers, D. C. (1948) *Anal. Chem.* **20**, 30–39
Stanier, R. Y., Palleroni, N. J. & Doudoroff, M. (1966) *J. Gen. Microbiol.* **43**, 159–272
Thompson, E. W., Notton, B. A., Richardson, M. & Boulter, D. (1971) *Biochem. J.* **124**, 787–791
van Niel, C. B. & Allen, M. B. (1952) *J. Bacteriol.* **64**, 413–422
Waley, S. G. & Watson, J. (1953) *Biochem. J.* **57**, 529–538