Biochem. J. (1973) **135**, 751–758
Printed in Great Britain

The Amino Acid Sequence of Cytochrome *c'* from *Alcaligenes* sp. N.C.I.B. 11015

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

(*Received* 28 *June* 1973)

The amino acid sequence of the cytochrome *c'* from *Alcaligenes* sp. N.C.I.B. 11015 (Iwasaki's 'Pseudomonas denitrificans') has been determined. This organism is the only non-photosynthetic bacterium in which the protein has been found. The protein consists of a single polypeptide chain of 127 residues, with a single haem covalently attached to two cysteines. Unlike normal cytochromes *c*, the haem attachment site is very close to the *C*-terminus. The amino acid sequence around the haem attachment site is very similar to that of *Chromatium vinosum* D cytochrome *c'*. Detailed evidence for the amino acid sequence of the protein has been deposited as Supplementary Publication SUP 50022 at the British Library (Lending Division), (formerly the National Lending Library for Science and Technology), Boston Spa, Yorks. LS23 7BQ, U.K., from whom copies may be obtained on the terms given in *Biochem. J.* (1973) **131**, 5.

Cytochromes *c'* (Kennel *et al.*, 1972) are a class of *c*-type cytochrome in which the haem prosthetic groups are bound covalently as in other cytochromes *c*, but in which the haem groups (under normal physiological conditions) are in a ligand field of high-spin character (Kamen, 1963), like haemoglobin or myoglobin, and unlike normal cytochromes *c*. Cytochromes *c'* have been found in several photosynthetic bacteria (Vernon & Kamen, 1954; Meyer, 1970), and also in a single strain of denitrifying bacterium (Iwasaki & Mori, 1955). Owing to their unusual spectral properties, the proteins have been called by an extraordinary variety of names since their first recognition (e.g. 'pseudohaemoglobin', Vernon & Kamen, 1954; 'RHP', Bartsch & Kamen, 1958; 'cryptocytochrome *c*', Suzuki & Iwasaki, 1962), but now cytochrome *c'* has been proposed (Kennel *et al.*, 1972) as a proper and consistent name, and will be employed in the text that follows.

The denitrifying organism isolated by Iwasaki & Mori (1955) from soil in Japan was tentatively identified as *Pseudomonas denitrificans* (Iwasaki *et al.*, 1963), but further examination (M. Doudoroff & R. Y. Stanier, personal communication; A. J. Holding, personal communication) has shown that it should rather be placed in the ill-defined *Alcaligenes/Achromobacter* group. This organism also produces an azurin (Suzuki & Iwasaki, 1962; Strahs, 1969) which has been shown to be closely related in amino acid sequence to the azurins from pseudomonads and from recognized strains of *Alcaligenes* (Ambler, 1971*a*). The cytochrome *c'* has been purified by Suzuki & Iwasaki (1962) and by Cusanovich *et al.* (1970), and both groups have examined the physical properties of the protein. It has been shown to consist of a dimer with subunits of molecular weight approximately 14000, each containing a single haem group. Cusanovich *et al.* (1970) have investigated the amino acid sequence around the haem attachment site, and have shown that the arrangement of the cysteine and histidine residues is as in most normal low-spin cytochromes *c* (Tuppy & Bodo, 1954).

In the present paper the elucidation of the complete amino acid sequence of the *Alcaligenes* cytochrome *c'* is reported. The sequence near the haem group is nearly the same as that tentatively proposed by Cusanovich *et al.* (1970).

Experimental

Growth of organism

The organism was obtained from Dr. G Strahs. It has been deposited at the National Collection of Industrial Bacteria, Aberdeen, U.K., as N.C.I.B. 11015. The experimental conditions used for the growth of the organism were very similar to those used for denitrifying pseudomonads (Ambler, 1963*a*; Ambler & Wynn, 1973). The organism was grown at 30°C under nearly anaerobic conditions in 35-litre batches on a nitrate–citrate–yeast extract medium. After the organism was harvested, an acetone-dried cell powder was prepared, with a yield of about 0.9 g/l of culture.

Preparation of cytochrome c'

The purification method used was basically the same as has been used for purifying cytochromes *c* and azurins from strains of *Pseudomonas* (Ambler, 1963*a*; Ambler & Wynn, 1973). Steps 1–5, omitting step 2 (extraction, acidification to pH4.1 and adsorption

and chromatography on CM-cellulose) were as described for the *Pseudomonas* proteins. At step 5 (chromatography on CM-cellulose), the azurin was eluted with 0.05M-ammonium acetate, pH4.7, a very small amount of cytochrome *c'* with 0.05M-ammonium acetate, pH5.1, and the bulk of the cytochrome *c'* with 0.1M-ammonium acetate, pH7. At step 6, the main cytochrome *c'* component was precipitated by $(NH_4)_2SO_4$ (at pH6) at between 50 and 65% saturation. Step 7 (DEAE-cellulose chromatography) was omitted. At step 8 (gel filtration through Sephadex G-75 in 0.05M-ammonium acetate, pH5.1) the cytochrome *c'* was eluted as the only haem-containing component at a volume that corresponded to a molecular weight of about 25000. This coloured fraction was freeze-dried, dissolved in a small volume of water, and $\frac{1}{5}$ volume of 50% (v/v) formic acid was added. The acid solution was then subjected to gel filtration through Sephadex G-75 equilibrated with 5% (v/v) formic acid, column etc. conditions being otherwise as normally used for step 8. Two well-resolved protein components were eluted and were present in approximately equal amounts. The component with the apparently higher molecular weight was colourless and was discarded. The smaller component was the cytochrome *c'*. It was freeze-dried and dissolved in water at a concentration of about 1mM. The final yield was about 1μmol from 30g of dry cells, the product of 35 litres of culture.

Properties of cytochrome c'

The general spectral properties of the cytochrome *c'* prepared by this method agreed well with those of the protein purified by Cusanovich *et al.* (1970) by a different method. In particular, the ratio of absorption at 280nm to that of the Soret maximum (Fe^{3+} form) at 400nm was 0.39, indicating an absence of impurities that were u.v.-absorbing but colourless.

Amino acid sequence determination

The amino acid sequence was determined by a study of the peptides produced by enzymic hydrolysis of cytochrome *c'* with trypsin, chymotrypsin and thermolysin, and by CNBr cleavage. Before enzymic digestion the haem was removed from the protein by treatment with $HgCl_2$ in 0.1M-HCl–8M-urea. The methods used were as described by Ambler (1963*b*) and Ambler & Wynn (1973).

Papain was used to digest one peptide further. Mercuripapain [Sigma (London) Chemical Co., London S.W.6, U.K.] was suspended in water (at 5mg/ml), centrifuged to wash it free of amino acids and peptides, and then dissolved (to 1mg/ml) in 5% (v/v) pyridine adjusted to pH5.5 with acetic acid, and containing 1% (w/v) dithiothreitol. The peptide (400nmol) was dissolved in 0.2ml of the same buffer and 0.2ml of the enzyme solution added. The mixture was incubated at 37°C for 5h, and then dried under vacuum.

Peptide fragments were separated by gel filtration, followed by high-voltage paper electrophoresis and paper chromatography, analysed quantitatively for amino acids, and examined for sequence by the Dns–phenyl isothiocyanate method. In some cases peptides were further degraded by other proteolytic enzymes.

Quantities of material used

The sequence-determination experiments were carried out with two completely independent preparations of the protein. From the first batch (4.3μmol), tryptic peptides were characterized from a digest of 1.8μmol, chymotryptic peptides from 1.6μmol, and a preliminary CNBr-cleavage experiment performed on 0.4μmol. From the second batch (6.3μmol), thermolysin peptides were characterized from a digest of 1.9μmol, selected chymotryptic peptides from 0.9μmol, and the two smaller CNBr fragments purified from an experiment with 1.8μmol.

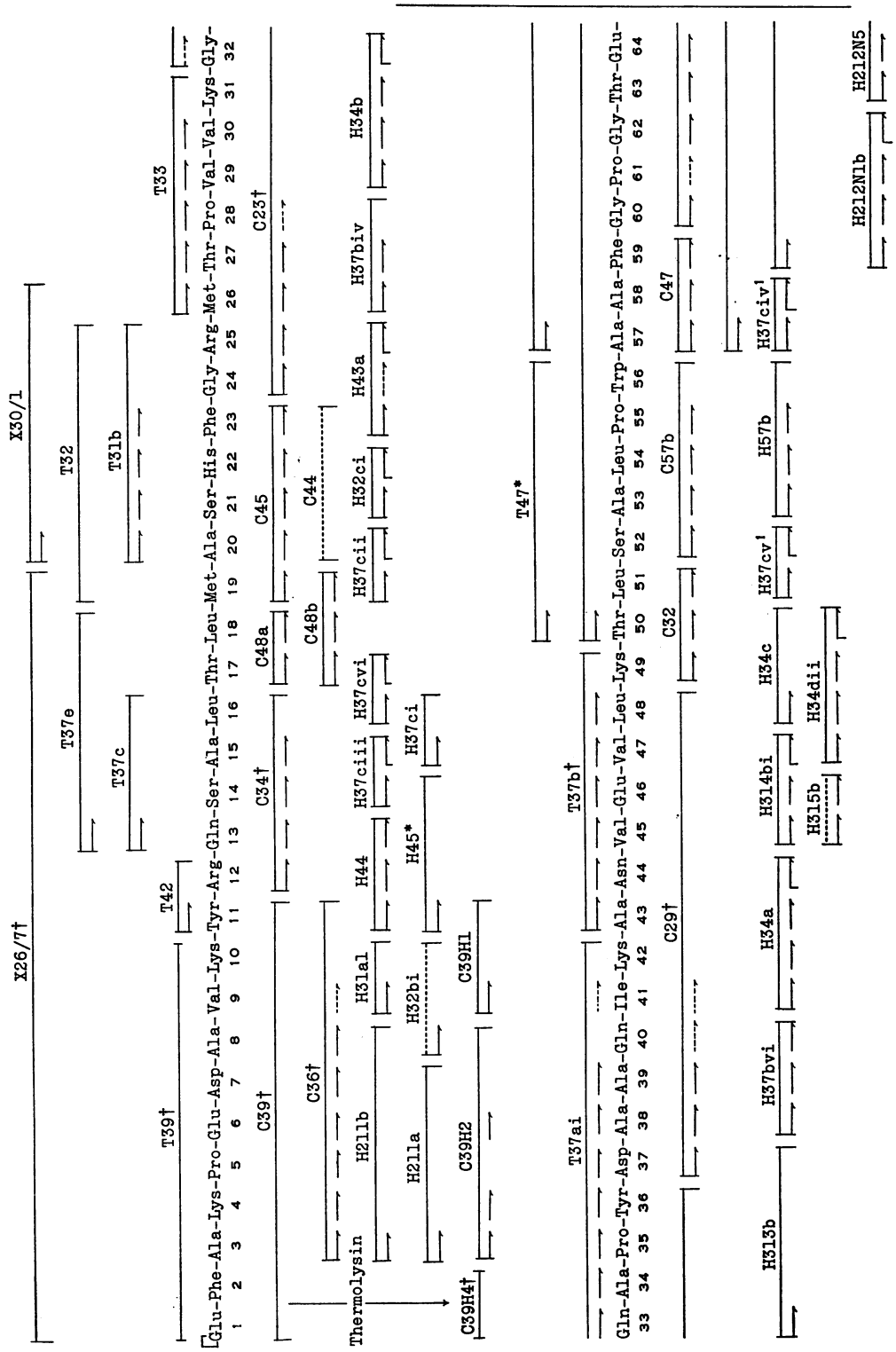
Results

The amino acid composition of *Alcaligenes* cytochrome *c'* is shown in Table 1, and is in good agreement with that determined by Cusanovich *et al.* (1970).

In agreement with Cusanovich *et al.* (1970), no free *N*-terminal group could be detected in the protein (by a Dns method), even after removal of the haem. Sequence investigation has shown that the *N*-terminus is a modified glutamic acid residue, and the evidence is compatible with it being pyrrolidonecarboxylic acid. After removal of the haem, carboxypeptidase A treatment released 2.4mol of lysine/mol of protein, but little of any other amino acid; treatment with carboxypeptidase A together with a small amount (0.6 unit/60nmol of protein) of carboxypeptidase B released nearly 3mol of lysine/mol of protein, together with nearly molar amounts of arginine and tyrosine. Full quantitative details of these experiments are given in the Supplementary Publication (SUP 50022).

The blocked *N*-terminus has prevented any use of the automatic sequenator (Edman & Begg, 1967) in this investigation. The sequence has been determined by the classical method of characterizing overlapping peptide fragments.

The evidence for the proposed amino acid sequence is summarized in Fig. 1. This shows all the peptides that were isolated from each digest, together with the peptides formed by secondary digestions. Symbols show how much of the sequence of each peptide has been determined by the Dns–phenyl isothiocyanate


X30/1
X26/7†
T32
T37e
T31b
T37c
T33
T42
T39†
C23†
C45
C48a
C34†
C44
C48b
C39†
C36†
H37biv
H34b
H43a
H32ci
H37cii
H37cvi
H37ciii
H44
H45*
H37ci
H31a1
H21b
H21a
H32bi
C39H1
C39H2
C39H4†
Thermolysin
T47*
T37b†
T37ai
C47
C57b
C32
C29†
H37civ1
H57b
H37cv1
H34c
H34dii
H314bi
H315b
H34a
H37bvi
H313b
H212N1b
H212N5
Glu-Phe-Ala-Lys-Pro-Glu-Asp-Ala-Val-Lys-Tyr-Arg-Gln-Ser-Ala-Leu-Thr-Leu-Met-Ala-Ser-His-Phe-Gly-Arg-Met-Thr-Pro-Val-Val-Lys-Gly-
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
Gln-Ala-Pro-Tyr-Asp-Ala-Ala-Gln-Ile-Lys-Ala-Asn-Val-Glu-Val-Leu-Lys-Thr-Leu-Ser-Ala-Leu-Pro-Trp-Ala-Ala-Phe-Gly-Pro-Gly-Thr-Glu-
33 34 35 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

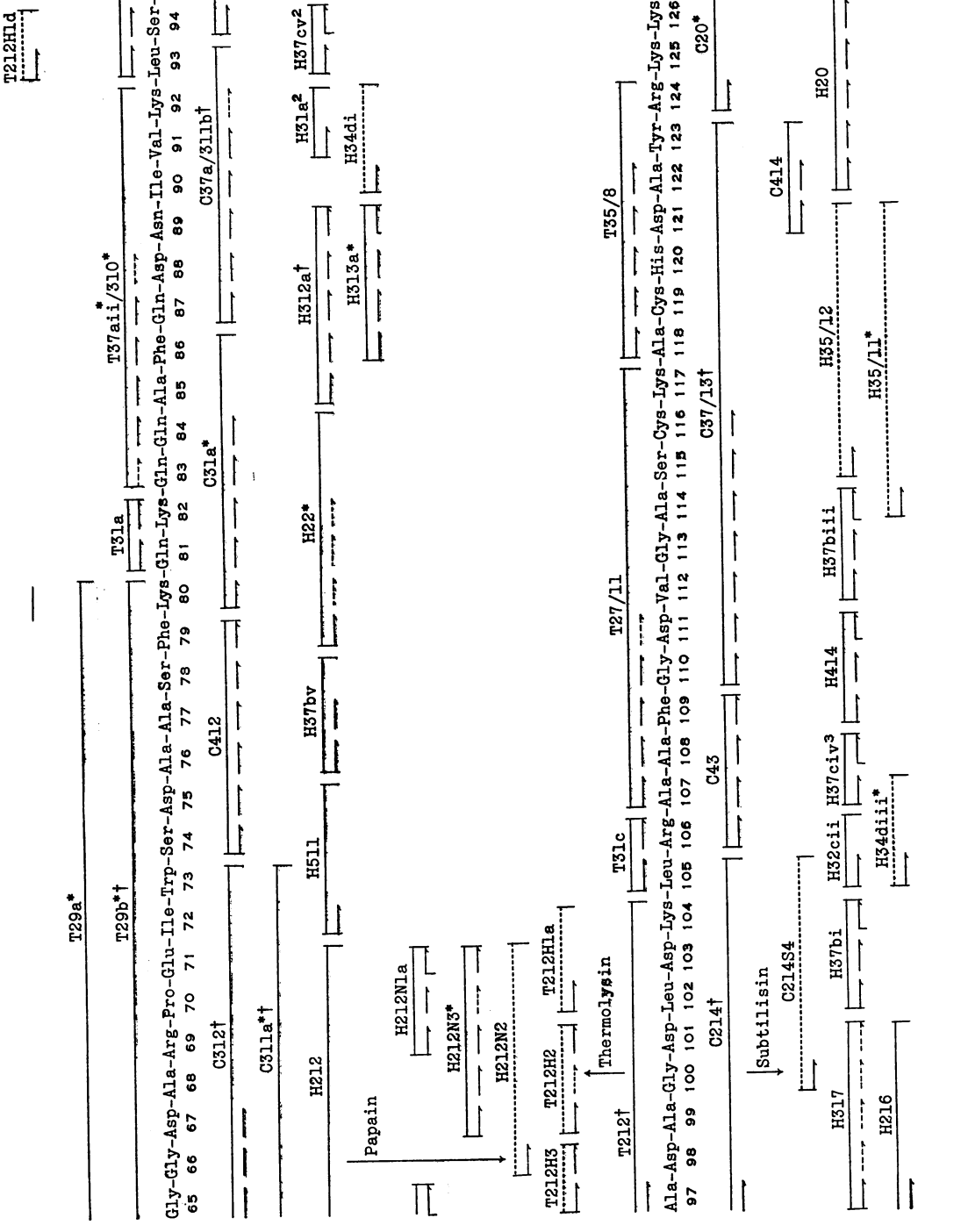

T212H1d
T29a*
T29b*†
T31a
T37aii/310*
Gly-Gly-Asp-Ala-Arg-Pro-Glu-Ile-Trp-Ser-Asp-Ala-Ala-Ser-Phe-Lys-Gln-Gln-Lys-Gln-Gln-Ala-Phe-Gln-Asp-Asn-Ile-Val-Lys-Leu-Ser-Ala-Ala-
65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96
C31a*
C37a/311b†
C412
C312†
C311a*†
H22*
H312a†
H31a2
H37cv2
H37civ2
H34di
H313a*
H212
H511
H37bv
Papain
H212N1a
H212N3*
H212N2
T212H3
T212H2
T212H1a
Thermolysin
T51c
T27/11
T35/8
T212†
Ala-Asp-Ala-Gly-Asp-Leu-Asp-Lys-Leu-Arg-Ala-Ala-Phe-Gly-Asp-Val-Gly-Ala-Ser-Cys-Lys-Ala-Cys-His-Asp-Ala-Tyr-Arg-Lys-Lys-Lys
97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127
C37/13†
C20*
C43
C214†
Subtilisin
C214S4
C414
H35/12
H20
H414
H37bili
H37bi
H32cii
H37civ3
H35/11*
H317
H34diii*
H216

Table 1. *Amino acid composition of Alcaligenes cytochrome c'*

Results are shown as residues per molecule. Hydrolysis and analysis samples were: (1), native protein, 24h hydrolysis; (2), native protein, 96h hydrolysis; (3), haem-free protein, 24h hydrolysis; (4), haem-free protein, oxidized with performic acid, 24h hydrolysis. Results for (1)–(4) were calculated on the basis that the total of all residues except tryptophan and cysteine was 123.

	Hydrolysis and analysis					
	(1)	(2)	(3)	(4)	Sequence	Cusanovich *et al.* (1970)
Glycine	9.3	9.5	8.6	9.9	9	10
Alanine	24.5	24.7	24.9	24.3	24	25
Valine	6.6	7.2	6.4	6.2	7	6
Leucine	8.2	8.1	8.4	8.4	8	7
Isoleucine	2.8	3.2	2.7	2.6	3	2
Serine	6.5	5.7	6.3	6.9	7	7
Threonine	3.9	3.7	4.0	3.9	4	4
Aspartic acid	12.3	12.4	12.9	12.2	10	12
Asparagine					2	
Glutamic acid	12.7	12.5	12.8	12.6	4	12
Glutamine					8	
Phenylalanine	5.9	5.9	6.2	6.1	6	6
Tyrosine	2.9	3.0	3.2	2.7	3	3
Tryptophan			0.9		2	1
Cysteine	0.7	1.4	0.9	1.5*	2	4
Methionine	2.1	1.5†	1.2†	2.3‡	2	2
Proline	6.5	5.7	5.7	5.9	6	6
Lysine	11.2	13.3	12.7	12.0	13	12
Histidine	2.1	2.0	1.9	1.9	2	1
Arginine	4.8	4.7	5.1	5.0	5	5
					127	126

* As cysteic acid.
† Some methionine sulphoxides present also, but not included in total.
‡ As methionine sulphone; peak not well resolved from aspartic acid.

method, as well as peptides that were examined with carboxypeptidase. Cases in which the results of Dns–phenyl isothiocyanate degradation or quantitative amino acid analysis were not considered wholly satisfactory are indicated.

Dns–phenyl isothiocyanate degradation results were considered unsatisfactory when more than one Dns-amino acid was identified as present in significant amount (subjective visual estimate of more than 20%) at a step of the degradation, or when for some reason or another (such as accident or shortage of material) the Dns-amino acid was not positively identified.

The criteria for satisfactory amino acid analyses are: (1) that no impurity should be present in amount larger than 0.20mol/mol and (2) that when the relative amounts of amino acids present are calculated on the basis that the average amount is integral, no

Fig. 1. *Amino acid sequence of Alcaligenes cytochrome c'*

Peptides derived by tryptic digestion (T) and CNBr cleavage (X) are shown above the sequence, and by chymotryptic (C) and thermolysin (H) digestion below the sequence. Vertical arrows show peptides produced by further digestion. Full lines indicate quantitative amino acid analyses (substandard if marked *). Broken lines indicate qualitative analyses. $\rightarrow$ indicates end groups and subsequent residues revealed by phenyl isothiocyanate degradation, and identified by the Dns method (substandard if shown as $\dashrightarrow$). $\rightarrowtail$ indicates final residues of peptides identified as the free amino acid after the removal of the remainder by phenyl isothiocyanate degradation. Peptides marked † were examined by carboxypeptidase A digestion. The *N*-terminal residue is probably pyrrolidonecarboxylic acid ($\square$Glu) derived from glutamine (see the text).

values should fall outside the limits 0.8–1.2, 1.8–2.2, 2.7–3.3 or 3.7–4.3. Nevertheless values as low as 0.7 (=1) are considered acceptable for tyrosine, as low yields of this amino acid from pure peptides are often found.

The peptides that did not meet these analytical criteria (marked * on Fig. 1) failed for a variety of reasons.

(1) Peptides T37aii and T310: both were derived from exactly the same region of sequence, but in peptide T310 (which was continually formed from peptide T37aii during attempts to purify the latter) the *N*-terminal glutamine residue had cyclized. The region contains an Ile–Val peptide bond stable to acid hydrolysis, and since only 24h hydrolyses were performed for each peptide the isoleucine and valine values are impermissibly low. Peptide T37aii was very difficult to separate from peptide T37ai, and as the amount remaining decreased with each purification step, the final product was contaminated with the common impurity amino acids from paper (Ser 0.34, Gly 0.29).

(2) Peptides T29a and T29b: these large tryptophan-containing peptides were difficult to purify, and elution recovery from paper was particularly bad (a notorious property of tryptophan peptides) so final yields were very low. The analyses were satisfactory except that peptide T29a showed serine 2.4 (=2) and peptide T29b threonine 1.7 (=2). In the latter case, threonine was *N*-terminal, and slight but significantly lower recovery of the *N*-terminal residue is frequently noticed after acid hydrolysis.

(3) Peptide C20: the analysis results for this peptide are shown as lysine 3.0, arginine 0.6, though they could obviously be recalculated as (say) lysine 5.0, arginine 1.0. Other evidence (e.g. the analysis values for peptide H20) indicate three residues of lysine, and if this is so, the low recovery of arginine would be another instance of the *N*-terminal effect (see above). Recoveries of this peptide (and of peptide H20) from paper were very poor, even when the strip was eluted with 1 M-acetic acid rather than the usual 0.1 M-aq. NH_3.

(4) Peptide C31a: lysine recovery was low (1.7 = 2); lysine was *N*-terminal, but the peptide may not have been quite pure.

(5) Peptides T47, C311a and H45 were peptides produced by minor splits and isolated in low yield. On analysis they showed contamination with moderate amounts (0.2–0.35 mol/mol) of unexpected amino acids. In three other cases, similar levels of contamination could be explained by the presence of appreciable amounts of recognizable peptides with similar purification properties. These cases were peptide H34diii contaminated by H34di, peptide H212N3b by H212N2, and peptide H22 by an expected peptide of the same sequence except that it would contain an additional residue of alanine at the *C*-terminus.

Supplementary publication

Detailed evidence for the amino acid sequence of the protein has been deposited as Supplementary Publication SUP 50022 with the British Library (Lending Division) (formerly the National Lending Library for Science and Technology) for storage on microfiche.

This evidence comprises:

(1) Tables showing the properties of all the peptides shown in Fig. 1. The successive steps used in the purification of each peptide are given, and where applicable the values of V/V_0 [elution volume/void volume for gel filtration through Sephadex G-25 in 5% (v/v) formic acid] and electrophoretic mobility on paper at pH 6.5. Absolute percentage yields are given for peptides from primary digests, and relative yields for peptides from secondary digests. Amino acid analyses are given for all peptides, which show all impurities present in amounts greater than 0.1 mol/mol of peptide.

(2) Further tables show the individual sequence evidence for each peptide. This consists of the *N*-terminal group analysis results, and the results of Dns–phenyl isothiocyanate-degradation experiments. Details are given when these results are substandard.

(3) The evidence for the presence or absence of amide groups on each aspartic acid and glutamic acid residue in the protein is given.

(4) Details are given of experiments with carboxypeptidase A on the peptides marked † in Fig. 1, and of the effect of carboxypeptidases A and B on the whole protein.

Discussion

Accuracy of proposed amino acid sequence

The evidence for the amino acid sequence proposed is summarized in Fig. 1. The strength of the evidence varies between different parts of the proposed sequence, with the weakest parts being considered to be those discussed in detail below.

There is no good evidence that the blocked *N*-terminal residue is pyrrolidone-carboxylic acid instead of some other derivative such as *N*-acetylglutamine. The observed electrophoretic mobility of peptide C39H4 at pH 6.5 was −0.70. The predicted mobility for [illegible]Glu-Phe would be −0.64 (Offord, 1966) whereas Ac-Gln-Phe would be expected to have a mobility of −0.52. The reproducibility of observed mobilities is not good enough for much weight to be given to these values. Definitive evidence could be obtained by mass spectroscopy, or by testing with an enzyme known to remove pyrrolidone-carboxylic acid specifically.

The evidence for the number of successive lysine residues at the *C*-terminus is weak. No evidence was obtained from the tryptic digest, as the poly-lysine

peptide (if not degraded to free lysine) was lost because of its high electrophoretic mobility. The amino acid composition of peptide C20 (the relevant chymotryptic peptide) could be interpreted as containing five lysines (see above). The evidence from the thermolysin peptide H20, from the action of carboxypeptidases on the whole protein, and of the lysine content of the whole protein (Table 1) is all compatible with there being three lysines in this tail.

There is no quantitative evidence for the tryptophan content of the protein, as experience with other *c*-type cytochromes (see, e.g., Ambler, 1963*a*) has shown the difficulty of getting reliable results either before or after haem removal. The presence of tryptophan residues in peptides was recognized by the Ehrlich reaction and by the characteristic fluorescences on paper. All these putative tryptophan-containing peptides were also characteristically retarded during gel filtration. The anomalous V/V_0 values for these peptides are recorded in the Supplementary Publication (SUP 50022). The values of V/V_0 do not suggest that any of the small peptides contain two residues of tryptophan. The u.v.-absorption profile (at 280nm) of the eluate obtained on gel filtration of the chymotryptic digest, although only semi-quantitative, suggested that approximately equal amounts of tryptophan were present in peptides C57b and C311a/C312. The final yield of peptide C57b was very low, probably owing to the very poor recovery from paper.

Peptides T212 and C214 are located in positions 94–104 (Fig. 1) on the negative evidence that there is nowhere else in the sequence that these peptides could be located, and location here is completely compatible with expected proteolytic action. No simple experiment (such as trypsin digestion after selective modification of lysine residues) has been thought of that was likely to give conclusive evidence. The peptides concerned were both isolated in good yield, and the sequence proposed in Fig. 1 to include them agrees well in amino acid composition with that obtained experimentally for the whole protein (Table 1).

The sequence proposed in this paper differs from that proposed for the haem region by Cusanovich *et al.* (1970). These workers included the sequence -Glu-Ala-Asp- between residues 115 and 116 (Fig. 1), and -Ser-Gly- between residues 120 and 121. They also suggested that the residue corresponding to 121 was asparagine, not aspartic acid. Otherwise their sequence for residues 107–124 is exactly the same as that now proposed. Cusanovich *et al.* (1970) reported difficulty in getting enough of their peptides for good amino acid analyses, and their extra residues were inserted in their sequence to account for what were probably free amino acids that occurred as impurities in their peptide preparations. They isolated the region of residues 107–124 with the haem still attached, and so there can be little doubt that the haem is covalently attached to the cysteine residues 116 and 119. Direct evidence that this region of sequence containing the haem attachment site is close to the *C*-terminus of the molecule is given by the results from carboxypeptidase treatment of the whole protein.

Comparison of sequence with that of other c-type cytochromes

The only similarity in sequence between this cytochrome *c'* and other *c*-type cytochromes that has been noticed is in the covalent haem attachment site, where the -Cys-X-Y-Cys-His- sequence is common. In all the monohaem *c*-type cytochrome sequences (except chromes *c'*) known to the author, the haem attachment site is close to the *N*-terminus of the polypeptide chain (Ambler, 1971*a*). In the polyhaem cytochromes c_3 (Ambler *et al.*, 1971*a*) and c_7 (Ambler, 1971*b*) one of the haem attachment sites is, as in cytochrome *c'*, close to the *C*-terminus. Preliminary results for the cytochrome c_3 from *Desulfovibrio salexigens* (R. P. Ambler, M. Bruschi & J. Le Gall, unpublished work) suggests that it also has a *C*-terminal poly-lysine tail.

No similarity can be seen between the sequence proposed for the *Alcaligenes* cytochrome *c'* (Fig. 1) and that proposed by Dus *et al.* (1962) for the haem-binding region of *Chromatium vinosum* cytochrome *c'*.

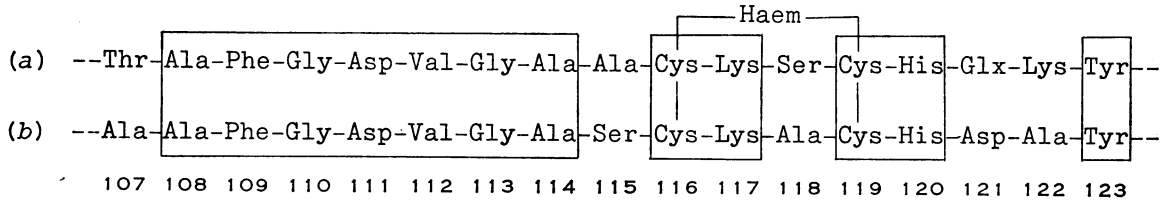


Fig. 2. *Amino acid sequences around the haem-binding site of the cytochromes c' of Chromatium vinosum D (a); (Kennel et al., 1972) and Alcaligenes sp. N.C.I.B. 11015 (b)*

Residues that are the same in the two sequences are enclosed by the boxes. The *C*-terminus of the *Alcaligenes* protein is residue 127.

This protein was then called *Chromatium* RHP, and believed to be a dihaem protein of about 26000 molecular weight. The cytochromes *c'* from photosynthetic bacteria have recently been re-examined (Kennel *et al.*, 1972) and shown to be monohaem proteins of size very similar to the *Alcaligenes* cytochrome *c'*. The amino acid sequence of the haem-binding site in *C. vinosum* cytochrome *c'* has been redetermined (Kennel *et al.*, 1972), and shows very considerable similarity with the corresponding sequence proposed here for the *Alcaligenes* cytochrome *c'* (Fig. 2). It has also been shown (T. E. Meyer, R. P. Ambler & M. D. Kamen, unpublished work) that this similarity extends right through the sequences, and is also to be seen in the sequences of the cytochromes *c'* of *Rhodospirillum rubrum*, *Rhodopseudomonas palustris* and *Rhodopseudomonas gelatinosa*.

The publication of the incorrect sequence of the *C. vinosum* cytochrome *c'* (Dus *et al.*, 1962) has served a useful purpose, as this sequence has been compared with many other presumably correct amino acid sequences by workers searching for significant structural similarities between proteins. The significant similarities claimed to be seen between this mythical sequence and that of cytochromes c_3 (Ambler *et al.*, 1971*b*) or *Pseudomonas* cytochrome *c*-551 (Sackin, 1969) is a warning of the danger of trying to overinterpret sequence results.

Distribution of respiratory proteins in bacteria

Cytochrome *c'* has an unusual distribution. It is found both in purple sulphur bacteria and in most purple bacteria (Meyer, 1970), being apparently absent from the purple *Rhodomicrobium vannieli*. In non-photosynthetic bacteria, despite fairly rigorous searches, it has only been found up until now in the present *Alcaligenes* strain. This organism also produces an azurin (Suzuki & Iwasaki, 1962), which has been shown to be clearly homologous to the azurins from other *Alcaligenes* species, *Bordetella bronchiseptica* and pseudomonads (Sutherland & Wilkinson, 1963; Ambler, 1971*a*). Azurins are not known in photosynthetic bacteria.

These findings could be explained by the present organism being a close symbiotic association of a micro-organism related to the purple photosynthetic bacteria and an *Alcaligenes* strain, comparable with the *Desulfovibrio/Chlorobium* ecosystem 'Chloropseudomonas ethylica' (Gray *et al.*, 1972; Van Beeumen & Ambler, 1973), although there is no direct bacteriological evidence in support of this whatsoever. A more likely explanation is that a process of this sort has gone further, a cytochrome *c'* gene from an organism like a purple photosynthetic bacterium having been incorporated by an azurin-possessing *Alcaligenes* strain. Since the incorporation, the two proteins would have co-evolved, and are now integrated into the same electron-transport system. There are suggestions from amino acid sequence results that comparable transfers have taken place between other groups of bacteria, such as the penicillinase gene between *Staphylococcus aureus* and *Bacillus licheniformis* (Ambler & Meadway, 1969). If this phenomenon can be shown to occur frequently, it will imply that no general phylogenetic classification is possible for bacteria. It will be possible to construct phylogenetic trees from amino acid sequence results for single genes (or parts of genes), but the trees for different genes will have different conformations.

I thank Mrs. Elizabeth Taylor for her enthusiastic assistance, and Dr. T. E. Meyer for his interest and advice.

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. & Meadway, R. J. (1969) *Nature (London)* **222**, 24–26
Ambler, R. P. & Wynn, M. (1973) *Biochem. J.* **131**, 485–498
Ambler, R. P., Bruschi, M. & Le Gall, J. (1971*a*) *FEBS Lett.* **18**, 347–350
Ambler, R. P., Bruschi, M. & Le Gall, J. (1971*b*) *Proc. Int. Congr. Microbiol. 19th* pp. 25–34
Bartsch, R. G. & Kamen, M. D. (1958) *J. Biol. Chem.* **230**, 41–63
Cusanovich, M. A., Tedro, S. M. & Kamen, M. D. (1970) *Arch. Biochem. Biophys.* **141**, 557–570
Dus, K., Bartsch, R. G. & Kamen, M. D. (1962) *J. Biol. Chem.* **237**, 3083–3093
Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
Gray, B. H., Fowler, C. F., Nugent, N. A. & Fuller, R. C. (1972) *Biochem. Biophys. Res. Commun.* **47**, 322–327
Iwasaki, H. & Mori, T. (1955) *J. Biochem. (Tokyo)* **42**, 375–380
Iwasaki, H., Shidara, S., Suzuki, H. & Mori, T. (1963) *J. Biochem. (Tokyo)* **53**, 299–303
Kamen, M. D. (1963) *Acta Chem. Scand.* **17**, S41–S46
Kennel, S. J., Meyer, T. E., Kamen, M. D. & Bartsch, R. G. (1972) *Proc. Nat. Acad. Sci. U.S.* **69**, 3432–3435
Meyer, T. E. (1970) Ph.D. Thesis, University of California, San Diego
Offord, R. E. (1966) *Nature (London)* **211**, 591–593
Sackin, M. J. (1969) in *Numerical Taxonomy* (Cole, A. J., ed.), pp. 241–256, Academic Press, London
Strahs, G. (1969) *Science* **165**, 60–61
Sutherland, I. W. & Wilkinson, J. F. (1963) *J. Gen. Microbiol.* **30**, 105–112
Suzuki, H. & Iwasaki, H. (1962) *J. Biochem. (Tokyo)* **52**, 193–199
Tuppy, H. & Bodo, G. (1954) *Monatsch. Chem.* **85**, 807–821
Van Beeumen, J. & Ambler, R. P. (1973) *Antonie van Leeuwenhoek; J. Microbiol. Serol.* **39**, 355–356
Vernon, L. P. & Kamen, M. D. (1954) *J. Biol. Chem.* **211**, 643–662