Biochem. J. (1986) 233, 333–337 (Printed in Great Britain)

The amino acid sequence of cytochrome *c*-555 from the methane-oxidizing bacterium *Methylococcus capsulatus*

R. P. AMBLER,* H. DALTON,† T. E. MEYER,‡§ R. G. BARTSCH‡ and M. D. KAMEN‡||

*Department of Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland, U.K., †Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, U.K., and ‡Department of Chemistry, University of California at San Diego, La Jolla, CA 92093, U.S.A.

The amino acid sequence of the cytochrome *c*-555 from the obligate methanotroph *Methylococcus capsulatus* strain Bath (N.C.I.B. 11132) was determined. It is a single polypeptide chain of 96 residues, binding a haem group through the cysteine residues at positions 19 and 22, and the only methionine residue is at position 59. The sequence does not closely resemble that of any other cytochrome *c* that has yet been characterized. Detailed evidence for the amino acid sequence of the protein has been deposited as Supplementary Publication SUP 50131 (12 pages) at the British Library Lending Division, Boston Spa, West Yorkshire LS23 7BQ, U.K., from whom copies are available on prepayment.

INTRODUCTION

Cytochromes *c* are widely distributed metalloproteins that are recognizable by their possession of an intense and characteristic u.v.–visible-absorption spectrum. This spectrum is associated with the presence of a haem group covalently attached to a polypeptide chain. Cytochromes of the *c*-type are commonly both soluble and small, in contrast with *a*-type and *b*-type cytochromes, and are often present in high concentrations in actively metabolizing cells. Because of their spectra they can be recognized in intact cells or in crude extracts, and so their protein moiety provides a gene product that can be isolated without any preconceptions as to function or to sequence-related properties such as charge or size. A comparative study of cytochrome *c* sequences can therefore be used to explore genetic relationships between very diverse organisms, and is effective for identifying similarities that are beyond the range of detection of nucleic acid hybridization.

The amino acid sequences of many different cytochromes *c* have been determined, from mitochondria and chloroplasts as well as from a wide range of bacteria, and can be divided into several structural classes (Ambler, 1982; Meyer & Kamen, 1982). Cytochromes of related sequence occur in organisms as diverse as purple phototrophic bacteria and eukaryotes (Ambler *et al.*, 1976), in cases where the cellular metabolism, and hence the functional role of the cytochrome, must be quite distinct. As part of an investigation of cytochrome *c* diversity, we are characterizing the soluble haem proteins of a selection of bacteria that possess unusual features in their energy metabolism.

Bacteria that can grow on C_1 compounds such as methane and methanol have attracted much attention because of their potential as sources of single-cell protein, of overproducing metabolites, and as biocatalysts for oxidizing hydrocarbons. Their biochemistry has been reviewed by Anthony (1982). Many different types of micro-organism have been isolated that are able to grow on reduced carbon compounds that contain no carbon–carbon bonds, but apart from their methylotrophy they have few unifying taxonomic features. Metabolically they can be divided into obligate and facultative methanotrophs, into those that can grow on methanol or methylamine but not methane, and by the different pathways that they use for carbon assimilation. In the present paper we report studies on electron-transport-chain components from *Methylococcus capsulatus* (Foster & Davis, 1966), an obligate methylotroph that grows best on methane, which is incorporated after oxidation to formaldehyde through the ribulose monophosphate cycle (Ferenci *et al.*, 1974).

EXPERIMENTAL

Purification of soluble cytochromes *c*

Soluble cytochromes *c* were found to be present at high concentrations in *Methylococcus capsulatus* during characterization of the methane mono-oxygenase (Colby & Dalton, 1978). *Methylococcus capsulatus* strain Bath (N.C.I.B. 11132) [this organism was found in the water from the hot spring (45 °C) in the Roman Baths at Bath, U.K., and is the only methanotroph that has so far been isolated from this source] was grown aerobically at 45 °C on a methane/ammonia/mineral salts medium (Colby *et al.*, 1977). The cytochrome *c*-555 used in the sequence determination was purified from freeze-dried cells, which were suspended in 0.1 M-Tris/HCl buffer, pH 7.3, and homogenized in a Ribi Cell Fractionator operated at 137 MPa (20000 lb/in²). Membranes were removed by centrifugation at 235000 *g* for 3 h, and the extract was desalted by gel filtration through Sephadex G-25 into 1 mM-Tris/HCl buffer, pH 8. This solution was passed through a column of Whatman DE-52 DEAE-cellulose equilibrated with the same buffer, to which about one-third of the soluble cytochrome was absorbed. The unretained cytochrome fraction, containing the cytochrome *c*-555 and a cytochrome *c'*, was then passed through

§ Present address: Department of Biochemistry, University of Arizona, Tucson, AR 85721, U.S.A.
|| Present address: Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, U.K.

a column of CM-cellulose equilibrated with 1 mM-potassium phosphate buffer, pH 6, which did not retain them. These cytochromes were then concentrated by $(NH_4)_2SO_4$ precipitation, and subjected to gel filtration through Sephadex G-75 in 20 mM-Tris/HCl buffer, pH 7.3, containing 0.1 M-NaCl. Traces of membranes and large-M_r cytochromes were eluted in the void volume, followed by the cytochrome *c'*, which was eluted before the cytochrome *c*-555. The cytochrome *c*-555 was adsorbed directly on to home-made (Bernardi, 1971) hydroxyapatite, in a 5 cm × 8 cm diam. column, equilibrated with 1 mM-potassium phosphate buffer, pH 7, containing 0.1 M-NaCl. The column was eluted with a 1-litre linear gradient of 1–100 mM-potassium phosphate buffer, pH 7, containing 0.1 M-NaCl, and the cytochrome *c*-555 was eluted with about 50 mM buffer. Final purification was by $(NH_4)_2SO_4$ fractionation when cytochrome *c*-555 was precipitated between 60 and 90% saturation. The protein migrated as a single band on sodium dodecyl sulphate/polyacrylamide-gel electrophoresis, with an M_r estimated to be 11000. (The recipe used for electrophoresis was supplied by Bethesda Research Laboratories, Gaithersburgh, MD, U.S.A., along with their 'low-molecular-weight' protein standards kit; it used 15% cross-liked gels containing 6 M-urea, 0.1% sodium dodecyl sulphate and 1% mercaptoethanol.) About 8 μmol of pure protein was recovered in 50% yield from 56g dry wt. of cells, an amount comparable with that obtained of the predominant cytochrome *c* in favourable cases from phototrophic or denitrifying bacteria.

In addition to the cytochrome *c*-555, at least three other soluble cytochromes *c* were present in *Methylococcus capsulatus* extracts, each at levels of total haem of about one-third of that of the major cytochrome. These were a neutral cytochrome *c'*, which behaved similarly on gel filtration and sodium dodecyl sulphate/polyacrylamide-gel electrophoresis to the dimeric cytochromes *c'* found primarily in purple phototrophic bacteria (Bartsch, 1978), and two acidic cytochromes *c*. These spectrally similar proteins (α-band maxima 551–552 nm) appeared to be easily denatured, as judged by slow alteration of spectral properties and a partial CO-binding capability, were precipitated at low concentrations of $(NH_4)_2SO_4$ (0–50% saturation), but have not yet been isolated in a homogeneous form. They are clearly different in M_r.

Table 1. Amino acid composition of *Methylococcus capsulatus* cytochrome *c*-555

Results were calculated on the basis that Σ(Gly + Ala + Leu + Phe + His + Arg) = 33. The samples were hydrolysed at 105 °C with 6 M-HCl. Sample (B) was hydrolysed for 96 h, the others for 24 h. The protein had the haem removed before hydrolysis, and sample (C) was also oxidized with performic acid before hydrolysis. The 'best value' is the average of the accepted estimates from analyses (A)–(C). Values discarded because the hydrolysis conditions were not optimal for the recovery of the amino acid concerned are shown in parentheses.

	Amino acid composition (mol of residue/mol of protein)				
	Hydrolysis and analysis				
	(A)	(B)	(C)	'Best value'	Sequence
Gly	9.1	9.1	9.2	9.1	9
Ala	9.4	8.7	9.2	9.1	9
Val	(4.4)	4.6	(3.5)	4.6	5
Leu	7.2	6.9	6.9	7.0	7
Ile	(2.4)	2.9	(2.4)	2.9	3
Ser	4.1	(3.9)	4.2	4.1	4
Thr	4.6	4.5	4.7	4.6	5
Asp	10.9	10.7	10.8	10.8	5
Asn					6
Glu	12.3	11.9	11.9	12.0	7
Gln					5
Phe	4.1	3.9	3.9	4.0	4
Tyr	3.6	3.6	(0.1)	3.6	4
Trp				(0.07)¶	0
Cys			1.9*	1.9	2
Met	0.9	0.9	0.9†	0.9	1
Pro	6.2	5.8	5.4	5.8	6
Lys	9.2	10.5	9.5	9.7	10
His	1.0	1.1	1.1	1.1	1
Arg	2.7	3.2	3.0	3.0	3
Total					96

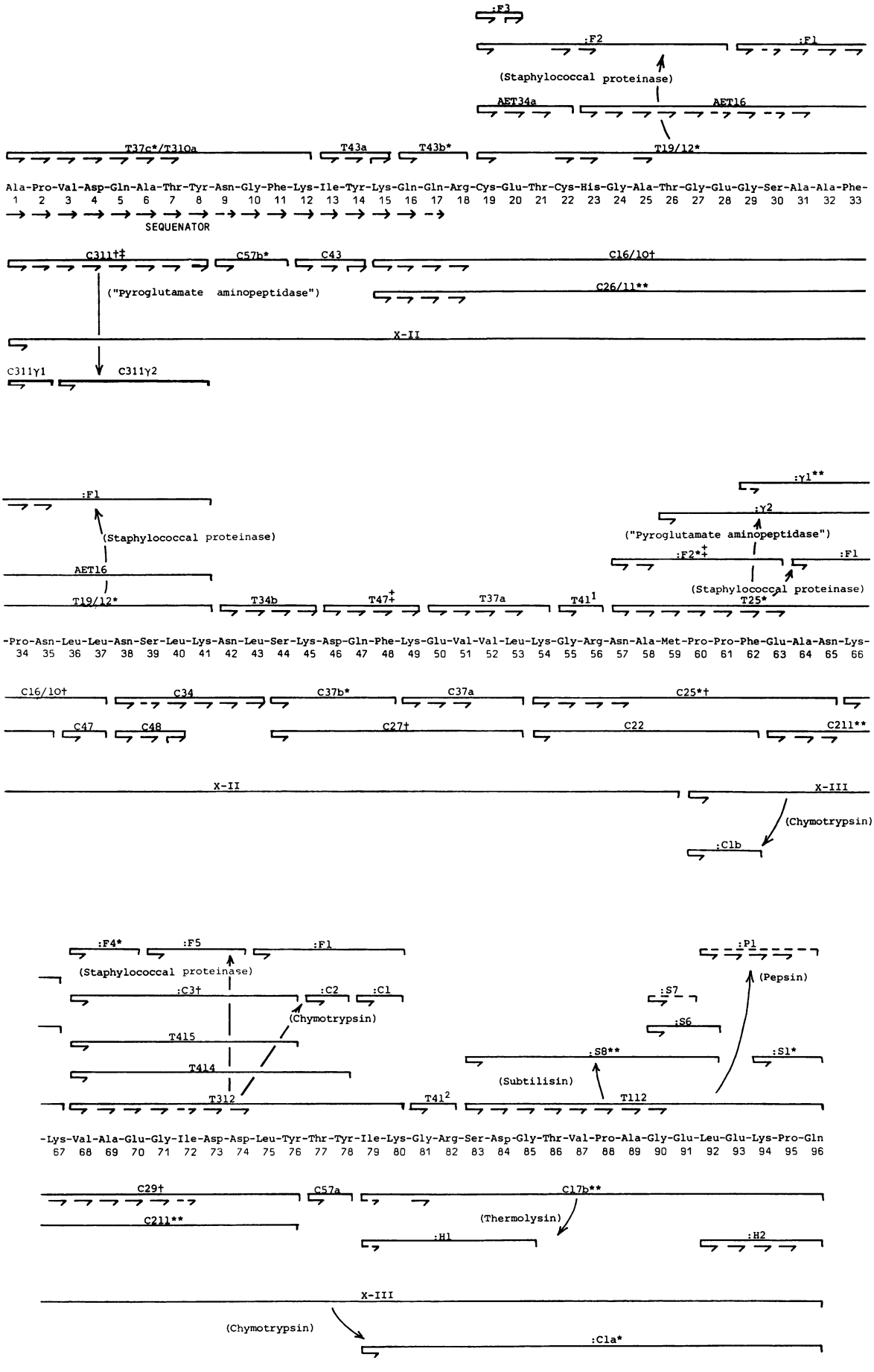
* As cysteic acid.
† As methionine sulphone.
¶ A fourth sample, of the native protein, was hydrolysed with 3 M-mercaptoethanesulphonic acid for 96 h at 105 °C. Tryptophan would have been detected if it had been present at a level greater than 0.07 mol of residue/mol of protein.

Amino acid sequence determination

The amino acid sequence was determined by the general methods that have been used for other bacterial cytochromes *c* (Ambler & Wynn, 1973; Ambler *et al.*, 1979), and to similar standards. Protein (about 2 μmol) was treated with 20 mg of $HgCl_2$ in 5 ml of 8 M-urea/0.1 M-HCl at 37 °C for 16 h to remove the haem moiety, and after gel filtration and freeze-drying the sample was digested with a proteinase. In other experiments the native protein was cloven with CNBr or the protein *S*-aminoethylated after haem removal and before

Fig. 1. Amino acid sequence of *Methylococcus capsulatus* cytochrome *c*-555

Peptides derived by digestion with trypsin (T) or by trypsin after *S*-aminoethylation (AET) are shown above the sequence, and by digestion with chymotrypsin (C) or by cleavage with CNBr (X) below the sequence. Peptides from subdigests are labelled with another letter indicating the second method of cleavage [F, staphylococcal proteinase; P, pepsin; S, subtilisin; γ, 'pyroglutamate aminopeptidase' (see the text)]. All peptides except T112P1 and T112S7 were analysed quantitatively for amino acid composition, but those marked * gave substandard and those marked ** particularly bad analyses. See Ambler & Wynn (1973) for a quantitative definition of these symbols. Under peptide lines, ⟝ indicates end groups and subsequent residues revealed by phenyl isothiocyanate degradation and identified by the dansyl (5-dimethylaminonaphthalene-1-sulphanyl) method, substandard if shown ⟝. *C*-Terminal residues identified as free amino acids after removal of the remainder by phenyl isothiocyanate degradation are shown ⟝. Peptides marked † were examined by carboxypeptidase A digestion, and those marked ‡ by aminopeptidase M digestion. Under the sequence, → indicates residues identified as phenylthiohydantoins after degradation of the whole protein in an automatic sequenator, only tentatively identified if marked ⇢.


:F3
:F2
:F1
(Staphylococcal proteinase)
AET34a
AET16
T37c*/T31Oa
T43a
T43b*
T19/12*
Ala-Pro-Val-Asp-Gln-Ala-Thr-Tyr-Asn-Gly-Phe-Lys-Ile-Tyr-Lys-Gln-Gln-Arg-Cys-Glu-Thr-Cys-His-Gly-Ala-Thr-Gly-Glu-Gly-Ser-Ala-Ala-Phe-
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
SEQUENATOR
C311†‡
C57b*
C43
C16/1O†
("Pyroglutamate aminopeptidase")
C26/11**
X-II
C311γ1
C311γ2
:F1
(Staphylococcal proteinase)
AET16
T19/12*
T34b
T47‡
T37a
T41¹
:γ1**
:γ2
("Pyroglutamate aminopeptidase")
:F2*‡
:F1
(Staphylococcal proteinase)
T25*
-Pro-Asn-Leu-Leu-Asn-Ser-Leu-Lys-Asn-Leu-Ser-Lys-Asp-Gln-Phe-Lys-Glu-Val-Val-Leu-Lys-Gly-Arg-Asn-Ala-Met-Pro-Pro-Phe-Glu-Ala-Asn-Lys-
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 65 66
C16/1O†
C34
C37b*
C37a
C25*†
C47
C48
C27†
C22
C211**
X-II
X-III
(Chymotrypsin)
:C1b
:F4*
:F5
:F1
(Staphylococcal proteinase)
:C3†
:C2
:C1
(Chymotrypsin)
T415
T414
T312
T41²
:P1
(Pepsin)
:S7
:S6
:S8**
:S1*
(Subtilisin)
T112
-Lys-Val-Ala-Glu-Gly-Ile-Asp-Asp-Leu-Tyr-Thr-Tyr-Ile-Lys-Gly-Arg-Ser-Asp-Gly-Thr-Val-Pro-Ala-Gly-Glu-Leu-Glu-Lys-Pro-Gln
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
C29†
C57a
C17b**
C211**
(Thermolysin)
:H1
:H2
X-III
(Chymotrypsin)
:C1a*

```
(1)  -CYS-Glu-Thr-CYS-HIS-GLY-Ala-Thr-GLY-Glu-GLY-Ser-ALA-Ala-----Phe-PRO-ASN-LEU-
           20                  25                  30                      35

(2)  -CYS-Val-Ala-CYS-HIS-GLY-Thr-Asp-GLY-Gln-GLY-Leu-ALA-Pro-Ile-Tyr-PRO-ASN-LEU-
           15                  20                  25                  30


(1)  -GLY-Arg-ASN-ALA-MET-PRO-PRO-PHE-
       55                  60

(3)  -GLY-Lys-ASN-ALA-MET-PRO-Ala-PHE-
            55                  60

(4)  -GLY-Ile-ASN-ALA-MET-PRO-PRO-Lys-
            60                  65
```

Fig. 2. Comparison of amino acid sequences of cytochromes *c* around the haem-attachment site (above) and the presumed sixth ligand methionine residue (below)

The sequences are: (1), *Methylococcus capsulatus* cytochrome *c*-555; (2), halophilic *Paracoccus* sp. N.C.I.B. 8669 cytochrome *c*-554(548) (see Ambler *et al.*, 1985); (3), *Porphyra tenera* cytochrome *c*-553 (Ambler & Bartsch, 1975); (4), *Pseudomonas mendocina* cytochrome c_5 (Ambler & Taylor, 1973). Matching residues are shown in capitals.

proteolysis. The peptides were fractionated by gel filtration followed by high-voltage paper electrophoresis and chromatography, and then analysed quantitatively for amino acid composition and purity. Peptide sequences were investigated by the dansyl/phenyl isothiocyanate method, and by exo- and endo-peptidase digestion. In two cases peptides were subdigested with commercial 'pyroglutamate aminopeptidase' (Boehringer Mannheim), which in addition to its official activity also has endopeptidase activity, particularly on bonds on the *C*-terminal side of proline residues. The *N*-terminal sequence of the whole native protein was investigated with an automatic sequenator (Beckman model 890A) with a DMBA program (Beckman 102473), the phenylthiohydantoin derivatives being identified by t.l.c. and by hydrolysis to the amino acid. Amide groups were assigned from peptide electrophoretic mobilities and exopeptidase analysis.

RESULTS

The amino acid composition of the cytochrome *c*-555 is shown in Table 1. The evidence for the proposed amino acid sequence is summarized in Fig. 1. Details of the purification, analysis and sequence determination experiments on the peptides are given in Supplementary Publication SUP 50131. The criteria for satisfactory results, and the nature and the format of the supplementary publication, are given in previous papers (Ambler & Wynn, 1973; Ambler *et al.*, 1979).

The amino acid sequence was deduced by characterizing peptides from four separate digests. These were: (*a*) 2.1 μmol digested with trypsin; (*b*) 0.8 μmol of native protein cloven with CNBr; (*c*) 1.9 μmol digested with chymotrypsin; (*d*) 1.4 μmol digested with trypsin after treatment with ethyleneimine under reducing conditions.

DISCUSSION

Accuracy of proposed sequence

The peptides are linked into a single sequence by overlaps, and peptides accounting for the complete postulated sequence were found in each digest. The evidence for the sequence of the *C*-terminal region is weaker than for the rest of the protein molecule. For residues 83–90 the evidence is from a single dansyl/phenyl isothiocyanate degradation experiment, and the amide ascriptions for residues 84, 91, 93 and 96 shown in Fig. 1 are based on the observed neutral mobilities of large peptides, as it proved difficult to obtain small proteolytic fragments from this region in pure form. No evidence was noticed to suggest that residue 84 was ever present in the protein as asparagine, but the well-known lability of -Asn-Gly- sequences makes it possible that the protein when synthesized does contain asparagine at this position. The amide group on the -Asn-Gly- sequence at position 9–10 was observed to be labile during peptide purification.

Structural affinities of *Methylococcus capsulatus* cytochrome *c*-555

The results shown in Fig. 1 indicate that the cytochrome *c*-555 is a cytochrome *c* of class I (Ambler, 1982), with a Cys-Xaa-Yaa-Cys-His- haem-binding site near the *N*-terminus, and the only methionine residue (position 59) is in a location where it is likely to be the sixth iron ligand. The sequence does not closely resemble that of any other cytochrome *c* that has yet been characterized, although the glycine residues close after the haem-attachment site are positioned as in the cytochrome *c*-554(548) from the halophilic *Paracoccus* sp. N.C.I.B. 8669 (see Fig. 2 in Ambler *et al.*, 1985). The sequence around methionine-59 is closely matched in the chloroplast cytochrome *c*-553 from the red alga *Porphyra tenera* (Ambler & Bartsch, 1975) and in *Pseudomonas mendocina* cytochrome c_5 (Ambler & Taylor, 1973) (Fig. 2). None of these organisms has any obvious metabolic or genetic affinity with *Methylococcus capsulatus*. Preliminary studies of the sequences of the cytochromes *c* from the methylotrophs *Pseudomonas* AM1 and *Methylophilus methylotrophus* have not revealed in them any proteins like the *Methylococcus capsulatus* cytochrome *c*-555.

We thank Margaret Daniel, Lesley McLellan and Siv Tedro for skilled and enthusiastic technical assistance. The work was supported by grants from the Medical Research Council to

R.P.A., from the Science Research Council to H.D., and from the National Institutes of Health (GM 18528) and the Department of Energy (ATO3-78ER-70293) to M.D.K.

REFERENCES

Ambler, R. P. (1982) in From Cyclotrons to Cytochromes (Robinson, A. R. & Kaplan, N. O., eds.), pp. 263–280, Academic Press, New York
Ambler, R. P. & Bartsch, R. G. (1975) Nature (London) **253**, 285–288
Ambler, R. P. & Taylor, E. (1973) Biochem. Soc. Trans. **1**, 166–168
Ambler, R. P. & Wynn, M. (1973) Biochem. J. **131**, 485–498
Ambler, R. P., Meyer, T. E. & Kamen, M. D. (1976) Proc. Natl. Acad. Sci. U.S.A. **73**, 472–475
Ambler, R. P., Daniel, M., Meyer, T. E., Bartsch, R. G. & Kamen, M. D. (1979) Biochem. J. **177**, 819–823
Ambler, R. P., Meyer, T. E., Trudinger, P. A. & Kamen, M. D. (1985) Biochem. J. **227**, 1009–1013
Anthony, C. (1982) The Biochemistry of Methylotrophs, Academic Press, London
Bartsch, R. G. (1978) in The Photosynthetic Bacteria (Clayton, R. K. & Sistrom, W. R., eds.), pp. 249–279, Plenum, New York
Bernardi, G. (1971) Methods Enzymol. **22**, 325–339
Colby, J. & Dalton, H. (1978) Biochem. J. **171**, 461–468
Colby, J., Stirling, D. I. & Dalton, H. (1977) Biochem. J. **165**, 395–402
Ferenci, T., Strøm, T. & Quayle, J. R. (1974) Biochem. J. **144**, 477–486
Foster, J. W. & Davis, R. H. (1966) J. Bacteriol. **91**, 1924–1931
Meyer, T. E. & Kamen, M. D. (1982) Adv. Protein Chem. **35**, 105–212

Received 29 May 1985/23 July 1985; accepted 9 September 1985