THE AMINO ACID SEQUENCE OF CYTOCHROME *c*-551.5 (CYTOCHROME c_7) FROM THE GREEN PHOTOSYNTHETIC BACTERIUM *CHLOROPSEUDOMONAS ETHYLICA*

R.P. AMBLER

Department of Molecular Biology, University of Edinburgh, Edinburgh, EH9 3JR, Scotland

Received 10 June 1971

Some photosynthetic bacteria contain a *c*-type cytochrome [1, 2] that is similar in absorption spectra and oxidation–reduction potential to the very low potential cytochrome c_3 of sulphate reducing bacteria [3]. Cytochrome *c*-551.5 was isolated from the green photosynthetic bacterium *Chloropseudomonas ethylica* 2K by Olson and Shaw [2], and has been further purified by Meyer, Bartsch and Kamen [4], who have shown it to be a small multi-haem protein, similar in some features of amino acid composition to the *Desulfovibrio* cytochromes c_3. In the present paper the amino acid sequence of the *Cps. ethylica* cytochrome is reported, and the structural features that it shares with cytochrome c_3 [5–7] discussed.

Cps. ethylica cytochrome *c*-551.5 was prepared by Meyer, Bartsch and Kamen [4]. The preparation contained no impurities detectable by starch-gel electrophoresis (at pH 4.0 or 8.5) or by terminal group analysis, but slight contamination with other proteins was indicated by trace amounts of arginine and methionine in acid hydrolysates of the protein. The protein does not contain tryptophan or trimethyl lysine. The haem was removed by treatment with mercuric chloride in 0.1 N HCl/8 M urea [6]. On treatment of the apoprotein with carboxypeptidase A, equal quantities of lysine, histidine and isoleucine were released, but under similar conditions only lysine and isoleucine were released from performic-acid-oxidized apoprotein. Alanine was the only *N*-terminal amino acid that could be detected in the apoprotein by the dansyl-chloride method. The peptides produced by tryptic digestion of 4 μmole apoprotein were oxidized with performic acid, and fractionated by gel-filtration, high-voltage paper electrophoresis and paper chromatography, and their amino acid sequences determined by standard methods (fig. 1) [8, 9]. The protein lacks methionine, and contains only small amounts of leucine and the aromatic amino acids [4], so neither cyanogen bromide cleavage nor digestion with chymotrypsin would be likely to be helpful in attempts to determine the order of the tryptic peptides. Thermolysin digestion of 4 μmole performic-acid-oxidized apoprotein produced a very complex mixture of peptides, including some that were very large, and many that were different but originated from the same part of the sequence (fig. 1). The data from the amino acid compositions, terminal groups and partial sequence of the thermolysin peptides provided positive evidence for linking all the tryptic peptides except T13a and T5b (fig. 1), as the Lys/Ile bond (36/37, fig. 1) was split by both trypsin and thermolysin. The peptides produced from this region by digestion of 2 μmole performic-acid-oxidized apoprotein with subtilisin B were isolated and fully characterized (fig. 1).

The proposed amino acid sequence is shown in fig. 1. It is considered extremely unlikely that there is any major error, since the thermolysin and trypsin digests of the protein have been examined very carefully, and no imcompatible peptides detected down to a level of a 1% molar yield. The sequence is in good agreement with the reported amino acid composition of the whole protein [4], and agrees with the terminal group data.

Meyer, Bartsch and Kamen [4] have shown that upon hydrolysis the cytochrome yields 23–24 amino acid residues per mole of haem, but estimated the molecular weight by gel-filtration to be about 11,000. The proposed amino acid sequence contains 68 residues, and three cysteine-histidine clusters similar to

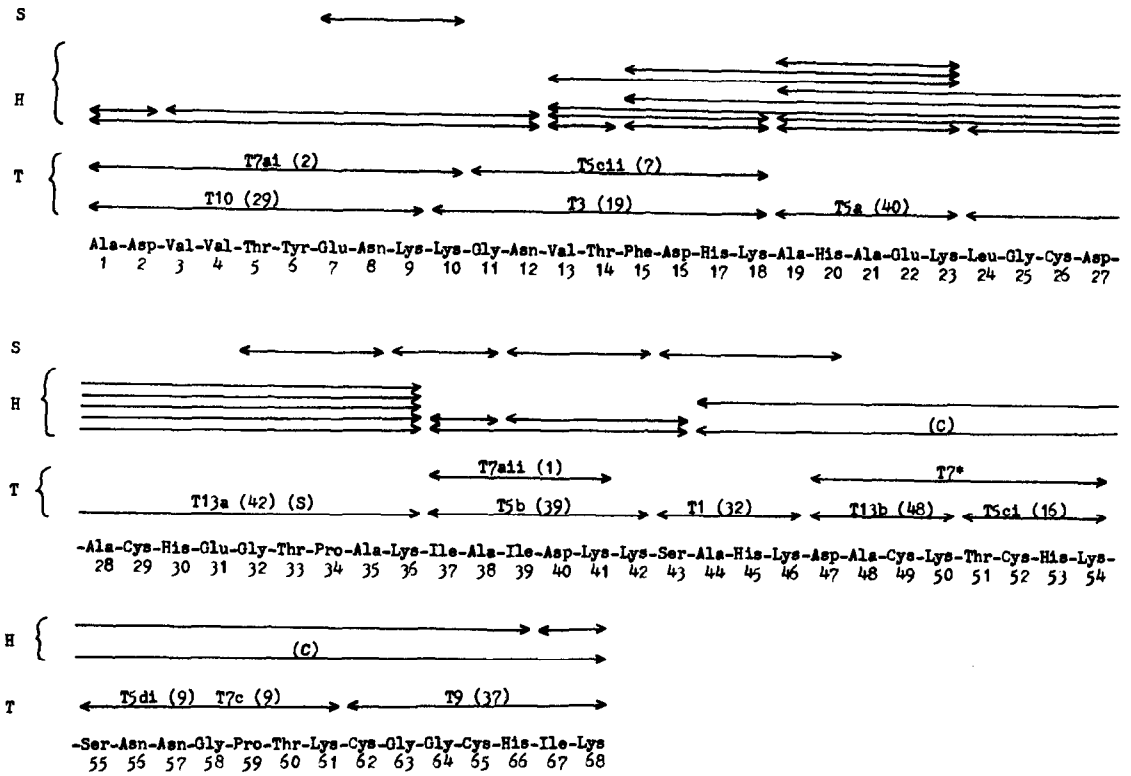

Fig. 1. Amino acid sequence of *Chloropseudomonas ethylica* cytochrome *c*-551.5 (cytochrome c_7). The peptides isolated from trypsin (T), thermolysin (H) and subtilisin B (S) digests are shown by bars. (C), (S), above bars indicate that chymotrypsin or subtilisin B was used for further degradation of the peptide. Where a digest contained two or more peptides derived from the same region of the cytochrome, the lower bar represents the peptide recovered in highest yield.

The yields of tryptic peptides are shown, as mole/100 mole protein. Peptides T5ci, T5cii, T5di, T7ai, T7aii and T7c require two steps of purification more than the other peptides, a cause of lower yields. Peptide T5di readily lost an amide group to form peptide T7c. The yield of free lysine in the tryptic digest was 11 mole/100 mole protein, and was presumably formed from residues 10 and 42. Peptide T7* was formed instead of peptides T5ci and T13b in another experiment, in which performic-acid-oxidized apoprotein was digested with trypsin.

those known to bind haem in mitochondrial cytochrome *c*, giving a formula weight (including three haems) of about 9,100. It is considered possible that this molecular weight discrepancy is due to anomalous behaviour of multi-haem proteins during gel-filtration. The possession of three haems per peptide chain makes *Cps. ethylica* cytochrome *c*-551.5 the first representative of a new and structurally distinct type of cytochrome *c*, and approval is being sought from the cytochrome sub-committee of the Enzyme Commission to call the protein cytochrome c_7.

The *Desulfovibrio* cytochromes c_3 contain four haems [4] attached to peptide chains of 102–110 residues [7]. The three known sequences are clearly homologous, and show similarity to the sequence of the *Cps. ethylica* cytochrome. This similarity is most marked when comparison is made with the cytochrome c_3 of *D. desulfuricans* [7] (fig. 2). Six out of the nine residues in the segment 13–21 of the *Cps. ethylica* sequence are identical to residues in the segment 22–30 of the *D. desulfuricans* sequence, and four of these six residues are constant in all three known cytochromes c_3. At the C-terminus of the molecules, residues 49–68 of the *Cps. ethylica* sequence are identical in fourteen out of twentyone positions to residues 81–101 of the cytochrome c_3, if a single

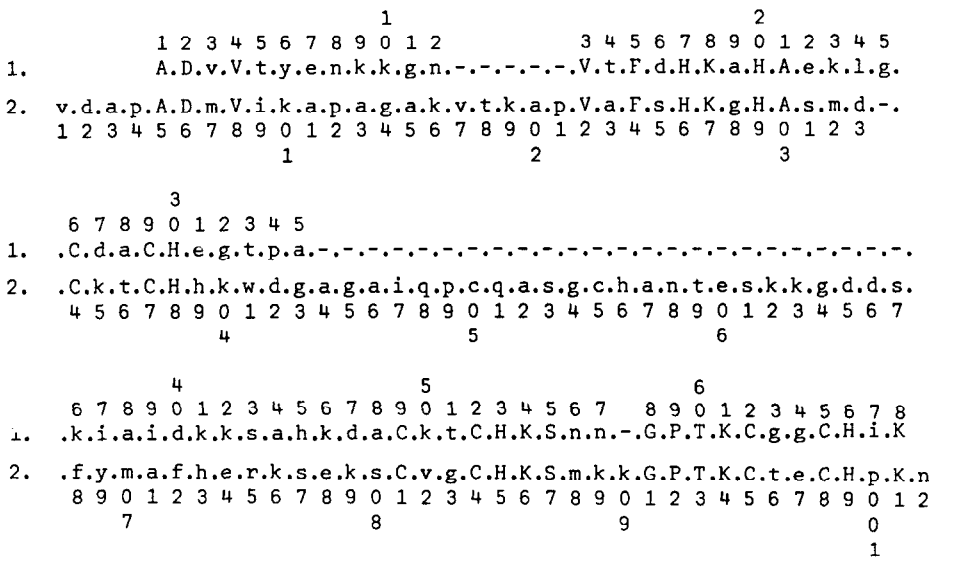

Fig. 2*. Amino acid sequences of cytochrome *c*-551.5 (cytochrome c_7) from *Chloropseudomonas ethylica* (1) and cytochrome c_3 from *Desulfovibrio desulfuricans* (2) [7]. Matching residues are shown in capitals, and residues that are common to all the three cytochromes c_3 are underlined. The putative haem binding sites are boxed. No attempt has been made to juggle deletions in the central section so as to 'improve' the matching.

* A one-letter notation recommended by IUPAC-IUB Commission on Biochemical Nomenclature (1968).

deletion is allowed in the former. This segment includes two of the putative haem binding sites of each molecule, but it is a region that shows considerable variability in different *Desulfovibrio* species [7]. The extent of the sequence similarity demonstrated in fig. 2, is so great that the proteins must be considered homologous, indicating an evolutionary connection between sulphate reducing bacteria and green photosynthetic bacteria.

Acknowledgements

I would like to thank Drs. Terry Meyer, J. Olson and E. Shaw for the gift of protein and for encouragement to work on this problem, and Miss Elizabeth Clark for technical assistance.

References

[1] R.W. Holton and J. Meyers, Biochim. Biophys. Acta 131 (1967) 362.
[2] J.M. Olson and E.K. Shaw, Photosynthetica 3 (1969) 288.
[3] J.R. Postgate, J. Gen. Microbiol. 14 (1956) 545.
[4] T.E. Meyer, R.G. Bartsch and M.D. Kamen, submitted to Biochim. Biophys. Acta (1971).
[5] R.P. Ambler, Biochem. J. 109 (1968) 47P.
[6] R.P. Ambler, M. Bruschi and J. Le Gall, FEBS Letters 5 (1969) 115.
[7] R.P. Ambler, M. Bruschi and J. Le Gall, FEBS Letters 2 (1971).
[8] R.P. Ambler and L.H. Brown, Biochem. J. 104 (1967) 784.
[9] W.R. Gray, Methods in Enzymology 11 (1967) 469.