Amino Acid Sequence of Cytochrome c' from the Purple Photosynthetic Bacterium *Rhodospirillum rubrum S1*

(Received for publication, September 26, 1974)

TERRANCE E. MEYER, RICHARD P. AMBLER,‡ ROBERT G. BARTSCH, AND MARTIN D. KAMEN

From the Department of Chemistry, University of California, San Diego, La Jolla, California 92033, and the Department of Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland

The amino acid sequence of cytochrome c' from the purple photosynthetic bacterium *Rhodospirillum rubrum* S1 has been determined and is consistent with homology to cytochrome c' from the nonphotosynthetic bacterium *Alcaligenes* sp. NCIB 11015. There is 29% identity in the chosen alignment of these two proteins. *R. rubrum* cytochrome c' is composed of a single peptide chain of 126 amino acid residues with a single heme covalently bound near the COOH terminus. There is no sequence similarity to mitochondrial cytochrome c, except at the heme binding site.

Cytochromes c' form a class of high spin heme proteins, related to other cytochromes c in solubility, in the covalent mode of attachment of the heme via cysteine thioether bonds, and possibly in having a histidine as fifth ligand to the iron. The high spin electronic absorption spectra of cytochromes c' differ in major features from those of other c-type cytochromes (which have typically low spin spectra) and in fine structure from those of the high spin heme proteins: myoglobin, hemoglobin, peroxidase, and catalase. Cytochromes c' bind carbon monoxide and nitric oxide with perturbation of the heme absorption spectra (1), whereas the globins bind a variety of small molecules in addition to nitric oxide and carbon monoxide. The cytochromes c' are isolated primarily as dimers (molecular weight 28,000) of identical monoheme subunits (2, 3). They exhibit midpoint oxidation-reduction potentials at pH 7 of 0 to 150 mv (2) and possess isoelectric points ranging from 4.6 to 9.6 (2).

The cytochromes c' were first isolated from photosynthetic bacteria, and more recently from some nonphotosynthetic bacteria. They have been found in most, although not all, purple and purple sulfur bacteria, but not in the green sulfur bacteria or the blue-green ("algae") bacteria (4). Cytochromes c' are notably absent from *Rhodopseudomonas viridis* NTHC 133 (6), and *Rhodopseudomonas palustris* strain 6 (ATCC 17001) (7), while present in all other purple bacteria examined, including two other *R. palustris* strains (4, 8). Cytochromes c' have been found in the strictly aerobic bacterium *Azotobacter vinelandii* (9), in the denitrifying bacterium *Alcaligenes* sp. NCIB 11015 (previously designated *Pseudomonas denitrificans*) (10), and as the hydroxylamine reductase of a halotolerant "micrococcus" ATCC 12084 (11). They have not been reported in such major groups as clostridia, coliforms, or higher organisms.

The results of the amino acid sequence of the nonphotosynthetic *Alcaligenes* sp. NCIB 11015 cytochrome c' have been published (12). This report is the first in a projected series on the sequences of several purple and purple sulfur photosynthetic bacterial cytochromes c'.

**RESULTS**

*Rhodospirillum rubrum* S1 (ATCC 11170) cells were grown, and cytochrome c' was isolated and purified according to Bartsch (8). Materials and procedures for determining the amino acid sequence were as described by Ambler and Wynn (13) and by Gray (14). Five micromoles of protein were used for each primary enzyme digest, except *Staphylococcus aureus* protease digestion, for which 3.7 μmole of cytochrome were used. Peptides were purified by chromatography on Sephadex G-25-F or G-50-F, by paper electrophoresis at pH 6.5, paper chromatography in butanol/acetic acid/water/pyridine, and dansyl amino acids were identified by chromatography on polyamide sheets (14). The proposed amino acid sequence of the protein is shown in Fig. 1, along with peptide overlaps.

1 The abbreviation used is: dansyl, 5-dimethylaminonaphthalene-1-sulfonate.

2 Uniform standards for evaluation of sequence data are difficult to establish, inasmuch as a variety of methodologies exist, automatic as well as manual. One recommendation holds that in the absence of sufficient quantitative data specific for each residue in a given peptide, the residues should be determined by at least two independent routes. However, an alternative procedure based on the Edman approach with dansylation—a new in these researches—is to ensure definitive assignments by use of many peptides which need be examined only once, assurance of reliability in assignment being based on quantitative amino acid composition data for each peptide with confirmation by comparison of the composition of the summed peptides with that of the whole protein. In these researches we have used three independent

---

*This research was supported by Grant #BMS72-02409 from the National Science Foundation and Grant GM 18528 from the National Institutes of Health.

†T. E. Meyer and M. D. Kamen, unpublished results.
Peptide designation in the supplementary tables and the following text corresponds to position in the sequence.

The amino acid sequences of the tryptic peptides were determined (Tables V and VII). They were aligned in a unique manner with the aid of the chymotryptic and Staphylococcus aureus protease (E) peptides for which the amino acid compositions and NH$_2$ termini were determined (Tables II and III). In addition, the amino acid sequences of several chymotryptic peptides (Table VI) and the amino acid compositions and NH$_2$ termini of the thermolytic and tryptic fragments of S. aureus protease peptides 31-50 and 78-92, respectively, were determined. There is some precedent (17) for chymotryptic cleavage at threonine (bond 66-67), lysine (bond 84-85), methionine (bonds 21-22 and 49-50), asparagine (bond 36-37), and leucine (bonds 14-15, 28-29, 91-92, 93-94, 109-110, and 112-113), and for a small amount of chymotryptic-like activity in commercial trypsin (18) resulting in cleavage at Leu-Ser 14-15 and Tyr-Met 20-21.

The low yields of tryptic peptides 12-22, 11-22, 21-22, and 15-22, and of chymotryptic peptides 21, 21-28, and 22-28, and 29-49 and 29-36, 50-73 and 50-66, 80-84, 85-93, 80-91, and 85-91, 92-109, and 110-112, and of S. aureus protease peptides 51-58 and 59-67 may be accounted for by partial digestion at less favorable peptide bonds. The low yields of the tryptic and chymotryptic cysteine peptides were likely due to partial air oxidation during purification, because these peptides were not immediately chemically oxidized after enzymic digestion. The cysteines were oxidized chemically at an early stage of purification of the S. aureus protease peptides and of duplicate tryptic peptides (2.9 Fmol of protein), resulting in much higher yield.

Two peptides T30-43, of identical amino acid composition but different mobilities, were isolated. The major peptide T30-43, when run on pH 6.5 paper electrophoresis a second time, separated into three components, one of which had the same mobility as the minor T30-43; the other was only slightly different from the minor component. These results show that the minor peptide T30-43 was an artifact of isolation. The major peptide T30-43 could be subdigested with chymotrypsin, whereas the minor T30-43 could not (some deamidation-labile asparagines form chymotryptic cleavage sites (19)). This result along with the difference in mobilities (corresponding to one extra charge in T30-43 minor) suggests, although it does not prove, that the modification was the result of deamidation at Asn-36. Since two minor bands with similar mobility were observed on repetition of electrophoresis, it is also possible that both amides were labile under the experimental conditions or

---

**Fig. 1.** Proposed amino acid sequence of Rhodospirillum rubrum Sl cytochrome c' showing tryptic (T), chymotryptic (C), Staphylococcus aureus protease (E), and thermolysin (H) peptides. Narrow lines indicate composition data has been collected and heavy lines denote extent of sequence data. The one-letter code is that recommended by IUPAC-IUB (15): A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

---

*Some of the data are presented as a miniprint supplement immediately following this paper. Tables I to XI will be found on pp. 8420-8421. Material published in miniprint form can be easily read with the aid of a large-field reading glass of a type readily available at most opticians. For the convenience of those who prefer to obtain supplementary material in the form of full size photocopies of 12 pages, these same data are available as JBC Document No. 74M-1498. Orders for supplementary material should specify the title, authors, and reference to this paper and the JBC Document Number, and the number of copies desired. Orders should be addressed to The Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, Maryland 20014, and must be accompanied by a remittance to the order of the Journal in the amount of $1.80 for each set of photocopies.*
that both deamidation and rearrangement (20) of Asn-36 took place. The presence of histidine in peptide T30-43 casts some doubt on interpretation of the mobility measurements, since the side chain of this residue has an ionization constant near the pH of the electrophoresis buffer. However, treatment of peptide EH33-38 (mobility, 0) with 2 M NH₄ for 5 hours at 37° resulted in production of a peptide (mobility, -0.37) in about 10% yield with the remainder unchanged, also suggesting deamidation.

Three procedures were utilized to determine amide positions in the sequence. Peptide mobilities on pH 6.5 paper electrophoresis (21) were considered to provide unambiguous evidence for the presence or absence of amides at the following positions in the sequence: Asp-2, Glu-8, Asn-19, Glu-30, Asp-31, Asn-36, Gln-37, Asp-40, Glu-50, Glu-58, Glu-67, Asp-75, Glu-77, and Asp-126. The specificity (16) of S. aureus protease for peptide bonds involving glutamic acid, but not glutamine, was consistent with the above assignments and confirmed the presence of glutamic acid at positions 8, 30, 50, 58, 67, 77, 92, 98, and 122, and glutamine at positions 37, 87, and 110. The peptide bond Glu-Gly 58-59 was only partially hydrolyzed on digestion of the whole protein but was readily cleaved upon redigestion of peptide E51-67. Thermolytic subdigestion of S. aureus protease peptide E31-50 was made primarily to confirm amide assignments.

Mobility (m) measurements at pH 6.5 taken in conjunction with successive steps of the Edman degradation gave the following results: fragments 86-93 (m, -0.59), 87-93 (m, -0.63), 88-93 (m, -0.70), and 89-93 (m, -0.39). Thus residue 87 is glutamic acid, residue 88 is aspartic acid, and residue 92 is glutamic acid. The mobility of fragments 37-43 (m, 0.38) and 38-43 (m, 0.40) further confirm that residue 37 is glutamine.

The amino acid composition derived from the sequence is in excellent agreement with that derived from hydrolysis and analysis of the whole protein reported previously (22). However, the NH₂-terminal sequence of the protein proposed previously (22) to be Ala-Asn-Val-Ala-Gly . . . , appears to have been only partially correct. The sums of acidic and basic residues are equal, providing a net protein charge of -2 (provided by heme), in agreement with the isoelectric point of the protein, reported to be 5.4 (2).

**DISCUSSION**

The amino acid sequence of *Rhodospirillum rubrum* cytochrome c' (19) is similar to that of *Alcaligenes* sp. cytochrome c' (22), as shown in Fig. 2, although there are postulated to be only 35 identical amino acid residues out of 122 compared (29%). It is recognized that there is considerable uncertainty in the quantitative comparison of amino acid sequences, especially when insertions or deletions are involved. These have arbitrarily been kept to a minimum, but if greater leeway is accepted, larger apparent identity results. In the restricted area of the heme binding region, there is stronger similarity between *Alcaligenes* sp. and *Chromatium vinosum* D cytochrome c' than either shows to the *R. rubrum* protein.

The sequence of *R. rubrum* cytochrome c' lends further support to the notion that cytochromes c' in general are dimeric proteins (molecular weight 28,000) made up of identical monoheme subunits (molecular weight 14,000) (3). Both *R. rubrum* and *C. vinosum* cytochromes c' were originally proposed to be monomeric diheme proteins, and some sequence evidence was offered in support of this proposal. It has already been shown that this is not the case for *Chromatium* (3) and *Alcaligenes* (12) cytochromes c'. Such an idea should be completely abandoned now that the sequence of *R. rubrum* cytochrome c' has been completed. The fragmentation pattern obtained by tryptic digestion of trifluoroacetylated *R. rubrum* cytochrome c' (22) must be considered erroneous just as the "*Chromatium* diheme peptide" sequence (22) was shown to be (3).

The heme in cytochrome c' is bound to a pair of cysteine residues, which, however, are located near the COOH terminus rather than near the NH₂ terminus as in all other cytochromes c for which sequence data are available. A similarity to cytochrome c structure is the presence of a histidine residue adjacent to the carboxyl side of the heme-binding cysteine pair, which in eukaryotic cytochrome c (23) and in *R. rubrum* cytochrome c₂ (24) is the fifth coordination position. In *Chromatium* and in cytochrome b₅ (26), the fifth coordination position is also taken by a histidine. The observation that histidine is the fifth iron ligand in the four classes of heme proteins, for which the three-dimensional structures are known, strongly suggests by analogy that the histidine adjacent to the heme-binding cysteines in cytochrome c' is probably the fifth iron ligand as well.

The sixth iron ligand (unlike the fifth) is known to be variable and apparently is a major determinant of the properties of the protein. In cytochrome b₅ (26) it is histidine and in cytochromes c₆ (23) and c₇ (24) it is methionine; therefore both amino acids appear to be strong field ligands within the protein.

**Fig. 2.** Comparison of *Rhodospirillum rubrum* (I) and *Alcaligenes* sp. (II) (12) cytochrome c' sequences, and the *Chromatium vinosum* D cytochrome c' heme peptide fragment (III) (3). The one-letter code is that recommended by IUPAC-IUB (15); see legend to Fig. 1.
environment and confer typically low spin electronic absorption spectra on the heme. The sixth iron ligand in the high spin cytochromes c' is probably neither histidine nor methionine, but most likely some weak field ligand such as water or a carboxyl group which might be provided by aspartic or glutamic acids. Further amino acid sequences for the cytochromes c' may reveal a constant residue, which could provide the sixth ligand to the iron, or alternatively, show that there are no constant residues which could fulfill that role.

The three-dimensional structures of cytochromes c (23) and c (24) on the one hand, and the globins (25) on the other, show these two groups to be unrelated in the manner of peptide chain folding. The feature most characteristic of the globins, and which indeed sets them apart from all other proteins for which the three-dimensional structure is known, is the predominance of the α helix as the basic structural element. In cytochromes c and c' the α helix appears to play a less important role, while the heme itself appears to be the dominant structural feature—providing a nucleus around which the peptide chain may fold. The helix content of cytochrome c' obtained by circular dichroism measurements is 63% (27), which is similar to that of myoglobin (64%) (28) and significantly higher than in cytochromes c and c (27% and 36%) (29). The lack of sequence similarity between cytochromes c and c' as well as the large difference in helix content, indicates a probable dissimilarity in three-dimensional structure.

The possibility of structural similarities between cytochromes c', E. coli cytochrome b-562, and the globins may be entertained. However, sequence similarities already noted between cytochrome b-562 and globins (30) may be coincidental, as has recently been shown for cytochrome b (26, 31). Unfortunately, the high variability in globin and cytochrome c' sequences precludes any meaningful sequence comparison among the globins, cytochromes c', and b-562. X-ray analyses in progress (32, 33) will hopefully provide structures, which will allow a more definitive comparison of the three proteins.

The role of cytochrome c' in cellular metabolism remains obscure. The distribution of the protein argues against a direct role in a photosynthetic electron transport pathway, not only because a protein of similar amino acid sequence is found in a protein of which three R. palustris strains examined completely lacks cytochrome c (7), yet seemingly is not impaired in growth on "halotolerant micrococcus," hydroxylamine reductase activity (11), and which indeed sets them apart from all other proteins for which the three-dimensional structure is known, is the predominance of the α helix as the basic structural element. In many purple photosynthetic bacteria have been shown to grow aerobically, none has yet been shown to grow in the dark, anaerobically, in the presence of nitrate (36). In conclusion, the role of cytochrome c' in photosynthetic bacteria is likely to be less than essential and may be different from that in the nonphotosynthetic bacteria.

Acknowledgements.—We thank the late Harold W. Smith for technical assistance and Richard Liu for prompt and efficient Spino service. T. E. M. thanks R. P. A. and the Department of Molecular Biology for hospitality and introduction to sequence methodology.

REFERENCES
Amino acid sequence of cytochrome c' from the purple photosynthetic bacterium *Rhodospirillum rubrum* S1.

T E Meyer, R P Ambler, R G Bartsch and M D Kamen


Access the most updated version of this article at http://www.jbc.org/content/250/21/8416

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/250/21/8416.full.html#ref-list-1