

The Primary Structures of the Low-Redox Potential Diheme Cytochromes *c* from the Phototrophic Bacteria *Rhodobacter sphaeroides* and *Rhodobacter adriaticus* Reveal a New Structural Family of *c*-Type Cytochromes

Isabel Vandenberghe,[‡] David Leys,[‡] Hans Demol,^{‡,§} Gonzalez Van Driessche,[‡] Terrance E. Meyer,^{||} Michael A. Cusanovich,^{||} and Jozef Van Beeumen^{*,‡}

Department of Biochemistry, Physiology and Microbiology, Laboratory of Protein Biochemistry and Protein Engineering, University of Gent, 9000 Gent, Belgium, and Department of Biochemistry, University of Arizona, Tucson, Arizona 85721

Received May 11, 1998; Revised Manuscript Received July 1, 1998

ABSTRACT: The complete amino acid sequence of the low-redox potential cytochrome *c*-551.5 from *Rhodobacter sphaeroides* was determined by automated Edman degradation combined with mass spectroscopy. There are 139 residues and two typical Cys-X-X-Cys-His heme-binding sites. A homologous low-redox potential cytochrome was also sequenced from *Rhodobacter adriaticus* and was found to contain 126 residues. It is 53% identical to that of *Rb. sphaeroides* and has two internal deletions of one and five residues. The *Rhodobacter* diheme cytochromes are 21–24% identical to the translated open reading frame SLL1886 from *Synechocystis* sp. PCC6801. There are at least two deletions of five and eight residues in the 188-residue cyanobacterial protein. Each of the three cytochromes has more histidines than it needs to bind the two hemes, but conserved histidines located 23 residues after the first heme and 14–19 residues before the second heme are likely to be the sixth heme ligands. There is no evidence for gene doubling and no similarity to any other known cytochromes. The measured helix content of 24% is much less than normal for *c*-type cytochromes. These proteins thus appear to be representative of an entirely new class of *c*-type cytochromes.

Six soluble heme proteins have been purified from the purple phototrophic bacterium *Rhodobacter sphaeroides* by Meyer and Cusanovich (1): cytochromes *c*₂, *c'*, *c*-551.5, and *c*-554, the oxygen-binding *c*-type heme protein known as SHP, and the protoheme-containing bacterioferritin. An alternate oxidase has been found to have two cytochrome *c* subunits, neither of which is like any of the soluble cytochromes (2). The cytochrome *bc*₁ complex has also been cloned and characterized (3). Cytochrome *c*-551.5 is more abundant in anaerobically grown cells. Michalski et al. (4) found that cytochrome *c*-551.5 levels were depressed in nitrate-grown cells of the denitrifying strain IL 106. They concluded that low-redox potential pathways are shut down when cells are grown on high-redox potential electron acceptors such as nitrate and oxygen.

Cytochrome *c*-551.5 is the only soluble cytochrome *c* from *Rb. sphaeroides* which had not been sequenced until now. It has a mass of 16 kDa and contains two hemes (1). On long-term storage, the protein appears to be partially degraded

to an 8 kDa species, although the spectral properties are unaffected. The redox potential is –254 mV as contrasted with that of cytochrome *c*₂ which is 356 mV. On the basis of these properties, cytochrome *c*-551.5 is likely to be structurally and functionally unrelated to cytochrome *c*₂. We now report the complete amino acid sequence of cytochrome *c*-551.5, which confirms that it is structurally unrelated not only to cytochromes *c*₂ but also to the other cytochromes which have been characterized in *Rb. sphaeroides*. In fact, it is representative of a new class of *c*-type cytochromes. We also found a diheme protein from *Rhodobacter adriaticus* which is homologous to the *Rb. sphaeroides* cytochrome. The cytochromes of *Rb. adriaticus* had not been reported until now, but we found the usual cytochromes *c*₂ and *c'* in addition to the diheme cytochrome *c*.

MATERIALS AND METHODS

Protein Isolation. The *Rb. sphaeroides* strain 2.4.1 cytochrome *c*-551.5 was prepared as described by Meyer and Cusanovich (1). *Rb. adriaticus* strain DSM 2781 was grown photosynthetically essentially as described by Neutzing et al. (5). Cells (350 g) were suspended in 2 L of 100 mM Tris-HCl buffer (pH 7.5) and broken in the Ribi cell fractionator (an automated French press). Membranes were removed by centrifugation in the Spinco Ti45 rotor for 3 h. The supernatant was desalted on Sephadex G25; 10 mM Tris-HCl (pH 8) was added, and the extract was adsorbed to Whatman DE52 DEAE-cellulose. The column was washed with 20 mM increments of NaCl in the Tris buffer. Approximately 2.5 μmol of cytochrome *c'* was eluted at 80–

[†] This work was supported in part by the Fund for Scientific Research FWO-Vlaanderen (Contract 3G005497 to J.V.B.) and by a grant from the National Institutes of Health (GM 21277) to M.A.C.

^{*} To whom correspondence should be addressed: Laboratorium voor Eiwitbiochemie en Eiwitengineering, Universiteit Gent, Ledeganckstraat 35, 9000 Gent, Belgium. Phone: 32 9 264 5109. Fax: 32 9 264 5338. E-mail: jozef.vanbeeumen@rug.ac.be.

[‡] University of Gent.

[§] Present address: Department of Medical Protein Chemistry, Flanders Interuniversity Institute for Biotechnology (VIB), University of Gent, 9000 Gent, Belgium.

^{||} University of Arizona.

120 mM NaCl and 1.5 μ mol of cytochrome c_2 eluted at 140–200 mM NaCl. The column was then washed with 0.5 M NaCl, which eluted residual cytochromes. The cytochrome fractions were concentrated by pressure filtration using a YM5 membrane (Amicon, Beverly, MA) and chromatographed on Sephadex G100. The cytochrome c_2 fractions were passed through a hydroxylapatite column which adsorbed about 0.25 μ mol of diheme cytochrome c , while the cytochrome c_2 went through. The diheme cytochrome was eluted at about 210 mM phosphate when a 100 to 300 mM phosphate linear gradient was applied (pH 7.0). It was essentially pure at this stage because very few proteins were adsorbed so tightly to the hydroxylapatite. Nevertheless, the diheme cytochrome was chromatographed further on DEAE-cellulose with a 40 to 240 mM NaCl gradient in 20 mM Tris-HCl (pH 8.0). The diheme cytochrome c eluted at 164 mM NaCl at nearly the same concentration as the cytochrome c_2 on a separate column. The whole procedure was completed in 1 week, but the length of the process might be shortened if unnecessary steps are eliminated. In hindsight, the diheme cytochrome c yield might have been better if hydroxylapatite had been used at the beginning of the preparation. The diheme cytochrome was not very soluble and did not completely elute from DEAE-cellulose. It also sticks to glass and plastic, resulting in large losses. The wavelength maxima of the oxidized protein were at 278 and 408 nm and of the dithionite-reduced protein at 418, 522, and 552 nm. The best ratio of 278 nm/408 nm absorbance was 0.15. The protein was completely oxidized as isolated and is likely to have a low redox potential.

Final Protein and Peptide Purification. Native protein from *Rb. adriaticus* was purified on a C₂–C₁₈ column, positioned in a SMART chromatographic system (Pharmacia, Uppsala, Sweden). The solvents used were 0.05% TFA¹ in water (solvent A) and 0.05% TFA in 80% acetonitrile (solvent B). The same chromatographic equipment was also used for purification of peptides obtained after enzymatic digestions of the apoprotein from *Rb. adriaticus*. Peptides generated by different digestions of the apoprotein from *Rb. sphaeroides* were purified by RP-HPLC on a C₁₈ column (214TP54, Vydac, Hesperia, CA). The chromatographic equipment consisted of a model 870 three-headed piston pump, a model 8800 system controller, a variable-wavelength detector set at 220 nm (DuPont, Wilmington, DE), and a model 7120 Rheodyne injector equipped with a loop of 100 μ L. The solvents used were 0.1% TFA in Milli-Q water (solvent A) and 0.1% TFA in 70% acetonitrile or 100% acetonitrile (solvent B) and were administered at a flow rate of 1 mL/min.

Protein Modification and Chemical Cleavage. The hemes of the native protein from *Rb. sphaeroides* and *Rb. adriaticus* were removed by treatment with mercuric chloride in 8 M urea containing 0.1 N HCl. After incubation overnight at 37 °C, the hemes were removed by gel filtration over a SG-25f column (1.5 cm \times 23 cm, Bio-Rad, Eke, Belgium) or a fast desalting column (Pharmacia), equilibrated and eluted with 5% formic acid.

The *Rb. sphaeroides* apoprotein (100 nmol) was dissolved in 70% formic acid containing 2 mmol of tryptamine and

20 mmol of cyanogen bromide and the mixture incubated for 20 h at room temperature under argon.

Enzymatic Digestions. Digestion with Glu-C endoproteinase (Boehringer, Mannheim, FRG) was performed on 295 nmol of *Rb. sphaeroides* apocytochrome in 0.05 M ammonium acetate (pH 4) for 12 h at 37 °C using an enzyme/substrate ratio (E/S , w/w) of 1/40. The apoprotein (17 nmol) was digested with Lys-C endoproteinase (Wako, Osaka, Japan) for 2 h at 37 °C in 0.05 M Tris-HCl buffer (pH 9.0, $E/S = 1/90$).

Lys-C endoproteinase digestion was performed on 7 nmol of *Rb. adriaticus* apocytochrome in 100 mM Tris-HCl (pH 8.0) for 3 h at 37 °C ($E/S = 1/40$). About 3.7 nmol of apoprotein was digested with Glu-C endoproteinase in 20 mM ammonium acetate (pH 4.1) at room temperature overnight ($E/S = 1/37$).

Sequence and Mass Analysis. N-Terminal sequence analyses were performed on a 477 or 476 pulsed-liquid sequenator (Applied Biosystems Division, Perkin-Elmer, Foster City, CA), equipped with an on-line 120A PTH analyzer. Mass analyses were done on a BIO-Q quadrupole (MicroMass, Altricham, U.K.) and a TofSpec SE Laser desorption mass spectrometer (Micromass, Wyntenshaw, U.K.). Conditions for the mass analyses were like those described by Devreese et al. (6).

RESULTS

The evidence for the sequences of *Rb. sphaeroides* and *Rb. adriaticus* diheme cytochromes is shown in Figure 1. The composition of the *Rb. sphaeroides* protein is in substantial agreement with that of Meyer and Cusanovich (1). The sequences were determined as follows.

For *Rb. sphaeroides*, the native protein was subjected to automated N-terminal sequence analysis for 30 cycles, during which 21 residues could be determined unambiguously. There was no evidence for endogenous proteolysis as observed with aged proteins (1). Cleavage of the apoprotein with cyanogen bromide resulted in 14 peaks. Peptide CN7, which resulted from the specific cleavage of a Met–Ser peptide bond, extended the sequence information of fraction CN8 by 14 residues. Sequence analyses of peptides obtained after enzymatic cleavage of the apoprotein with Glu-C (Figure 2A) and Lys-C endoproteinases were useful in completing the sequence. Fraction S11 contained two fragments, in which fraction S11B showed an overlap with CN9 and CN10. Peptide S12 extended the sequence information of peptide CN14 and overlapped with peptide S9 which originated from nonspecific cleavage of an Arg–Gly peptide bond. Peptides K5 and K1 completed the sequence by overlapping CN10 and CN14, and S9 and CN8, respectively. The masses of all sequenced peptides are given in Table 1. Mass analysis of the native protein revealed that the measured mass of 16 242.1 Da (Figure 3A) is consistent with the calculated mass of 16 244.6 Da, including two covalently bound hemes.

There is also no evidence from mass spectrometry for any endogenous proteolysis such as that reported previously (1). The end groups supposedly released by endogenous proteolysis, Gly and Ala, may be accounted for in part if there had been nonspecific cleavages at several positions within the segment of residues 71–80. This would result in two

¹ Abbreviations: ESMS, electrospray ionization mass spectrometry; LDMS, laser desorption mass spectrometry; TFA, trifluoroacetic acid.

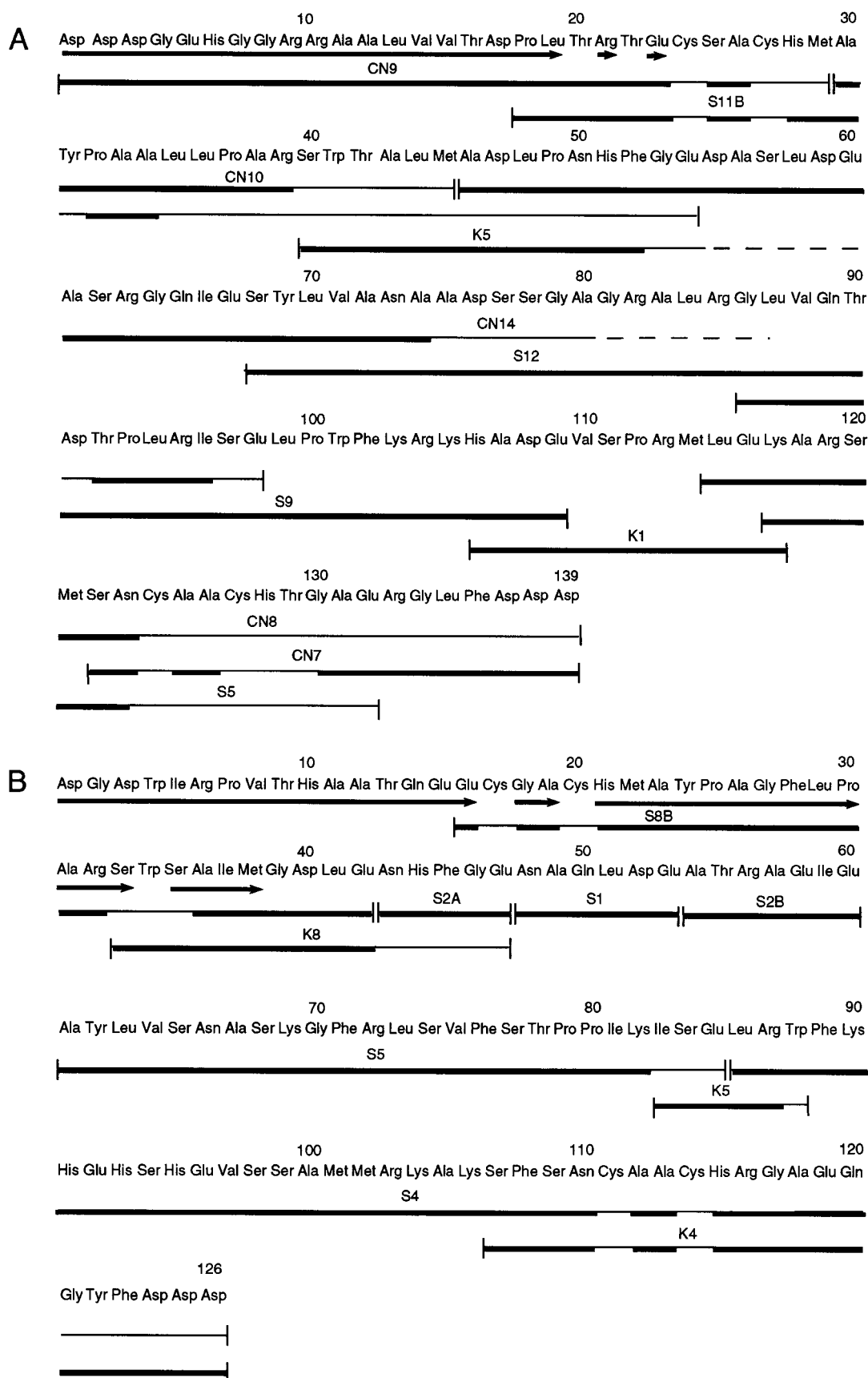


FIGURE 1: Evidence for the amino acid sequence of *Rb. sphaeroides* cytochrome *c*-551.5 (A) and *Rb. adriaticus* cytochrome *c*-552 (B). Peptides obtained after digestion of the apocytochrome by Lys-C and Glu-C endoproteinase are indicated by K and S, respectively, and cleavage is indicated with cyanogen bromide by CN. Amino acids identified during the Edman degradation of the native protein are shown by bold arrows and those of peptides by bold lines.

equally sized halves of the protein as previously reported (1). This furthermore suggests that the region of residues

71–80 may be disordered in the three-dimensional structure. Most cytochromes *c* have 50% or more helical secondary

Table 1: Mass Analysis Data of Peptides Obtained after Enzymatic Digestion with Glu-C (S) and Lys-C (K) Endoproteinases on the Apoprotein from *Rb. sphaeroides* and *Rb. adriaticus*^a

Rb. sphaeroides				Rb. adriaticus			
peptide	measured mass (Da)	calculated mass (Da)	sequence position	peptide	measured mass (Da)	calculated mass (Da)	sequence position
S1	759.2 ^b	759.8	61–67	<i>S1</i>	NR	(688.7)	48–53
S2	830.4 ^b	831.0	110–116	<i>S2</i>	A NR	(602.6)	43–47
S3	1778.4 ^b	1778.9	68–85	<i>S3</i>	B 788.4 ^b	788.9	54–60
S4	1129.9 ^b	1228.3	82–91		1694.9 ^b	1695.8	1–15
S5	2453.8 ^c	2455.7	117–139		1824.9 ^b	1824.9	1–16 ^d
	2653.9 ^c	2655.7	117–139 ^e	<i>S4</i>	4770.3 ^c	4771.3	86–126 ^d
S6	2098.1 ^b	2100.2	120–139		4970.2 ^c	4971.3	86–126 ^{d,e}
	2299.0 ^b	2300.0	120–139 ^e	<i>S5</i>	2711.8 ^b	2712.1	61–85
S7	1415.4 ^b	1415.6	42–54	<i>S8</i>	A NR	(2755.2)	17–42
S8	1198.4 ^b	1198.4	34–44		B 3082.0 ^c	3084.3	16–42 ^e
S9	2835.6 ^b	2837.2	85–109				
<i>S11</i>	A 1687.8 ^b	1688.9	40–54				
	B 3127.1 ^b	3131.4	18–44 ^e	K2	A 1757.8 ^b	1762.0	21–36 ^f
<i>S12</i>	4598.3 ^c	4598.1	68–109		B 1957.7 ^b	1962.0	21–36 ^d
S13	A 4598.3 ^b	4598.1	68–109	K3	1195.3 ^b	1195.6	1–10 ^f
	B 4895.4 ^b	4896.3	1–44 ^e	<i>K4</i>	2193.3 ^b	2193.3	107–126
S14	5411.6 ^b	5411.1	68–116		2393.7 ^b	2393.3	107–126 ^e
				<i>K5</i>	803.0 ^b	802.9	83–88 ^f
<i>K1</i>	1411.0 ^c	1411.6	106–117	<i>K6</i>	A 1078.2 ^b	1078.3	83–90
K2	2325.9 ^c	2327.5	118–139		B 1193.4 ^b	1194.3	56–69 ^f
	2526.1 ^c	2527.5	118–139 ^e		C 1723.0 ^b	1722.9	54–69 ^f
K3	2100.7 ^c	2100.4	86–103	K7	A 1448.7 ^b	1448.7	70–82
K4	9140.8 ^c	9141.8	1–85 ^e	<i>K8</i>	1692.6 ^b	1692.8	33–47 ^f
K6	11240.4 ^c	11240.3	1–103 ^{e,g}	<i>K9</i>	4069.2 ^b	4068.4	33–69 ^f
K7	11224.2 ^c	11224.3	1–103 ^e				

^a Fractions are numbered following their chromatographic elution. Peptides mentioned in Figure 1 are italicized. ^b Measured by LDMS. ^c Measured by ESMS. ^d The Glu-X peptide bond was not cleaved. ^e Fraction contained the Hg²⁺ adduct to the heme binding region Cys-X-X-Cys-His. ^f Nonspecific cleavage of a peptide bond. ^g Presence of oxygen adduct to methionine. NR means there was no mass detected during mass measurements.

structure and are generally fairly stable. We thus measured the secondary structure as shown in Figure 4. The 24% helix content is significantly lower than for other cytochromes and may account in part for the susceptibility to endogenous proteolysis.

For *Rb. adriaticus*, N-terminal sequence analysis of native protein resulted in 35 out of 40 residues being established unambiguously. We digested the apoprotein with Glu-C (Figure 2B) and Lys-C endoproteinases. By comparison with the complete sequence of the diheme cytochrome from *Rb. sphaeroides*, all the sequenced peptides could be fitted into the complete sequence. Overlaps were established by sequence analysis of peptides S8B, K5, and K8. The correctness of the positioned peptides was further supported by mass analysis of different peptides as presented in Table 1. The measured mass of the native protein was 15 266.9 Da (Figure 3B), consistent with the sum of the masses of the amino acids supplemented with two covalently bound hemes (15 268.6 Da). The other two species observed in mass analysis are 17 and 33 Da larger than native protein which suggests that they represent oxidized forms of the diheme cytochrome. The obvious sites of modification are at one or more of the four methionines in the protein.

DISCUSSION

Comparison with Other Cytochromes. There are 139 amino acid residues in *Rb. sphaeroides* cytochrome *c*-551.5. Although the cysteines were not identified by chemical modification, the heme binding sites were recognized by their classical heme binding pattern Cys-X-X-Cys-His. Another indication of the presence of such a pattern is the presence

of the Hg²⁺ adduct to the cysteines of the apoprotein as detected during mass determination of appropriate peptides by LDMS or ESMS (7) (Table 1). The four cysteines are at positions 24, 27, 125, and 128. Cytochrome *c*-551.5 has a low redox potential of –254 mV (*I*) and is spectrally similar to *Desulfovibrio* cytochrome *c*₃ which has histidines at both the fifth and sixth heme ligand positions (8). In addition to the fifth heme ligand histidines at positions 28 and 129, cytochrome *c*-551.5 has histidines at positions 6, 51, and 106 which could be ligands to the hemes.

The amino acid sequence of the *Rb. adriaticus* cytochrome contains 126 residues and also has two heme binding sites. The protein is 53% identical to that of *Rb. sphaeroides* with two internal deletions of one and five residues. There are eight histidines in the *Rb. adriaticus* diheme cytochrome, twice as many as are needed to ligate the hemes.

A BLAST search (9) showed that the translated *Synechocystis* PCC 6803 SLL1886 gene product (10) is 21–24% identical to the diheme cytochromes of *Rb. sphaeroides* and *Rb. adriaticus*. It contains 188 residues and has at least two internal deletions of five and eight residues as shown in the alignment of Figure 5. An interesting feature of the *Synechocystis* sequence is the presence of a signal peptide which may direct the protein to the periplasmic space. However, it is not obvious whether the signal may be cleaved and if the protein is soluble or membrane-bound. There are five histidines in the *Synechocystis* cytochrome, but only two of them apparently are conserved in all three proteins. These residues are likely to be the heme sixth ligands. One His is located 23 residues after the first heme, and the other His is found 14–18 residues before the second heme. The first

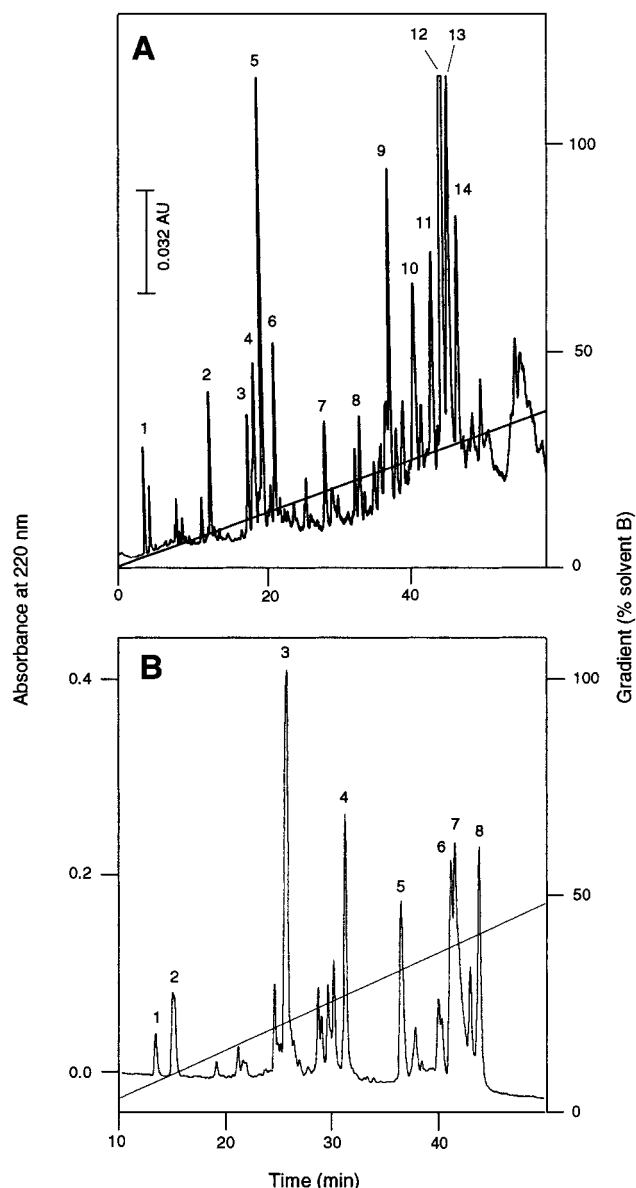


FIGURE 2: Reversed phase HPLC separation of peptides obtained after digestion of the apoprotein from *Rb. sphaeroides* (A) and *R. adriaticus* (B) with Glu-C endoproteinase. Conditions for the separation are given in Materials and Methods.

half of the diheme cytochrome is more highly conserved than is the second half, and the deletions have also occurred in the second half. This suggests that the first half of the sequence is also functionally more important.

Several diheme cytochromes are known to be the product of gene doubling, such as cytochrome *c*₄, bacterial cytochrome *c* peroxidase, and flavocytochrome *c* sulfide dehydrogenase (11–13) which are known to be related to class I cytochromes such as mitochondrial cytochrome *c*. The positions of the hemes near the N terminus and the C terminus and the locations of the probable sixth heme ligands between the hemes indicate that the *Rhodobacter* and *Synechocystis* cytochromes did not result from gene doubling and are not related to the class I cytochromes.

The secondary structure of the *Rb. sphaeroides* diheme cytochrome indicates approximately 24% α -helix as calculated. Chou–Fasman secondary structure analysis (14) predicts seven short helices totaling 41 residues (30%) as

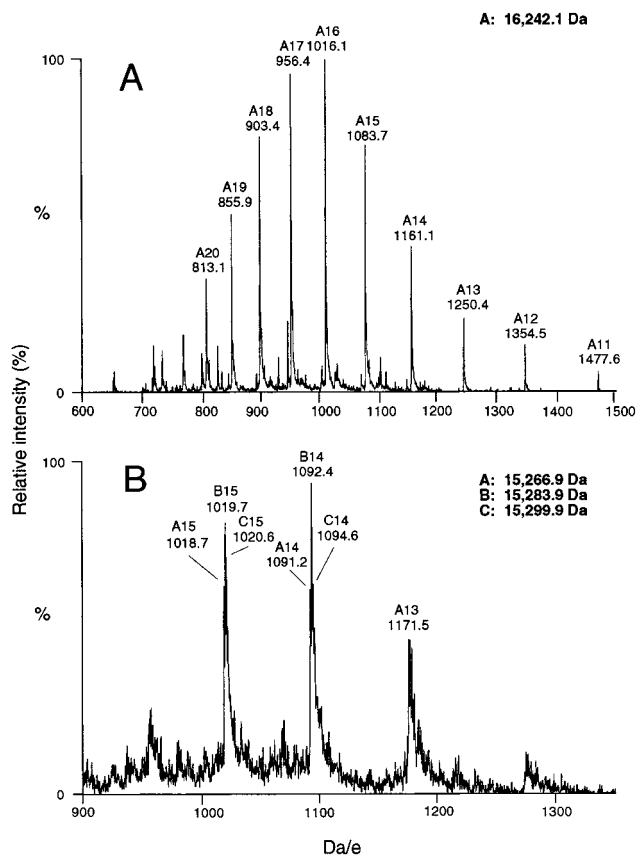


FIGURE 3: Electrospray ionization mass spectrometry of the native protein from *Rb. sphaeroides* (A) and *R. adriaticus* (B). The number at the top of each peak represents the number of positive charges for the particular *m/z* peak. The B and C fragments of the *Rb. adriaticus* protein represent the presence of some oxidized methionine residues.

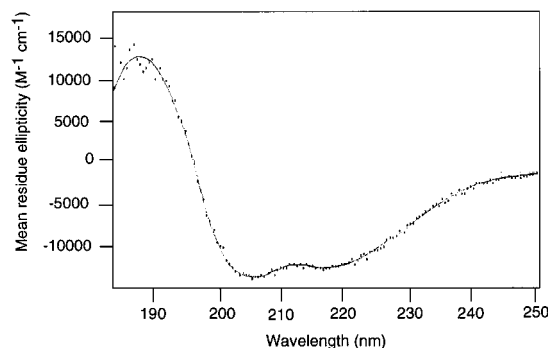
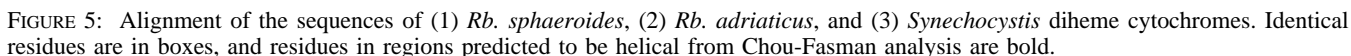


FIGURE 4: Circular dichroism spectrum for *Rb. sphaeroides* cytochrome *c*-551.5 in the far-UV region.

indicated in Figure 5. There is some overlap in the predicted helical content of the three diheme cytochromes, thus increasing the confidence that these segments of secondary structure do actually exist. Although one heme is bound near the C terminus as in class II cytochromes *c* (including cytochromes *c*' and *c*-554), the low helix content of cytochrome *c*-551.5 rules out the possibility that there is any structural homology with cytochrome *c*' or other class II cytochromes.

The *Desulfovibrio* cytochromes *c*₃ or class III cytochromes *c* have low redox potentials and multiple hemes and also have very little secondary structure, as exemplified by the three-dimensional structure of *Desulfovibrio vulgaris* cytochrome *c*₃ (8). However, the four hemes in cytochromes *c*₃



The presence of the diheme cytochrome in two species of *Rhodobacter* and in *Synechocystis* suggests that the gene is more widespread than it currently appears from the limited number of examples. It is not abundant in the *Rhodobacter* species and has not been observed as a protein in *Synechocystis* or any other cyanobacterium. We have not found it in either *Rhodobacter capsulatus* or *Rhodobacter sulfidophilus*, although it may have been overlooked if it is present at very low concentrations. It is also likely that it may be membrane-bound in some species. The genetic context of the *Synechocystis* diheme gene gives no clue as to the

REFERENCES

1. Meyer, T. E., and Cusanovich, M. A. (1985) *Biochim. Biophys. Acta* 807, 308–319.
2. Garcia-Horsman, J. A., Berry, E., Shapleigh, J. P., Alben, J. O., and Gennis, R. B. (1994) *Biochemistry* 33, 3113–3119.
3. Yun, C. H., Beci, R., Crofts, A. R., Kaplan, S., and Gennis, R. B. (1990) *Eur. J. Biochem.* 194, 399–411.
4. Michalski, W. P., Miller, D. J., and Nichols, D. J. D. (1986) *Biochim. Biophys. Acta* 849, 304–315.
5. Neutzling, O., Imhoff, J. F., and Truper, H. G. (1984) *Arch. Microbiol.* 137, 256–261.
6. Devreese, B., Costa, C., Demol, H., Papaefthymiou, V., Moura, I., Moura, J. J., and Van Beeumen, J. (1997) *Eur. J. Biochem.* 248, 445–451.
7. Samyn, B., Berks, B. C., Page, M. D., Ferguson, S. J., and Van Beeumen, J. J. (1994) *Eur. J. Biochem.* 219, 585–594.

8. Higuchi, Y., Kusunoki, M., Yasuoka, N., Kakudo, M., and Yagi, T. (1981) *J. Biochem.* **90**, 1715–1723.
9. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
10. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996) *DNA Res.* **3**, 109–136.
11. Kadziola, A., and Larsen, S. (1996) *Structure* **5**, 203–216.
12. Fulop, V., Ridout, C. J., Greenwood, C., and Hajdu, J. (1995) *Structure* **3**, 1225–1233.
13. Chen, Z., Koh, M., Van Driessche, G., Van Beeumen, J. J., Bartsch, R. G., Meyer, T. E., Cusanovich, M. A., and Mathews, F. S. (1994) *Science* **266**, 430–432.
14. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* **13**, 222–245.

BI981076Z