Ligand Binding and Covalent Structure of an Oxygen-Binding Heme Protein from *Rhodobacter sphaeroides*, a Representative of a New Structural Family of *c*-Type Cytochromes†

Klaus Klarskov,‡ Gonzalez Van Driessche,‡ Katrien Backers,‡ Chantal Dumortier,§ Terrance E. Meyer,§ G. Tollin,§ Michael A. Cusanovich,§ and Jozef J. Van Beeumen*,‡

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

Received October 9, 1997; Revised Manuscript Received February 25, 1998

ABSTRACT: The amino acid sequence of an oxygen-binding heme protein (SHP) from *Rhodobacter sphaeroides* has been determined. The cysteines, which bind the single heme group in the 112-residue protein, are located at positions 43 and 46. SHP is similar in size to the large membrane-bound form of the class I cytochrome c_5 of *Azotobacter vinelandii* (116 residues) and in the location of the heme binding site at positions 48 and 51. Two extra cysteines in SHP (residues 89 and 97) are located in positions similar to those of cytochrome c_5 (residues 98 and 101) and form a disulfide bridge in both proteins. In total, four regions of α-helix are predicted, covering 46% of the protein, which is comparable to that in other small cytochromes. SHP is thus distantly related to small class I *c*-type cytochromes but is representative of a distinct family by virtue of its high-spin nature, the lack of a strong sixth ligand, and its capacity to bind oxygen. Potentially, the most important characteristic of SHP is its ability to transiently bind oxygen during autoxidation, which occurs with a half-life of 3 min with a 4-fold excess of O_2. SHP also binds carbon monoxide, azide, and cyanide. The kinetics of reduction by free flavins indicate that SHP is less reactive than other class I cytochromes *c* and that the heme is less accessible to solvent. There is localized positive charge (+3) at the site of reduction of SHP, although the overall protein charge is −2. This may account in part for the ability of SHP to bind anions.

Meyer and Cusanovich (*1*) showed that there are at least six different soluble heme proteins in phototrophically grown *Rhodobacter sphaeroides*. The major proteins are cytochromes c_2 and *c*′, but there are also smaller quantities of cytochromes *c*-551 and *c*-554, an oxygen-binding heme protein (SHP[1]), and a bacterioferritin. In addition, a cytochrome c_2 isozyme was discovered following deletion of the wild-type cytochrome (*2*−*4*). The sequences of the cytochrome c_2 isozymes, *c*′, and *c*-554 were previously reported (*4*−*7*).

SHP has a molecular mass of about 12 kDa; the redox potential is −22 mV, and there is a single high-spin heme (*1*). It is spectrally distinct, however, from cytochrome *c*′, which also has a high-spin heme. SHP has an oxidized Soret peak with a very large extinction coefficient (170 mM^{-1} cm^{-1}) and narrow half-bandwidth (25 nm), whereas cytochromes c′ generally have a Soret peak with an absorptivity of 85 mM^{-1} cm^{-1} and a half-width of about 62 nm. SHP transiently binds oxygen during slow autoxidation (*1*), whereas cytochromes *c*′ are rapidly oxidized in the presence of oxygen, with no obvious intermediates. Absorption spectra of cytochrome *c*′, partially oxidized by oxygen, show the same isosbestic points as the fully reduced and oxidized protein (unpublished). However, SHP does not show isosbestic points during oxidation but has a distinct intermediate with a 5 min half-life previously estimated at an unknown oxygen concentration (*1*). Although the physiological significance of oxygen binding is unknown, SHP provides a simple system for comparison to other oxygen-binding heme proteins such as the globins and may shed additional light on the factors which determine oxygen binding. SHP is not abundant in phototrophically grown cells and is not induced by aerobic growth (*7*).

There are two other sources of SHP-like oxygen-binding heme proteins, from *Chromatium vinosum* (*8*) and *Rhodocyclus gelatinosus* (T. E. Meyer, unpublished), and both are minor soluble components. The *C. vinosum* heme protein has the same molecular mass as *Rb. sphaeroides*, but it has a basic isoelectric point and a redox potential of −110 mV. The N-terminal sequence of *Rb. sphaeroides* SHP was previously reported (*1*). We have now determined the complete amino acid sequence of *Rb. sphaeroides* SHP. We

† This work was supported in part by a grant from the National Institutes of Health (GM 21277), by the Concerted Research Action of the Flemish Government (Contract 12052293), and by the Fund for Scientific Research-Flanders (Contract G.0054.97).

* Author to whom reprint requests should be addressed: Laboratory of Protein Biochemistry and Protein Engineering, State University of Gent, Ledeganckstraat 35, B-9000 Gent, Belgium. Phone: 32 9 246 5109. Fax: 32 9 264 5338. E-mail: jozef.vanbeeumen@rug.ac.be.

‡ State University of Gent.

§ University of Arizona.

[1] Abbreviations: SHP, *Rb. sphaeroides* heme protein; FMN, flavin mononucleotide; $FMNH_2$, fully reduced FMN; SQ, semiquinone; CO, carbon monoxide; EDTA, ethylenediaminetetraacetic acid; V_{ii}, interaction energy.

S0006-2960(97)02498-7 CCC: $15.00
Published on Web 04/11/1998

have also examined ligand binding in greater detail and have measured the kinetics of reduction by free flavin semiquinones which further emphasize the unique properties of SHP.

EXPERIMENTAL PROCEDURES

Preparation of the Apoprotein. SHP was purified by the method of Meyer and Cusanovich (*1*). Native SHP (500 nmol) was submitted to the deheming procedure of Ambler and Wynn (*9*). The apoprotein was separated from the heme and salts by gel filtration through a Sephadex column (22 × 1.5 cm, SG-25 fine, Pharmacia, Uppsala, Sweden) eluted with 5% formic acid.

Enzymatic Digestions and Chemical Cleavage. One digest of the apoprotein was carried out for 10 h at 37 °C in 50 mM ammonium acetate buffer (pH 4) using *Staphylococcus aureus* V8 protease (Miles, Slough, U.K.) at an enzyme/substrate ratio (*E*/*S*) of 1/30. A digest with Arg-C endoproteinase (Boehringer Mannheim, Mannheim, Germany) was carried out for 2 h at 37 °C in 100 mM ammonium bicarbonate buffer (pH 8.4, $E/S = 1/30$). Peptides were also prepared by partial acid hydrolysis in 2% formic acid for 4 h at 106 °C. The amounts used for each of these three cleavages were the same (23 nmol) as those deduced from amino acid analysis.

Peptide Purification. Peptides were in all cases separated by reversed-phase HPLC using a C-18 column (214TPS4, 4.6 × 250 mm, Vydac, Hesperia, CA). The HPLC setup included a model 870 three-headed plunger pump, a model 8800 system controller, and a UV spectrophotometer set at 220 nm (all from Dupont, Wilmington, DE). The solvents used for gradient elution were 0.1% TFA in water (solvent A) and 100% acetonitrile (solvent B).

Determination of Cysteines and the Disulfide Bridge. Cysteine residues were pyridylethylated prior to sequence analysis essentially as described by Amons (*10*). Disulfide bonds in SHP were identified by dissolving 5 nmol of native protein in 50 mM Tris-HCl (pH 7.8) containing 20% acetonitrile. Enzymatic digestion was performed with trypsin ($E/S = 1/22$, w/w). After 20 min of incubation at room temperature, 5 μL was withdrawn and diluted with 15 μL of argon-flushed 75 mM Tris-HCl buffer containing 1 mM EDTA (pH 8.0). One microliter of this solution was analyzed by LC−ESMS as described below.

Sequence and Amino Acid Analysis. Compositional analyses of protein and peptides were carried out on a 420 Derivatizer coupled to a model 130A Separation System (Perkin-Elmer, Applied Biosystems Division, Foster City, CA). Hydrolysis was performed at 106 °C under vacuum using 6 N HCl.

Sequence analysis was carried out on a model 477A pulsed liquid sequenator with detection of the PTH derivatives on a model 120A Separation System (Perkin-Elmer, Applied Biosystems Division).

C-Terminal Analysis. Sixteen nanomoles of apoprotein was digested with carboxypeptidase P at an enzyme/substrate ratio of 1/200 and incubated at room temperature. At time intervals, fractions were withdrawn for determination of released amino acids on a model 420A derivatizer equipped with a model 130A Separation System.

Mass Spectrometry. Electrospray ionization mass spectrometry (ESMS) was performed on a Bio-Q triple-quadrupole mass spectrometer equipped with a pneumatically assisted electrospray source (Micromass, Altrincham, U.K.). All analyses were carried out using a liquid sheath probe supplied with the instrument. The probe was slightly modified in the laboratory by replacing the original inner metal capillary with a 105 cm long fused silica tube (50 μm inside diameter, 150 μm outside diameter). The sheath liquid consisted of 2/1 2-methoxyethanol/2-propanol (v/v) and was delivered at a flow rate of 4.2 μL/min by a model 140A HPLC pump (Perkin-Elmer, Applied Biosystems Division). Samples were dissolved in 1% formic acid to a concentration of 10−20 pmol/μL and introduced at a flow rate of 6 μL/min using a syringe pump (Harvard, South Natick, MA). Spectra were acquired in the *m*/*z* range of 600−1600 at a scan rate of 12 s/scan. The mass spectrometer was calibrated using horse heart myoglobin adjusted so that the peak at *m*/*z* 998 was 0.8 amu wide 50% above baseline.

Liquid Chromatography−Mass Spectrometry (LC−ESMS). On-line LC−ESMS analysis was, with minor modifications, carried out essentially as previously described (*11*). The system consisted of a model 140A HPLC pump equipped with two 40 mL syringes. Reproducible solvent gradients at microliter flow rates were obtained by a preinjection flow mixing and splitting device (split ratio of 1/50) similar to the model AC-30 instrument available from LC-packings (Amsterdam, The Netherlands). The capillary column (0.32 × 200 mm) was packed in the laboratory with an ODS-AQ C18 packing (5 μm particle size, 120 Å pore diameter) (YMC, Schermbeck, Germany) according to ref *12*. The column outlet and the mass spectrometer probe were connected to a Kontron 332 variable UV detector (Kontron Instruments, Milan, Italy) equipped with a U/Z-shaped capillary flow cell by a zero-dead volume peak union (Upchurch Scientific, Oak Harbor, WA). The solvents used for separations consisted of 0.1% TFA in 5% aqueous acetonitrile (solvent A) and 0.1% TFA in 80% aqueous acetonitrile (solvent B). Peptides obtained from tryptic digestion of SHP were separated using the following solvent conditions: 5% solvent B for 5 min followed by a linear increase to 40% solvent B after 35 min and to 100% solvent B after another 10 min. The mass spectrometer was calibrated prior to connecting the HPLC system by direct injection of a mixture of polyethylene dissolved in a 1/1 solution of 10 mM ammonium bicarbonate and acetonitrile. The resolution was set so that the peak at *m*/*z* 995 from polyethylene glycol was 2 amu 50% above baseline. LC−ESMS scanning was performed from *m*/*z* 375 to 1800 over the course of 6 s, and the total ion current was recorded.

Circular Dichroism Analysis. Circular dichroism was performed on an Aviv Instruments modified Cary model 60 spectropolarimeter (Lakewood, NJ). The concentration of protein was 1 μM in water. Five scans were averaged and smoothed prior to analysis.

Ligand Binding Analysis. Ligand binding was measured using a Hi-Tech SF-51 stopped flow spectrometer. Experiments were performed at 25 °C, with 100 mM phosphate or HEPES buffers at pH 7−7.5. The protein (20−30 mM) was photoreduced with 0.5 mM lumiflavin or 5-deazariboflavin for 30 min in the presence of 5 mM EDTA or with a stoichiometric amount of dithionite under an argon atmosphere before mixing with oxygen.

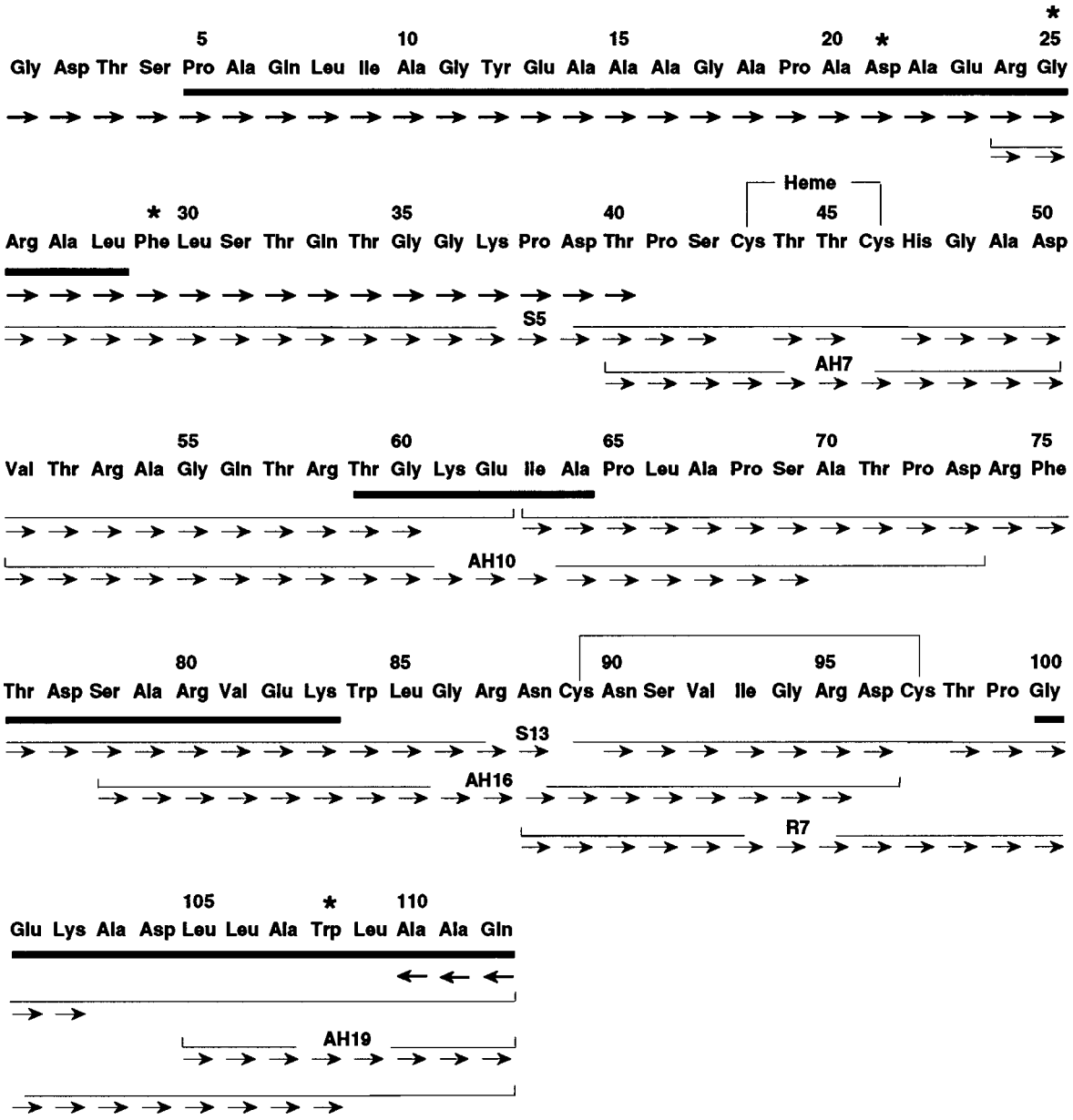

FIGURE 1: Complete amino acid sequence of SHP from *Rh. sphaeroides*. It was determined by Edman degradation of peptides obtained after digestion of apo-SHP with Glu-C (S) and Arg-C (R) endopeptidases and partial cleavage by acid hydrolysis (AH). Amino acids identified during Edman degradation of the native protein are shown by bold arrows and those of peptides by plain arrows. The reversed arrows indicate the residues released by a time course analysis with carboxypeptidase P treatment of the apoprotein. The bold underlines show regions predicted to be helical, and starred residues indicate class I consensus residues.

Measurement of the kinetics of reduction of SHP by laser flash photolysis of flavins was carried out as previously described (*13*, *14*). The triplet state of free flavin abstracts an electron from EDTA, producing the flavin semiquinone. The flavin semiquinone either reacts with the protein or disproportionates into oxidized and fully reduced flavin, which is also capable of reacting with protein. All experiments were performed under anaerobic conditions at ambient temperature with 50−100 mM flavin in addition to 10 mM EDTA and 20 mM phosphate buffer (pH 7.0) bubbled with argon. The reaction was initiated by a laser flash at the absorption maximum of the flavin (450 or 400 nm), and SHP reduction was followed by the increase in absorbance at 550 nm. The experiments with FMN were performed using 100 mM FMN, 1 mM EDTA, and 5 mM phosphate buffer at pH 7, and the ionic strength was varied by addition of 4 M NaCl. The charge at the active site was calculated according to Watkins et al. (*15*), using a radius of interaction ρ of 4.5 Å, a distance of closest approach r_{12} of 3.5 Å, a dielectric constant of 50, and charges on FMN SQ of −1.9 and $FMNH_2$ of −2.9.

RESULTS

Amino Acid Sequence. The proposal for the amino acid sequence of *Rb. sphaeroides* SHP (Figure 1) is based upon the results of Edman degradation analysis of peptides obtained by cleavage of the apoprotein with three endoproteinases and one chemical cleavage. The separation of the peptides generated by Glu-C endoproteinase resulted in most of the sequence information (Figure 2A). The fact that several Glu−X peptide bonds were not cleaved, such as Glu82−Lys83 and Glu101−Lys102, was very beneficial for the sequence determination. It is also noteworthy that peptides S4 and S5, although eluted separately, cover the

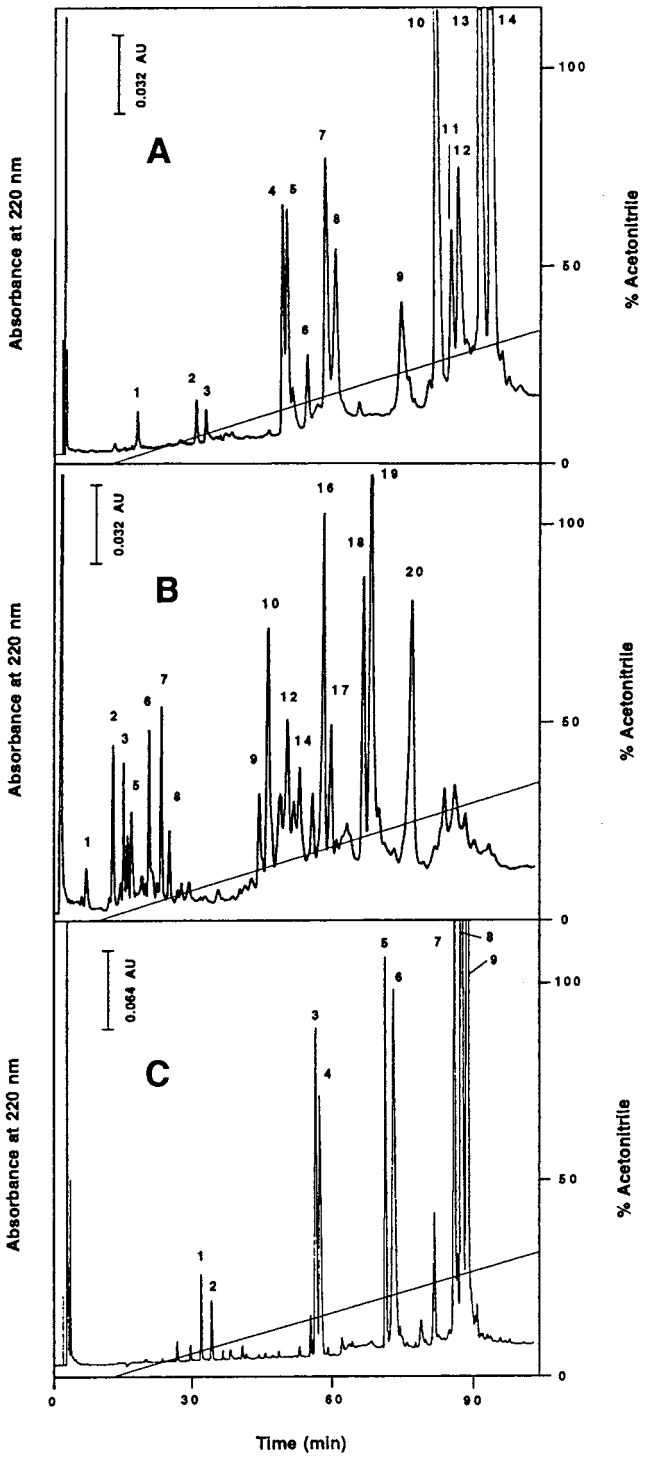

FIGURE 2: Reversed-phase liquid chromatography of peptides obtained after digestion with Glu-C protease (A), partial acid hydrolysis (B), and digestion with Arg-C endoproteinase (C). The same separation conditions were used for the three analyses (see Experimental Procedures).

same sequence region (Arg24−Glu62). The same is true for peptides S13 and S14 (Ile63−Gln112). This phenomenon of different elution times is clearly associated with the presence of free or linked cysteine residues in each of these sequence regions. The overlap between both regions was best provided by peptide AH10 (Val51−Asp73) from the digest mixture obtained by partial acid hydrolysis (Figure 2B). This cleavage procedure also provided clear evidence, after pyridylethylation of peptide AH7, that residues 43 and

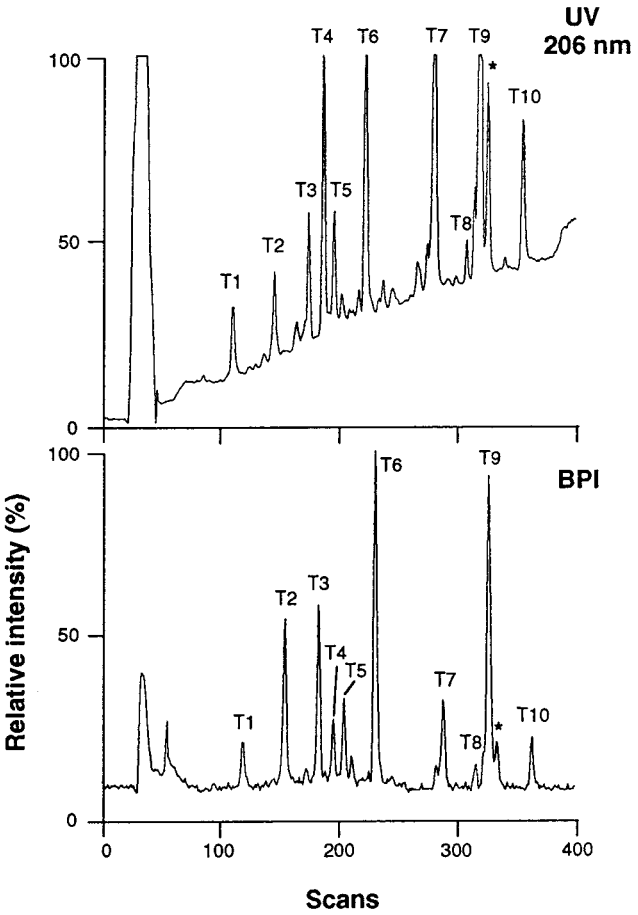

FIGURE 3: LC−ESMS analysis of tryptic peptides from native SHP. Peptide fractions are numbered following their retention times during liquid chromatography (top). The fraction marked with an asterisk could not be assigned. The *y*-axes represent the absorbances measured during chromatography (UV, 0.2 unit full scale) and the base peak intensities measured during mass spectrometry (BPI). The *x*-axis gives the total number of mass spectra scanned during chromatographic analysis.

46 are cysteines. The proof that residues 89 and 97 are also cysteines was obtained through sequence analysis, after pyridylethylation, of peptide R7 which was obtained after cleavage of the apoprotein with Arg-C protease (Figure 2C). This peptide also provided the overlap between peptides AH16−S13 and AH19, the latter being the C-terminal partial acid hydrolysate peptide of the protein. Although Edman degradation analysis showed glutamine at cycle 9, amino acid compositional analysis of peptide AH19 revealed that this peptide contained only one Glx residue. This conclusion was also confirmed by treatment of the apoprotein with carboxypeptidase P. After incubation for 10 min, only a single residue of Gln was cleaved.

Mass spectrometry was used to determine whether the cysteines at positions 89 and 97 in native SHP are involved in a disulfide bridge. Figure 3 shows the LC−ESMS base peak intensities (BPI) chromatogram of peptides generated by tryptic digestion of native SHP. The choice of this protease was made due to the presence of a cleavable Arg−Asp bond occurring between Cys89 and Cys97. The individual masses of the peptides are given in Table 1. Not only do they confirm the proposal for the amino acid sequence of SHP (Figure 1), but they also show that the two tryptic peptides (Asn88−Arg95 and Asp96−Lys102) are

Table 1: Capillary LC−ESMS Analysis of Tryptic Peptides from Native SHP

peptide	measured mass (Da)	calculated mass (Da)	position
T1	695.4	695.7	Phe75−Arg80
T2[a]	1608.9	1608.8	Asn88−Arg95 Asp96−Lys102
T3[a,b]	1590.7	1590.8	Asn88−Lys102
T4	530.4	530.6	Trp84−Arg87
T5	1337.4	1337.5	Glu62−Arg74
T6	2015.2	2015.2	Glu62−Arg80
T7	2302.3	2302.4	Gly1−Arg24
T8[b,c]	3593.5	3593.9	Gly25−Arg53
T9[c]	3380.5	3379.6	Ala27−Arg53
T10	1070.9	1071.2	Ala103−Gln112

[a] Cysteines 89 and 97 linked by a disulfide bridge. [b] Trypsin sensitive peptide bond not cleaved. [c] Peptide containing one heme group, covalently linked to cysteines at positions 43 and 46.

linked to each other by a disulfide bridge between Cys89 and Cys97. Fragment T9 which also contains two cysteines, at positions 43 and 46, was found to carry a heme group as expected from the classical heme binding pattern Cys-X-Y-Cys-His.

The mass of the native SHP measured by ESMS was 12 169.8 Da which corresponds very well to the calculated mass of 11 556.8 Da, supplemented with one heme (616.5 Da) and taking into account the formation of one disulfide bridge (−2.02 Da) (Figure 4A). ESMS analysis of the apoprotein gave masses of 11 553.0, 11 753.4, and 11 954.0 Da (Figure 4B). These values are in good agreement with the calculated masses of 11 552.8, 11 753.4, and 11 954.0 Da for the apoprotein containing two disulfide bridges, one disulfide bridge and one Hg adduct, and two Hg adducts, respectively. Such mercury adducts have already been described for several apocytochromes obtained after heme removal of the native protein with acidic $HgCl_2$ (*16*).

Secondary Structure. The far-UV circular dichroism spectrum of SHP was determined as shown in Figure 5. The helical content was calculated to be 50%. β-Structure and -turns are far less reliably extracted from these data and were ignored. Normally, *c*-type cytochromes do not contain β-structure, and that is also likely to be the case with SHP. When we performed Chou−Fasman (*17*) analysis, we also ignored the predictions of β-structure and concentrated on α-helix which is more reliably predicted as well. The agreement between measurement (50% helix) and prediction (46% helix) was quite good.

Reduction by Free Flavin Semiquinone. The reactions of heme proteins with free reduced flavins can provide information on access to the heme and charge at the site of reduction (*18*). In general, flavin reduction can distinguish between families of redox proteins and can be used to infer structural information with unknown proteins. Thus, the rate constant for reaction of SHP with lumiflavin semiquinone was found to be 2.1×10^6 M^{-1} s^{-1}, and with fully reduced lumiflavin (produced through semiquinone disproportionation), it was 5.3×10^5 M^{-1} s^{-1}. Similarly, the rate constant for reduction by 5-deazariboflavin semiquinone was 1.2×10^7 M^{-1} s^{-1}. These data were compared to the corresponding data for myoglobin, cytochrome *c*, cytochrome *c*′, and yeast cytochrome *c* peroxidase. For the purpose of this work, the cytochrome *c* and cytochrome *c*′ data were normalized to a

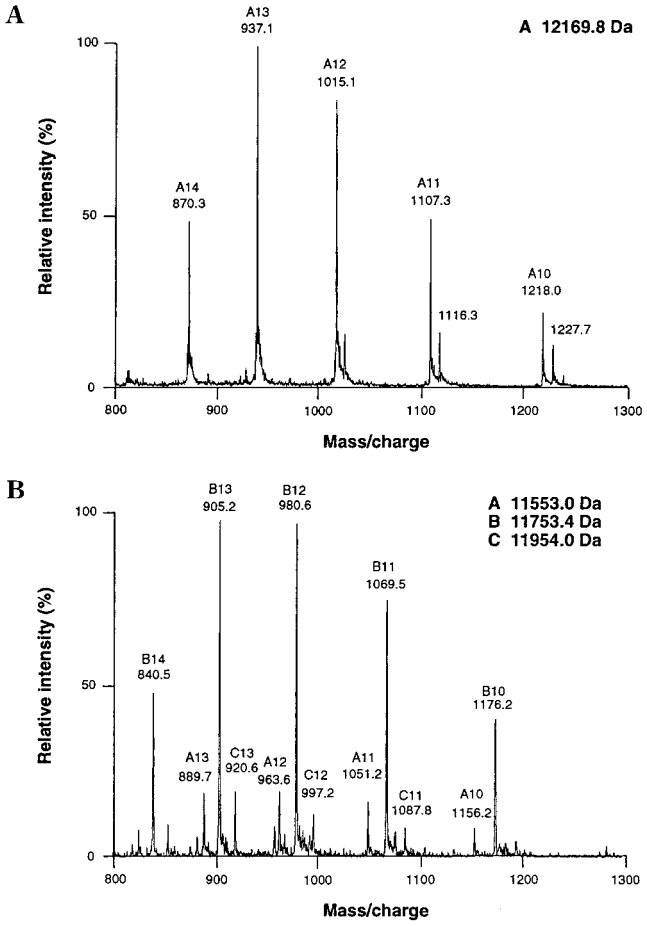

FIGURE 4: Electrospray ionization mass spectrum of native SHP (A) and apo-SHP (B) from *Rb. sphaeroides*. The number at the top of each peak represents the number of positive charges for the particular *m*/*z* peak.

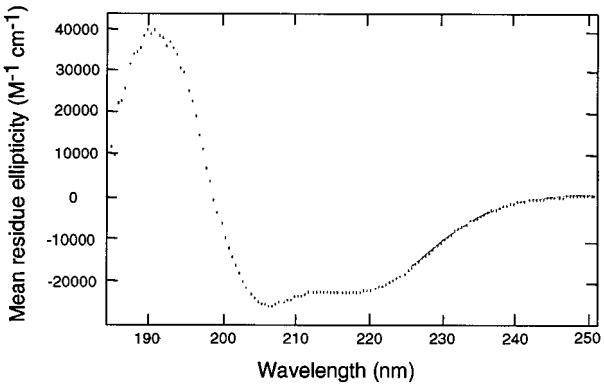

FIGURE 5: Far-UV circular dichroism spectrum of *Rb. sphaeroides* SHP.

redox potential of −22 mV, that of SHP under these solvent conditions (*13*, *18*). SHP was less reactive with lumiflavin than homologous class I cytochromes *c* by about 8-fold, and about 24-fold less reactive than cytochromes *c*′, but was more or less comparable to myoglobin (*13*). SHP was about 2 orders of magnitude less reactive with 5-deazariboflavin than are the cytochromes *c*, but its reactivity is comparable to that of yeast cytochrome *c* peroxidase (*19*), a protein which has a much lower redox potential. An edge of the heme is exposed in class I cytochromes *c*, and the heme face is exposed in cytochromes *c*′. However, the heme is buried in the potein interior of myoglobin and peroxidase. We

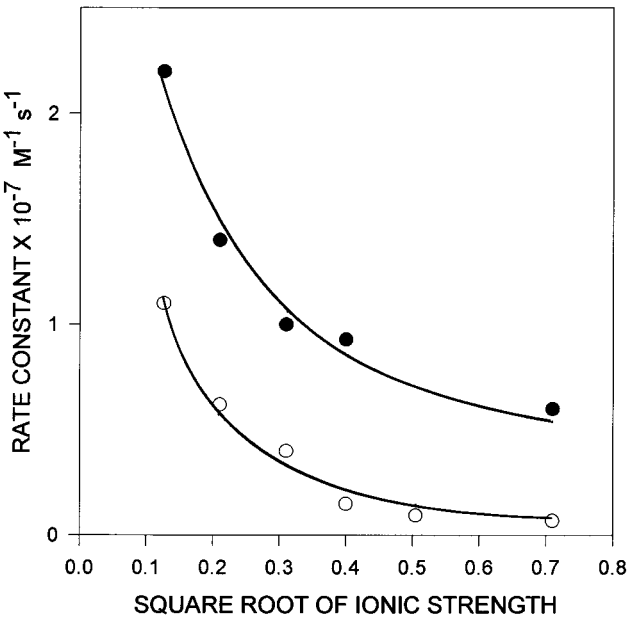

FIGURE 6: Effect of ionic strength on the kinetics of reduction of SHP by FMN semiquinone (●) or by fully reduced FMN (○). Theoretical curves were generated according to Watkins et al. (*15*) using a radius of interaction ρ of 4.5 Å, resulting in V_{ii} values of −2.6 and −4.75 kcal/mol, and rate constants at infinite ionic strength k_∞ of 3.5×10^6 and 4.1×10^5 M^{-1} s^{-1} for FMN SQ and fully reduced FMN, respectively.

conclude that the heme is also inaccessible to solvent in SHP.

Flavin mononucleotide (FMN) has a negative charge due to the ribityl phosphate side chain, and thus, the effect of ionic strength on the kinetics of reduction of heme proteins allows some estimate of the magnitude of localized charge at the active site (*14*). The reaction of SHP with FMN semiquinone or with fully reduced FMN decreases with ionic strength as shown in Figure 6, consistent with a positive charge localized at the site of reduction. The data were fit using the parallel plate model of Watkins et al. (*15*), resulting in a localized charge of +3.1 or +3.7 depending upon which species of flavin was used, semiquinone or fully reduced. Significantly, the interaction site charge is the opposite of the net charge of the protein which was found to be −2 from the amino acid sequence.

Ligand Binding. Generally, high-spin heme proteins such as SHP have a weak field ligand such as water or have no sixth ligand at all. Thus, they are capable of binding the appropriate exogenous ligands in both redox states. These would be anionic ligands such as azide or cyanide in the oxidized state and carbon monoxide in the reduced state. Table 2 summarizes kinetic parameters determined from the data shown in Figure 7. In all three cases, the kinetics were pseudo-first-order at all ligand concentrations used and the intercepts of the k_{obs} versus ligand concentration were not zero, consistent with a measurable dissociation rate constant. As can be seen from the data presented, SHP is substantially less reactive with exogenous ligands than is myoglobin. In fact, SHP has reactivities more like these of the cytochromes *c*′, which are notoriously unreactive with exogenous ligands, as compared to typical high-spin heme proteins, because the ligand binding site is completely buried. These results suggest that the SHP heme environment is not optimized for ligand binding and that, due to steric hindrance, it has low affinities in either redox state for exogenous ligands.

Table 2: Kinetics and Equilibrium for Ligand Binding in *Rb. sphaeroides* SHP (this work) and *C. vinosum* SHP (*30*) Compared with Those of Horse (*Equus caballus*) Myoglobin (*33*) and *C. vinosum* Cytochrome *c*′ (*31*, *32*)

		k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_d
Rs SHP	azide	2.1	0.6	280 mM
	cyanide	37	0.024	650 μM
	CO	1.5×10^4	0.031	33 μM
Cv SHP	CO	5.4×10^3	0.031	5.7 μM
Ec Mb	azide	7×10^3	0.49	70 μM
	cyanide	170	0.003	18 μM
	CO	5×10^5	0.017	34 nM
	O_2	1.9×10^7	37	1.9 μM
cyt *c*′	cyanide			48 μM
	CO	140	0.0018	13 μM

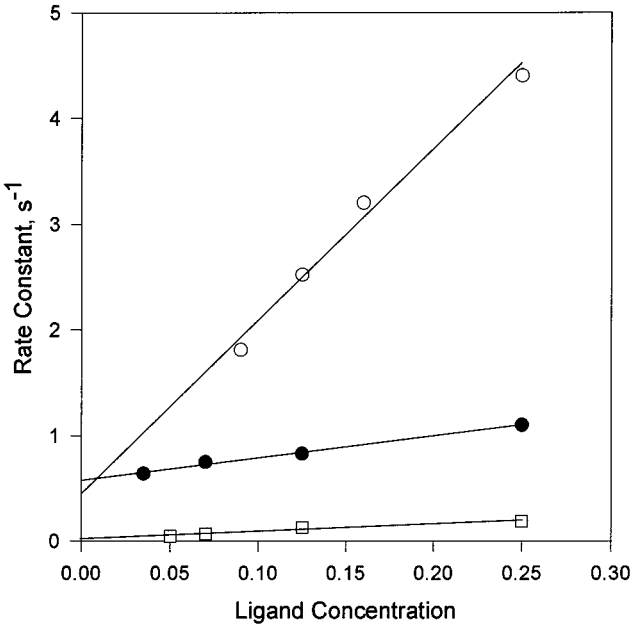

FIGURE 7: Kinetics of ligand binding by SHP. Carbon monoxide is shown in open circles on a millimolar concentration scale. Azide is shown in filled circles on a molar scale and cyanide in open squares with total cyanide on a molar scale.

SHP, like the cytochromes *c*′, can be autoxidized and is slowly converted to the oxidized form when exposed to molecular oxygen. However, in sharp contrast to the cytochromes *c*′, it has been shown that a molecular oxygen−SHP complex is transiently formed in the presence of molecular oxygen (*1*). Unfortunately, efforts to quantify the kinetics of the formation and breakdown of the SHP−O_2 complex were unsuccessful because of technical reasons, including the relatively rapid autoxidizability and limited quantities of protein available. However, we find that, in the presence of 125 μM oxygen (a 4-fold excess), the SHP−oxygen complex has a half-life of about 3 min and is oxygen concentration-dependent. SHP has been cloned, and we anticipate overproducing the protein to obtain sufficient material for oxygen binding and oxidation studies in the future.

DISCUSSION

There are 112 amino acid residues in *Rb. sphaeroides* SHP, including four cysteine residues as shown in Figure 1. Cysteine residues 43 and 46 along with histidine 47 constitute a typical heme binding site as found in all *c*-type cyto-

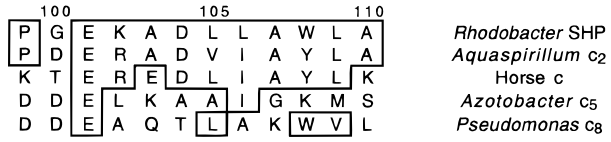

FIGURE 8: Alignment of the C-terminal helices of some class I cytochromes with *Rb. sphaeroides* SHP residues 99−110.

chromes. In class I cytochromes, the heme is bound near the N terminus in the vicinity of residues 10−20 and in class II cytochromes near the C terminus in the vicinity of residues 120−130. SHP fits neither pattern as the heme is near the center of the protein. However, there is a possibility that SHP is related to type I *c*-type cytochromes but may have an N-terminal extension of about 30 residues. The C-terminal fragment would then be about the same size as in small cytochromes such as *Pseudomonas* and *Azotobacter* cytochromes c_5 and c_8. However, when we performed a BLAST search for homologues, we did not find any significant similarities. This does not exclude the possibility that SHP is related to other cytochromes, only that a three-dimensional structure will be necessary to prove it.

Cytochrome c_5 is actually found in two forms. The water-soluble form has 83 amino acid residues; its sequence and three-dimensional structure are known (*20*, *21*). The membrane-bound form can be solubilized with butanol/water and has about 33 extra residues at the N terminus (*21*). Thus, SHP might be like the membrane form of cytochrome c_5. There is an additional similarity between SHP and cytochrome c_5 as they both have extra cysteines not required to bind the heme and which are of extremely rare occurrence in *c*-type cytochromes. Cys68 in soluble cytochrome c_5 or Cys101 in the membrane c_5 appears to be equivalent to Cys97 in SHP. Cys68 forms a disulfide with Cys65 in cytochrome c_5 (*20*), but the other extra cysteine in SHP is at position 89 which is at about the same location as the Met sixth ligand in cytochrome c_5. We have shown here by mass spectrometric analysis that Cys89 and Cys97 are also linked by a disulfide bridge. There are thus seven instead of two intervening residues between the cysteine partners.

Although tryptophan is generally a rare amino acid in proteins, there are two of these residues in SHP (positions 84 and 108). Neither is located in the same region as the sole Trp in cytochrome c_5, but when the heme binding sites are aligned, the two Trps in SHP are in approximately the same location as the two in *Pseudomonas* cytochrome c_8 (*21*). A sequence region in SHP that is similar to class I cytochromes *c* also is situated near Trp108. In fact, there is strong similarity to *Aquaspirillum itersonii* class I cytochrome c_2, which is a typical class I protein (Figure 8). Trp108 corresponds to the consensus aromatic residue in this helix. There are fewer identities to horse cytochrome *c* and *Pseudomonas* cytochrome c_8. Moreover, those substitutions which do occur are conservative in nature. Thus, Trp108 is replaced by Tyr, Leu106 by Ile, and Lys102 by Arg.

Class I cytochromes *c* generally have the following consensus sequence at the N-terminal helix, DXXXGXXXFXXXC. These consensus residues also occur in SHP, but there are 13 residues between the Phe and Cys instead of the usual 3. This in not unheard of with class I cytochromes since it is known that the second heme domain of *Pseudomonas stutzeri* cytochrome c_4 has 11 residues (*22*) and the first heme domain of bacterial cytochrome *c* peroxidase from *Pseudomonas aeruginosa* has 12 residues between the consensus Phe and the heme (*23*). These insertions bulge out in loops and do not affect the orientation of the N-terminal helices. On the basis of this and the above-mentioned comparisons, we believe that SHP is distantly related to class I cytochromes *c*.

The region of similarity around Trp108, shown in Figure 8, comprises the C-terminal helix in class I cytochromes *c*. In the small form of cytochrome c_5, there are three other helices which altogether amount to 49% of the protein. *Pseudomonas* cytochrome c_8 is 52% helical, and tuna cytochrome *c* contains 51% helix (*24*, *25*).

The secondary structure of SHP measured by circular dichroism spectroscopy shows a helical content of 50%. Thus, SHP appears to be similar to class I cytochromes *c* in secondary structure as well. Chou−Fasman analysis of SHP results in the prediction of four helical regions (residues 5−28, 59−64, 76−83, and 100−112), totaling 46% of the protein (see Figure 1). Thus, the segment which appears to be homologous to the C-terminal helix in class I cytochromes *c* is also predicted to be helical in SHP. The region near the N terminus which contains the three consensus residues also is predicted to be helical as in class I cytochromes *c*.

Since SHP binds oxygen, we looked for a possible structural similarity to the oxygen binding proteins such as hemoglobin and myoglobin. Bacterial hemoglobins have been sequenced from *Vitreoscilla* (*26*), *Escherichia coli* (*27*), and *Alcaligenes eutrophus* (*28*). *Vitreoscilla* hemoglobin has 146 residues and is thus about 30% larger than SHP. The proteins also differ in that the bacterial hemoglobin has a noncovalently bound protoheme and there is clear homology with eukaryotic hemoglobins in terms of both sequence and crystal structure (*29*). *A. eutrophus* and *E. coli* hemoglobins are larger than *Vitreoscilla* hemoglobin and contain a flavoprotein reductase domain in addition to the oxygen-binding heme domain. The hemoglobins are virtually fully helical structures (80−90%) which fold in a manner different from that of the class I cytochromes and SHP. We conclude that SHP is not related to hemoglobin. It eventually will be interesting to see, however, if there are any convergent features in the three-dimensional structure near the heme in SHP which might stabilize the oxygen complex. X-ray analysis of SHP is being carried out in our laboratory.

Gaul et al. (*30*) measured CO binding in a protein spectrally similar to SHP, but which was isolated from another phototrophic bacterium, *C. vinosum*. The values they obtained for the rate constants and dissociation constant are similar to those we obtained for SHP. Ligand binding in cytochrome *c*′ (*31*, *32*) is slightly stronger than that for SHP. However, the affinities of SHP for carbon monoxide, azide, and cyanide are more 1−3 orders of magnitude weaker than that for myoglobin (*33*). The most dramatic difference between the two proteins is that myoglobin is not very autoxidizable at pH 7 and forms a relatively stable oxygen complex. SHP, on the other hand, only transiently forms an oxygen complex during autoxidation. The half-life for autoxidation of myoglobin is a few hours at pH 5 but is about 1 day at pH 7 (*34*). This is to be contrasted with a few minutes for SHP autoxidation at pH 7. There are two known factors which contribute to a stable oxygen complex in the globins, a hydrophobic binding site and H bonding to the

distal histidine. The fact that SHP binds oxygen suggests that the heme environment may also be hydrophobic. However, the oxygen complex may not be as stable in SHP as in myoglobin due to the lack of a distal histidine. The sequence shows that there is only one histidine in SHP and it forms the proximal heme ligand.

The kinetics of reduction of SHP by free flavins allows two important conclusions: that there is appreciable steric hindrance at the heme site relative to that of *c*-type cytochromes and that the active site has a strong positive charge. The reactivity of SHP is actually between that of myoglobin and yeast cytochrome *c* peroxidase (YCCP), both of which are less reactive than is mitochondrial cytochrome *c*. In YCCP, the heme is deeply buried in the protein interior, whereas one edge of the heme is exposed to solvent in cytochrome *c*. Thus, we propose that the edge of the heme which is normally exposed to solvent in class I cytochromes *c* is covered by a loop of peptide chain in SHP. Cytochromes which do not have charge localized at the active site show a small electrostatic effect in reaction with reduced FMN which reflects the influence of the net protein charge. However, because the net charge of SHP from the sequence is −2, the +3 charge experienced by reduced FMN must reflect localized charge at the site of reduction. All mitochondrial cytochromes *c* and bacterial cytochromes c_2 have positive charge at the site of electron transfer, which is in the region of the exposed heme edge and which is independent of the net protein charge (*14*).

REFERENCES

1. Meyer, T. E., and Cusanovich, M. A. (1985) *Biochim. Biophys. Acta 807*, 308−319.
2. Fitch, J., Cannac, V., Meyer, T. E., Cusanovich, M. A., Tollin, G., Van Beeumen, J., Rott, M. A., and Donohue, T. J. (1989) *Arch. Biochem. Biophys. 271*, 502−507.
3. Rott, M. A., and Donohue, T. J. (1990) *J. Bacteriol. 172*, 1954−1961.
4. Rott, M. A., Witthuhn, V. C., Schilke, B. A., Soranno, M., Ali, A., and Donohue, T. J. (1993) *J. Bacteriol. 175*, 358−366.
5. Ambler, R. P., Hermoso, J., Meyer, T. E., Bartsch, R. G., and Kamen, M. D. (1979) *Nature 278*, 659−660.
6. Ambler, R. P., Bartsch, R. G., Daniel, M., Kamen, M. D., McLellan, L., Meyer, T. E., and Van Beeumen, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6854−6857.
7. Bartsch, R. G., Ambler, R. P., Meyer, T. E., and Cusanovich, M. A. (1989) *Arch. Biochem. Biophys. 271*, 433−440.
8. Gaul, D. F., Gray, C. O., and Knaff, D. B. (1983) *Biochim. Biophys. Acta 723*, 333−339.
9. Ambler, R. P., and Wynn, M. (1973) *Biochem. J. 131*, 485−498.
10. Amons, R. (1987) *FEBS Lett. 212*, 69−72.
11. Klarskov, K., Roecklin, D., Bouchon, B., Sabatié, J., Van Dorsselaer, A., and Bischoff, R. (1994) *Anal. Biochem. 216*, 127−134.
12. Davis, M. T., and Lee, T. D. (1992) *Protein Sci. 1*, 935−944.
13. Meyer, T. E., Przysiecki, C. T., Watkins, J. A., Bhattacharyya, A., Simondsen, R. P., Cusanovich, M. A., and Tollin, G. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6740−6744.
14. Meyer, T. E., Watkins, J. A., Przysiecki, C. T., Tollin, G., and Cusanovich, M. A. (1984) *Biochemistry 23*, 4761−4767.
15. Watkins, J. A., Cusanovich, M. A., Meyer, T. E., and Tollin, G. (1994) *Protein Sci. 3*, 2104−2114.
16. Samyn, B., Berks, B. C., Page, M. D., Ferguson, S. J., and Van Beeumen, J. J. (1994) *Eur. J. Biochem. 219*, 585−594.
17. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry 13*, 222−245.
18. Tollin, G., Meyer, T. E., and Cusanovich, M. A. (1986) *Biochim. Biophys. Acta 853*, 29−41.
19. Hazzard, J. T., Poulos, T. L., and Tollin, G. (1987) *Biochemistry 26*, 2836−2848.
20. Carter, D. C., Melis, K. A., O'Donnell, S. E., Burgess, B. K., Furey, W. F., Jr., Wang, B. C., and Stout, C. D. (1985) *J. Mol. Biol. 184*, 279−295.
21. Ambler, R. P. (1991) *Biochim. Biophys. Acta 1058*, 42−47.
22. Kadziola, A., and Larsen, S. (1997) *Structure 5*, 203−216.
23. Fulop, V., Ridout, C. J., Greenwood, C., and Hajdu, J. (1995) *Structure 3*, 1225−1233.
24. Mandel, N., Mandel, G., Trus, B. C., Rosenburg, J., Carlson, G., and Dickerson, R. E. (1977) *J. Biol. Chem. 252*, 4619−4636.
25. Matsuura, Y., Takano, T., and Dickerson, R. E. (1982) *J. Mol. Biol. 156*, 389−409.
26. Wakabayashi, S., Matsubara, H., and Webster, D. A. (1986) *Nature 322*, 481−483.
27. Vasudevan, S. G., Armarego, W. L. F., Shaw, D. C., Lilley, P. E., Dixon, N. E., and Poole, R. K. (1991) *Mol. Gen. Genet. 226*, 49−58.
28. Cramm, R., Siddiqui, R. A., and Friedrich, B. (1994) *J. Biol. Chem. 269*, 7349−7354.
29. Ermler, U., Siddiqui, R. A., Cramm, R., and Friedrich, B. (1995) *EMBO J. 14*, 6067−6077.
30. Gaul, D. F., Ondrias, M. R., Findsen, E. W., Palmer, G., Olson, J. S., Davidson, M. W., and Knaff, D. B. (1987) *J. Biol. Chem. 262*, 1144−1147.
31. Cusanovich, M. A., and Gibson, Q. H. (1973) *J. Biol. Chem. 248*, 822−834.
32. Kassner, R. J., Kykta, M. G., and Cusanovich, M. A. (1985) *Biochim. Biophys. Acta 831*, 155−158.
33. Antonini, E., and Brunori, M. (1971) *Hemoglobin and Myoglobin in their reactions with ligands*, Elsevier North Holland, New York.
34. Brown, W. D., and Mebine, L. B. (1969) *J. Biol. Chem. 244*, 6696−6701.

BI972498W