Protein Science (2000), *9*:587–597. Cambridge University Press. Printed in the USA.

Cloning, expression, purification, and preliminary characterization of a putative hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803

NANCY L. SCOTT AND JULIETTE T.J. LECOMTE

Department of Chemistry and the Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802

(RECEIVED November 3, 1999; FINAL REVISION January 3, 2000; ACCEPTED January 13, 2000)

Abstract

The genome of the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 contains a gene (slr2097, *glbN*) encoding a 123 amino-acid product with sequence similarity to globins. Related proteins from cyanobacteria, ciliates, and green algae bind oxygen and have a pronounced tendency to coordinate the heme iron with two protein ligands. To study the structural and functional properties of *Synechocystis* sp. PCC 6803 hemoglobin, slr2097 was cloned and overexpressed in *Escherichia coli*. Purification of the hemoglobin was performed after addition of hemin to the clarified cell lysate. Recombinant, heme-reconstituted ferric *Synechocystis* sp. PCC 6803 hemoglobin was found to be a stable helical protein, soluble to concentrations higher than 500 μM. At neutral pH, it yielded an electronic absorption spectrum typical of a low-spin ferric species, with maxima at 410 and 546 nm. The proton NMR spectrum revealed sharp lines spread over a chemical shift window narrower than 40 ppm, in support of low-spin hexacoordination of the heme iron. Nuclear Overhauser effects demonstrated that the heme is inserted in the protein matrix to produce one major equilibrium form. Addition of dithionite resulted in an absorption spectrum with maxima at 426, 528, and 560 nm. This reduced form appeared capable of carbon monoxide binding. Optical data also suggested that cyanide ions could bind to the heme in the ferric state. The spectral properties of the putative *Synechocystis* sp. PCC 6803 hemoglobin confirmed that it can be used for further studies of an ancient hemoprotein structure.

Keywords: cyanoglobin; GlbN; hemichrome; hemoglobin; *Synechocystis* sp. PCC 6803

The amino acid sequences and gene structures of hemoglobins found in various kingdoms of organisms suggest that all globins have evolved from an ancestral set of hemoproteins. These existed before atmospheric O_2 became available and the physiological roles they assumed are open to speculation (Hardison, 1998). Certain cyanobacteria, the oxygenic photosynthetic organisms thought to be responsible for the modification of the atmosphere 3.5 to 2 billion years ago (Schopf & Walter, 1982), synthesize a hemoprotein belonging to the globin family (cyanoglobin, GlbN; Potts et al., 1992; Hill et al., 1996). GlbN and its relatives are expected to provide insights into the evolution of globin functions and to reveal new aspects of heme chemistry, in particular the mechanism of discrimination among ligands to the iron such as NO, CO, and O_2.

Cyanoglobin was first discovered in *Nostoc commune* UTEX 584, a filamentous heterocyst-forming cyanobacterium capable of aerobic N_2 fixation. In *N. commune*, GlbN is the 12.5 kDa product of the *glbN* gene, which forms a part of two contiguous operons encoding proteins of the nitrogenase complex (Potts et al., 1992; Angeloni & Potts, 1994; Hill et al., 1996). GlbN was originally proposed to scavenge oxygen to protect the nitrogenase complex from inactivation. A participation in nitrogen fixation has since been supported through in vivo experiments, and the location of GlbN on the cytosolic face of the cell membrane further suggests that GlbN presents O_2 to a terminal oxidase complex (Hill et al., 1996). Purified *N. commune* GlbN has a high affinity for O_2 and is capable of rapid O_2 release (Thorsteinsson et al., 1999) in agreement with the proposed functions.

Reprint requests to: Juliette T.J. Lecomte, Chemistry Department, The Pennsylvania State University, 152 Davey Laboratory, University Park, Pennsylvania 16802; e-mail: jtl1@psu.edu.

Abbreviations: bp, base pair; CD, circular dichroism; DQF-COSY, double-quantum-filtered correlated spectroscopy; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; EDTA, ethylenediaminetetraacetic acid; EPR, electron paramagnetic resonance; GlbN, cyanoglobin; Hb, hemoglobin; HbCO, carbonmonoxyhemoglobin; MALDI, matrix-assisted laser desorption ionization; MCOSY, magnitude two-dimensional correlated spectroscopy; MRE, molar residual ellipticity; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; PAGE, polyacrylamide gel electrophoresis; ppm, parts per million; SDS, sodium dodecylsulfate; TOCSY, totally correlated two-dimensional spectroscopy; TPPI, time-proportional phase incrementation; Tris, tris(hydroxymethyl)-aminomethane.

```
                  AAAAAAAAAAA      BBBBBBBB BBBBBBBCCC CCCC  EEEE EEEEEEEEEE EEEEE         FFFFFFFF
                     1                                                                        70
GLB_SYN/1-124     ---MSTLYEKLGG TTAVDLAVDK FYERVLQDDR IKHFFADVDM AKQRAHQKAF LTYAFGGTDK YDGRYMREAH
GLB_TETPY/1-121   MNKPQTIYEKLGG ENAMKAAVPL FYKKVLADER VKHFFKNTDM DHQTKQQTDF LTMLLGGPNH YKGKNMTEAH
GLB_TETTH/1-121   MRKQPTVFEKLGG QAAMHAAVPL FYKKVLADDR VKHYFKNTNM EHQAKQQEDF LTMLLGGPNH YKGKNMAEAH
GLB1_CHLEU/39-161 RKCPSSLFAKLGG REAVEAAVDK FYNKIVADPT VSTYFSNTDM KVQRSKQFAF LAYALGGASE WKGKDMRTAH
GLB2_CHLEU/39-161 KKCPFSLFAKLGG REAVEAAVDK FYNKVVADPT VSVFFSKTDM KVQRSKQFAF LAYALGGAAE WKGKDMRTAH
S60030/1-117      ----MTLFEQLGG EAAVTAVTTQ FYANIQADAT VANFFNGINM ADQTNKTASF LCAALGGPKA WGGRNLKEVH
S60031/1-117      ----MTLFEQLGG EAAVHAVTTQ FYANIAADAT VANFFNGINM PTQTDKTAAF LCAALGGPNA WAGRNLKEVH
GLB_PARCA/1-117   ----MSLFEQLGG QAAVQAVTAQ FYANIQADAT VATFFNGIDM PNQTNKTAAF LCAALGGPNA WTGRNLKEVH
GLBN_NOSSN/1-118  ---MSTLYDNIGG QPAIEQVVDE LHKRIATDSL LAPIFAGTDM AKQRNHLVAF LGQIFEGPKQ YGGRPMDKTH
GLBN_NOSCO/1-118  ---MSTLYDNIGG QPAIEQVVDE LHKRIATDSL LAPVFAGTDM VKQRNHLVAF LAQIFEGPKQ YGGRPMDKTH
Y04N_MYCTU/9-134  KREPISIYDKIGG HEAIEVVVED FYVRVLADDQ LSAFFSGTNM SRLKGKQVEF FAAALGGPEP YTGAPMKQVH

                  FF       G GGGGGGGGGG GGGGGGGG    HHHHHHHHH HHHHHHHHHHHHH HHHH
                  71                                    ---          124
GLB_SYN/1-124     KELVENHGLN GEHFDAVAED LLATLKEMGV PEDLIAEVAA VAG---APAHKRD VLNQ
GLB_TETPY/1-121   KG----MNLQ NLHFDAIIEN LAATLKELGV TDAVINEAAK VI-----EHTRKD MLGK
GLB_TETTH/1-121   KG----MNLQ NSHFDAIIEN LAATLKELGV SDQIIGEAAK VI-----EHTRKD CLGK
GLB1_CHLEU/39-161 KDLVP--HLS DVHFQAVARH LSDTLTELGV PPEDITDAMA VV-----ASTRTE VLNM
GLB2_CHLEU/39-161 KDLVP--HLT DVHFQAVVRH LSDTLAELGV TPGDIADAMA VV-----ASTKTE VLNM
S60030/1-117      AN----MGVT NAQFTTVIGH LRSALTSAGV AADLVEQTVA VA-----ETVRGD VVTA
S60031/1-117      AN----MGVT NAQFTTVIGH LRSALTGAGV AAALVEQTVA VA-----ETVRGD VVTV
GLB_PARCA/1-117   AN----MGVS NAQFTTVIGH LRSALTGAGV AAALVEQTVA VA-----ETVRGD VVTV
GLBN_NOSSN/1-118  AG----LNLQ QPHFDAIAKH LGEAMAVRGV SAEDTKAALD RV-----TNMKGA ILNK
GLBN_NOSCO/1-118  AG----LNLQ QPHFDAIAKH LGERMAVRGV SAENTKAALD RV-----TNMKGA ILNK
Y04N_MYCTU/9-134  QG----RGIT MHHFSLVAGH LADALTAAGV PSETITEILG VIAPLAVDVTSGE STTA
```

Fig. 1. Primary structure of the globin from *Synechocystis* sp. PCC 6803 (GLB_SYN; gene: slr2097, *glbN*) and its ClustalW alignment (Thompson et al., 1994) with other globins. The proteins are: GLB_TETPY, myoglobin (hemoglobin) from *Tetrahymena pyriformis*; GLB_TETTH, myoglobin (hemoglobin) from *Tetrahymena thermophila*; GLB1_CHLEU, globin LI637 {gene: *LI637*} from *Chlamydomonas eugametos*; GLB2_CHLEU, globin LI410 {gene: *LI410*} from *C. eugametos*; S60030, hemoglobin from *Paramecium jenningsi*; S60031, hemoglobin from *Paramecium multimicronucleatum*; GLB_PARCA, myoglobin (hemoglobin) from *P. caudatum*; GLBN_NOSSN, cyanoglobin {gene: *glbN*} from *Nostoc* sp. (strain mun 8820); GLBN_NOSCO, cyanoglobin. {gene: *glbN*} from *Nostoc commune* UTEX 584; Y04N_MYCTU, hypothetical 14.5 kd protein cy48.23 {gene: mtcy48.23, *glbN*} from *Mycobacterium tuberculosis*. The numbers following the denomination refer to the portion of the sequence used in the alignment. The numbers above the *Synechocystis* sp. PCC 6803 sequence refer to that structure alone. The top line contains the helical portions as established with vertebrate and nonvertebrate sequence templates (Moens et al., 1996). The exact helical boundaries remain to be determined with three-dimensional structural data. According to the alignment, this group of globins lacks the beginning of the A-helix as well as the D-helix. The proximal histidine (F8) is His70 (bold and underlined) and the distal position (E7) is occupied by a Gln (underlined). Strictly conserved residues in this group are in bold.

In contrast to *N. commune*, *Synechocystis* sp. PCC 6803 is a unicellular cyanobacterium that does not produce the nitrogenase complex and thus utilizes only fixed nitrogen sources such as ammonium or nitrate ions for growth. The genome of *Synechocystis* sp. PCC 6803 has been sequenced in its entirety (Kaneko et al., 1996). The sequence revealed the presence of a gene (slr2097, *glbN*), encoding a globin sharing 55% sequence identity with the cyanoglobin from *N. commune* (Fig. 1). Several other organisms that do not fix nitrogen synthesize a related globin (Fig. 1; Thompson et al., 1994; Moens et al., 1996). These include the gram-positive bacteria *Bacillus subtilis* (Kunst et al., 1997) and *Mycobacterium tuberculosis* (Cole et al., 1998), which actually encodes two globin-like proteins (GlbN and GlbO), as well as eukaryotes of the genera *Tetrahymena* (Iwaasa et al., 1990), *Paramecium* (Iwaasa et al., 1989; Shikama et al., 1995; Yamauchi et al., 1995), and *Chlamydomonas* (Couture et al., 1994).

Limited functional information is available for this disparate group of proteins. *Chlamydomonas eugametos* LI637 hemoglobin releases its bound O_2 three orders of magnitude more slowly than horse heart myoglobin (Couture et al., 1999a). The slow O_2 release and the sub-μM hemoglobin concentration in the chloroplast preclude a transport role. *Mycobacterium tuberculosis* GlbN forms a dimer binding oxygen cooperatively but its primary function may be to scavenge reactive nitrogen compounds (Couture et al., 1999b). The functions of hemoglobin in the ciliates have not been determined.

In *Synechocystis* sp. PCC 6803, the organism of interest here, the location of the slr2097 gene in the genome does not immediately suggest a role for the protein. It is not known under which conditions the synthesis of the gene product is induced, where this product is localized, and whether it binds dioxygen reversibly. If not expressed under the range of conditions typically encountered by this organism, it is possible that slr2097 is a cryptic gene, one that can be reactivated under extreme occasional environmental pressure (Li, 1984). To uncover the role of the putative hemoglobin in *Synechocystis* sp. PCC 6803 and to produce a tractable hemoprotein for the study of structure-reactivity relationships in ancient globin folds, the slr2097 gene was cloned and its protein product overexpressed in *Escherichia coli*. The purified recombinant protein was subjected to a preliminary set of optical and nuclear magnetic resonance spectroscopy experiments. It was found that the protein folds into a helical conformation and forms a stable hexacoordinate complex with iron(III)-protoporphyrin IX.

Results

Expression, extraction, and purification of Synechocystis *sp. PCC 6803 hemoglobin*

The gene for *Synechocystis* sp. PCC 6803 globin (Kaneko et al., 1996) was cloned from purified chromosomal DNA as described in the Materials and methods. The gene was inserted in a pET3c

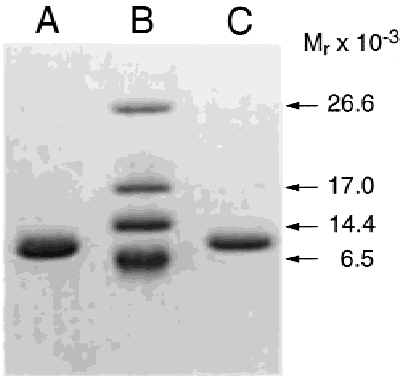

Fig. 2. SDS-PAGE of the purified recombinant S6803 Hb. **B:** The *central lane* contains the molecular weight markers, with masses indicated to the right of the figure. **C:** The *right-hand lane* contains 2.5 μg of the purified protein for comparison with the markers. **A:** The *left-hand lane* is overloaded with ~10 μg protein. This gel and a silver-stained gel indicate a purity higher than 95%.

vector and *E. coli* BL21(DE3) cells were used for expression. In initial tests, samples were visualized on gels stained with Coomassie blue and showed that S6803 Hb[1] was expressed as at least 60% of total cellular protein by 4–5 h postinduction at 37 °C. Gel electrophoresis of clarified cell lysates and pellets demonstrated the production of protein in a soluble form as well as the accumulation of insoluble protein in inclusion bodies. Although both sources yield pure S6803 Hb with the same properties, only results obtained with the soluble fraction are presented here.

Hemin-reconstituted S6803 Hb eluted from the ion-exchange column between ~0.13–0.22 M of the sodium chloride gradient. Elution from the size exclusion column occurred in an almost entirely resolved peak, and pooled fractions yielded a preparation of red-orange-colored protein greater than 95% pure as determined by Coomassie blue- and silver-stained gels. The former is shown in Figure 2. This purification method yielded 5.2 mg of pure protein per liter of bacterial culture. As evidenced by the pinkish color of harvested cells, ferrous holoHb is likely produced in vivo to an extent limited by the level of heme biosynthesis in M-9 cultured BL21(DE3) cells. However, throughout this aerobic preparation, no separate band corresponding to oxyHb was detected. At each chromatographic step, the Hb concentration estimated by gel electrophoresis and by electronic absorption at 410 nm confirmed that the heme group remained associated with the protein matrix throughout the purification.

The identity of the purified material was ascertained by MALDI mass spectrometry of the intact and trypsin-digested polypeptide. The $[M + H]^+$ peak due to the undigested protein corresponded to a mass of 13,730 Da, within experimental error of the theoretical 13,736 Da mass calculated for complete post-translational removal of the N-terminal methionine and heme loss during sample preparation. The tryptic digest and further analysis by liquid chromatography yielded a series of fragments whose molar masses were consistent with the expected primary structure.

The extinction coefficient of the Soret band of S6803 metHb was determined to be 1.0×10^2 mM^{-1} cm^{-1} by the hemochromogen method (de Duve, 1948). The visible spectrum of the reduce pyridine heme complex indicated that no chemical modification of the heme vinyl group or the porphyrin ring occurred during the aerobic preparation of the protein. Nevertheless, the properties of the hemoglobin synthesized by *Synechocystis* sp. PCC 6803 may be different from those of the recombinant material. Where the distinction may matter, the prepared protein is qualified with the prefix "r" to specify its recombinant nature and the suffix "-R" to indicate that the heme was inserted through an in vitro reconstitution reaction with ferric protoporphyrin IX.

Electronic absorption spectra

Figure 3 presents the visible absorption spectra of oxidized and dithionite-reduced S6803 rHb-R. The spectral properties of these forms are summarized in Table 1. The oxidized protein (metHb; solid line) showed a broad absorption at 546 nm and a shoulder at 578 nm. There was no maximum near 630 nm, where high-spin ferric species generally show a charge-transfer band (Hanania et al., 1966). The spectrum suggested that the recombinant metHb prepared by reconstitution had an endogenous sixth ligand to the iron and a low-spin electronic structure. This particular coordination scheme for the oxidized form is termed a hemichrome (Rachmilewitz et al., 1971). Unlike for vertebrate metmyoglobins and methemoglobins (Antonini & Brunori, 1971), this species existed at neutral pH.

The spectrum of the reduced S6803 rHb-R (Fig. 3, dotted lines) also contrasts with those of the deoxy states of vertebrate myoglobins and hemoglobins. Under physiological conditions, the latter proteins have no sixth ligand or a weak sixth ligand to the ferrous iron; they are in a high-spin state and show a single broad

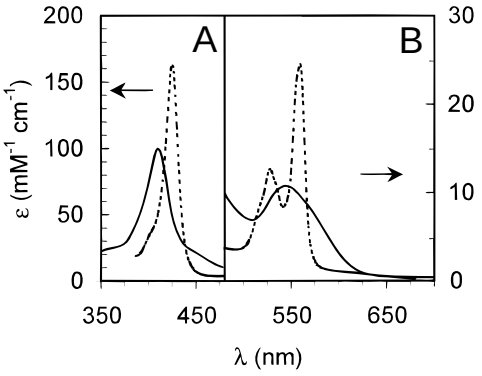

Fig. 3. Optical absorption spectrum of S6803 rHb-R in the visible region. The protein (8.5 μM heme) was maintained at pH 7.4 and 298 K. The ferric form is represented by a solid line, whereas the ferrous form (reduced with dithionite) is shown in a dotted line. The arrows point to the relevant abscissae, **(A)** to the left and **(B)** to the right. Extinction coefficients and maxima are listed in Table 1.

[1]The protein poses a nomenclature problem. Its relative from *Nostoc commune* is called GlbN, as per the name of the gene, *glbN*. In *Synechocystis* sp. PCC 6803, slr2097 was named *glbN*, although it is not closely related to the *Nostoc commune glbN* gene (Hill et al., 1996) and its product may not serve the same function. On the other hand, the generic name "cyanoglobin" (for cyanobacterial globin) applies to the *Synechocystis* sp. PCC 6803 protein but is confusing because of the frequent use of cyanide as a ligand in hemoglobin and myoglobin studies. Here, both GlbN and cyanoglobin are avoided to describe the product of slr2097. The protein is referred to as S6803 hemoglobin (S6803 Hb); this allows the unambiguous use of metHb for the oxidized (ferric) form and cyanometHb for its cyanide adduct.

Table 1. *Spectral properties of S6803 rHb-R*

Hb derivative	λ (nm)	ϵ (mM^{-1} cm^{-1})[a]
Ferric	362	25
	410	100
	546 (sh: 578)	11 (sh: 7)
Ferrous	426	162
	528 (sh: 518)	13 (sh: 9)
	560 (sh: 556)	24 (sh: 23)
Ferrous + CO	420	184
	544 (sh: 568)	14 (sh: 11)

[a]Refer to Figures 3 and 4 for the spectral appearance. Values were determined by the hemochromogen method (de Duve, 1948) with samples of the ferric protein; values in parentheses and marked "sh" denote a shoulder to the main peak.

band near 556 nm (Antonini & Brunori, 1971). Dithionite-reacted S6803 rHb-R presented two resolved bands at 528 nm (β) and 560 nm (α), and its sharp Soret (γ) band was maximal at 426 nm. These features are reminiscent of low-spin hexacoordinated species such as ferrocytochrome b_5 (Strittmatter, 1960).

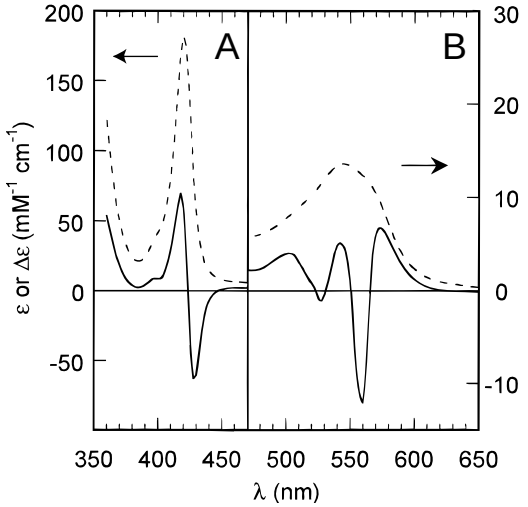

Fig. 4. Optical absorption spectrum of S6803 rHb-R in the visible region. The protein (8.5 μM heme) was maintained at pH 7.4 and 298 K. The dashed line represents the spectrum obtained by reduction in the presence of CO. Extinction coefficients and maxima are listed in Table 1. The solid line represents the difference spectrum obtained by subtracting the dithionite-reduced spectrum (shown Fig. 3) from the reduced-CO spectrum. The extrema in the difference spectrum are at 418, 428, and 558 nm. The arrows point to the relevant abscissae, (**A**) to the left and (**B**) to the right.

Ligand binding properties

To assess whether one of the endogenous axial ligands could readily be displaced by an exogenous ligand in the ferric state, an excess of cyanide was added to a solution of S6803 metHb at pH 7.4. The absorption spectrum, which was monitored over the course of hours, exhibited minor spectral perturbations. At pH 4.4, the metHb spectrum was practically identical to that obtained near neutral pH. Addition of cyanide at this low pH exaggerated the spectral changes seen at pH 7.4, which consisted of a small red shift of the Soret band, a red shift and symmetrization of the absorbance at 546 nm, and the emergence of a shoulder near 360 nm at the base of the Soret peak. Similar small changes are seen in the absorption spectrum of *C. eugametos* LI637 hemoglobin upon cyanide binding (Couture & Guertin, 1996).

The dithionite-reacted form of S6803 rHb-R was used in attempts to generate the oxygen-bound complex. The spectrum obtained after brief passage of air through the solution resulted in marked spectral changes compared to the reduced form. However, the α/β bands remained unresolved and did not appear as in typical O_2-bound spectra of hemoglobins and myoglobins. A comparison of the ferric form and the oxygen-exposed reduced form showed slight differences in the breadth of the Soret band and in the 584 nm shoulder. This result suggested that the protein underwent a rapid reaction in vitro, likely back to an oxidized state. Data were also acquired at pH 6, where the behavior was identical to that at neutral pH. At approximately pH 5 (theoretical pI), the reduced protein precipitated and its oxygenation was not monitored.

A similar procedure was applied to produce the carbon monoxide complex of S6803 rHb-R. When a sample of metHb was saturated with CO and subsequently reduced with excess dithionite, the Soret band increased, sharpened up, and exhibited a red shift (Fig. 4). In contrast with the spectra of CO-bound hemoglobins and myoglobins (Antonini & Brunori, 1971), which show resolved α and β bands, S6803 HbCO exhibited a single band at 544 nm with a shoulder at 568 nm. No further spectral changes were observed after prolonged CO bubbling through the solution. The difference spectrum of the CO-bound reduced protein and the reduced protein (Fig. 4) displayed the features of CO-bound forms of terminal oxidases (Becker & Schafer, 1991): γ peak at 418 nm, γ trough at 428 nm, α trough at 560 nm, and α peak at 574 nm. The depth of the α trough and the ratio $\epsilon(\gamma$ peak $-$ γ trough$)/\epsilon(\alpha$ peak $-$ α trough$) = 7$ both suggested a predominantly low-spin reduced species (Wood, 1984).

Heme release

The transfer of heme from S6803 metHb-R to the apoprotein of equine myoglobin was monitored optically through the change in absorption at 635 nm in a solution containing both proteins. When the molar ratio of apomyoglobin to S6803 metHb was ~3 and the pH was 7.34 (comparable to the cytoplasmic pH), several hours were needed to obtain a detectable absorption attributable to holomyoglobin. Thus, partitioning of the heme from the hemichrome was slow under these conditions.

Circular dichroism spectrum

To assess the extent of secondary structure in the hemichrome, a far-ultraviolet (UV) CD spectrum of S6803 metHb was recorded at pH 7.4 (not shown). The trace displayed the double minima at 209 nm (molar residual ellipticity of $\sim -19{,}000$ mdeg cm^2/dmol) and 222 nm (molar residual ellipticity of $\sim -17{,}500$ mdeg cm^2/dmol) characteristic of a protein containing a predominance of α helices. The molar residual ellipticity was decomposed in contributions from α, β, and random structure with a neural network treatment (Andrade et al., 1993; Merelo et al., 1994). The algo-

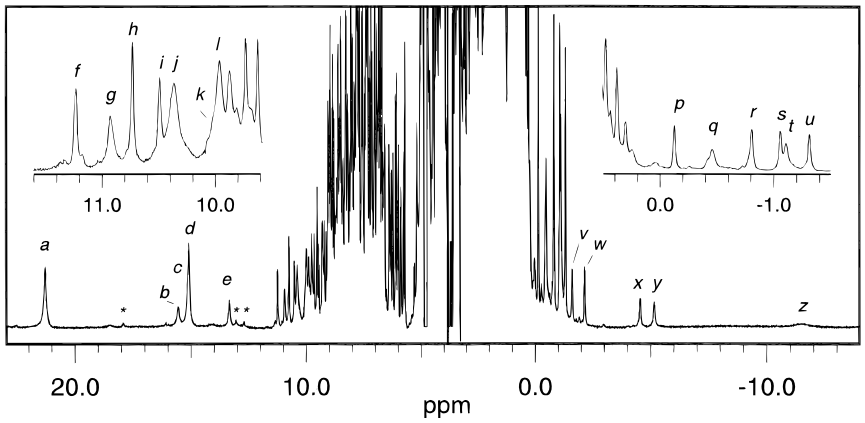

Fig. 5. 600 MHz ^{1}H NMR spectrum of recombinant S6803 metHb-R. The protein (600 μM heme) was in 95% H_2O/5% ^2H_2O, buffered at pH 6.88 with Tris/EDTA; the probe temperature was 298 K. The two insets expand 2-ppm wide regions from the same spectrum. The letters indicate hyperfine-shifted signals (Table 2). Examples of minor peaks are marked with stars.

rithm returned a helical content of over 60% and less than 10% β structure.

NMR spectroscopy

The one-dimensional ^{1}H NMR spectrum of S6803 metHb is shown in Figure 5. Protein and heme resonances were moderately broad and were contained between 24 and -14 ppm. Both observations were consistent with a low-spin ferric species, as inferred from the optical data. The range of chemical shift was narrower than observed for the cyanide complex of *Physeter catodon* (sperm whale) metmyoglobin (Emerson & La Mar, 1990) and for two well-studied oxidized *b* cytochromes, cytochrome b_5 (Keller & Wüthrich, 1980), which has bis-histidine heme ligation, and cytochrome b_{562} (Wu et al., 1991), which has histidine and methionine heme ligation. The range was comparable to that in the cyanide complex of *Glycera dibranchiata* methemoglobin (Alam & Satterlee, 1994) and indicated low magnetic anisotropy. The relative intensity of the resolved resonances showed that the protein adopted mostly a single form. Minor peaks were reproducibly observed that were likely to arise from a small extent of equilibrium heme orientational heterogeneity within the cavity (La Mar et al., 1983) and possibly distinct heme attachment scheme(s). The excellent dispersion observed in the diamagnetic region of the spectrum demonstrated that the protein was folded under those conditions. Chemical shift changes indicating the formation of a distinct low-spin species were observed upon cyanide addition to concentrated neutral solutions of S6803 metHb (not shown). These converted NMR samples displayed the optical features noted above upon cyanide addition at neutral and acidic pH. This confirmed that the small changes in the absorption spectrum were characteristic of cyanide binding and that the cyanometHb complex could be made over a range of pH. In contrast to the metHb form, the reduced protein prepared by anaerobic addition of excess dithionite exhibited limited solubility and its NMR study was not attempted.

Two-dimensional NMR data were collected on S6803 metHb-R to ascertain that the heme group was bound specifically in the protein matrix, and to initiate the structural determination of the hemichrome. Two three-spin systems corresponding to the heme vinyl substituents (see Fig. 6 for the heme structure) were recognized in the DQF-COSY, TOCSY, and MCOSY data (V1: signals *b*, *x*, and *y*; V2: signals at 6.84 ppm, *v*, and *w*). The chemical shifts of these and other signals are listed in Table 2. The *cis* and *trans* β protons of the vinyl side chains were distinguished by the magnitude of the coupling constant to the α proton ($J_{\alpha-\beta t} \sim 15$ Hz $>$ $J_{\alpha-\beta c} \sim 10$ Hz). NOESY data allowed other heme signals to be recognized. Thus, Figure 7 illustrates that the *trans*-β-vinyl protons *v* and *x* were found in dipolar contact with sets of protons at 9.98 ppm (l) and 15.07 ppm (d), respectively. These signals with relative intensity 3 could be attributed to two of the four heme methyl groups. The three-proton signals at 21.27 ppm (a) and at 10.37 ppm (j) had the width and shift of heme methyl groups as well.

Figure 7 shows that both heme vinyl side chains had additional NOEs. The β protons of vinyl V1 were found in contact with protons resonating at 7.38 and 7.29 ppm. Analysis of the correlated data revealed that these, along with proton(s) resonating at 6.75 ppm, arose from a phenylalanine ring. The β protons of vinyl V2 had NOEs to one of the γ methyl groups of a valine at -0.81 ppm (r).

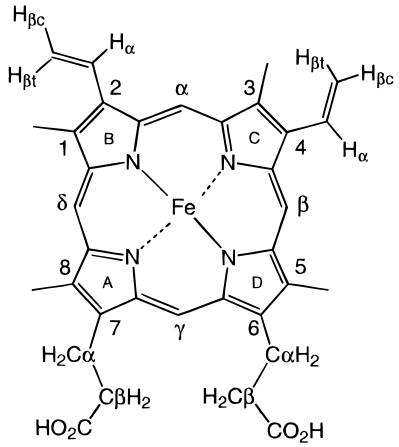

Fig. 6. The structure of the heme group (iron-protoporphyrin IX) used in the reconstitution of S6803 rHb. Methyl groups are located at positions 1, 3, 5, and 8. The meso protons are labeled α through γ.

Table 2. *^{1}H NMR parameters for selected resonances in the spectrum of ferric S6803 rHb-R*

Signal[a]	δ (ppm)[b]	Assignment	δ (ppm) at $T = \infty$[c]	T_1 (ms)[d]
a	21.27	Heme methyl	6.3	130
b	15.54	α-Vinyl (V1)	11	210
d	15.07	Heme methyl	2.3	190
j	10.37	Heme methyl	11	230
l	9.98	Heme methyl	6	260
	6.84	α-Vinyl (V2)		
v	−1.67	*trans*-β-vinyl (V2)	6.2	230
w	−2.14	*cis*-β-vinyl (V2)	5.7	260
x	−4.54	*trans*-β-vinyl (V1)	7	470
y	−5.15	*cis*-β-vinyl (V1)	7	410
z	−11.6		7	3.7

[a]Refer to Figures 5 and 7 for resonance labeling and 6 for heme structure. V1 and V2 distinguish two different sets of signals arising from vinyl groups.

[b]In 95% ^1H_2O/5% ^2H_2O, at 298 K and pH 6.88, with water resonance set at 4.766 ppm with respect to DSS.

[c]Intercept at $1/T = 0$, from the Curie plot presented in Figure 8.

[d]Measured by nonselective inversion recovery in 95% ^1H_2O/5% ^2H_2O, at 298 K and pH 6.88 on a 300 μM sample in 30 mM Tris. The error is 10%.

The same side-chain methyl was in contact with heme methyl *j*, whereas the second γ-methyl (−1.31 ppm, *u*) interacted with an upfield-shifted phenylalanine ring at 5.15, 5.74, and 6.00 ppm. These and other contacts between the heme moiety and the protein matrix demonstrated that the heme was bound in a pocket lined with several hydrophobic side chains.

The properties of the paramagnetic form were characterized further with the temperature dependence of the chemical shift (related to the electronic structure of the heme) and the nonselective spin-lattice relaxation times (related to the distance to the iron). The observed chemical shift of a resonance contains a diamagnetic contribution (δ_{dia}) and a hyperfine contribution made of a contact term and a dipolar term ($\delta_{hf} = \delta_{con} + \delta_{dip}$). In simple compounds with a unique and isolated spin state, δ_{hf} vanishes at infinite temperature, decreasing linearly with $1/T$ as T is raised (Jesson, 1973). A Curie plot, or chemical shift vs. reciprocal absolute temperature (Jesson, 1973), is presented in Figure 8 for selected resonances of S6803 metHb-R. The majority of the paramagnetically shifted signals moved toward diamagnetic values as the temperature was raised from 10 to 31 °C. The behavior was linear, leading to the intercepts at infinite temperature listed in Table 2. The diamagnetic shift of heme methyl groups ranges between 2.5 and 4 ppm (La Mar et al., 2000). Heme methyl *d* yielded an intercept at the edge of this range, whereas heme methyls *a* and *l* extrapolated to ~6 ppm in a slight hypo-Curie response. Heme methyl *j* had near anti-Curie behavior, remaining practically at the same shift as T was raised. Nonideal response of this order has been observed with other low-spin iron(III) hemes (Nguyen et al., 1999). The mean heme methyl shift of 14.2 ppm at 298 K was comparable to that in *G. dibranchiata* cyanometHb (Alam & Satterlee, 1994) and in model compounds (La Mar, 1979).

The nonselective spin-lattice relaxation rate of a proton *i* under the influence of the unpaired spin(s) also contains a contact contribution and a dipolar contribution. The dipolar effects depend on the distance to the iron according to $(R_{i-Fe})^{-6}$ and can be used for structural purposes. The relaxation times of a few resolved resonances are listed in Table 2. The values were moderately short and again comparable to those in *G. dibranchiata* cyanometHb (Alam & Satterlee, 1994). Of note was signal *z* (Fig. 5), which had a T_1 of 3.7 ms. When this T_1 value was compared to that of a heme methyl experiencing a small hyperfine shift and therefore a small scalar contribution ($T_1 \sim 200$ ms), proton *z* appeared located at ~3.2 Å from the iron atom. This distance, as well as the chemical shift and the broad linewidth, suggested assignment to a proton belonging to one of the axial residues, possibly one of the CH protons of the proximal histidine (La Mar et al., 1981).

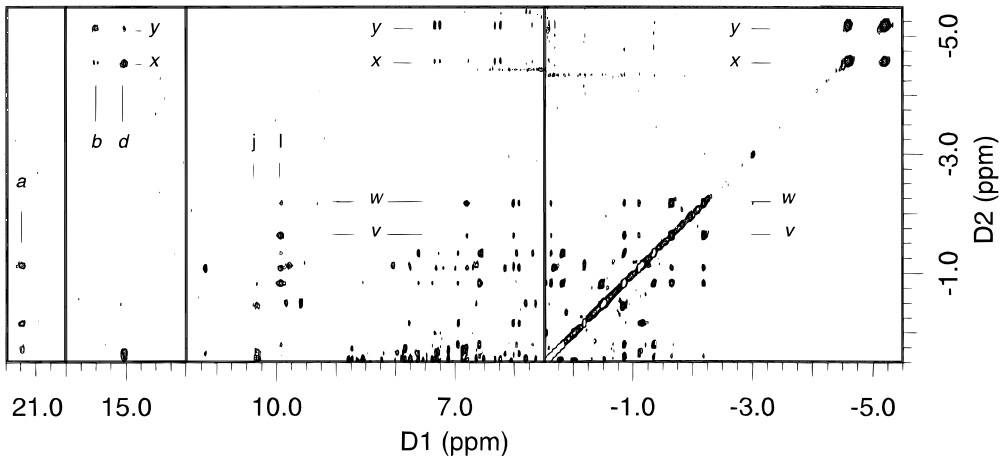

Fig. 7. Portion of the 600 MHz ^{1}H NMR NOESY spectrum of recombinant S6803 metHb-R. The sample was as described in Figure 5. The mixing time was 100 ms. The letters follow the nomenclature of Figure 5 and Table 2. Cross peaks between the heme methyl at 15.07 ppm (*d*) and the β protons of vinyl V1 (*x* and *y*) are marked. A similar set of peaks for the resonance at 9.98 ppm (*l*) and β protons of vinyl V2 (*v* and *w*) provided a tentative assignment of a second heme methyl. Selected NOEs between the heme group and the protein are discussed in the text.

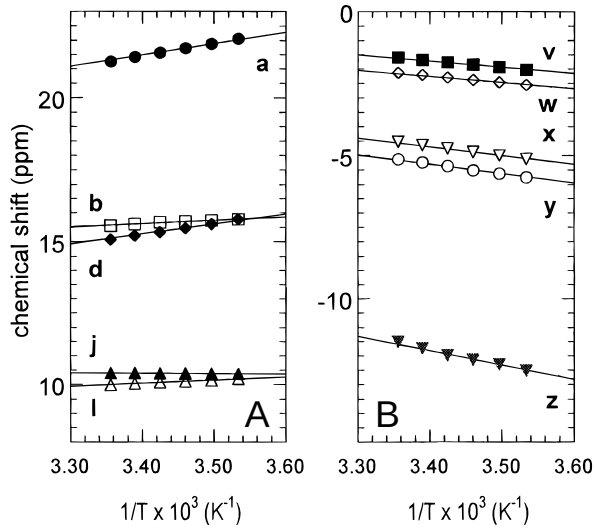

Fig. 8. Curie plot for selected resonances observed in the 600 MHz ^{1}H NMR spectrum of recombinant S6803 metHb-R. **A:** Low field resonances. **B:** High field resonances. The protein (600 μM heme) was in 95% H_2O/5% ^2H_2O, buffered at pH 6.88 with Tris/EDTA. The letters indicate hyperfine-shifted signals as labeled in Figure 5. Over this narrow temperature range (10–31 °C), only the heme methyl signal at 10.37 ppm (*j*) exhibited non-Curie response with a practically invariant chemical shift.

Discussion

The oxidized state of recombinant and hemin-reconstituted *Synechocystis* sp. PCC 6803 hemoglobin was found to have an electronic absorption spectrum and a proton NMR spectrum characteristic of a low-spin, hexacoordinated *b* hemoprotein (Antonini & Brunori, 1971; Wood, 1984). A low-spin ferric coordination scheme is observed at basic pH in the related *N. commune* GlbN (Thorsteinsson et al., 1996), mature *C. eugametos* LI637 hemoglobin (Couture et al., 1999a), and *Paramecium caudatum* hemoglobin generated by autoxidation of the oxy form (Tsubamoto et al., 1990). The absorption spectrum of these hemoglobins depends strongly on pH and reflects a reduction in the number of ligands at acidic values. The S6803 rHb-R hemichrome did not show a pronounced pH response between pH 8 and pH 5, and the endogenous hexacoordination scheme of this hemoprotein apparently persisted through a broader range of conditions than in its closest relatives.

The identity of the side chains coordinating the iron atom in the cyanoglobin group of proteins is a matter for speculation. Sequence alignments (Fig. 1) and comparison of vertebrate and nonvertebrate globin templates (Moens et al., 1996) predict the A helix to be truncated by several residues and the D helix to be missing. The alignments identify with confidence the conserved His at position 70 in the *Synechocystis* sp. PCC 6803 hemoglobin sequence as the proximal residue. However, a sixth ligand is not readily proposed on the basis of the primary structures and the available hemichrome stability data. A priori, functional groups provided by the amino terminus, Cys, His, Lys, and Met are capable of binding to the iron to form low-spin species, and it is likely that different proteins recruit different side chains.

The possibility of alternative binding schemes is reinforced by experiments that Couture and coworkers performed on *C. eugametos* LI637 Hb. They subjected the protein to amino acid replacements at positions B10 (Tyr → Leu), E4 (conserved Met → Ala), and E10 (Lys → Ala) (Couture et al., 1999a). These residues are in contact with the distal side of the heme in the myoglobin model and are plausible candidates for heme coordination. None of the individual replacements prevents the formation of a six-coordinate species at basic pH. Recent resonance Raman and EPR spectroscopy data point to a novel Tyr B10/Lys E10 ligation scheme, where the lysine serves to stabilize the tyrosinate ion, thereby favoring a low-spin state (Das et al., 1999a). When Tyr B10 is replaced by a leucine, Lys E10 becomes the sixth ligand. The proteins of the cyanoglobin group therefore appear prone to undergo the conformational distortions necessary to achieve various patterns of endogenous hexacoordination and may serve to illustrate heme ligand-switching reactions akin to those observed during the folding of cytochrome *c* (Yeh et al., 1997; Telford et al., 1999).

The stability of the S6803 rHb-R hemichrome is interesting in view of results obtained with sperm whale myoglobin with a histidine introduced at position E11 (Qin et al., 1994). In this variant, the temperature dependence of the chemical shifts and the nonselective spin-lattice relaxation times clearly indicate a thermal equilibrium between a low-spin state and a high-spin state. This property arises from coordination of His E11, but with nonideal geometry allowing only a weak bond. S6803 metHb displayed no such mixture at room temperature, a mark of unstrained bonding. S6803 Hb contains a tyrosine at B10 and a histidine at E10; this pair of residues may confer properties similar to those reported in *C. eugametos* or in the hemophore from *Serratia marcescens*, a nonglobin hemoprotein that appears to modulate its affinity for the heme through interactions between a coordinated tyrosinate and a nearby histidine group (Arnoux et al., 1999).

N. commune GlbN in the ferric state at pH 7 exhibits an enhanced tendency to lose its heme group compared to sperm whale myoglobin (Thorsteinsson et al., 1999). Heme loss in myoglobin is dependent on the presence of the D helix (Whitaker et al., 1995) as well as the polarity and size of the heme cavity (Hargrove et al., 1996). The instability of ferric *N. commune* GlbN is attributed to these factors and, in particular, a glutamine at position FG5 (Thorsteinsson et al., 1999), where myoglobin typically has an isoleucine. The qualitative data on S6803 rHb-R heme retention suggest a heme release slower than from *N. commune* GlbN. This difference between the two proteins may be caused by the distinct length and composition of the FG loop (Fig. 1).

The preparation of the recombinant *Synechocystis* sp. PCC 6803 hemoglobin included a reconstitution step by which hemin was added to the overexpressed apoprotein. This reconstitution procedure may generate a reversible hemichrome even if the native protein does not adopt this coordination scheme (Rachmilewitz et al., 1971). In the highly evolved oxygen delivery proteins, hemichrome formation interferes with normal function (Rachmilewitz et al., 1971; Rifkind et al., 1994; Arnold et al., 1999). However, hemochromes (reduced hemichrome, retaining endogenous hexacoordination) may be nondetrimental or even desirable in other species. Barley (*Hordeum vulgare*) hemoglobin (Duff et al., 1997; Das et al., 1999b) and rice (*Oryza sativa* L.) hemoglobin Hb1 (Arredondo-Peter et al., 1997) have a high oxygen affinity due to stabilization of the bound oxygen by distal interactions. In these proteins, the ability to make the additional contacts that slow down oxygen release is accompanied with a pronounced propensity to form hemi- and hemochrome species through the distal histidine in the absence of exogenous ligand. The slow oxygen release indi-

cates that the function of these hemoglobins is most certainly not that of carrying oxygen or facilitating its diffusion. Thus, hemichrome formation may signify nontraditional roles, for example, oxygen scavenging or electron transfer. The genome of *Synechocystis* sp. PCC 6803 contains three sets of genes for terminal respiratory oxidases, including a functional cytochrome *bd* quinol oxidase (*Cyd*) localized mostly on the cytoplasmic membrane (Howitt & Vermaas, 1998). A possible function for *Synechocystis* sp. PCC 6803 hemoglobin in vivo could be that of presenting O_2 to this high affinity oxidase.

The present study determined that the slr2097 (*glbN*) gene in *Synechocystis* sp. PCC 6803 encodes a polypeptide sequence capable of folding, specific heme binding, and ligand binding to the heme. The potential for this hemoglobin to have an active role in *Synechocystis* sp. PCC 6803 is underscored by the observations that slr2097, the sole globin gene present in the organism, has undergone some selective pressure to maintain these features in its product over time. Experiments are underway to explore both physiological and functional properties of *Synechocystis* sp. PCC 6803 Hb. The favorable thermodynamic and spectral characteristics of the recombinant material will allow the detailed description of this ancestral hemoprotein.

Materials and methods

Materials

All chemicals were purchased from Sigma (St. Louis, Missouri) and enzymes from Promega Corporation (Madison, Wisconsin), except where otherwise indicated.

Amplification of glbN gene and construction of expression vector

Purified *Synechocystis* sp. PCC 6803 chromosomal DNA was provided by Dr. D.A. Bryant, Department of Biochemistry and Molecular Biology, The Pennsylvania State University. The slr2097 (*glbN*) gene was amplified and terminal restriction enzyme recognition sites added by the polymerase chain reaction (PCR) using Ready-To-Go PCR Beads (Pharmacia Biotech, Uppsala, Sweden). The reaction yielded the expected 395 bp product by comparison to a 0.07–12.2 kbp ladder (Boehringer-Mannheim, Mannheim, Germany). Only trace low molecular weight secondary products were observed. The major PCR product was separated from the reaction mixture by 1% agarose gel electrophoresis and purified with the JETsorb Gel Extraction Kit (PGC Scientifics, Gaithersburg, Maryland). PCR product and pET3c vector DNA (Novagen Inc., Milwaukee, Wisconsin) were doubly digested with endonucleases *Nde*I and *Bam*HI in separate reactions. Vector DNA was 5′ dephosphorylated with calf intestinal alkaline phosphatase. Digestion products were gel-purified as above. The *glbN* insert and pET3c vector (threefold molar excess of fragment) were ligated using T4 DNA ligase in duplicate, and these mixtures were used to transform competent DH5-α *E. coli* cells.

E. coli transformation and purification of expression vector

Competent DH5-α cells were prepared and transformed by a slightly modified procedure of the calcium chloride method of Sambrook et al. (1989). Here, 50 mM $CaCl_2$, 10 mM Tris-Cl, pH 8 was the transformation buffer and cells were incubated with 10 μL of the above ligation reactions for 1.5 h before a heat shock period of 2 min. LB medium was exclusively employed throughout the procedure. Cells were screened for presence of the construct on agar plates supplemented with 50 μg/mL ampicillin (Eastman Kodak Company, Rochester, New York). Select colonies were grown overnight in liquid LB/Amp medium, and purified plasmid DNA was prepared from cells with the QIAprep Spin Miniprep Kit (Qiagen Inc., Valencia, California). The correct sequence and incorporation of *glbN* between the *Nde*I and *Bam*HI recognition sequences downstream of the T7 promoter region of pET3c were verified by automated sequencing performed by the Nucleic Acid Facility of the Penn State Biotechnology Institute.

BL21(DE3) transformation and expression of glbN

E. coli BL21(DE3) host strain was chosen for its high-level inducible expression of recombinant proteins. The cells were made competent by a procedure developed by Simanis (Hanahan, 1985) with the following modifications: cells were initially cultured in TYM medium (2% tryptone and 0.5% (w/v) yeast extract (Difco Laboratories, Detroit, Michigan), 100 mM NaCl, 10 mM $MgSO_4$); potassium chloride was used in lieu of rubidium chloride; competent cells and plasmid DNA were incubated on ice for 90 min; cells were heat-shocked for only 45 s; and LB was used as recovery medium. Transformed cells were spread on M9 agar plates supplemented to final concentrations of 0.2% glucose, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$ and 50 μg/mL ampicillin (M9+). Single transformed colonies were used to inoculate 60 mL M9+ and cultured at 37 °C for 16 h. Ten milliliters of the overnight cultures were transferred into 500 mL fresh M9+ in 2 L baffled flasks. This culture was incubated at 37 °C with vigorous shaking until OD_{600} reached $\cong$0.9, at which point isopropyl β-thiogalactoside (IPTG) was added to a final concentration of 0.5 mM. Protein production continued for 5 h after which faintly pink cells were separated from the growth media by centrifugation and stored at −20 °C.

Protein extraction and purification

Cell pellets were thawed and resuspended in 20 mL 50 mM Tris, 1 mM EDTA pH 8.0 per 5 g cells. Phenylmethylsulfonyl fluoride (PMSF; 50 mM in 100% ethanol) was added to a concentration of 1 mM, and the ice-cold suspension was sonicated for six 10 s bursts at 70% power (Fisher Scientific Model 60 Sonic Dismembrator) to release soluble globin. Cellular debris was removed by centrifugation and the above procedure repeated twice. HoloHb was reconstituted by adding a saturated solution of bovine hemin chloride in 0.1 M NaOH to the pooled supernatants stirring in the dark at 4 °C. Free heme was precipitated by lowering the reconstituted protein solution pH to ~6 followed by centrifugation at 48,000 × *g* for 45 min. This crude mixture was passed through a 0.45 μm filter and applied to a DEAE Sephacel (Pharmacia, Uppsala, Sweden) anion exchange column (2.5 × 5 cm) equilibrated with the above buffer. A linear sodium chloride gradient (0 to 0.5 M) in buffer was used to elute the bound protein. Colored fractions were analyzed by SDS-PAGE and samples free of contaminating proteins less than twice the molecular weight of Hb were pooled and concentrated to a volume of ~12 mL. This solution was passed through a 2.5 × 96 cm Sephadex G-50 Fine (Sigma) size exclusion column. Fractions were analyzed by SDS-PAGE and those essentially free of impurity were pooled and concen-

trated to yield a stock solution of metHb at pH ≅ 7 for further analysis. SDS-PAGE gels of the purified protein were stained with Coomassie blue (Fig. 2) and silver nitrate (Silver Stain Plus; Bio-Rad Laboratories, Hercules, California).

Mass spectrometry analysis

Mass spectrometry and tryptic mapping data were collected at the Penn State Intercollegiate Center for Mass Spectrometry on a sample of purified metHb. All matrix-assisted laser desorption ionization (MALDI) mass spectra were generated on a Voyager-DE STR mass spectrometer (Perseptive Biosystems, Framingham, Massachusetts) using 337 nm light from a nitrogen laser for desorption. Samples were desalted using ZipTips-C18 (Millipore Corporation, Bedford, Massachusetts) and mixed with a solution of alpha-cyano-4-hydroxycinnamic acid matrix (in acetonitrile/water 50:50 v/v plus 0.3% trifluoroacetic acid). The spectrum of the whole protein was generated in linear mode.

Tryptic mapping was performed by dissolving the protein in 100 mM ammonium bicarbonate containing 10 mM $CaCl_2$ and incubating with trypsin (Sigma) at a trypsin:protein ratio of approximately 1:50. Incubations were carried out at 37 °C for 48 h. Digestion products were desalted using ZipTips-C18 as above, and MALDI spectra were acquired in reflector mode. This procedure did not provide information on fragments with masses lower than 750 Da.

Determination of extinction coefficient

The millimolar extinction coefficient of the metHb Soret band was determined according to the hemochromogen method of de Duve (1948). At minimum, quadruplicate readings of reduced pyridine-heme were averaged to determine heme concentration using ϵ_{557} = 32.0 mM^{-1} cm^{-1} and ϵ_{525} = 16.0 mM^{-1} cm^{-1}, and duplicate absorbance readings of the original metHb solution at 410 nm were averaged to determine the extinction coefficient at that wavelength. The hemochromogen assay was performed with two separate protein samples, and ϵ_{410} is reported as an average of the values obtained from each. All S6803 Hb concentrations reported in this work are expressed on a heme basis.

CN^-, O_2, and CO binding

The ability of S6803 rHb-R to bind various ligands was tested at a heme concentration of ~8.5 μM in 30 mM Tris, 0.6 mM EDTA, pH 7.4 or as noted. In the case of CN^-, a small grain of solid potassium cyanide was added directly to the oxidized sample and spectra recorded over a period of several hours. Cyanide binding was also monitored at pH 4.4. In the case of O_2, the solution was first exposed to nitrogen bubbling for 5 min, then a small amount of solid sodium dithionite was added to convert all metHb to the reduced form. An excess of dithionite was immediately confirmed by UV absorbance at 390 nm. Filtered air was bubbled through the solution for 1 min, after which the spectrum was recorded. Air was bubbled through the sample for an additional minute and the spectrum was again recorded. No residual dithionite was left in solution. To produce the carbon monoxide adduct, pure CO gas was passed through the reduced solution for 10 min prior to recording the spectrum. In a second procedure, the metHb solution was saturated with CO before the reduction step.

Heme release assay

The ability of metHb-R to release the heme prosthetic group was verified by competition with equine apomyoglobin. Briefly, apomyoglobin prepared by the butanone method of Teale (1959) and S6803 metHb-R were mixed in 30 mM Tris, 0.6 mM EDTA, pH 7.4 with an approximate ratio of 3:1. The solution was monitored by visible spectroscopy for the appearance of a charge transfer band at 635 nm to indicate the formation of metaquomyoglobin (Antonini & Brunori, 1971). Readings were taken upon mixing and at 10 to 15 min intervals for the first hour, 30 min intervals for the next hour, and at various intervals over the next two days.

Optical spectroscopy

All electronic absorption spectra were collected on an Aviv model 14 DS spectrophotometer at 298 K. Raw data were corrected for background by subtracting a matched spectrum of 18 MΩ cm water collected on the same day in the same cuvette pair.

CD measurements

S6803 metHb was diluted to a concentration of 15.6 μM and duplicate sets of data collected from 192–300 nm on an Aviv Model 62 DS Circular Dichroism spectrometer. Raw data were corrected by subtracting values collected for buffer over the same interval. The molar residual ellipticity (MRE in mdeg cm^2/dmol) was calculated at each wavelength according to the equation:

$$\mathrm{MRE} = \frac{\Theta \times 100 \times M_r}{c \times d \times N_{AA}}$$

where Θ is rotation of circularly polarized light in degrees, M_r is the molar mass of S6803 rHb-R, c is its concentration in mg/mL, d is the pathlength in centimeters, and N_{AA} is the number of amino acids. Values of M_r and N_{AA} did not include the initial methionine as per mass spectroscopic data. An estimate of the extent of secondary structure was evaluated with the neural net program *k2d* (Andrade et al., 1993; Merelo et al., 1994).

NMR spectroscopy

^{1}H NMR spectra were collected at 600 MHz on a Bruker DRX spectrometer. The samples (heme concentration ~600 μM) were prepared in 95% ^1H_2O/5% ^2H_2O and contained Tris (~20 mM) and EDTA (~0.4 mM). The temperature of the probe was 298 K or set at values between 283 and 304 K. The water signal was suppressed by low-power presaturation. NOESY (Kumar et al., 1980), DQF-COSY (Rance et al., 1983), MCOSY (Aue et al., 1976), and TOCSY (Braunschweiler & Ernst, 1983) data were collected on metHb with TPPI quadrature detection in the indirect dimension (Drobny et al., 1979). For each of these, 512 increments were used and the recycling time was 1 s. Other parameters were: NOESY, mixing time of 100 and 50 ms; TOCSY, mixing time of 45 ms with DIPSI-2 scheme (Shaka et al., 1988) and a power of 8.4 kHz. Spectral width in the direct (D1) dimension was 24,038 Hz collected over 2,048 complex points. Spectral width in the indirect dimension (D2) was either 24,038 or 13,514 Hz. Felix (Molecular Simulations Inc., San Diego, California) was used for Fourier transformation of the data. Squared sine-bell windows shifted by 45° (direct dimension) and 60° (indirect dimension) were applied to the

NOESY data before transformation into a 2,048 × 2,048 real matrix. The MCOSY data were redundant with the DQF-COSY and TOCSY data, an indication that cross correlation effects between ^{1}H–^{1}H dipolar and ^{1}H–Fe Curie relaxation (Qin et al., 1993) are not interfering in this small low-spin protein. NOESY mixing times were set at 100 ms to emphasize the diamagnetic connectivities and at 50 ms to verify the specificity of the effects involving the protons influenced by paramagnetism. None of the effects discussed here appeared relayed. Complete analysis including mixing time dependence as necessary in paramagnetic systems (La Mar et al., 2000) is in progress.

The temperature dependence of the chemical shift was studied between 283 and 304 K by collecting data every 3 K. The probe temperature was calibrated by using either ethylene glycol or methanol and the equations given in Cavanagh et al. (1996). Chemical shifts were referenced indirectly to DSS through the shift of the water line whose temperature dependence is −11.9 ppb/K (Wishart et al., 1995).

Nonselective T_1 relaxation times were obtained with an inversion recovery sequence, presaturation of the water line, and a recycle time of 3.5 s on a 300 μM heme sample. A minimum of 18 data points was collected throughout the recovery. The data were analyzed with a three-parameter fit accounting for incomplete inversion of the signal at zero recovery time. If the relaxation of proton i is dominated by the paramagnetic dipolar contribution, $(T_{1,i})^{-1}$ is proportional to $(R_{i-\mathrm{Fe}})^{-6}$, where $R_{i-\mathrm{Fe}}$ is the distance to the iron. An estimate of $R_{i-\mathrm{Fe}}$ can be obtained by comparing the T_1 value of the proton to that of a methyl group endowed with a small scalar contribution to the relaxation. A heme methyl group (Me) with hyperfine shift <15 ppm provides an acceptable reference. The distance to the iron is $R_{\mathrm{Me\text{-}Fe}} \sim 6.2$ Å and $R_{i-\mathrm{Fe}}$ in Å is obtained through $(T_{1,i}/T_{1,\mathrm{Me}})^{1/6} = R_{i-\mathrm{Fe}}/6.2$.

Acknowledgments

The authors are particularly grateful for Dr. D.A. Bryant's help and advice throughout the project and for the gift of purified *Synechocystis* sp. PCC 6803 chromosomal DNA. Dr. C.J. Falzone is thanked for critical reading of the manuscript, Dr. A.D. Jones for the collection of mass spectrometry data, and Dr. J.D. Satterlee for his guidance. This work was supported by National Institutes of Health Grant GM-54217. The Voyager mass spectrometer (Penn State Intercollegiate Center for Mass Spectrometry) was purchased in part using NIH funds (Grant S10 RR11318); the Bruker DRX-600 NMR spectrometer was purchased in part using NIH funds (Grant S10 RR10524).

References

Alam SL, Satterlee JD. 1994. Complete heme proton hyperfine resonance assignments of the *Glycera dibranchiata* component IV metcyano monomer hemoglobin. *Biochemistry 33*:4008–4018.

Andrade MA, Chacon P, Merelo JJ, Moran F. 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng 6*:383–390.

Angeloni SV, Potts M. 1994. Analysis of the sequences within and flanking the cyanoglobin-encoding gene, *glbN*, of the cyanobacterium *Nostoc commune* UTEX 584. *Gene 146*:133–134.

Antonini E, Brunori M. 1971. *Hemoglobin and myoglobin in their reactions with ligands.* Amsterdam: North-Holland.

Arnold EV, Bohle DS, Jordan PA. 1999. Reversible and irreversible hemichrome generation by the oxygenation of nitrosylmyoglobin. *Biochemistry 38*:4750–4756.

Arnoux P, Haser R, Izadi N, Lecroisey A, Delepierre M, Wandersman C, Czjzek M. 1999. The crystal structure of HasA, a hemophore secreted by *Serratia marcescens*. *Nat Struct Biol 6*:516–520.

Arredondo-Peter R, Hargrove MS, Sarath G, Moran JF, Lohrman J, Olson JS, Klucas RV. 1997. Rice hemoglobins. Gene cloning, analysis, and O_2-binding kinetics of a recombinant protein synthesized in *Escherichia coli*. *Plant Physiol 115*:1259–1266.

Aue WP, Bartholdi E, Ernst RR. 1976. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J Chem Phys 64*:2229–2246.

Becker M, Schafer G. 1991. Purification and spectral characterization of a *b*-type cytochrome from the plasma membrane of the archaebacterium *Sulfolobus acidocaldarius*. *FEBS Lett 291*:331–335.

Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson 53*:521–528.

Cavanagh J, Fairbrother WJ, Palmer AGI, Skelton NJ. 1996. *Protein NMR spectroscopy. Principles and practice*. San Diego: Academic Press.

Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III, et al. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature 393*:537–544.

Couture M, Chamberland H, St-Pierre B, Lafontaine J, Guertin M. 1994. Nuclear genes encoding chloroplast hemoglobins in the unicellular green alga *Chlamydomonas eugametos*. *Mol Gen Genet 243*:185–197.

Couture M, Das TK, Lee HC, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB, Guertin M. 1999a. *Chlamydomonas* chloroplast ferrous hemoglobin. Heme pocket structure and reactions with ligands. *J Biol Chem 274*:6898–6910.

Couture M, Guertin M. 1996. Purification and spectroscopic characterization of a recombinant chloroplastic hemoglobin from the green unicellular alga *Chlamydomonas eugametos*. *Eur J Biochem 242*:779–787.

Couture M, Yeh SR, Wittenberg BA, Wittenberg JB, Ouellet Y, Rousseau DL, Guertin M. 1999b. A cooperative oxygen-binding hemoglobin from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA 96*:11223–11228.

Das TK, Couture M, Lee HC, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB, Guertin M. 1999a. Identification of the ligands to the ferric heme of *Chlamydomonas* chloroplast hemoglobin: Evidence for ligation of tyrosine-63 (B10) to the heme. *Biochemistry 38*:15360–15368.

Das TK, Lee HC, Duff SM, Hill RD, Peisach J, Rousseau DL, Wittenberg BA, Wittenberg JB. 1999b. The heme environment in barley hemoglobin. *J Biol Chem 274*:4207–4212.

de Duve C. 1948. A spectrophotometric method for the simultaneous determination of myoglobin and hemoglobin in extracts of human muscle. *Acta Chem Scan 2*:264–289.

Drobny G, Pines A, Sinton S, Weitekamp DP, Wemmer D. 1979. Fourier transform multiple quantum nuclear magnetic resonance. *Symp Faraday Soc 13*:49–55.

Duff SM, Wittenberg JB, Hill RD. 1997. Expression, purification, and properties of recombinant barley (*Hordeum* sp.) hemoglobin. Optical spectra and reactions with gaseous ligands. *J Biol Chem 272*:16746–16752.

Emerson SD, La Mar GN. 1990. NMR determination of the orientation of the magnetic susceptibility tensor of cyano metmyoglobin: A new probe of steric tilt of bound ligand. *Biochemistry 29*:1556–1566.

Hanahan D. 1985. Techniques for transformation of *E. coli*. In: Glover DM, ed. *DNA cloning, A practical approach.* Washington DC: IRL Press. pp 109–135.

Hanania GI, Yeghiayan A, Cameron BF. 1966. Absorption spectra of sperm-whale ferrimyoglobin. *Biochem J 98*:189–192.

Hardison R. 1998. Hemoglobins from bacteria to man: Evolution of different patterns of gene expression. *J Exp Biol 201*:1099–1117.

Hargrove MS, Wilkinson AJ, Olson JS. 1996. Structural factors governing hemin dissociation from metmyoglobin. *Biochemistry 35*:11300–11309.

Hill DR, Belbin TJ, Thorsteinsson MV, Bassam D, Brass S, Ernst A, Boger P, Paerl H, Mulligan ME, Potts M. 1996. GlbN (cyanoglobin) is a peripheral membrane protein that is restricted to certain *Nostoc* spp. *J Bacteriol 178*:6587–6598.

Howitt CA, Vermaas WF. 1998. Quinol and cytochrome oxidases in the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochemistry 37*:17944–17951.

Iwaasa H, Takagi T, Shikama K. 1989. Protozoan myoglobin from *Paramecium caudatum*. Its unusual amino acid sequence. *J Mol Biol 208*:355–358.

Iwaasa H, Takagi T, Shikama K. 1990. Protozoan hemoglobin from *Tetrahymena pyriformis*. Isolation, characterization, and amino acid sequence. *J Biol Chem 265*:8603–8609.

Jesson JP. 1973. The paramagnetic shift. In: La Mar GN, Horrocks WD, Holm RH, eds. *NMR of paramagnetic molecules*. New York: Academic Press. pp 1–52.

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, et al. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res 3*:109–136.

Keller RM, Wüthrich K. 1980. Structural study of the heme crevice in cytochrome b_5 based on individual assignments of the ^{1}H-NMR lines of the heme group and selected amino acid residues. *Biochim Biophys Acta 621*:204–217.

Kumar A, Ernst RR, Wüthrich K. 1980. A 2D NOE experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1–6.

Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessieres P, Bolotin A, Borchert S, et al. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature 390*: 249–256.

La Mar GN. 1979. Model compounds as aids in interpreting NMR spectra of hemoproteins. In: Schulman RG, ed. *Biological application of magnetic resonance*. New York: Academic Press. pp 305–343.

La Mar GN, Burns PD, Jackson JT, Smith KM, Langry KC, Strittmatter P. 1981. Proton magnetic resonance determination of the relative heme orientations in disordered native and reconstituted ferricytochrome b_5. Assignment of heme resonances by deuterium labeling. *J Biol Chem 256*:6075–6079.

La Mar GN, Davis NL, Parish DW, Smith KM. 1983. Heme orientational disorder in reconstituted and native sperm whale myoglobin. Proton nuclear magnetic resonance characterizations by heme methyl deuterium labeling in the met-cyano protein. *J Mol Biol 168*:887–896.

La Mar GN, Satterlee JD, de Ropp JS. 2000. Nuclear magnetic resonance of hemoproteins. In: Kadish K, Smith K, Guilard R, eds. *The porphyrins handbook*. Burlington, MA: Academic Press. pp 185–298.

Li WH. 1984. Retention of cryptic genes in microbial populations. *Mol Biol Evol 1*:213–219.

Merelo JJ, Andrade MA, Prieto A, Morán F. 1994. Proteinotopic feature maps. *Neurocomputing 6*:443–454.

Moens L, Vanfleteren J, Van de Peer Y, Peeters K, Kapp O, Czeluzniak J, Goodman M, Blaxter M, Vinogradov S. 1996. Globins in nonvertebrate species: Dispersal by horizontal gene transfer and evolution of the structure-function relationships. *Mol Biol Evol 13*:324–333.

Nguyen BD, Xia Z, Yeh DC, Vyas K, Deaguero H, La Mar GN. 1999. Solution NMR determination of the anisotropy and orientation of the paramagnetic susceptibility tensor as a function of temperature for metmyoglobin cyanide: Implications for the population of excited electronic states. *J Am Chem Soc 121*:208–217.

Potts M, Angeloni SV, Ebel RE, Bassam D. 1992. Myoglobin in a cyanobacterium. *Science 256*:1690–1691.

Qin J, Delaglio F, La Mar GN, Bax A. 1993. Distinguishing the effects of cross correlation and J coupling in COSY spectra of paramagnetic proteins. *J Magn Reson (Ser B) 102*:332–336.

Qin J, La Mar GN, Dou Y, Admiraal SJ, Ikeda-Saito M. 1994. ^{1}H NMR study of the solution molecular and electronic structure of engineered distal myoglobin His64(E7) Val/Val68(E11) His double mutant. Coordination of His64(E11) at the sixth position in both low-spin and high-spin states. *J Biol Chem 269*:1083–1090.

Rachmilewitz EA, Peisach J, Blumberg WE. 1971. Studies on the stability of oxyhemoglobin A and its constituent chains and their derivatives. *J Biol Chem 246*:3356–3366.

Rance M, Sørensen OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K. 1983. Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering. *Biochem Biophys Res Commun 117*:479–485.

Rifkind JM, Abugo O, Levy A, Heim J. 1994. Detection, formation, and relevance of hemichromes and hemochromes. *Methods Enzymol 231*:449–480.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning. A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schopf JW, Walter MR. 1982. Origin and early evolution of cyanobacteria: The geological evidence. In: Carr NG, Whitton BA, eds. *The biology of cyanobacteria*. Oxford, UK: Blackwell Scientific. pp 543–564.

Shaka AJ, Lee C, Pines A. 1988. Iterative schemes for bilinear operators; application to spin decoupling. *J Magn Reson 77*:274–293.

Shikama K, Matsuoka A, Iwaasa H. 1995. The unique structures of protozoan myoglobin and yeast hemoglobin: An evolutionary diversity. *Int J Biochem Cell Biol 27*:1107–1115.

Strittmatter P. 1960. The nature of the heme binding in microsomal cytochrome b_5. *J Biol Chem 235*:2492–2497.

Teale FWJ. 1959. Cleavage of heme-protein link by acid methylethylketone. *Biochim Biophys Acta 35*:543.

Telford JR, Tezcan FA, Gray HB, Winkler JR. 1999. Role of ligand substitution in ferrocytochrome *c* folding. *Biochemistry 38*:1944–1949.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res 22*:4673–4680.

Thorsteinsson MV, Bevan DR, Ebel RE, Weber RE, Potts M. 1996. Spectroscopical and functional characterization of the hemoglobin of *Nostoc commune* UTEX 584 (Cyanobacteria). *Biochim Biophys Acta 1292*:133–139.

Thorsteinsson MV, Bevan DR, Potts M, Dou Y, Eich RF, Hargrove MS, Gibson QH, Olson JS. 1999. A cyanobacterial hemoglobin with unusual ligand binding kinetics and stability properties. *Biochemistry 38*:2117–2126.

Tsubamoto Y, Matsuoka A, Yusa K, Shikama K. 1990. Protozoan myoglobin from *Paramecium caudatum*. Its autoxidation reaction and hemichrome formation. *Eur J Biochem 193*:55–59.

Whitaker TL, Berry MB, Ho EL, Hargrove MS, Phillips GN Jr, Komiyama NH, Nagai K, Olson JS. 1995. The D-helix in myoglobin and in the beta subunit of hemoglobin is required for the retention of heme. *Biochemistry 34*:8221–8226.

Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. 1995. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J Biomol NMR 6*:135–140.

Wood PM. 1984. Bacterial proteins with CO-binding b- or c-type haem. Functions and absorption spectroscopy. *Biochim Biophys Acta 768*:293–317.

Wu JZ, La Mar GN, Yu LP, Lee KB, Walker FA, Chiu ML, Sligar SG. 1991. ^{1}H NMR study of the solution molecular and electronic structure of *Escherichia coli* ferricytochrome b_{562}: Evidence for S = 1/2 in equilibrium S = 5/2 spin equilibrium for intact His/Met ligation. *Biochemistry 30*:2156–2165.

Yamauchi K, Tada H, Usuki I. 1995. Structure and evolution of *Paramecium* hemoglobin genes. *Biochim Biophys Acta 1264*:53–62.

Yeh SR, Takahashi S, Fan B, Rousseau DL. 1997. Ligand exchange during cytochrome *c* folding. *Nat Struct Biol 4*:51–56.