Structural and Dynamic Repercussions of Heme Binding and Heme−Protein Cross-linking in *Synechococcus* sp. PCC 7002 Hemoglobin†

David A. Vuletich, Christopher J. Falzone, and Juliette T. J. Lecomte*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

Received July 28, 2006; Revised Manuscript Received September 21, 2006

ABSTRACT: The recombinant two-on-two hemoglobin from the cyanobacterium *Synechoccocus* sp. PCC 7002 (S7002 rHb) is a bishistidine hexacoordinate globin capable of forming a covalent cross-link between a heme vinyl and a histidine in the C-terminal helix (H helix). Of the two heme axial histidines, His46 (in the E helix, distal side) and His70 (in the F helix, proximal histidine), His46 is displaced by exogenous ligands. S7002 rHb can be readily prepared as an apoglobin (apo-rHb), a non-cross-linked hemichrome (ferric iron and histidine axial ligands, rHb-R), and a cross-linked hemichrome (rHb-A). To determine the effects of heme binding and subsequent cross-linking, apo-rHb, rHb-R, and rHb-A were subjected to thermal denaturation and $^{1}H/^{2}H$ exchange. Interpretation of the latter data was based on nuclear magnetic resonance assignments obtained with uniformly ^{15}N- and $^{13}C,^{15}N$-labeled proteins. Apo-rHb was found to contain a cooperative structural core, which was extended and stabilized by heme binding. Cross-linking resulted in further stabilization attributed mainly to an unfolded-state effect. Protection factors were higher at the cross-link site and near His70 in rHb-A than in rHb-R. In contrast, other regions became less resistant to exchange in rHb-A. These included portions of the B and E helices, which undergo large conformational changes upon exogenous ligand binding. Thus, the cross-link readjusted the dynamic properties of the heme pocket. $^{1}H/^{2}H$ exchange data also revealed that the B, G, and H helices formed a robust core regardless of the presence of the heme or cross-link. This motif likely encompasses the early folding nucleus of two-on-two globins.

Over the past several years, the number of sequenced globin genes has increased dramatically. With the new data, a comprehensive phylogenetic analysis has revealed three lineages of globins: one domain, two domain, and truncated or two-on-two (*1*). Many of these proteins are expected to provide insight into the functional roles of ancient globins, which were likely different from oxygen delivery and storage. Among the most promising in this regard are the two-on-two globins, for which genes are found in all three kingdoms of life. A majority of identified examples occur in pathogenic bacteria. In some cases, the products have been implicated in the defense of the organism from the NO liberated by the host immune system (*2*−*4*). Two-on-two globins are also found in cyanobacteria, and in instances of nitrogen-fixing species, a function as an oxygen scavenger has been proposed (*5*).

The two-on-two globin domain has a primary structure shorter than that of the canonical vertebrate globin by 20−40 residues, which led to the original name of "truncated hemoglobin" (*6*). Vertebrate globins contain up to eight helices (A−H). Sequence alignments show that the truncation in the two-on-two proteins occurs through a missing D helix and an abbreviated A helix (*7*, *8*). The first three-dimensional structures of representative members of the family revealed a two-on-two orthogonal bundle (*9*) reminiscent of the three-on-three topology of the vertebrate globins (*10*). In the bound state, the heme iron has one axial ligand from the protein (proximal histidine, in the F helix) and one exogenous ligand on the opposite side of the heme (distal side, defined with the E helix). All investigated examples of two-on-two globins contain one heme group tightly associated with the protein, and the affinity for O_2 and CO is high (*11*).

Two-on-two globins present noteworthy structural features. These include distal hydrogen-bond networks stabilizing the bound state of O_2 or cyanide (*12*, *13*), ligand tunnels facilitating exogenous ligand access to the heme group (*14*, *15*), bishistidine hexacoordination of the heme iron (*16*−*18*), and heme−protein covalent cross-linking (*19*, *20*). Sequence comparisons divide two-on-two globins in three orthologous groups (I or N, II or O, and III or P) (*8*, *21*) and suggest that distal networks exist throughout the two-on-two globin family. Access tunnels, in contrast, are conspicuous principally in group I proteins. The last two properties, which have been observed only in certain cyanobacterial proteins, highlight the chemical diversity of this ancient family.

Synechococcus sp. PCC 7002 (S7002)[1] and *Synechocystis* sp. PCC 6803 (S6803) are non-nitrogen-fixing cyanobacteria that contain group I two-on-two globins capable of cross-linking the heme group. This linkage occurs through the reaction of His117, a residue located in the H helix, and the α carbon of the heme 2-vinyl group. The cross-link increases the resistance of the protein to denaturation to both high temperature and low pH (*20*). As with cytochrome *c*, which

† This study was supported by National Science Foundation Grant MCB-0349409 and NASA Grant NNG04GN33H (to D.A.V.).

* To whom correspondence should be addressed. Telephone: (814) 863-1153. E-mail: jtl1@psu.edu.

10.1021/bi061532g CCC: $33.50
Published on Web 11/02/2006

forms two heme−protein thioether bridges via the vinyl groups and cysteines, the functional importance of the cross-link in cyanobacterial hemoglobins (Hbs) has not been established. Aside from simply preventing cofactor loss, a cross-link may protect a protein from deleterious modification, as proposed for certain mammalian peroxidases (*22*); it is also possible that cross-links serve to modulate exogenous ligand binding by adjusting the properties of the axial histidines, an effect worth investigating for its potentially broad implications in hemoprotein chemistry.

Three-dimensional structures of the non-cross-linked S6803 protein [recombinant heme-reconstituted hemoglobin (rHb-R) (*23*)] and cross-linked S6803 protein [recombinant hemoglobin with covalently linked heme (rHb-A) (*24*)] have been solved by solution and solid-state methods, respectively. A comparison of nuclear magnetic resonance (NMR) spectral properties supports that covalent heme attachment caused local structural perturbations (*20*). The minor differences noted between the two states, however, provide only partial insight into the consequences of the linkage. The sequence of S7002 recombinant hemoglobin (rHb) is 59% identical to that of S6803 rHb. Data collected thus far on holoproteins from both sources support similarities (*13*, *18*, *20*) that allow the structure of S6803 rHb to be used as a low-resolution model for S7002 rHb. Distinct properties are also apparent, in particular with respect to ligand binding (*13*), which warrants a separate inspection of S7002 rHb.

The goal of the present work was to investigate the consequences of heme binding and post-translational heme modification on individual structural elements in S7002 rHb, the globin for which we are pursuing in vivo functional characterization. We based our study on the premise that apo-Hb, Hb, or Hb with covalently linked heme (Hb-A) may play a role in the synechococcal cell. We used hydrogen−deuterium ($^1H/^2H$) exchange monitored by NMR spectroscopy (*25*). Under conditions of high protein stability, as chosen here, the exchange studies report on local unfolding events (*26*). This technique was well suited to examine the dynamics of specific regions of S7002 rHb. In addition, native-state exchange identified the most stable substructures of the two holoproteins. The comparison of slow-exchange cores was extended to the apoprotein. A common nucleus composed of elements of the B, G, and H helices suggested a relationship between the folding process of two-on-two globins and that of other members of the globin superfamily.

MATERIALS AND METHODS

Protein Preparation and Purification. Plasmid and unlabeled protein preparations were performed as previously described (*18*). In brief, unlabeled S7002 apo-rHb was purified from inclusion bodies and reconstituted with hemin if holoprotein was desired. Cross-linked samples were prepared by treatment with dithionite (*20*). To the freshly prepared rHb-A solution was added a 20-fold molar excess of deferoxamine mesylate to remove iron resulting from heme breakdown; the mixture was allowed to incubate overnight and was purified over a Sephadex G-25 column. Isotopically labeled proteins were also prepared as previously described (*20*), with minor protocol modification. For ^{15}N growths, the apoprotein was purified from the inclusion bodies, whereas a mixture of apo- and holoproteins was extracted from the soluble cell fraction and fully reconstituted with hemin. In $^{13}C,^{15}N$ growths, minimal protein was found in inclusion bodies and, therefore, extraction was limited to the soluble cell fraction. The apoprotein was separated from the holoprotein on a (diethylamino)ethyl (DEAE) anion-exchange column with 0.4 M NaCl gradient elution and further purified on a size-exclusion column. The yield of uniformly $^{13}C,^{15}N$-labeled preparations was 8 mg of protein for a 2 L growth, whereas that of uniformly ^{15}N-labeled preparations was 12−14 mg for the same volume.

Holoprotein extinction coefficients were determined with the hemochromogen assay (*27*) as previously described (*18*). All holoprotein concentrations are reported on a heme basis using extinction coefficients of 96 mM^{-1} cm^{-1} at 411 nm for rHb-R and 87 mM^{-1} cm^{-1} at 409 nm for rHb-A. The apoprotein extinction coefficient was estimated with ProtParam from the ExPASy molecular biology server (*28*). The value of 4.47 mM^{-1} cm^{-1} at 280 nm was used without adjustment.

NMR Spectroscopy. NMR samples were in 20 mM phosphate buffer at pH 7.1−7.7. All NMR data were collected at 25 °C on a 14.1 T Bruker DRX 600 spectrometer operating at a ^1H frequency of 600.05 or 600.18 MHz. ^1H chemical shifts were referenced to DSS through the (residual) water line set a 4.76 ppm. Indirect referencing was applied with TMS Ξ values to obtain ^{15}N and ^{13}C shifts (*29*). Two- and three-dimensional experiments were collected on unlabeled, uniformly ^{15}N-labeled, and uniformly $^{13}C,^{15}N$-labeled samples for the purpose of resonance assignments as described for S6803 rHb-R (*30*); acquisition and processing details are provided in Table S1 in the Supporting Information. The data were processed with NMRPipe (*31*). Cross-peak volumes were obtained with SPARKY (*32*) using the Gaussian fit routine.

$^1H/^2H$ Exchange Experiments. NMR samples were prepared by lyophilization of solutions of uniformly ^{15}N-labeled protein in 20 mM phosphate at the appropriate pH. The lyophilized protein was first dissolved in 20 μL of doubly distilled (DD) ^1H_2O; 580 μL of ^2H_2O was then added to the mixture to start the exchange. The final holoprotein concentration was 1 mM on a heme basis; the apoprotein concentration was 1 mM according to the estimated extinction coefficient. The time from initial addition of ^2H_2O to NMR data collection was ~12 min. ^1H-^{15}N heteronuclear single-quantum coherence (HSQC) spectra were recorded with 4096 complex points in the direct dimension and 128 real points in the indirect dimension (States-TPPI quadrature detection), using a WATERGATE solvent suppression scheme with a 3-9-19 binomial sequence (*33*, *34*). The spectral width for rHb-A samples was 8250 Hz (8013 Hz for rHb-R) in the ^1H dimension, and 2500 Hz (1800 Hz for rHb-R) in the ^{15}N dimension (centered at 114 ppm in rHb-A and 116.5 ppm in rHb-R). The spectral width for apo-rHb was 5682 Hz in

[1] Abbreviations: CD, circular dichroism; $\Delta\delta$, chemical-shift difference in ppm; DD, doubly distilled; Hb, hemoglobin; Hb-A, hemoglobin with covalently linked heme; HSQC, heteronuclear single-quantum coherence; MRE, molar residual ellipticity; NOE, nuclear Overhauser effect; *P*, protection factor; pH*, pH uncorrected for isotope effects; rHb, recombinant hemoglobin; rHb-A, recombinant hemoglobin with covalently linked heme; rHb-R, recombinant heme-reconstituted hemoglobin; S6803, *Synechocystis* sp. PCC 6803; S7002, *Synechococcus* sp. PCC 7002; TOCSY, totally correlated spectroscopy.

the ^{1}H dimension and 1900 Hz in the ^{15}N dimension (centered at 114 ppm). For holoprotein samples, sequential ^{1}H-^{15}N HSQC data sets were recorded first with 4 transients per t_1 increment. This number was increased to 8 and then to 16 transients for the remaining time points to achieve a useful signal-to-noise figure. Within the first 24 h, approximately 32 spectra were collected; after this period, one spectrum was collected daily or once every 2 days until the end of the study. In cases of overlapping peaks, the first plane of a ^{15}N-separated totally correlated spectroscopy (TOCSY) data set was collected for signal identification. A typical exchange study was completed in 10 days. Samples were kept at 25 °C in a water bath when not in the spectrometer. The same procedure was applied to the apoprotein, although in this case, only 4 data sets could be collected because exchange was complete after 2 h. The pH uncorrected for the isotope effect (pH*), which was read at the end of the experiment, was 7.40 for rHb-A and rHb-R and 7.2 for apo-rHb.

For the scope of this paper, knowledge of the exchange regime (EX1 or EX2) is not necessary. Protection factors (*P* values) were determined by fitting cross-peak volumes versus time (taken as the start of the HSQC experiment) to a single exponential and dividing the determined rate by the calculated rates of model peptides under the same condition of pH, ionic strength, and temperature (*35*, *36*). In one of two ^{15}N-labeled rHb-R samples, a few HSQC cross-peaks decayed in a biexponential fashion from an abnormally high initial value (125%) compared with the majority of cross-peaks. This raised the possibility that biphasic behavior was systematic across the protein, although lost in the dead time of the experiment for most amides. The exchange reaction was therefore followed with the rapid succession of 1D spectra collected on an unlabeled sample (1 mM protein and 20 mM phosphate at final pH* 7.45). With this sample, it was possible to compare the initial intensity (6-min delay) of exchangeable signals downfield from 9.3 ppm with that of heme signals and the entire exchangeable region to that in a spectrum collected in 95% ^1H_2O. The short acquisition time allowed for a more frequent sampling of the intensity than with HSQC spectra. Heme signal intensity was invariant with time, whereas the resolved exchangeable signals decayed with a single exponential from a normal initial intensity (100%). This experiment confirmed that occasional biphasic behavior in the labeled sample was caused by overlap with a non-native form of rHb-R. This phase is not discussed further; the reported data are from the sample that yielded monophasic behavior. In addition, a few resonances arising from the ~10% heme disorder (see the Results) could be followed for a short period of time. To a first approximation, these signals disappeared at the same rate as their major counterparts.

Circular Dichroism (CD) and Thermal Denaturation. The secondary structure of S7002 rHb and its resistance to temperature were analyzed with CD on a Jasco J-810 spectrometer. The raw data were converted to molar residual ellipticity (MRE). Thermal denaturation experiments were performed with the aid of a Peltier thermoelectric device. The temperature ranged between 25 and 95 °C. Data were collected every 2 °C with a 5-min equilibration time. Reversibility was tested by returning the samples to initial conditions in 3- or 2-°C step with a 5-min equilibration time. The samples were apo-rHb, rHb-R, or rHb-A in 20 mM phosphate buffer with a protein concentration of <15 μM in a 2-mm cell. The data were analyzed at 222 nm and fit to a Gibbs−Hemlholtz equation (*37*).

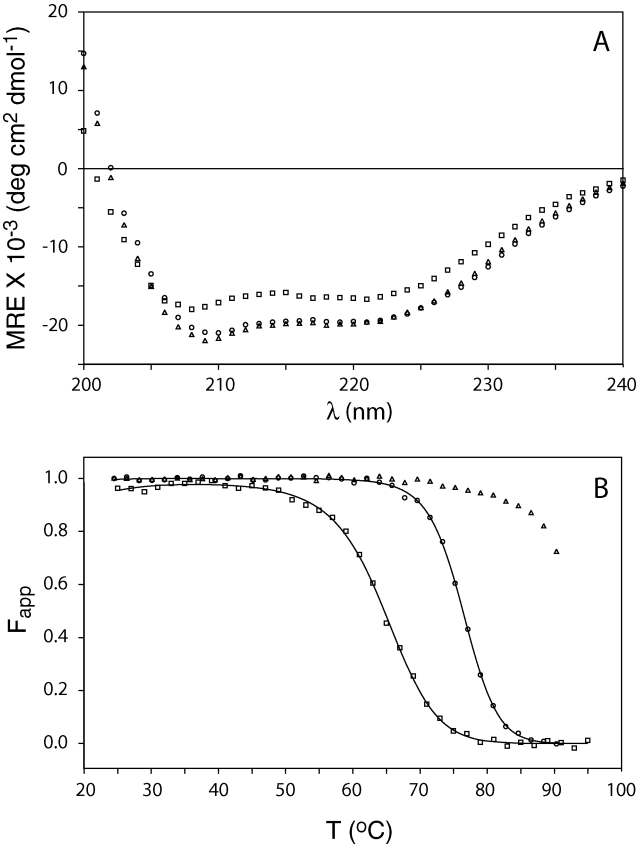

FIGURE 1: (A) Far-UV CD spectra of S7002 rHb-R (○), rHb-A (△), and apo-rHb (□). Conditions were 10 μM protein in 20 mM phosphate buffer at pH 7.4 and 25 °C. (B) Apparent fraction of folded protein (F_{app}) versus temperature. Data for S7002 rHb-R (○), rHb-A (△), and apo-rHb (□) are shown along with solid lines representing the results of two-state fits. The data for rHb-A were normalized by assuming an unfolded-state spectrum identical to that of rHb-R.

RESULTS

Secondary Structure by CD. The secondary structure of S7002 rHb-R, rHb-A, and apo-rHb was examined with far-UV CD. The spectra obtained at 25 °C and pH 7.2 are shown in Figure 1A after conversion to MRE. As expected, the holoprotein had a high α-helical content with characteristic minima at 208 and 222 nm. No significant difference was observed between rHb-R and rHb-A. The α-helical content of the apoprotein was slightly lower than that of the holoproteins.

Thermal Stability. Thermal denaturation experiments monitored by far-UV CD were performed on all three forms of the protein. When possible, the data were fit with a two-state model returning T_m, the temperature at the midpoint of the transition. The T_m for rHb-R was 76.4 ± 0.2 °C. The T_m of rHb-A could not be reached but appeared to be higher than 95 °C. Similar results were observed for the two corresponding forms of the S6803 protein (*20*). The apoprotein of the latter undergoes a broad thermal denaturation profile (*17*), whereas S7002 apo-rHb displayed a cooperative transition with a well-defined T_m of 64.6 ± 0.3 °C. The apparent fraction of folded protein is plotted versus temper-

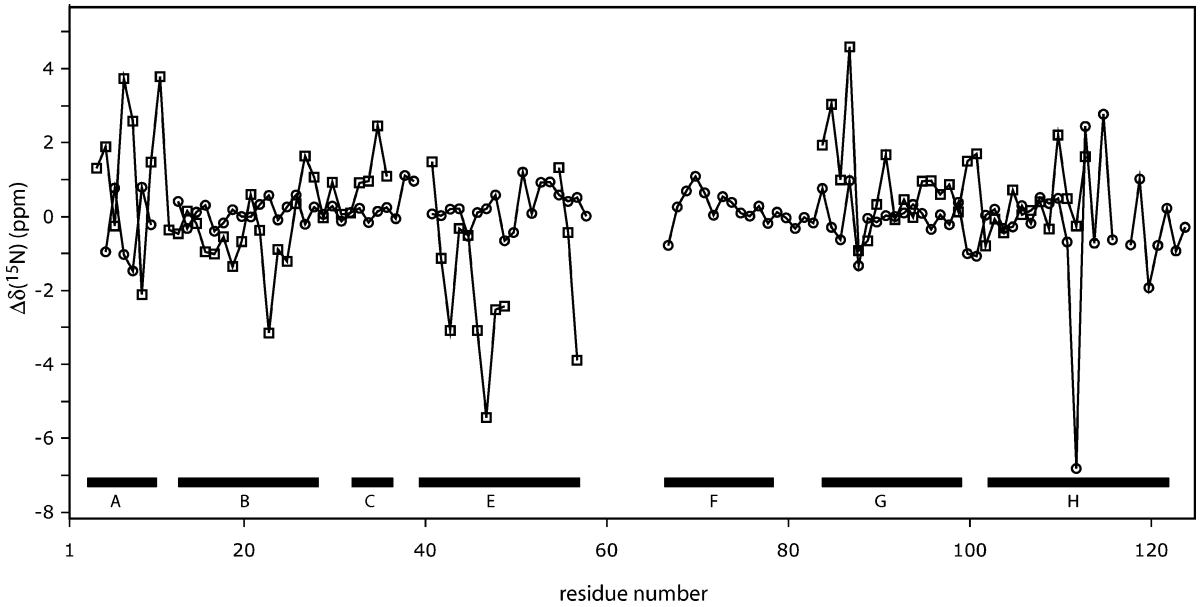

FIGURE 2: Backbone amide $\Delta\delta(^{15}N)$ for S7002-rHb: rHb-R−rHb-A (○) and apo-rHb−rHb-A (□). The secondary structure of S6803 rHb-R is indicated for reference. The areas of largest difference between the two holoproteins is the second half of the H helix and the beginning of the A helix. Apo-rHb shifts are most like the holoprotein shifts in the B and G helices.

ature in Figure 1B. Although reliable standard free energies of denaturation could not be extracted from the data because of partial reversibility, it was evident that the stability of the protein increased upon heme binding and further upon heme cross-linking. The observation of a cooperative transition with high T_m for the apoprotein implied stable tertiary interactions, as observed in the native state of sperm whale apomyoglobin (*38*).

NMR Assignments and Secondary Structure by NMR Spectroscopy. To interpret the $^1H/^2H$ exchange data, it was necessary to assign backbone NH resonances in all three forms of the protein. Standard assignment methods were applied to uniformly ^{15}N- and $^{13}C,^{15}N$-labeled S7002 rHb, as previously described for the related S6803 rHb-R (*30*). Figures S1−S3 in the Supporting Information show annotated ^1H-^{15}N HSQCs for rHb-R, rHb-A, and apo-rHb, respectively. Table S2 in the Supporting Information lists the chemical shifts, and Figure S4 in the Supporting Information portrays the assigned regions on the NMR structure of S6803 rHb-R.

The heme-containing proteins, S7002 rHb-R and rHb-A, were studied in the ferric state. These species are paramagnetic ($S = {}^1/_2$) and give rise to well-resolved spectra. The amino acid sequence (123 residues) is such that a total of 121 backbone NH resonances are expected. In rHb-A, all detectable ^1H-^{15}N HSQC cross-peaks (113) were assigned. The missing signals were located in the EF loop, a flexible region of the S6803 rHb-R structure (*23*) and presumably the S7002 rHb structure as well. In S7002 rHb-R, two sets of peaks in a 9:1 ratio were clearly distinguished. The minor form was attributed to the protein containing the heme group rotated by 180° along the $\alpha-\gamma$ meso axis (*18*). This isomerism was not present in rHb-A because the reaction of the 2-vinyl group drove the equilibrium to a single orientation. A total of 111 cross-peaks were resolved for the major form of rHb-R, and 108 of these were assigned. The signals of the EF loop were absent in this form of the protein as well. In addition, the backbone resonances of residues 4 (A helix), 11 and 12 (in the AB turn), 40 (CE turn), and 117 (H helix) were identified in rHb-A but not in rHb-R. Helical secondary structure was assigned tentatively through the dipolar connectivities between subsequent amide protons (NH_i-NH_{i+1}) (*39*). The data for rHb-R and rHb-A supported nearly identical helical boundaries for each structural element.

S7002 apo-rHb yielded well-resolved NMR data. Inspection of the ^1H-^{15}N HSQC spectrum indicated that ~95 cross-peaks were detectable under a normal condition of temperature and pH. Of these, 88 were unambiguously assigned. NH_i-NH_{i+1} nuclear Overhauser effect (NOE) connectivities were strong throughout the B, C, and G helices and first half of the H helix. Weaker connectivities were found in the N-terminal portion of the E helix, for amides 40−47. Additional evidence for helical structure was noted in the C-terminal turn of the A helix. The only helix for which no signal was detected was the F helix.

Low-Resolution Structural Comparison. ^{15}N chemical-shift differences [$\Delta\delta(^{15}N)$] were calculated to assess structural perturbation along the sequence. Figure 2 illustrates $\Delta\delta$ between the two holoproteins. The magnitude of most deviations was smaller than 1 ppm. Several outliers were found among residues 38−88, a region that was likely affected by changes in the electronic structure of the heme group. The largest exceptions were found in the second half of the H helix (near the cross-link), where a combination of paramagnetic and coordinate alterations was at work. The chemical shifts were also perturbed in the A helix; because this region is remote from the heme group, this was interpreted as a long-range effect of the cross-link. The regions least affected were the B and G helices. $\Delta\delta(^1H)$ paralleled the $\Delta\delta(^{15}N)$ trends.

$\Delta\delta(^{15}N)$ values calculated with the apoprotein lost their direct significance in close proximity of the heme group owing to both paramagnetic and porphyrin ring current effects. At locations remote from the heme, the $\Delta\delta(^{15}N)$ values shown in Figure 2 indicated the same pattern as for the holoprotein: the smallest deviations were found in the

Table 1: Classification of Amides According to Their Exchange Rate[a]

	rHb-R	rHb-A
fast[b]	63	59
intermediate[c]	37 (24 fitted)	43 (28 fitted)
slow[d]	11 (2 fitted)	11 (4 fitted)

[a] At 25 °C and pH* ~7.4. [b] NHs detected in the HSQC that have completely disappeared within the first 13 min of exchange. [c] NHs that were detected after 13 min and had disappeared by 10 days. "Fitted" indicates that a rate constant could be obtained. [d] NHs that were present with appreciable intensity (>20%) after 10 days of exchange.

B and G helices; sizable differences were noted in the 40−90 region; and the A helix was perturbed more strongly than the B and G regions. The same trends were observed for $\Delta\delta(^1H)$ and $\Delta\delta(^{15}N)$. Overall, the $\Delta\delta$ analysis emphasized the B and G helices as resilient structural motifs of the apoprotein.

$^1H/^2H$ Exchange in rHb-R and rHb-A. At neutral pH, both rHb-R and rHb-A were found to display a wide range of exchange rates. For the purpose of analysis, three categories of backbone amide NHs were arbitrarily distinguished. The "fast" category included those amide NHs that were completely exchanged within the first 13 min of exposure to ^2H_2O at neutral pH. This corresponded to a *P* value below 1000. Over half of the amide NHs were in this group. Amide NHs from the "intermediate" category exchanged in greater than 13 min but fewer than 10 days ($\sim 3 < \log P < \sim 6$). Finally, the amides that retained more than 20% of their intensity after 10 days were placed in the "slow" category ($\log P > \sim 6$). Table 1 provides the number of residues in each category.

Within the intermediate group, 24 rHb-R and 28 rHb-A backbone amide NHs had exchange profiles that could be used for curve fitting. Other resonances overlapped or disappeared too rapidly for reliable evaluation. Numerical values for the exchange rate constants are given with the fitted curves in Figure S5 in the Supporting Information. *P* values are listed in Table S3 in the Supporting Information. A total of 19 of the locations for which quantitative assessment was made were common to the two forms of the protein; the ratio of *P* values, *P*(rHb-R)/*P*(rHb-A), for these 19 NHs is listed in Table 2 along with the ratio for the single common fitted NH in the slow category (Ala88). The data and fits for one of these residues, Ala54 in the E helix, are shown in Figure 3 as representative of the difference between the two forms of the protein. In some cases, a smaller number of cross-peak volumes could be obtained reliably. This was evident with Thr80 in rHb-R. As a control, this signal was monitored in a separate experiment with a rapid succession of 1D spectra (Figure S2 in the Supporting Information). The discrepancy between 1D and 2D rate constants was ~14%. In other instances of sparse data, the 1D resolution was insufficient; there was, however, no ambiguity regarding the ranking of the rates (rHb-R versus rHb-A), even though the *P* ratio contained a large error.

To complement the *P* values, the intermediate exchange category of rHb-R and rHb-A was examined qualitatively at two separate time points, 13 min and 20 h. After 13 min of exchange, 34 of the "intermediate" cross-peaks in the rHb-R and rHb-A spectra (shown in Figure 4) had identical origin. Numerical data are provided in Table S3 in the Supporting Information. The spectra in Figure 4 display notable differences. For instance, the B helix, E helix, and second half of the G helix contained amide NHs that were more protected in rHb-R than in rHb-A; of note is a longer-lived amide NH adjacent to the distal His46 (Gln47). In contrast, certain amide NHs persisted longer in rHb-A than in rHb-R. These were found mostly in the second half of the H helix (119−124), downstream from the cross-linking His117. Of note in this set is the backbone amide NH of the proximal His70. From these data, it appeared that the same regions of the fold exhibited different local stability in rHb-R and rHb-A, betraying specific cross-link effects throughout the structure.

After 20 h of exchange, 16 "intermediate" NH cross-peaks were detected in rHb-A and 12 in rHb-R (Table S4 in the Supporting Information). The amide NHs that exchanged more slowly in rHb-A than in rHb-R clustered to the H helix (residues 120−122) and the FG loop (residues 79 and 80) along with one amide NH (residue 84) in the G helix. In the structure of S6803 rHb-R and rHb-A, the NH of Asn80 is involved in a hydrogen bond with the side chain of His83. The same interaction between Thr80 and His83, resulting

Table 2: Protection Factors Measured in rHb-A and rHb-R

residue	*P*(rHb-R) $\times 10^{-6}$	*P*(rHb-A) $\times 10^{-6}$	*P*(rHb-R)/*P*(rHb-A)
Phe21 (B9)	0.527 ± 0.020	0.286 ± 0.017	1.84 ± 0.18
Tyr22 (B10)	9.1 ± 0.5	1.51 ± 0.05	6.06 ± 0.6
Asp28	1.94 ± 0.05	1.66 ± 0.05	1.17 ± 0.07
Val31 (C3)	0.611 ± 0.018	0.62 ± 0.04	0.98 ± 0.10
Met51 (E15)	8.3 ± 0.3	2.59 ± 0.17	3.2 ± 0.3
Thr52 (E16)	1.76 ± 0.07	4.07 ± 0.18	0.43 ± 0.04
Tyr53 (E17)	1.71 ± 0.07	1.45 ± 0.12	1.18 ± 0.14
Ala54 (E18)	1.31 ± 0.04	0.401 ± 0.016	3.27 ± 0.22
Val74 (F12)	0.0310 ± 0.0006	0.0706 ± 0.0025	0.439 ± 0.024
Thr80 (FG7)	0.089 ± 0.007	2.47 ± 0.10	0.036 ± 0.004
His83 (G3)	0.189 ± 0.008	0.48 ± 0.03	0.40 ± 0.04
Phe84 (G4)	0.127 ± 0.011	5.1 ± 0.5	0.025 ± 0.005
Ala86 (G6)	0.64 ± 0.03	0.368 ± 0.016	1.74 ± 0.15
Ala88 (G8)	11.7 ± 1.2	8.1 ± 0.9	1.5 ± 0.3
Glu96 (G17)	3.67 ± 0.11	1.60 ± 0.15	2.3 ± 0.3
Leu98 (G19)	0.80 ± 0.03	0.502 ± 0.024	1.59 ± 0.14
Leu105 (H5)	0.80 ± 0.03	0.365 ± 0.013	2.20 ± 0.16
Asp106 (H6)	0.188 ± 0.006	0.104 ± 0.003	1.81 ± 0.11
Val108 (H9)	0.64 ± 0.04	0.379 ± 0.019	1.70 ± 0.18
Val109 (H10)	0.143 ± 0.006	0.0438 ± 0.0019	3.3 ± 0.3

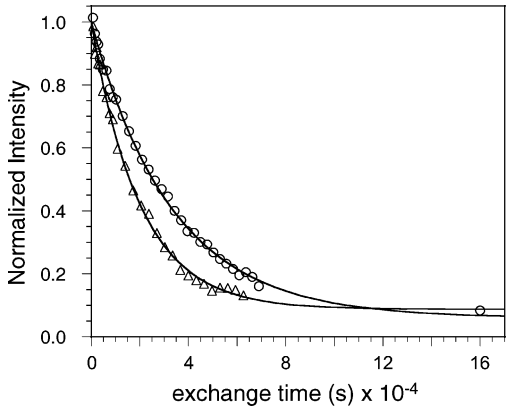

FIGURE 3: Exchange profile for Ala54 in rHb-R (○) and rHb-A (△) plotted as a fraction of the initial amplitude as a function of the exchange time. The solid lines represent fits to single exponentials with rate constants of $(3.00 \pm 0.08) \times 10^{-5}$ and $(5.03 \pm 0.19) \times 10^{-5}$ s^{-1}, respectively. The last point in the rHb-R data set was collected after ~45 h of exchange.

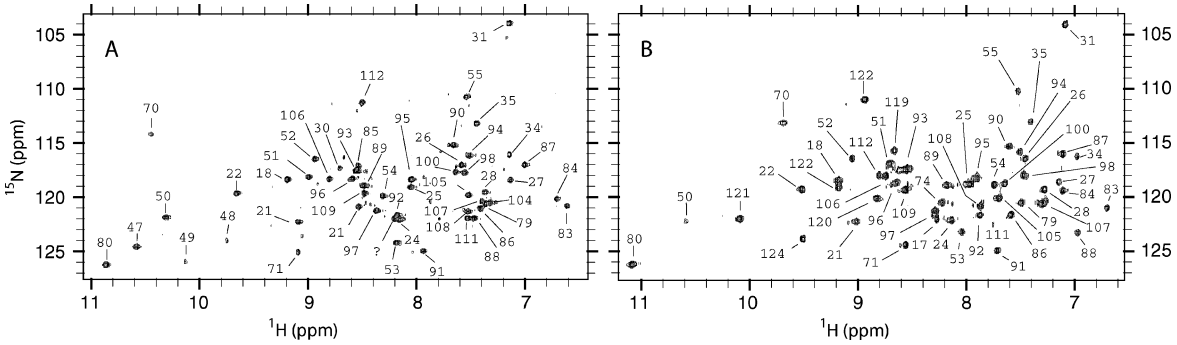

FIGURE 4: ^{15}N-^{1}H HSQC of S7002 rHb after ~13 min of ^{1}H/^{2}H exchange. (A) rHb-R and (B) rHb-A.

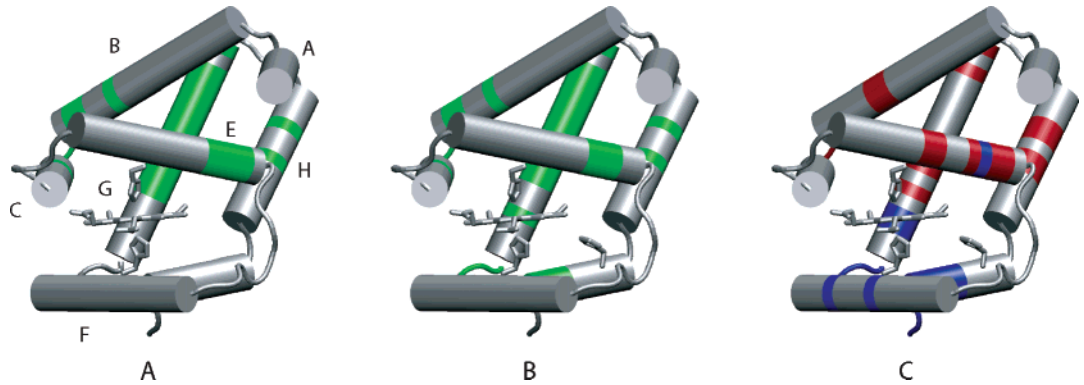

FIGURE 5: Location of the amide NHs remaining in S7002 rHb-R (A) and rHb-A (B) after 20 h of exchange, shown on a NMR structure of S6803 rHb-R (PDB 1MWB). His46 (distal), His70 (proximal), and the heme group are shown with licorice bonds. In B and C, the location of His117 (forming the cross-link to the heme 2-vinyl) is also indicated. Helices are labeled A−H according to the vertebrate nomenclature. C illustrates the differential effects of the cross-link, with the color blue for amide NHs exchanging more slowly in rHb-A than in rHb-R and red for amide NHs exchanging faster in rHb-A than in rHb-R.

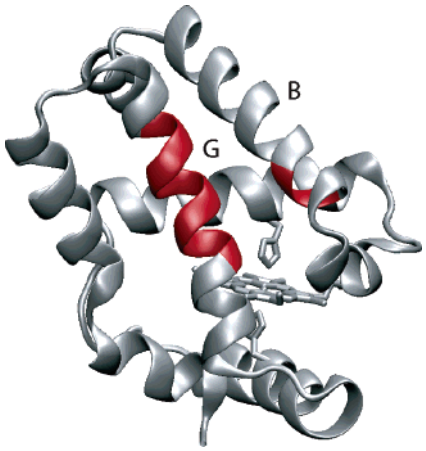

FIGURE 6: Slow-exchange core of S7002 rHb-R and S7002 rHb-A as defined by the criteria listed in the text and represented in red using the structure of S6803 rHb-R (PDB 1MWB).

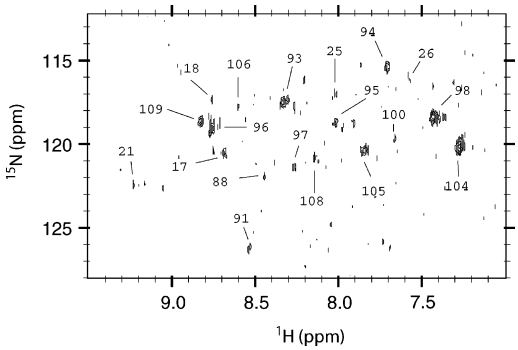

FIGURE 7: ^{15}N-^{1}H HSQC of S7002 apo-rHb after ~13 min of ^{1}H/^{2}H exchange. Some of the labeled peaks are weak but are not artifactual because they were readily detected in the spectrum collected in 95% $^{1}H_2O$ (see the Supporting Information) and did disappear in later spectra.

in a signature NH/CϵH NOE, occurred in both forms of S7002 rHb. Two NH cross-peaks were detected in rHb-R and not in rHb-A; these were located in the E and G helices, at positions 54 and 86. Parts A and B of Figure 5 illustrates the location of the amide NHs remaining in rHb-A and rHb-R after 20 h of exchange. The slow category of rHb-A and rHb-R (Figure 6) contained common residues: 25, 26, and 87−95. These are situated in the B and G helices.

^{1}H/^{2}H Exchange in Apo-rHb. The same experiments applied to the apoprotein revealed that most amide NHs belonged to the fast category as defined for the holoproteins. A total of 19 cross-peaks observed in the 13 min spectrum (Figure 7 and Table S4 in the Supporting Information) had disappeared after 2 h, and just one was detectable after 4 h (Ile105). The 13-min set was arbitrarily chosen to represent the slow-exchange core of the apoprotein. All amide NHs in this set were from the B, G, and H helices; their locations are displayed in Figure 8.

DISCUSSION

Local and Global Stability. The binding of a heme group to an apoprotein typically results in a pronounced stabilization of the fold. This is accompanied with conformational changes of various magnitudes. The complex, however, may not have the desired heme properties, and post-translational modifications may be required to tune them. For S7002 rHb, cross-link formation following heme binding is a particularly

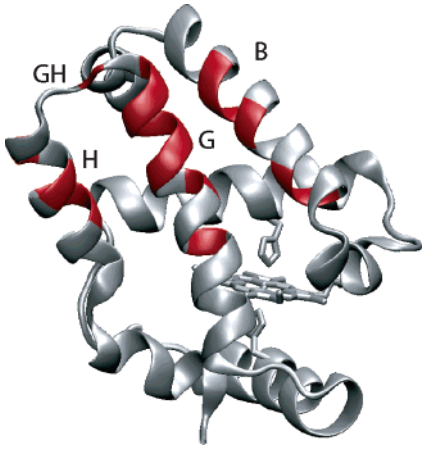

FIGURE 8: Slow-exchange core of S7002 apo-rHb as defined by the criteria listed in the text and represented using the structure of S6803 rHb-R (PDB 1MWB). The amides belonging to the core are colored in red: in the B helix, Ala17, Val18, Phe21, Val25, and Leu26; in the G helix, Ala88, Leu91, Leu93, Thr94, Leu95, Gln96, Glu97, and Leu98; in the GH loop, Val100; and in the H helix, Leu104, Ile105, Asp106, Val108, and Val109.

facile reaction. Although physiological data are not available, it is possible that this reaction occurs in the cell. The cross-link could adjust axial ligand strength, a plausible consequence supported by the observations that the rate of cyanide binding by ferric S6803 rHb is accelerated when the cross-link is present (*40*). Local structural stability is also of interest, because binding of exogenous ligands is accompanied by a large conformational change that may be facilitated or opposed in the cross-linked state. S7002 rHb-A is significantly more resistant to conditions of low pH and high temperature than rHb-R (*20*). According to the NMR and X-ray data, cross-link formation does not cause large-scale structural changes (*13*, *20*, *24*). The main contribution to the global stability effect can therefore be attributed to a loss of entropy in the unfolded state. The $^1H/^2H$ exchange studies described in this work allowed for an examination of local native-state features related to heme binding and its covalent attachment to the protein matrix.

A parallel can be drawn to cytochrome *c*. In this class of proteins, covalent heme attachment occurs through the reaction of two Cys side chains with the heme vinyl groups (*41*). When attachment through thioether bonds is prevented in the mitochondrial protein, folding does not take place even in the presence of the heme (*42*). This result illustrates the predisposition of the protein for disorder and an extreme consequence of cofactor attachment but provides little information on its chemical aspects. To date, there is still no detailed assessment of the advantages of thioether bond formation in *c* cytochromes despite extensive studies (*43*). Interestingly, one covalent linkage, the second in the sequence, is sufficient to establish the holoprotein fold (*44*). Only in the case of a thermophilic cytochrome c_{552} has it been possible to generate a stable *b* form of a *c* cytochrome by mutagenesis of the reactive cysteines (*45*). Not surprisingly, this protein shows an increase in stability upon cross-linking (a T_m change from 58 to over 100 °C). In addition, the apocytochrome was found to be 8 °C less stable than the holoprotein containing no thioether linkage. S7002 rHb, with its structured apo, non-cross-linked, and cross-linked states, offers an opportunity to study the effect of covalent attachment in a mesophilic protein without resorting to amino acid replacement.

Effect of Heme Binding. The apoprotein of S7002 rHb was found to be partially structured and stable up to temperatures higher than that of the normal environment of the cyanobacterium (~25 °C, with optimal growth at 38 °C). All holoprotein helices were represented in the apoprotein with at least a few residues, except for helix F. In the EF−F region of the protein and the C-terminal end of the H helix, no ^1H-^{15}N HSQC cross-peak was assigned and, given the total number of cross-peaks, most were not detected. For a comparison, the F helix and the end of the H helix are known to unfold in sperm whale apomyoglobin (*46*). Studies are in progress to improve the description of the residual structure in S7002 apo-rHb. Nevertheless, it is already clear that heme binding is responsible for major conformational restriction on the proximal side of the heme to lead to the completion of a stable two-on-two fold. In this, the S7002 apoglobin resembled many apo *b* hemoproteins.

Effect of the Cross-link: Site of Modification. The H helix contains a kink at position 115 (*23*). Near the site of the cross-link and at the end of the H helix (119−124), backbone NHs were more protected against exchange in rHb-A than in rHb-R, likely as a result of reduced conformational flexibility in this region. This interpretation was supported by the absence of the His117 NH cross-peak in the rHb-R ^1H-^{15}N HSQC data and its presence in rHb-A data. Prior to the kink, exchange was faster in rHb-A than in rHb-R, which indicated that this feature disrupted the dynamic behavior of the helix. The $\Delta\delta(^{15}N)$ between rHb-R and rHb-A were greatest in the latter half of the H helix (Figure 2). In the S6803 rHb-R NMR structure, the end of the H helix is conformationally flexible (*23*) and cross-linking is expected to diminish this flexibility. Although the primary structure of the two cyanobacterial globins differs in this region, our preliminary analysis supported similar behaviors. Cytochrome *c* also shows enhanced amide NH protection in the cross-linked area, specifically for those residues around the essential second link (*47*).

Effect of the Cross-link: Remote Repercussions. The changes associated with cross-linking were not restricted to the immediate surroundings of the attachment point in the H helix. In the FG turn, the NH of Thr80 exchanged more slowly in rHb-A than in rHb-R, which indicated a hydrogen bond of increased strength between Thr80 NH and the side chain of His83 in the cross-linked protein. The side chain of Asn80 acts as an N-terminal cap for the G helix in S6803 rHb (*23*); the capping motif is classified as type Ia (*48*). The same motif was expected in S7002 rHb and was supported by the NMR data. The reciprocal backbone/side chain hydrogen bonds between residues 80 and 83 serve as a start signal for the G helix. Inspection of the NOEs at these locations confirmed that the conformation of the protein was unchanged by the cross-link. This particular element of structure initiating the G helix appeared reinforced in rHb-A.

The proximal histidine (His70) backbone NH was also slower to exchange in rHb-A, as were other NHs in the F helix, for example, that of Val74 (Table 2). The slow exchange of His70 and Thr80 may have a common origin, illustrating transduction through the F helix. This helix and the first half of the G helix of rHb-A presented regions of

retarded exchange compared to rHb-R. Retardation, however, was not a generalized phenomenon. The reverse effect was observed for residues 21, 22, and 28 (B helix); 47, 48, 51, 53, and 54 (E helix); 86, 88, 96, and 98 (G helix); and 105, 106, 108, and 109 (H helix). The uneven distribution of responses, depicted in Figure 5C, demonstrated a complex influence of the cross-link resulting in acceleration mostly on the distal side of the heme, where exogenous ligand binding occurs. Complex distribution was also observed in cytochrome c_{552}, as opposite changes were measured in the region around the second cross-link and in the N terminus of the protein (*47*). In another example, it has been reported that binding of different ligands (azide and cyanide) to the ferric substrate of heme oxygenase from *Pseudomonas aeruginosa* leads to large changes in *P* values, distributed over the structure. This is attributed in part to an alteration of the distal hydrogen-bond network configuration (*49*) and illustrates a mechanism by which small protein structural perturbations can cause large differences in the exchange properties.

Effect of the Cross-link: Exogenous Ligand-Binding Site. The function of most two-on-two globins is unknown; all have been found to bind diatomic ligands with high affinity even when capable of endogenous hexacoordination, and this property has guided functional speculations. The data presented here showed that the cross-link affected indirectly the properties of the binding site. A possible interpretation of the E helix data was that this element of secondary structure was less tightly wound in rHb-A than in rHb-R because the repositioning of the heme group affected the His46−Fe ligation bond and, therefore, the exogenous ligand binding site. In S6803 rHb, the H helix contacts the E helix at Ala54. The same contact was detected in NOE data collected on S7002 rHb-A (not shown). Because the amide NH of Ala54 was more protected in rHb-R than in rHb-A, it is plausible that cross-linking at His117 perturbed the E−H contact, in turn propagating changes to the distal binding site. As mentioned above, an opposite result was observed on the proximal side, with the backbone NH of His70 disappearing nearly 24 h later in rHb-A than in rHb-R. Transmission from one side of the heme to the other can be mediated by ligation to the iron (trans effect). Thus, the weakening of the E helix may be related to a reinforcement of the F helix. It has also been shown that covalently linking His-coordinated heme to peptide models can result in strengthening an axial Fe−His bond (*50*), which could provide an additional contribution to the effect above.

Among two-on-two globin sequences and available structures, position B10 stands out as a determinant of ligand-binding properties. This has been demonstrated in S6803 rHb (*13*, *51*) and S7002 rHb (*13*), as Tyr22 (B10) participates in a hydrogen-bonding network with bound cyanide. The observation that the B helix underwent slower exchange in rHb-R than in rHb-A implied reduced conformational flexibility of rHb-R and a potential impact on ligand binding, since the structural rearrangement that promotes the formation of the distal hydrogen-bond network might require transient distortion of the B helix. Consequently, the data suggested that the distal pocket was more dynamic in rHb-A than in rHb-R. This may be related to our observation that, in S6803 rHb, the presence of the cross-link promotes faster cyanide binding (*40*).

Slow-Exchange Core. The slow-exchange core of rHb-A and rHb-R contained residues 25, 26 (in the B helix), and 87−95 (in the G helix). These two regions of the protein make contact with one another and therefore could represent a nucleation point early in protein folding. In support of this, the B−G helical contacts have been found to form a key motif in most globins (*52*), including two-on-two globins (*53*). Considering the extent of refolding that is coupled to heme binding, it is of interest to inspect the contribution of the cofactor in establishing the stable architecture of the protein.

Exchange data collected on apo-rHb revealed a cluster that involved B, G, and H segments. As illustrated on a model of the holoprotein structure (Figure 8), these three helices contact each other but not the heme group. NMR analysis of the native state further showed that these regions were helical in the apoprotein. A comparison of amide ^{1}H and ^{15}N chemical shifts (Figure 2) suggested only small perturbations between the apo- and holoproteins in the regions of the B and G helix. Although the H helix was not part of the slowest exchange core in the holoproteins by the criteria established above, elements of it that were part of the apoprotein core exhibited slow amide exchange (the NH of residues 105 and 108 persisted for over 5 days in rHb-R). In addition, all amides remaining at 5 days were located in the B, G, and H helices in both forms of the holoprotein (with the exception of Met51 in rHb-R). As a result, it was concluded that the core of the holoprotein was partially formed in the apoprotein. Heme binding stabilized and extended this existing nucleus.

The core of S7002 apo-rHb appeared better formed and stabilized than that of S6803 rHb under identical conditions (*17*, *40*). Among the 19 assigned amides that were slowest to exchange in S7002 apo-rHb, five amino acid replacements were noted between the two proteins, all located in the C-terminal region of the G helix and the N-terminal region of the H helix: Leu93Ala, Gln96Lys, Leu98Met, Asp106Ala, and Val109Ala. It is possible that these mostly nonconservative substitutions alter the ability of the GH turn to remain folded and serve as a scaffold for further lasting interactions in the absence of the heme group.

Implications for S7002 rHb Folding. The formation of a functional hemoglobin molecule is a complex process requiring the folding of the polypeptide chain and the binding of the cofactor. Although not much is known about the reaction in vivo, progress has been made in the description of the apoprotein folding step through extensive in vitro studies of sperm whale apomyoglobin as a model globin (*54*). For this particular protein, progression to the native state from a urea-unfolded state involves accretion of helices A, G, H, and part of helix B. The A, G, and H helices constitute the nucleus (*55*). This mechanism, however, may have limited applicability to other globins. For example, leghemoglobin, which has the same three-on-three fold, follows a different pathway involving the early collapse of the E, G, and H helices (*56*). The structural differences between the three-on-three and two-on-two folds, which encompass the nature of the interhelical contacts (*53*) as well as the relative orientation, number, and length of helices (*9*), lead to an expectation of yet a distinct energy landscape.

In many instances, the amide NHs most protected from exchange with solvent cluster with secondary structure, and

it is thought that these NHs belong to parts of the protein that are first to fold (*57*, *58*). The slow-exchange core of S7002 apo-rHb contained parts of the B, G, and H helices, three of the four helices defining the two-on-two fold. The last of the four, the E helix, became fully folded only upon heme binding. Kinetic studies of sperm whale apomyoglobin and apoleghemoglobin point to the importance of the G and H helices in full-length globin folding. G−H peptide studies have revealed an intrinsic propensity of the sperm whale sequence to sample a helical structure (*59*). G−H contacts are also expected to be important in the "minihemoglobin" of the nemertean worm *Cerebratulus lacteus*. This protein possesses no A helix and only the first half of the H helix (*60*), which aligns with the H helix residues participating in the slow-exchange core of S7002 apo-rHb. With regards to the two-on-two globins, a recent computational analysis supports a key role for the G−H segment (*61*). The native-state exchange studies of S7002 apo-rHb strongly suggested that these helices might play an early role in the folding process of two-on-two globins as well.

CONCLUSIONS

The two-on-two globin from *Synechococcus* sp. PCC 7002 offered an opportunity to inspect the large and small structural perturbations induced by heme binding and heme cross-linking. In the formation of rHb-R from apo-rHb, extensive conformational changes were noted that transformed a protein containing several partially formed helices and a folded exchange core into a stable, fully folded two-on-two orthogonal bundle. In rHb-R, the core remained defined by the same helices as in the apoprotein, namely, parts of B, G, and H helices. The observation reinforced the importance of the G−H motif toward globin folding and may represent important contacts in the progenitor globin fold. Further and more subtle perturbations were imposed by the formation of a covalent linkage between protein and heme periphery. This type of post-translational modification, which is common among heme proteins, brought uneven changes in dynamics throughout the protein and notably on the distal side of the heme, where exogenous ligands bind. In vivo cyanobacterial studies are in progress that, it is hoped, will provide direct information as to the function of the protein. The present in vitro study revealed three forms of the protein potentially viable in the cell and will guide the interpretation of the biochemical data.

ACKNOWLEDGMENT

The authors thank Nancy Scott for assistance throughout the project and Dr. Jane Knappenberger and Daniel Landfried for assistance with the thermal denaturation experiments.

SUPPORTING INFORMATION AVAILABLE

A table of parameters for heteronuclear NMR data acquisition; a table of backbone ^{1}H and ^{15}N chemical shifts for rHb-R, rHb-A, and apo-rHb; ^{1}H-^{15}N HSQC spectra of rHb-R, rHb-A, and apo-rHb; a table of calculated *P* values for selected amide NHs of rHb-R and rHb-A; a table of amide NHs detected after 13 min and 20 h of exchange; a figure of assigned backbone regions of rHb-R, rHb-A, and apo-rHb; and figures of exchange data with fitted parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Vinogradov, S. N., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Gough, J., Dewilde, S., Moens, L., and Vanfleteren, J. R. (2006) A phylogenomic profile of globins, *BMC Evol. Biol. 6*, 31.
2. Ouellet, H., Ouellet, Y., Richard, C., Labarre, M., Wittenberg, B., Wittenberg, J., and Guertin, M. (2002) Truncated hemoglobin HbN protects *Mycobacterium bovis* from nitric oxide, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5902−5907.
3. Mukai, M., Ouellet, Y., Ouellet, H., Guertin, M., and Yeh, S. R. (2004) NO binding induced conformational changes in a truncated hemoglobin from *Mycobacterium tuberculosis*, *Biochemistry 43*, 2764−2770.
4. Crespo, A., Marti, M. A., Kalko, S. G., Morreale, A., Orozco, M., Gelpi, J. L., Luque, F. J., and Estrin, D. A. (2005) Theoretical study of the truncated hemoglobin HbN: Exploring the molecular basis of the NO detoxification mechanism, *J. Am. Chem. Soc. 127*, 4433−4444.
5. Hill, D. R., Belbin, T. J., Thorsteinsson, M. V., Bassam, D., Brass, S., Ernst, A., Boger, P., Paerl, H., Mulligan, M. E., and Potts, M. (1996) GlbN (cyanoglobin) is a peripheral membrane protein that is restricted to certain *Nostoc* spp., *J. Bacteriol. 178*, 6587−6598.
6. Moens, L., Vanfleteren, J., van de Peer, Y., Peeters, K., Kapp, O., Czeluzniak, J., Goodman, M., Blaxter, M., and 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.
7. Iwaasa, H., Takagi, T., and Shikama, K. (1990) Protozoan hemoglobin from *Tetrahymena pyriformis*. Isolation, characterization, and amino acid sequence, *J. Biol. Chem. 265*, 8603−8609.
8. Wittenberg, J. B., Bolognesi, M., Wittenberg, B. A., and Guertin, M. (2002) Truncated hemoglobins: A new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes and plants, *J. Biol. Chem. 277*, 871−874.
9. Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., Moens, L., and Bolognesi, M. (2000) A novel two-over-two α-helical sandwich fold is characteristic of the truncated hemoglobin family, *EMBO J. 19*, 2424−2434.
10. Kendrew, J. C., Dickerson, R. E., Stranberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., and Shore, V. C. (1960) Structure of myoglobin. Three-dimensional Fourier synthesis at 2 Å resolution, *Nature 185*, 422−427.
11. Egawa, T., and Yeh, S. R. (2005) Structural and functional properties of hemoglobins from unicellular organisms as revealed by resonance Raman spectroscopy, *J. Inorg. Biochem. 99*, 72−96.
12. Milani, M., Savard, P. Y., Ouellet, H., Ascenzi, P., Guertin, M., and Bolognesi, M. (2003) A TyrCD1/TrpG8 hydrogen bond network and a TyrB10−TyrCD1 covalent link shape the heme distal site of *Mycobacterium tuberculosis* hemoglobin O, *Proc. Natl. Acad. Sci. U.S.A. 100*, 5766−5771.
13. Vu, B. C., Nothnagel, H. J., Vuletich, D. A., Falzone, C. J., and Lecomte, J. T. J. (2004) Cyanide binding to hexacoordinate cyanobacterial hemoglobins: Hydrogen bonding network and heme pocket rearrangement in ferric H117A *Synechocystis* Hb, *Biochemistry 43*, 12622−12633.
14. Milani, M., Pesce, A., Ouellet, Y., Ascenzi, P., Guertin, M., and Bolognesi, M. (2001) *Mycobacterium tuberculosis* hemoglobin N displays a protein tunnel suited for O_2 diffusion to the heme, *EMBO J. 20*, 3902−3909.
15. Milani, M., Pesce, A., Ouellet, Y., Dewilde, S., Friedman, J., Ascenzi, P., Guertin, M., and Bolognesi, M. (2004) Heme−ligand tunneling in group I truncated hemoglobins, *J. Biol. Chem. 279*, 21520−21525.
16. Couture, M., Das, T. K., Savard, P. Y., Ouellet, Y., Wittenberg, J. B., Wittenberg, B. A., Rousseau, D. L., and Guertin, M. (2000) Structural investigations of the hemoglobin of the cyanobacterium *Synechocystis* PCC 6803 reveal a unique distal heme pocket, *Eur. J. Biochem. 267*, 4770−4780.
17. Lecomte, J. T. J., Scott, N. L., Vu, B. C., and Falzone, C. J. (2001) Binding of ferric heme by the recombinant globin from the cyanobacterium *Synechocystis* sp. PCC 6803, *Biochemistry 40*, 6541−6552.

18. Scott, N. L., Falzone, C. J., Vuletich, D. A., Zhao, J., Bryant, D. A., and Lecomte, J. T. J. (2002) The hemoglobin of the cyanobacterium *Synechococcus* sp. PCC 7002: Evidence for hexacoordination and covalent adduct formation in the ferric recombinant protein, *Biochemistry 41*, 6902−6910.
19. Vu, B. C., Jones, A. D., and Lecomte, J. T. J. (2002) Novel histidine−heme covalent linkage in a hemoglobin, *J. Am. Chem. Soc. 124*, 8544−8545.
20. Vu, B. C., Vuletich, D. A., Kuriakose, S. A., Falzone, C. J., and Lecomte, J. T. J. (2004) Characterization of the heme−histidine cross-link in cyanobacterial hemoglobins from *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002, *J. Biol. Inorg. Chem. 9*, 183−194.
21. Vuletich, D. A., and Lecomte, J. T. J. (2006) A phylogenetic and structural analysis of truncated hemoglobins, *J. Mol. Evol. 62*, 196−210.
22. Huang, L., Wojciechowski, G., and de Montellano, P. R. (2006) Role of heme−protein covalent bonds in mammalian peroxidases: Protection of the heme by a single engineered heme−protein link in horseradish peroxidase, *J. Biol. Chem. 281*, 18983−18988.
23. Falzone, C. J., Vu, B. C., Scott, N. L., and Lecomte, J. T. J. (2002) The solution structure of the recombinant hemoglobin from the cyanobacterium *Synechocystis* sp. PCC 6803 in its hemichrome state, *J. Mol. Biol. 324*, 1015−1029.
24. Hoy, J. A., Kundu, S., Trent, J. T., III, Ramaswamy, S., and Hargrove, M. S. (2004) The crystal structure of *Synechocystis* hemoglobin with a covalent heme linkage, *J. Biol. Chem. 279*, 16535−16542.
25. Dempsey, C. E. (2001) Hydrogen exchange in peptides and proteins using NMR spectroscopy, *Prog. Nucl. Magn. Reson. Spectrosc. 39*, 135−170.
26. Miller, D. W., and Dill, K. A. (1995) A statistical mechanical model for hydrogen exchange in globular proteins, *Protein Sci. 4*, 1860−1873.
27. de Duve, C. (1948) A spectrophotometric method for the simultaneous determination of myoglobin and hemoglobin in extracts of human muscle, *Acta Chem. Scand. 2*, 264−289.
28. http://us.expasy.org/cgi-bin/protparam.
29. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR Spectroscopy. Principles and Practice*, Academic Press, San Diego, CA.
30. Falzone, C. J., and Lecomte, J. T. J. (2002) Assignment of the ^{1}H, ^{13}C, and ^{15}N signals of *Synechocystis* sp. PCC 6803 methemoglobin, *J. Biomol. NMR. 23*, 71−72.
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR. 6*, 277−293.
32. Goddard, T. D., and Kneller, D. G. (2006) SPARKY 3, University of California−San Francisco, San Francisco, CA.
33. Piotto, M., Saudek, V., and Sklenár, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR. 2*, 661−665.
34. Sklenár, V., Piotto, M., Leppik, R., and Saudek, V. (1993) Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity, *J. Magn. Reson. 102*, 241−245.
35. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins 17*, 75−86.
36. www.fccc.edu/research/labs/roder.
37. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989) *Protein Structure: A Practical Approach*, IRL Press, Oxford, U.K.
38. Griko, Y. V., Privalov, P. L., Venyaminov, S. Y., and Kutyshenko, V. P. (1988) Thermodynamic study of the apomyoglobin structure, *J. Mol. Biol. 202*, 127−138.
39. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
40. Knappenberger, J. A., Kuriakose, S. A., Vu, B. C., Nothnagel, H. J., Vuletich, D. A., and Lecomte, J. T. J. (2006) Proximal influences in two-on-two globins: Effect of the Ala69Ser replacement on *Synechocystis* sp. PCC 6803 hemoglobin, *Biochemistry 45*, 11401−11413.
41. Pettigrew, G. W., and Moore, G. R. (1991) pp 1−223, Springer-Verlag, Berlin, Germany.
42. Fisher, W. R., Taniuchi, H., and Anfinsen, C. B. (1973) On the role of heme in the formation of the structure of cytochrome *c*, *J. Biol. Chem. 248*, 3188−3195.
43. Barker, P. D., and Ferguson, S. J. (1999) Still a puzzle: Why is haem covalently attached in c-type cytochromes? *Struct. Fold. Des. 7*, R281−R290.
44. Tanaka, Y., Kubota, I., Amachi, T., Yoshizumi, H., and Matsubara, H. (1990) Site-directedly mutated human cytochrome *c* which retains heme *c* via only one thioether bond, *J. Biochem. 108*, 7−8.
45. Tomlinson, E. J., and Ferguson, S. J. (2000) Conversion of a *c* type cytochrome to a *b* type that spontaneously forms in vitro from apo protein and heme: Implications for *c* type cytochrome biogenesis and folding, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5156−5160.
46. Lecomte, J. T. J., Sukits, S. F., Bhattacharya, S., and Falzone, C. J. (1999) Conformational properties of native sperm whale apomyoglobin in solution, *Protein Sci. 8*, 1484−1491.
47. Wain, R., Redfield, C., Ferguson, S. J., and Smith, L. J. (2004) NMR analysis shows that a *b*-type variant of *Hydrogenobacter thermophilus* cytochrome c_{552} retains its native structure, *J. Biol. Chem. 279*, 15177−15182.
48. Aurora, R., and Rose, G. D. (1998) Helix capping, *Protein Sci. 7*, 21−38.
49. Rodriguez, J. C., Wilks, A., and Rivera, M. (2006) Backbone NMR assignments and H/D exchange studies on the ferric azide- and cyanide-inhibited forms of *Pseudomonas aeruginosa* heme oxygenase, *Biochemistry 45*, 4578−4592.
50. Cowley, A. B., Lukat-Rodgers, G. S., Rodgers, K. R., and Benson, D. R. (2004) A possible role for the covalent heme−protein linkage in cytochrome *c* revealed via comparison of *N*-acetylmicroperoxidase-8 and a synthetic, monohistidine-coordinated heme peptide, *Biochemistry 43*, 1656−1666.
51. Trent, J. T., III, Kundu, S., Hoy, J. A., and Hargrove, M. S. (2004) Crystallographic analysis of synechocystis cyanoglobin reveals the structural changes accompanying ligand binding in a hexacoordinate hemoglobin, *J. Mol. Biol. 341*, 1097−1108.
52. Lesk, A. M., and Chothia, C. (1980) Solvent accessibility, protein surfaces, and protein folding, *Biophys. J. 32*, 35−47.
53. Lecomte, J. T. J., Vuletich, D. A., and Lesk, A. M. (2005) Structural divergence and distant relationships in proteins: Evolution of the globins, *Curr. Opin. Struct. Biol. 15*, 290−301.
54. Jamin, M. (2005) The folding process of apomyoglobin, *Protein Pept. Lett. 12*, 229−234.
55. Nishimura, C., Lietzow, M. A., Dyson, H. J., and Wright, P. E. (2005) Sequence determinants of a protein folding pathway, *J. Mol. Biol. 351*, 383−392.
56. Nishimura, C., Prytulla, S., Dyson, H. J., and Wright, P. E. (2000) Conservation of folding pathways in evolutionarily distant globin sequences, *Nat. Struct. Biol. 7*, 679−686.
57. Li, R., and Woodward, C. (1999) The hydrogen exchange core and protein folding, *Protein Sci. 8*, 1571−1590.
58. Woodward, C., Carulla, N., and Barany, G. (2004) Native state hydrogen-exchange analysis of protein folding and protein motional domains, *Methods Enzymol. 380*, 379−400.
59. Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., and Wright, P. E. (1993) Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin, *Biochemistry 32*, 6337−6347.
60. Pesce, A., Nardini, M., Dewilde, S., Geuens, E., Yamauchi, K., Ascenzi, P., Riggs, A. F., Moens, L., and Bolognesi, M. (2002) The 109 residue nerve tissue minihemoglobin from *Cerebratulus lacteus* highlights striking structural plasticity of the α-helical globin fold, *Structure 10*, 725−735.
61. Nakajima, S., Alvarez-Salgado, E., Kikuchi, T., and Arredondo-Peter, R. (2005) Prediction of folding pathway and kinetics among plant hemoglobins using an average distance map method, *Proteins 61*, 500−506.

BI061532G