# X-ray and Visible Absorption Spectroscopy of Wild-Type and Mutant T4 Gene 32 Proteins: $His^{64}$, not $His^{81}$, Is the Non-Thiolate Zinc Ligand

**Juqian Guo,[†] Shengke Wang,[‡,§] Jun Dong,[‡] Huawei Qiu,[†] Robert A. Scott,*[,‡] and David P. Giedroc*[,†]**

*Contribution from the Department of Biochemistry and Biophysics, Center for Macromolecular Design, Institute for Biosciences and Technology, Texas A&M University, College Station, Texas 77843-2128, and Center for Metalloenzyme Studies, Chemistry Building, University of Georgia, Athens, Georgia 30602-2556*

*Received September 6, 1994*[⊗]

**Abstract:** In bacteriophage T4 gene 32 protein (gp32), three sulfur donor atoms, derived from $Cys^{77}$, $Cys^{87}$, and $Cys^{90}$, provide ligands to the Zn(II) ion. Each of the two histidines in the molecule, $His^{81}$ and $His^{64}$, has been variously proposed to donate the fourth non-thiolate ligand to complete the first coordination shell. In this paper, X-ray and visible absorption spectroscopies have been used to distinguish between these possibilities. The Zn X-ray absorption edge and EXAFS spectra of wild-type and H81A ($His^{81} \rightarrow$ Ala) gp32s are virtually identical and most consistent with one histidine imidazole ligand with a Zn−N bond distance of 2.06 Å and three S-containing ligands at 2.33 Å. In contrast, as expected for a change in the first coordination shell, the Zn X-ray absorption edge and EXAFS spectra of H64C gp32 are easily distinguished from the wild-type and H81A proteins. In particular the Fourier transform (FT) peaks at $R' \approx 3$ and 4 Å, attributed to the outer shell atoms of histidine imidazole rings, are absent in this mutant. Detailed curve fitting of the EXAFS data, however, do not readily distinguish between $S_3N_1$, $S_3imid_1$, and $S_4$ fits. Therefore, Co(II)-substituted H64C gp32 was also characterized. The ligand field transition envelope of the optical absorption spectrum of Co(II) H64C gp32 is significantly red-shifted relative to the wild-type protein, coupled with pronounced changes in the $-S^- \rightarrow$ Co(II) LMCT region, both consistent with an additional sulfur donor atom to the Co(II) ion. Other substitutions at residue position 64 (H64D and H64L gp32) do not give rise to the same spectral changes. The Co EXAFS of the Co(II) H64C derivative also reveals the apparent absence of outer shell imidazole scattering. These data strongly suggest that $His^{64}$ in the wild-type protein (and $Cys^{64}$ in H64C gp32) completes the zinc coordination in T4 gene 32 protein.

## Introduction

Gene 32 protein from bacteriophage T4 is the prototype member of a large class of single-strand-specific DNA binding proteins which play accessory roles in DNA replication, recombination, and repair.[1] These proteins generally bind nonspecifically to nucleic acids and the binding is moderately to highly cooperative. The above processes are all characterized by transient formation of single-stranded DNA. The cooperative binding of gene 32 protein monomers to regions of ssDNA results in formation of clusters which are thought to be the functional form of the molecule.

In 1986 it was discovered that gene 32 protein (gp32) isolated from T4-infected *E. coli* or from a recombinant source contained a single tightly bound Zn(II) ion per 33 kD monomer by atomic absorption spectroscopy.[2] Limited proteolysis by trypsin to remove the N-terminal 21 amino acids (B-domain) and the C-terminal 48 amino acids (A-domain) to yield the core fragment of the molecule [gp32-(A+B)] had no effect on Zn(II) coordination by gene 32 protein, thus confining the ligands to the amino acids 22−254.[2] The UV−visible absorption spectrum of the Co(II)-substituted gene 32 protein[2] and $^{113}Cd$ NMR studies of the $^{113}Cd$-substituted protein[3] were consistent with tetrahedral ligation by the side chain donor atoms of three cysteines and a fourth non-thiolate donor atom. The three cysteine ligand donors were unambiguously identified as $Cys^{77}$, $Cys^{87}$, and $Cys^{90}$.[3] Inspection of the primary structure of gene 32 protein[4] revealed that $His^{81}$ was found within this cluster of cysteines. On the basis of amino acid sequence comparisons with other structural zinc sites in zinc-finger nucleic acid binding proteins, in particular the retroviral-type zinc-finger formed by a Cys-$X_2$-Cys-$X_4$-His-$X_4$-Cys (where X = any amino acid) sequence, $His^{81}$ was proposed to complete the metal coordination complex in gp32.[2,5] This would form a $Cys^{77}$-$X_3$-$His^{81}$-$X_5$-$Cys^{87}$-$X_2$-$Cys^{90}$ zinc-liganding sequence derived from 14 contiguous amino acids in the primary structure, similar to the retroviral-type site. $His^{64}$, just N-terminal to this region and the only other histidine in the molecule, was not initially considered as a strong candidate for the non-thiolate donor group in gp32,[3] although it was clear that conformational changes extended to the side chain protons of $His^{64}$ upon zinc removal from gp32.[6]

In this study, we have applied X-ray absorption and UV−visible spectroscopies to probe the first coordination shell of

* To whom correspondence should be addressed.
† Texas A&M University.
‡ University of Georgia.
§ Current address: Department of Applied Science, Building 801, Brookhaven National Laboratory, Upton, NY 11973.
⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1995.

(1) Karpel, R. L. In *The Biology of Nonspecific DNA−Protein Interactions*; edited by Revzin, A., Ed.; CRC Press, Inc.: Boca Raton, FL, 1990; pp 103−130.
(2) Giedroc, D. P.; Keating, K. M.; Williams, K. R.; Konigsberg, W. H.; Coleman, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8452.
(3) Giedroc, D. P.; Johnson, B. A.; Armitage, I. M.; Coleman, J. E. *Biochemistry* **1989**, *28*, 2410.
(4) Williams, K. R.; LoPresti, M. B.; Setoguchi, M. *J. Biol. Chem.* **1981**, *256*, 1754.
(5) Berg, J. M. *Science* **1986**, *232*, 485.
(6) Pan, T.; Giedroc, D. P.; Coleman, J. E. *Biochemistry* **1989**, *28*, 8828.

0002-7863/95/1517-9437$09.00/0 

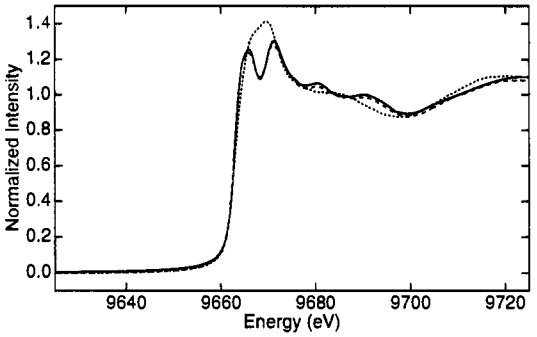


**Figure 1.** Comparison of Zn K-edge spectra for wild-type gp32-(A+B) (—), H81A gp32-(A+B) (— —), and H64C gp32 (- - -).

ligands around the Zn(II)- and Co(II)-substituted wild-type and two mutant gene 32 proteins. One mutant protein contains a nonconservative and nonliganding substitution of the previously proposed ligand $His^{81}$, $His^{81} \rightarrow$ Ala forming H81A gp32, while the other contains a substitution of $His^{64} \rightarrow$ Cys to form H64C gp32. Previous experiments have shown that contrary to simple predictions of the widely accepted original model, substitution of $His^{81}$ is remarkably silent.[7,8] For example, the UV−visible absorption spectrum of Co(II)-H81A gp32 and $^{113}Cd$ NMR spectrum of $^{113}Cd$-H81A gp32 are very nearly indistinguishable from that of the wild-type protein.[7] H81A gp32 is fully functional in an *in vitro* biochemical assay which probes early steps in homologous recombination.[8] In contrast, the reduced and metal-free protein is inactive in this assay, revealing that formation of the metal chelate in gene 32 protein is firmly coupled to biological function.[9] These findings for H81A gp32, while compelling, do not firmly rule out an unanticipated consequence of substitution of a liganding $His^{81}$, e.g., a new solvent ligand[7,8] or other nonnative protein ligand fulfilling the role played by $His^{81}$ in the wild-type protein. On the other hand, a number of $His^{64}$ mutants of gp32 have been characterized which lack Zn(II) upon purification and also bind to ssDNA substrates weakly.[7] X-ray absorption spectroscopic experiments presented here provide evidence for a novel Zn(II)-coordination complex formed by the $His^{64}$-$X_{12}$-$Cys^{77}$-$X_9$-$Cys^{87}$-$X_2$-$Cys^{90}$ sequence in gp32.

## Results and Discussion

Figure 1 compares the Zn K X-ray absorption edge spectra obtained for the wild-type gp32-(A+B), H81A gp32-(A+B), and H64C gp32. These spectra clearly show that the wild-type and H81A gp32s exhibit virtually indistinguishable edge spectra. In contrast, the edge spectrum obtained for H64C gp32 is markedly different. This spectrum alone can be taken to suggest that the first shell of ligands is clearly perturbed in the $His^{64}$ substitution mutant, rather than in the $His^{81}$ mutant.

The experimental $k^3$ EXAFS spectra and Fourier transforms (FTs) of the wild-type gp32-(A+B), H81A gp32-(A+B), and H64C gp32 are shown in Figure 2. Again, the H81A and wild-type proteins give rise to virtually identical EXAFS and FT spectra, indicating insignificant structural changes at the Zn(II) site upon alanine substitution for $His^{81}$. In particular, the FT peaks at ~3 and ~4 Å arising from the outer-shell atoms of histidine imidazole rings (indicated by the vertical arrows) remain in the H81A variant. We do not expect curve fitting of the first shell Zn EXAFS to be able to distinguish $ZnS_4$ and $ZnS_3N_1$ environments and the results of such an analysis (Table 1) confirm these expectations.

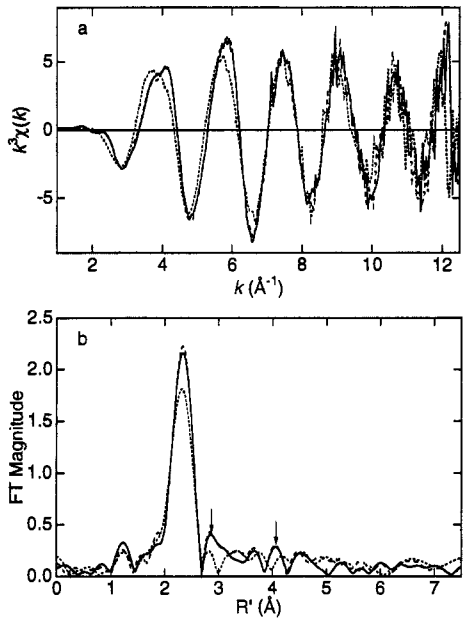


**Figure 2.** (a) Experimental $k^3$-weighted EXAFS spectra for wild-type gp32-(A+B) (—), H81A gp32-(A+B) (— —), and H64C gp32 (- - -). (b) Zn EXAFS Fourier transforms ($k^3$ weighting, $k = 1-12.7$ Å$^{-1}$) for wild-type gp32-(A+B) (—), H81A gp32-(A+B) (— —), and H64C gp32(- - -).

**Table 1.** Curve Fitting of First-Shell Filtered Zn and Co EXAFS for T4 Gene 32 Proteins[a]

| sample | fit | shell | $N_s$ | $R_{as}$ (Å) | $\sigma_{as}^2$ (Å$^2$) | $\Delta E_0$ (eV) | $f'$ [b] |
|---|---|---|---|---|---|---|---|
| [Zn] wild type gp32 | 1 | Zn−S | (4)[c] | 2.32 | 0.0060 | 0.6 | 0.027 |
| | 2 | Zn−S | (3) | 2.33 | 0.0041 | 1.5 | 0.031 |
| | | Zn−N | (1) | 2.06 | 0.0074 | (0.0) | |
| [Zn] H81A gp32 | 3 | Zn−S | (4) | 2.32 | 0.0058 | 0.7 | 0.034 |
| | 4 | Zn−S | (3) | 2.33 | 0.0037 | 0.8 | 0.030 |
| | | Zn−N | (1) | 2.04 | 0.0062 | (0.0) | |
| [Zn] H64C gp32 | 5 | Zn−S | (4) | 2.32 | 0.0071 | −2.8 | 0.052 |
| | 6 | Zn−S | (3) | 2.32 | 0.0045 | −3.1 | 0.038 |
| | | Zn−N | (1) | 2.03 | 0.0030 | (0.0) | |
| [Co] H64C gp32 | 7 | Co−S | (4) | 2.27 | 0.0066 | 1.4 | 0.032 |
| | 8 | Co−S | (3) | 2.27 | 0.0045 | 0.5 | 0.024 |
| | | Co−N | (1) | 1.98 | 0.0081 | (0.0) | |

[a] $N_s$ is the number of scatterers per metal. $R_{as}$ is the melt−scatterer distance, $\sigma_{as}^2$ is a mean-square deviation in $R_{as}$. $\Delta E_0$ is the shift in $E_0$ for the theoretical scattering functions. [b] $f'$ is a normalized error ($\chi^2$): $f' = \{\sum_i[k^3(\chi_{obs}(i) - \chi_{calc}(i))]^2/N\}^{1/2}/[k^3(\chi_{max} - \chi_{min})]$. [c] Numbers in parentheses were not varied during optimization.

The XAS data strongly suggest that the first coordination shells of ligands about the native Zn(II) ion in the wild-type and H81A proteins are identical, with both sites containing a single imidazole ligand (judging from the presence of the outer-shell FT peaks in Figure 2b). Since H81A gp32 contains only one His residue, $His^{64}$, this requires that $His^{64}$ provide a coordination bond to the native Zn(II) ion in gp32. Consistent with this interpretation, the EXAFS spectrum and FT obtained for H64C gp32 is markedly different (Figure 2). Although the major FT peak at 2.32 Å remains, indicative of Cys ligation, the weak outer-shell peaks attributable to imidazole coordination appear to be lacking in this derivative. This is consistent with the coordination complex incorporating $Cys^{64}$ in place of $His^{64}$ in H64C gp32 as a first-shell ligand to the Zn(II) ion.

We prepared Co(II)-substituted H64C gp32 to obtain an independent evaluation of the changes induced in the first coordination shell of Co(II) as a result of replacing $His^{64}$ with

(7) Giedroc, D. P.; Qiu, H.; Khan, R.; King, G. C.; Chen, K. *Biochemistry* **1992**, *31*, 765.
(8) Qiu, H.; Giedroc, D. P. *Biochemistry* **1994**, *33*, 8139.
(9) Qiu, H.; Kodadek, T.; Giedroc, D. P. *J. Biol. Chem.* **1994**, *269*, 2773.

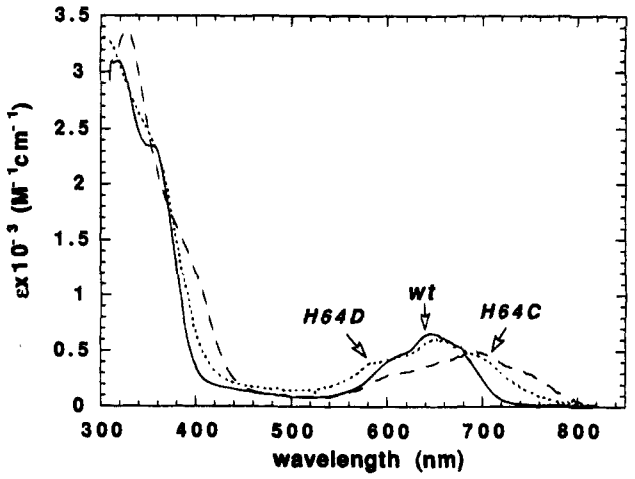


**Figure 3.** Visible absorption spectra of Co(II)-substituted wild-type gp32, H64C gp32, and H64D gp32 at 25.0 ± 0.1 °C and in 20 mM HEPES, pH 7.5, 0.1 M KCl.

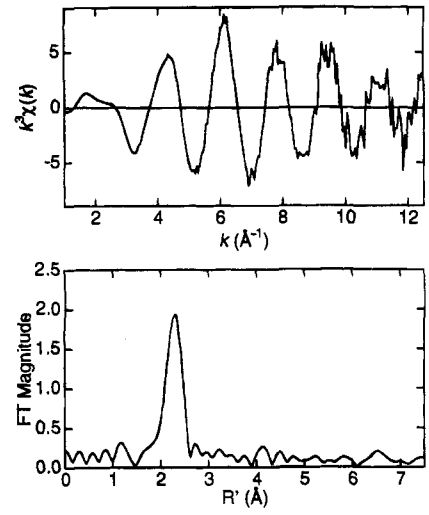


**Figure 4.** Co EXAFS (top) and Fourier transforms ($k^3$ weighting, $k$ = 1−12.7 Å$^{-1}$) (bottom) for Co(II)-substituted H64C gp32.

Cys. Co(II) is widely used as a structural surrogate for Zn(II) in zinc metalloproteins due to its favorable electronic properties.[10,11] The ligand-field transition region ($\lambda$ = 550−800 nm) of tetrahedral, high-spin Co(II) is extremely sensitive to the changes in the nature, number, and coordination geometry of the first shell of ligands.[2] In addition, Co EXAFS of the Co(II)-substituted protein can provide data which are complementary to UV−visible absorption spectroscopy.

Mutagenesis of His$^{64}$ to Asn, Asp, Gln, Leu, or Cys occurs concomitant with easily detectable changes in the conformation of the protein, since these mutants as a group are isolated in a zinc-free form when purified using conventional techniques.[7] It is now known that the metal-liganding cysteines (Cys$^{77}$, Cys$^{87}$, and Cys$^{90}$) are readily oxidized by molecular oxygen in the absence of bound metal, which irreversibly blocks metal binding.[8] However, the oxidized cysteines can be reduced upon incubation with a high concentration of reducing agent, and a metal complex can be formed in an anaerobic environment. The Co(II) derivatives of H64C and H64D (His 64 → Asp) gp32s were similarly prepared to evaluate whether spectral changes could be attributable to Cys−S$^-$ ligation or to more general alterations in the geometry of the first coordination shell.

The UV−visible absorption spectra of Co(II)-substituted H64C and H64D gene 32 proteins are shown in Figure 3, compared to that previously published for the wild-type protein.[2] The blue-green Co(II)-substituted H64C protein is characterized by a strongly red-shifted "d−d" visible absorption envelope (≈600−700 nm) and contains additional lower energy alterations in the S$^-$ → Co(II) ligand-to-metal charge transfer region ($\lambda \leq$ 450 nm). These pronounced spectral shifts are clearly indicative of a perturbation in the first coordination shell of ligands about the Co(II) ion in H64C gp32, as expected for the participation of His$^{64}$ in the wild-type complex.[11] In particular, the transition at 750 nm is nearly always associated with Co(II)−S$_4$ complexes.[11-14] These changes are consistent with the formation of a new tetrahedral, tetrathiolate coordination complex formed by the side chains of Cys$^{64}$, Cys$^{77}$, Cys$^{87}$, and Cys$^{90}$.[11-14]

In contrast, the ligand-field transition region of the spectrum of the H64D derivative is clearly perturbed relative to wild-type, but it is only modestly shifted to lower energy. The 750-nm transition is notably absent. The present data do not allow us to say with confidence that the side chain of Asp$^{64}$ forms a coordination bond to the Co(II) ion since a non-liganding amino acid substitution here (in H64L gp32) gives rise to the same general spectral perturbations (data not shown). They do show, however, that perturbations induced by Cys are easily distinguished from imidazole and other non-thiolate ligands and are therefore most consistent with a new Cys64−Co(II) bond in H64C gp32.[11]

As expected from the Zn EXAFS FT of Zn(II) H64C gp32, Co EXAFS FT of Co(II)-substituted H64C gp32 reflects the apparent absence of any outer-shell imidazole scattering (Figure 4). Curve-fitting analysis shows that a single shell of S-containing ligands at 2.27 Å gives a reasonably good fit to the EXAFS data (fit 7 in Table 1), although, again, as expected with the additional parameters, $CoS_3N_1$ fits (fit 8) give a modest improvement in the quality of the fit. The molar absorptivity of the ligand-field complex in the UV−visible spectrum and the magnitude of the 1s → 3d pre-edge peak in the Co edge spectrum (not shown) suggest that approximately tetrahedral coordination by four thiolates is the most probable geometry. The simplest interpretation of these data is that Cys$^{64}$ in the mutant protein forms a new coordination bond to the Co(II) ion and can replace the coordination bond provided by His$^{64}$ in the wild-type protein.

## Summary and Conclusions

Contrary to predictions made from amino acid sequence comparisons, we provide spectroscopic evidence that Zn(II) binds to the His$^{64}$-X$_{12}$-Cys$^{77}$-X$_9$-Cys$^{87}$-X$_2$-Cys$^{90}$ sequence in gp32, and therefore utilizes a novel HCCC coordination strategy not found in any other zinc-containing single-stranded or duplex nucleic acid zinc-finger protein. Zn−ligand bond lengths in gp32 of 2.06 Å (Zn−N) and 2.33 Å (Zn−S) are well within the range expected for a tetrahedral Zn(II) complex, anticipated from the optical spectrum of Co(II)-substituted gp32. We believe that these experiments provide an excellent illustration that X-ray absorption spectroscopy can be a powerful tool to assess the nature of changes in the first coordination shell of ligands to the native Zn(II) ion as a result of site-directed

(10) Coleman, J. E.; Giedroc, D. P. In *Metal ions in biological systems. Vol. 25. Interrelations among metal ions, enzymes, and gene expression*; (Sigel, H., Sigel, A., Eds.; Marcel Dekker, Inc.: New York and Basel, 1989; pp 171−234.

(11) Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937.

(12) Corwin, D. T., Jr.; Gruff, E. S.; Koch, S. A. *Inorg. Chim. Acta* **1988**, *151*, 5.

(13) Garmer, D. R.; Krauss, M. *J. Am. Chem. Soc.* **1993**, *115*, 10247.

(14) Formicka-Kozlowska, G.; Schneider-Bernlöhr, H.; von Wartburg, J.-P.; Zeppezauer, M. *Eur. J. Biochem.* **1988**, *173*, 281.

mutagenesis. In many ways, Zn XAS is complementary to the well-established use of Co(II) as an optical spectroscopic surrogate probe for the native Zn(II) ion in zinc metalloproteins, with the important distinction that the native Zn(II) complex can be investigated directly.

## Experimental Section

**Proteins.** Recombinant wild-type and mutant gp32s were expressed and purified as described previously.[7,8] Wild-type and H81A gp32s containing 0.97 and 0.93 g-atom of Zn(II) by atomic absorption were subjected to limited proteolysis by trypsin to obtain the gp32-(A+B) core fragment as described.[2] Wild-type and H81A gp32-(A+B) in 10 mM Tris−HCl, pH 8.1, 0.1 M NaCl, and 5% (v/v) glycerol were concentrated to ≈2 mM using an Amicon ultrafiltration cell and dialyzed exhaustively against 10 mM $Na_2HPO_4$, pH 8.0, 30 mM NaCl and 10% (v/v) glycerol. Glycerol (50 μL) was added to 150 μL of each protein solution to achieve 1.82 mM wild-type and 1.64 mM H81A Zn(II) gp32-(A+B) in 7.5 mM $Na_2HPO_4$, 22 mM NaCl, 35% (v/v) glycerol, pH 8.0, and the solutions were loaded into two XAS cuvettes. H64C gp32 was prepared as described previously.[8] H64D gp32 was prepared from an overproducing strain of *E. coli* TB1 transformed with $pP_L$øgp32.H64D by heat induction of the phage $\lambda P_L$ promoter.[8] Both proteins were purified through the ssDNA-cellulose chromatography step and concentrated to ≈1 mM, dithiothreitol was added to 0.1 M, and the sample was incubated overnight at 4 °C. This protein solution was then degassed and brought into an anaerobic glovebox and dialyzed exhaustively (4 changes × 500 mL) against 7.0 mM $Na_2HPO_4$, 20 mM NaCl, 31% (v/v) glycerol, pH 8.0. This procedure afforded 1 to 1.5 mM metal-free reduced H64C gp32. Next, 5.0 μL of 64.5 mM specpure (Johnson-Matthey) $CoCl_2$ (1.0 molar equiv) was added to 200 μL of this protein to create the Co(II)-substituted protein. The Zn(II) derivative was similarly prepared.

**X-ray Absorption Spectroscopy.** For XAS, the protein was loaded into a capped XAS cuvette, removed from the glovebox, and immediately frozen in liquid $N_2$ to exclude oxygen. X-ray absorption spectroscopic (XAS) data were collected as fluorescence excitation spectra on frozen (10 K) samples of Zn(II)-wild-type gp32-(A+B), Zn(II)-H81A gp32-(A+B), Zn(II)-H64C gp32, and Co(II)-H64C gp32 at the Stanford Synchrotron Radiation Laboratory, beam line 7−3, with the SPEAR ring running at 3.0 GeV, 50−90 mA. A 13-element array of intrinsic germanium solid-state fluorescence detectors (Canberra) was used and individual detector channel count rates were kept below 35 kcps. The data were reduced and analyzed using the EXAFSPAK programs written by G. George, SSRL. Single-scattering first-shell fits made use of theoretical scattering parameters.[15] Refinement consisted of optimizing Debye−Waller $\sigma_{as}^2$ values and distances for each shell.

**Visible Absorption Spectroscopy.** Co(II)-substituted gp32s were similarly prepared as described for XAS samples, loaded into an anaerobic quartz cuvette in the glovebox, and subjected to UV−visible absorption spectroscopy on a Hewlett-Packard 7452 spectrophotometer at 25.0 ± 0.1 °C. The spectrum of the metal-free apo protein was subtracted from the spectrum obtained for the Co(II)-saturated gp32s.

**Acknowledgment.** D.P.G. acknowledges support from National Institutes of Health (GM 42569). XAS studies in the R.A.S. group are supported by the National Institutes of Health (GM 42025) and by the NSF Research Training Group Award to the Center for Metalloenzyme Studies (DIR 90-14281). The XAS data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL), which is operated by the Department of Energy, Division of Chemical Sciences. The SSRL Biotechnology Program is supported by the National Institutes of Health, Biomedical Resource Technology Program, Division of Research Resources. Support for the X-ray fluorescence detector is from NIH BRS Shared Instrumentation Grant RR05648. D.P.G. thanks Dr. Paul Lindahl, Department of Chemistry, Texas A&M University for use of his glovebox and many helpful suggestions.

JA942956D

(15) McKale, A. G.; Veal, B. W.; Paulikis, A. P.; Chan, S.-K.; Knapp, G. S. *J. Am. Chem. Soc.* **1988**, *110*, 3763.

(16) Rehr, J. J.; Zabinsky, S. I.; Alber, R. C. *Phys. Rev. Lett.* **1992**, *69*, 3397.