Zinc Site Redesign in T4 Gene 32 Protein: Structure and Stability of Cobalt(II) Complexes Formed by Wild-Type and Metal Ligand Substitution Mutants†

Juqian Guo‡ and David P. Giedroc*

Department of Biochemistry and Biophysics, Center for Macromolecular Design, Institute of Biosciences and Technology, Texas A&M University, College Station, Texas 77843-2128

Received July 18, 1996; Revised Manuscript Received October 21, 1996⊗

ABSTRACT: Phage T4 gene 32 protein (gp32) is a zinc metalloprotein which binds cooperatively and preferentially to single-stranded nucleic acids and functions as a replication and recombination accessory protein. Zn(II) coordination by gp32 employs a His-Cys_3 metal ligand donor set derived from the His^{64}-X_{12}-Cys^{77}-X_9-Cys^{87}-X_2-Cys^{90} sequence in the ssDNA-binding core domain of the molecule. Crystallographic studies reveal that His^{64} and Cys^{77} are derived from two independent β-strands within a distorted three-stranded β-sheet and are relatively more buried from solvent than are Cys^{87} and Cys^{90}, which are positioned immediately before and within, respectively, an α-helix. In an effort to understand the origin of the stability of the metal complex, we have employed an anaerobic optical spectroscopic, competitive metal binding assay to determine the coordination geometry and association constants (K_a) for the binding of Co(II) to wild-type gp32 and a series of zinc ligand substitution mutants. At pH 7.5, 25 °C, wild-type gp32 binds Co(II) with a $K_a \approx 1 \times 10^9$ M^{-1}. Competition experiments reveal that K_a for Zn(II) is 3.0 ($\pm$1.0) $\times 10^{11}$ M^{-1}. We find that all non-native metal complexes retain tetrahedral or distorted tetrahedral coordination geometry but are greatly destabilized in a manner essentially independent of whether a new protein-derived coordination bond is formed (e.g., in H64C gp32) or not. Co(II) binding isotherms obtained for three His^{64} substitution mutants, H64C, H64D, and H64N gp32s, suggest that each mutant forms a dimeric Cys_4 tetrathiolate intermediate complex at limiting $[Co(II)]_f$; each then rearranges at high $[Co(II)]_f$ to form a monomolecular site of the expected geometry and $K_a \approx 1 \times 10^4$ M^{-1}. Like the His^{64} mutants, C77A gp32 appears to form at least two types of complexes over the course of a Co(II) titration: one with octahedral coordination geometry formed at low $[Co(II)]_f$, with a second tetrahedral or five-coordinate site formed at higher $[Co(II)]_f$. Apo C87S and C90A gp32s, in contrast, each form a single complex at all $[Co(II)]_f$, consistent with Cys_2-His-H_2O tetrahedral geometry of $K_a \approx (1-2) \times 10^5$ M^{-1}. These studies reveal that the local protein structure restricts accommodation of a non-native metal complex in a ligand-specific manner. The implications of this work for *de novo* design of zinc complexes in proteins are discussed.

Zinc is the second most abundant trace metal in higher animals (Coleman, 1992). Since the discovery of the first zinc enzyme, carbonic anhydrase, about 300 zinc-containing proteins with enzymatic activity have been identified. The subsequent discovery and characterization of the first non-enzymatic Zn(II)-containing nucleic acid binding protein, transcription factor IIIA (TFIIIA) from *Xenopus* oocytes (Miller et al., 1985; Klug & Rhodes, 1987), clearly established that the biological role of zinc proteins extends to the regulation of gene expression (Berg & Shi, 1996).

TFIIIA contains nine independently folded, structurally conserved ≈30-amino acid minidomains, each characterized by tetrahedral ligation of a Zn(II) ion by two cysteine (Cys) and two histidine (His) side chains (Pavletich & Pabo, 1991, 1993; Rebar & Pabo, 1994). This structure was dubbed a "zinc finger" (Miller et al., 1985). Using TFIIIA as a model, a large number of structurally diverse transcription factors and gene regulatory proteins were subsequently identified which contained a defined arrangement of conserved Cys or Cys/His residues clustered in the primary structure which would serve as potential ligands to Zn(II). Any protein fulfilling these criteria was called a zinc finger protein. Using this broad definition, zinc finger proteins are a collection of ubiquitous and structurally diverse proteins involved in the regulation of gene expression in many biological contexts (Schmiedeskamp & Klevit, 1994).

Although individual zinc finger families differ in metal coordination strategy, secondary structure, and modularity, they all incorporate one or more structural Zn(II) ions as a means to stabilize functionally important region(s) of the molecule. As a corollary to this role, Zn(II) chelates provide a favorable thermodynamic driving force which may mediate or direct the folding of the polypeptide chain to the native state (Giedroc, 1994). It is generally agreed that a chelate of high stability and specificity for Zn(II) is critical for function (Krizek & Berg, 1992; Berg & Merkle, 1989). The high stability observed for naturally occurring metal sites in proteins (of 10^9-10^{12} M^{-1}) is determined by metal ligand bonding as well as secondary hydrogen bonding and van der Waals interactions which occur as a result of formation of

† This work was supported by grants from the NIH (GM42569) and the Robert A. Welch Foundation (A-1295) and is in partial fulfillment of the requirements of the M.S. degree at Texas A&M University (to J.G.).

* Address all correspondence to this author. Telephone: 409-845-4231. FAX: 409-862-4718. e-mail: giedroc@bioch.tamu.edu.

‡ Present address: Metabolex, Inc., Hayward, CA 94545.

⊗ Abstract published in *Advance ACS Abstracts,* January 1, 1997.

S0006-2960(96)01776-X CCC: $14.00

the chelate (Kiefer et al., 1993a,b; Alexander et al., 1993; Kiefer & Fierke, 1994; Ippolito et al., 1995). The chelate must also be relatively exchange inert or kinetically stable, thereby limiting ligand exchange reactions which might lead to zinc expulsion by biological chelators like metallothionein, or oxidation of liganding residues in the case of cysteine thiolates (Zeng et al., 1991).

How an appropriate degree of thermodynamic and kinetic stability of zinc complexes is obtained in proteins is largely unresolved. Only a few systems have been systematically investigated by redesigning or reengineering a naturally occurring metal site to one with an altered first shell ligand donor set and determining how the structural, chemical, and functional properties of the protein change upon ligand substitution. Carbonic anhydrase II (CAII), the extensively studied zinc metalloenzyme that catalyzes the reversible hydration of CO_2 to HCO_3^-, is one such example. The zinc cofactor lies at the bottom of the active site cleft where it is tetrahedrally coordinated to His^{94}, His^{96}, His^{119}, and a hydroxide ion at physiological pH. To probe the functional importance of the three histidine residues in CAII which coordinate the zinc, Fierke and co-workers have substituted these amino acids with Cys, Asp, Glu, and Ala and characterized the catalytic parameters and zinc and inhibitor binding properties of these variants (cf. Keifer & Fierke, 1994). They found that replacement of any one of the three His ligands with Cys, Asp, or Glu, or removal of a ligand by replacement with Ala, results in a $\approx 10^4$-fold or a $\approx 10^5$-fold decrease in zinc affinity, respectively. Although the CAII variants with three potential zinc ligands bind zinc more tightly than Ala mutants (by $\approx$5−20-fold), the differences are small compared to the attenuation caused by any substitution of the three His ligands with a liganding *or* nonliganding side chain. This work, in combination with the three-dimensional structures of these variants (Kiefer et al., 1993a; Ippolito & Christianson, 1994; Ippolito et al., 1995), reveals that the metal binding site of CAII is sufficiently plastic to accommodate binding of one zinc ion when any one of the three natural His zinc ligands is replaced with an alternative ligand, but that this substitution drastically reduces zinc affinity and catalytic efficiency in a manner relatively independent of the nature of the substitution.

Like CAII, TFIIIA-like zinc finger sites have also been used as a model for metal site design (Regan & Clarke, 1990; Klemba et al., 1995) or redesign (Krizek et al., 1993) experiments. TFIIIA-like zinc finger proteins are characterized by tetrahedral Cys_2His_2 zinc coordination sites. Sequence variants have been prepared in which one or both liganding His residues have been substituted with Cys. Krizek et al. (1993) investigated the metal binding affinity of a consensus zinc finger peptide and two His→Cys variants and found that all of the ligand mutants bind Co(II), Zn(II), and Cd(II) with high affinity. In striking contrast to CAII, there were only modest differences in the affinity for Zn(II) among the native Cys_2His_2 and mutant Cys_3His and Cys_4 sites ($\leq$5-fold) (Krizek et al., 1993). Considering the fact that the zinc finger is an independently folded domain in which folding is tightly coupled to Zn(II) coordination, these findings suggest that structural rearrangements that are introduced as a result of ligand substitution are small relative to the extent of conformational change required to coordinate any metal.

CA(II) and the TFIIIA-like zinc finger may represent two extreme cases of biological Zn(II) chelates. CAII is an example of a highly optimized binding site which is largely constructed by the folded conformation of the apoprotein; this site coordinates zinc in a strongly ligand-specific manner. On the contrary, TFIIIA-like zinc finger may be relatively less selective on ligands since folding is entirely coupled to zinc binding. What role do zinc ligands play in other proteins? We have used the zinc site in the prototype replicative SSB,[1] bacteriophage T4 gene 32 protein (gp32) (Giedroc et al., 1986), as a model system to address this question. Like CAII, the metal site is part of much larger protein, but like TFIIIA, plays a structural role in forming a binding surface for another macromolecule, ssDNA in this case (Giedroc et al., 1991; Shamoo et al., 1995). In contrast to both systems, spectroscopic experiments suggest there is some, but not extensive, induced polypeptide folding upon metal binding (Giedroc et al., 1987; Qiu et al., 1994). The formation of the Zn(II) chelate is directly linked to single-stranded DNA binding and functional efficacy of gp32 in DNA metabolism (Qiu et al., 1994). Zn(II) and Co(II) X-ray absorption spectroscopy, Co(II) UV−visible absorption spectroscopy of the Co(II)-substituted gp32, and ^{113}Cd NMR of the ^{113}Cd(II)-substituted wild-type and mutant gp32s reveal that Zn(II) is coordinated to a single His (His^{64}) and three Cys derived from the sequence His^{64}-X_{12}-Cys^{77}-X_9-Cys^{87}-X_2-Cys^{90} with distorted tetrahedral ligand symmetry (Giedroc et al., 1986, 1989, 1992; Qiu & Giedroc, 1994; Guo et al., 1995).

The crystal structure of gp32-(A+B) bound to ssDNA ligand and solved to 2.2 Å resolution is consistent with our spectroscopic findings (Shamoo et al., 1995). As can be seen (Figure 1), the zinc chelate may play a major role in determining the fold of the globular domain I of the molecule. His^{64} and Cys^{77} are derived from two sequential β-strands within a three-stranded β-sheet while Cys^{87} and Cys^{90} are from an α-helical region. The β-strand containing His^{64} residue links subdomain I and subdomain II and crosses the DNA binding cleft of the molecule. Although the globular structure of the metal-free apoprotein is not known to high resolution, far-UV circular dichroism (CD) spectroscopy suggests that there are only small and possibly localized changes in the structure of the molecule upon zinc removal (Giedroc et al., 1987; Qiu et al., 1994). The subtle differences among gp32 derivatives in protein conformation might lead to distinct degrees of perturbation of ssDNA binding equilibria. In addition, the local structure controlled by metal coordination is critical for both function and stability of the protein (Qiu & Giedroc, 1994). In this paper, we present systematic studies which address to what degree metal ligand substitutions are accommodated in gp32. Results from these studies may provide insight for the *de novo* design of zinc complexes in proteins.

MATERIALS AND METHODS

Materials

All buffers were prepared with doubly distilled and deionized Milli-Q water. ssDNA-cellulose was prepared as

[1] Abbreviations: gp32, T4 gene 32 protein; gp32-(A+B), gene 32 protein residues 22−253 (lacking the A and B domains); ss, single-stranded; SSB, single-strand binding protein.

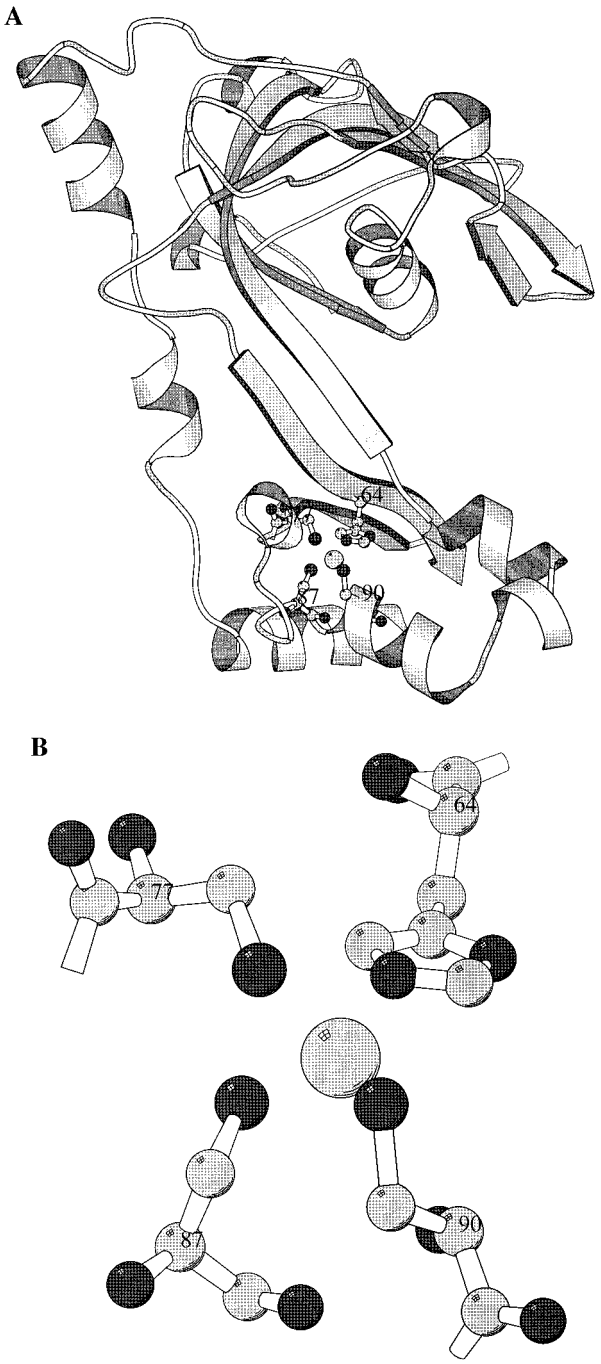

FIGURE 1: (A) Ribbon diagram of the crystallographic structure of the gp32 core fragment (ssDNA ligand not shown). In this orientation, globular subdomain I lies at the base of the molecule and contains the zinc coordination complex and subdomain II lies at the top of the molecule; these domains are connected by a connecting region (Shamoo et al., 1995). (B) A close-up of the Zn(II) coordination complex in subdomain I consistent with the metal−ligand coordination bond distances determined by EXAFS (Guo et al., 1995). The same orientation as is shown in panel A. Adapted from Shamoo et al. (1995).

described (Giedroc et al., 1990). DE-52 was obtained from Whatman while phenyl-Sepharose CL-4B was purchased from Sigma. Urea and ammonium sulfate were obtained from United States Biochemical Corp. (USB) and used without further purification. 5,5′-Dithionitrobenzoic acid (DTNB), HEPES, and chromatographically purified DNase were obtained from Sigma. Metal-free HEPES buffer was used in all metal reconstitution experiments and was prepared by passing stock solutions of 200 mM HEPES and 2 M KCl through a Chelex-100 column. Chelex-100 and dithiothreitol (DTT) were purchased from BioRad. Ultrapure cobalt chloride and zinc sulfate were from Johnson-Matthey.

Methods

Plasmid Constructions. All standard molecular biological methods were carried out according to Maniatus et al. (1982). The phagemid p32NB.6211 (Giedroc et al., 1992) has a 1.1 kb insert which is flanked by unique 5′ *Nhe*I and 3′ *Bam*HI restriction sites containing the T4 gene 32 coding sequences. This insert also contains a unique *Spe*I site at codon 60−61 and a unique *Nco*I site at codon 80−82, which was used for selection of the mutant plasmids. The parent gp32 expression plasmid is $pP_L\phi$g32.wt (Villemain & Giedroc, 1993), which when transformed into *Escherichia coli* TB1 provides for gp32 expression by derepression of λP_L cI857-mediated repression of the P_L promoter upon shift to nonpermissive (40 °C) temperature. The insert of this plasmid contains unique *Spe*I, *Nco*I, and *Bam*HI sites, which allows for subcloning of the mutant fragments from p32NB.6211 into the expression vector.

Site-directed mutagenesis was performed according to the protocol provided with the Transformer site-directed mutagenesis kit (CLONTECH Laboratories, Inc). The phagemid p32NB.6211 was used as the template for all mutagenesis experiments. The details of the construction of expression plasmids which direct the expression of individual gp32s are described below.

Construction of $pP_L\phi$g32.H64D (H64D gp32) and $pP_L\phi$g32.H64N (H64N gp32) Expression Plasmids. p32NB.6211 was subjected to site-directed mutagenesis with oligonucleotide primer JQ1 (5′-CCATTCGCAATTCTTG-TAAAT[A/G]ACGGTTTTAAAAAAAATGG-3′) which changes His^{64} codon (CAC) to Asn^{64} (AAC) or Asp^{64} (GAC), while removing the unique *Spe*I restriction site (5′-ACTAGT-3′) and adding a *Dra*I site (5′-TTTAAA-3′) by translationally silent mutations. The 210 bp *Sac*I−*Nco*I fragments from p32NB6211.H64D and H64N were subcloned into the same sites of $pP_L\phi$g32.wt to create each expression plasmid.

Construction of $pP_L\phi$g32.C77A (C77A gp32). p32NB.6211 was subjected to site-directed mutagenesis with oligonucleotide primer C77A-MS (5′-TATTGAAACAGCCTC GAGTACCCACGGTGATTACG-3′), which changes Cys^{77} codon (TGT) to Ala^{77} (GCC) while removing the unique *Nco*I site (5′-CCATGG-3′) and adding a translationally silent *Xho*I site (5′-CTCGAG-3′) for selection purposes. The 800 bp *Spe*I−*Bam*HI fragment from p32NB.6211.C77A was subcloned into the same sites of $pP_L\phi$g32.wt to create the expression plasmid.

Construction of $pP_L\phi$g32.C90A (C90A gp32). p32NB.6211 was subjected to site-directed mutagenesis with oligonucleotide primer C90A-MS (5′-TTGAAACATGCTC GAG-TACCCACGGTGATTACGATTCTTGCCCAGTA-GCTCAATACATCAG-3′), which changes Cys^{90} codon (TGT) to Ala^{90} (GCT) while removing the unique *Nco*I site (5′-CCATGG-3′) and adding a translationally silent *Xho*I site (5′-CTCGAG-3′) for selection. The 300 bp *Spe*I−*Sal*I fragment from p32NB.6211.C90A was subcloned into the same sites of $pP_L\phi$g32.wt.

Purification of gp32s. Gp32 and derivatives were inducibly expressed in Luria broth using methods outlined previously (Giedroc et al., 1989). Wild-type, H64C, and

C87S gp32s were purified from the low-speed supernatant lysis fraction using a combination of three chromatographic steps (DE-52, ssDNA-cellulose, and phenyl-Sepharose chromatographies). For the purification of H64D, H64N, C77A, and C90A gp32s, a similar procedure was followed except that ssDNA-cellulose was not routinely employed since these mutants exhibited dramatically weakened affinity for this matrix.[2] Instead, a 40% ammonium sulfate precipitation step was performed on the pool obtained prior to or following the DE-52 column chromatography step to effect removal of nucleic acids and other proteins from gp32. SDS−PAGE and UV-absorbance analysis of the gp32 derivatives used in the study indicate that all were homogeneous to an acceptable degree for the metal binding experiments (≥85%) and exhibited A_{280}/A_{260} ratios exceeding 1.7, indicating little or no contamination by endogenous nucleic acids. All mutant gp32s were obtained in a metal-free and oxidized form using this purification protocol.

Metal-free wild-type gp32 was prepared via *S*-methylation by reacting the Zn(II) protein with 50−70 molar equiv (12−15-fold over free Cys) of methyl methanethiolsulfonate (MMTS) at 4 °C (Qiu & Giedroc, 1994; Qiu et al., 1994). This protein was then subjected to exhaustive dialysis against HEPES buffer (20 mM HEPES, pH 7.5, 0.1 M KCl) at 4 °C to remove excess MMTS and Zn(II). The resulting protein was Zn(II)-free by atomic absorption spectroscopy (AAS) and contained no reactive Cys by DTNB under denaturing conditions (Qiu et al., 1994), revealing that the chemical modification reaction had gone to completion. Apo gp32 was recovered by the addition of 2 mM DTT (Qiu et al., 1994) followed by exhaustive dialysis against chelexed metal-free HEPES buffer in an anaerobic glovebox. The resulting protein was analyzed by AAS and DTNB reaction which showed that the Zn(II) content was less than 0.10 mol/monomer and the Cys residues were fully reduced. Mutant gene 32 proteins obtained in a Zn(II)-free and oxidized form were reduced by the addition of 0.1 M dithiothreitol, incubated at 4 °C overnight, and then brought into the anaerobic glovebox. Reduced, metal-free apoproteins were prepared by exhaustive dialysis against metal-free buffer in the anaerobic chamber. Following the dialysis, zinc content was determined by AAS and the number of reduced thiols quantitated by DTNB reactivity under denaturing conditions (Giedroc et al., 1992).

Atomic Absorption Spectroscopy. The Zn(II) content of gp32s was determined with a Perkin-Elmer Model 2380 atomic absorption spectrophotometer operating in the flame mode with detection at 213.7 nm (slit = 0.7 nm). The zinc content of reduced and "metal-free" gp32s used for Co(II)-binding isotherms was found to range from 0.02 to 0.31 mol of Zn(II) per mole of gp32 monomer (Table 1). The Co(II) concentration in the titrant or reconstituted proteins was measured in the same way as Zn(II) content except that a cobalt hollow cathode lamp (Perkin-Elmer) was used with detection wavelength at 240.7 nm (slit = 0.2 nm).

Table 1: Characterization of the Apo gp32s Used in This Study

gp32	stock concn (mM)	Zn(II) content (mol/monomer)	no. of free thiols per monomer (expected)
wild-type intact	0.12	0.05	3.9 (4)
wild-type core	0.22	0.03	3.5 (4)
H64C	0.45	n.d.[a]	5.6 (5)
H64D	0.28	0.02	3.3 (4)
H64N	0.26	0.18	3.8 (4)
C77A	0.28	0.24 ± 0.02	3.0 (3)
C87S	0.31	0.12	2.9 (3)
C90A	0.22	0.31	2.9 (3)

[a] Not determined.

Free Thiol Quantitation. The number of free thiols in wild-type and mutant gp32 preperations was determined using the DTNB colorimetric assay. For these reactions, protein samples were diluted to 10−15 μM in 400 μL of 7 M urea/2.5 mM EDTA solution in order to denature the protein, and 25 μL of 2.5 mM DTNB solution was added. For the apoproteins, the reaction was incubated in the glovebox for 30 min before being removed from the glovebox to ensure that the reaction was complete prior to exposure to air. The reaction of the protein with DTNB was monitored by recording the UV−visible spectrum from 240 to 600 nm. The molar concentration of liberated aromatic thiolate anion and thus reduced thiol groups in the protein was quantified at 412 nm using $\epsilon_{412} = 13\,600\ M^{-1}\ cm^{-1}$ after subtraction of the absorbance obtained with buffer alone. The analytical properties of the apoproteins used in this study are summarized in Table 1.

Co(II) Binding Titrations of Wild-Type, C87S, and C90A gp32s. All metal binding experiments were carried out in 20 mM HEPES, pH 7.5, and 0.1 M KCl, 25 °C, unless otherwise noted. Apo wild-type or mutant gp32 (0.8−1.0 mL of 100−200 μM) was prepared in the presence or absence of 1.0 mM potassium EGTA and loaded into an anaerobic cuvette. About 200 μL of Co(II) titrant was taken up into an adjustable volume Hamilton syringe. The syringe was attached to the cuvette to create an O_2-free sealed environment and removed from the glovebox. Optical spectra of the apoprotein and following each *i*th addition of a known aliquot of Co(II) titrant (10−15 μL) were collected on a Hewlett-Packard 8452A spectrophotometer at 25.0 ± 0.1 °C. Corrected spectra were obtained by subtraction of the starting apoprotein spectrum and correction for dilution. Following each experiment, the concentration of the protein in the cuvette was determined from $\epsilon_{280} = 4.13 \times 10^4\ M^{-1}\ cm^{-1}$ and Co(II) in the titrant by atomic absorption spectroscopy.

From the Co(II) binding titrations, the average number of moles of Co(II) bound per mole of protein, ν, was determined by

$$\nu = A_i\,(\max \lambda)/A_{\max}\,(\max \lambda) \qquad (1)$$

where $A_{\max}$ is the corrected absorption of fully substituted Co(II) protein at the absorption maximum in the ligand field envelope (e.g., max λ = 646 nm for wild type gp32) (Giedroc

[2] All mutant gp32s studied here are known or projected to bind weakly to single-stranded nucleic acids. The binding affinities of H64N gp32-(A+B) and C87S gp32 for poly(dT) are 3−4 orders of magnitude smaller than for wild-type gp32-(A+B) (Giedroc et al., 1992). Unlike Co(II)-reconstituted wild-type gp32 (Giedroc et al., 1987), freshly reconstituted Co(II)-substituted H64C and H64D gp32s are unable to destabilize the partially double-stranded alternating copolymer, poly-[d(A-T)], even under very low salt conditions where the duplex is strongly destabilized (15 mM KCl, pH 7.5) (J. Guo and D. P. Giedroc, unpublished observations), requiring that K_a be reduced by at 3 orders of magnitude from a comparison with other mutant gp32s (Villemain & Giedroc, 1993). Similar perturbations in ss nucleic acid binding affinity are projected to characterize C77A and C90A gp32s, although these have not been measured.

et al., 1986) and A_i is the corrected absorption following the ith addition of Co(II). This analysis assumes that the change in absorption is linearly related to a change in ν. At each ν, the concentration of free Co(II), $[Co(II)]_f$, is given by the positive root of the following equation (Mely et al., 1991):

$$K_E[Co(II)]_f^2 + \{K_E([E]_t - [Co(II)]_t + \nu[P]_t) + 1\}[Co(II)]_f + \nu[P]_t - [Co(II)]_t = 0 \quad (2)$$

where $[E]_t$, $[Co(II)]_t$, and $[P]_t$ are respectively the total concentrations of EGTA, Co(II), and active apoprotein at the ith addition, while K_E is the affinity of EGTA for Co(II) under these solution conditions, or 1.41×10^9 M^{-1} (Martell & Smith, 1974; Mely et al., 1991). The experimental values of [Co] and ν allow nonlinear least squares parameter estimation of K_a from the Langmuir equation

$$\nu = K_a[Co(II)]_f/(1 + K_a[Co(II)]_f) \quad (3)$$

For titrations carried out in the absence of chelator EGTA, [Co] was determined from

$$[Co(II)]_f = [Co(II)]_t - \nu[P]_t \quad (4)$$

and the data subjected to the same analysis to determine K_a.

For titrations carried out in the absence of EGTA, a second method of analysis was also employed (Klemba & Regan, 1995). First, A_{max} was determined by fitting the double-reciprocal plot of $1/[Co(II)]_{t,i}$ vs $1/A_i$, using the following equation

$$1/[Co(II)]_{t,i} = [(m_1)(1/A_i)(1 - m_2A_i)]/[(m_2 + m_1)(m_3)(1 - m_2A_i)] \quad (5)$$

where A_i is the corrected absorption at max λ, $[Co(II)]_{t,i}$ is the total Co(II) concentration present after each ith addition of Co(II) titrant, and m_1, m_2, and m_3 were left to vary during the parameter optimization. $m_2 = 1/A_{max}$, or the reciprocal of the absorption of fully substituted Co(II)−protein at max λ. A_{max} determined in this way was then used to calculate the fractional molar saturation, ν_i, at each ith addition of Co(II) during the titration from eq 1. This series of ν_i permits a plot of CV/ν_i vs $(V_0 + V)/(1 - \nu_i)$, where C is the concentration of Co(II) in the titrant, V_0 is the starting volume (in μL) of the protein solution, and V is the cumulative volume of titrant (in μL) added at each ν_i. This plot was subjected to linear least squares optimization, which returns a slope of $1/K_a$ and a y-intercept of P_tV_0, where P_t is the total concentration of active apoprotein competent to bind metal.

Zn(II) Titrations. Since Zn(II) is spectroscopically silent, a competition [Zn(II) vs Co(II)] experiment was employed to determine K_a of wild-type gp32 for Zn(II) (Berg & Merkle, 1989; Klemba & Regan, 1995). Apo wild-type gp32 (0.8 mL of 85 μM) was mixed with 206 μL of 13.7 mM $CoCl_2$ to form 67.6 μM Co(II)-substituted gp32 in an anaerobic glovebox with large excess of Co(II). Optical spectra were recorded following each addition of $ZnSO_4$ titrant delivered from an adjustable volume Hamilton syringe. The molar concentration of Co(II)−gp32 was calculated from the known extinction coefficient of 660 M^{-1} cm^{-1} at 646 nm (Giedroc et al., 1986) and Zn(II) in the titrant by atomic absorption spectroscopy. The data were then analyzed according to a competition model which assumes that 1:1 protein−Co(II) and protein−Zn(II) complexes are the only forms of gp32−metal complex present in solution (Berg & Merkle, 1989).

Co(II)-Binding Titrations of H64C, H64D, H64N, and C77A gp32s. Co(II) titrations were performed in the same way as described above for C87S and C90A gp32. In contrast to the wild-type, C87S, and C90A gp32s, the overall shape of the spectrum changed as the total Co(II) concentration increased during the titration, suggesting the formation of two or more distinct complexes during the course of a titration; this prevented the determination of the stability of the Co(II) complex from the direct Co(II) titrations. A third competitive binding assay was therefore employed to estimate K_a. For each of these four mutants, 0.8 mL of 100−300 μM Co(II)-reconstituted gp32 was prepared in 20 mM potassium HEPES, pH 7.5, 0.1 M KCl in the anaerobic chamber with a total Co(II) concentration of ≈4 molar equiv relative to the protein concentration. Optical spectra were collected as above at 25.0 ± 0.1 °C of the fully saturated Co(II)-coordinated protein and following each ith addition of an aliquot of chelator titrant delivered anaerobically from an adjustable volume Hamilton syringe. The fractional change in the absorption in visible region is directly proportional to the fraction of sites with bound metal, ν, which permits determination of K_a upon application of a 1:1 binding model. From the EGTA binding titrations, the average number of moles of Co(II) bound per mole of protein ν was determined from eq 1. The first spectrum collected before addition of the first aliquot of EGTA is assumed to be that of fully substituted Co(II) protein in each of these four cases. The remainder of the analysis was carried out exactly as described for the Co(II) titrations with EGTA.

RESULTS

The Co(II) binding properties of wild-type and metal ligand substitution mutants of gp32 are discussed in turn.

Wild-Type gp32. The thermodynamic stability of the native protein Co(II) complex provides a reference point with which to compare mutant complexes. Two Co(II)-binding isotherms obtained for apo intact and core wild-type gp32 collected at pH 7.5, 25 °C, and 0.1 M KCl are shown in Figure 2A. They are essentially stoichiometric and provide only a lower limit of K_a ($\geq 1 \times 10^7$ M^{-1}). In order to obtain a quantitative measurement, an appropriate metal complexant must be present in excess over protein concentration in order to buffer very low concentrations of free metal (cf. Mely et al., 1991). Co(II) titrations were therefore carried out in the presence of ethylene glycol bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA) as described in Materials and Methods to determine the K_a of intact wild-type and core gp32 for Co(II). Two representative EGTA competition binding isotherms collected under the same solution conditions of Figure 2A are shown in Figure 2B. From these data, the K_a for Co(II) of wild-type intact and core fragment gp32s estimated using eq 3 are $(1.26 \pm 0.05) \times 10^9$ M^{-1} and $(0.9 \pm 0.1) \times 10^9$ M^{-1}, respectively. These two values are quite close to one other, which is expected since the core domain of gp32 is an independently folded domain as revealed by the limited proteolysis (Giedroc et al., 1987).

A competition experiment was carried out to estimate the binding affinity of wild-type gp32 for the natural metal Zn(II). The protein was first loaded with saturating amounts

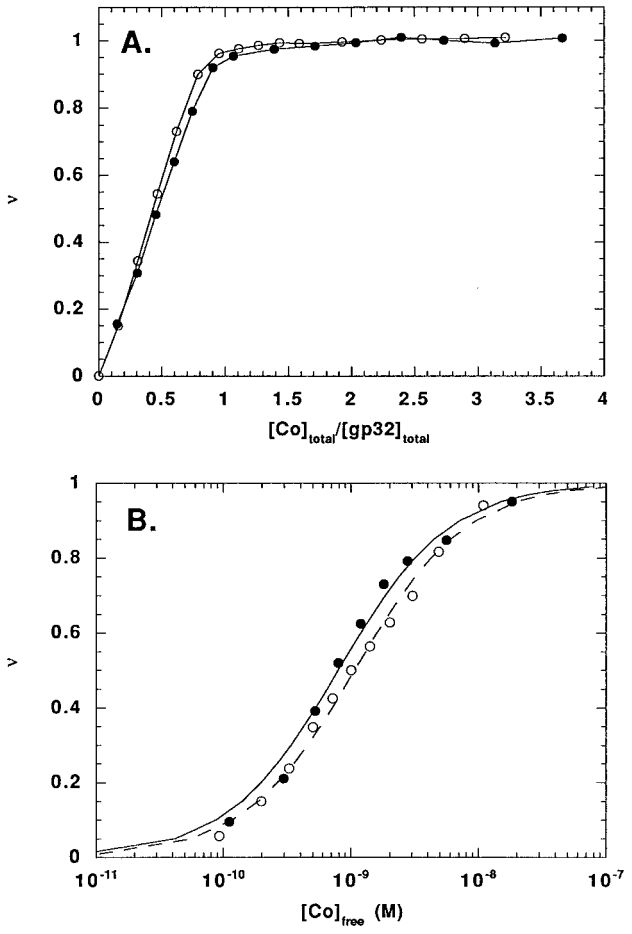

FIGURE 2: (A) Saturation of wild-type gp32 [intact protein (●); core protein (○)] with Co(II) as a function of the ratio of $[Co(II)]_{total}$: $[gp32]_{total}$ in the absence of chelator. The concentrations of active apoprotein are 85 μM intact gp32 and 217 μM gp32 core fragment as determined from the known extinction coefficient of Co(II)−gp32 complex (Giedroc et al., 1986). (B) Co(II)-binding isotherms obtained for wild-type intact (●) and core gp32 (○) in the presence of 1 mM EGTA (see Materials and Methods for details). The protein concentrations are as indicated in panel A.

(≈2 mM) of Co(II), and then was back-titrated with $ZnSO_4$, with the absorption spectra collected following each addition of Zn(II). The molar concentration of Co(II)−gp32 complex decreased during the titration as shown in Figure 3. Fitting these data to a simple competition model (Berg & Merkle, 1989) reveals that the apparent affinity of wild-type gp32 for Zn(II) is ≈250-fold higher than that for Co(II) or $(3.0 \pm 1.0) \times 10^{11}$ M^{-1}. A similar difference between the affinity of metal sites for Zn(II) and Co(II) has been observed in other systems and appears to be due to ligand-field stabilization energies which favor formation of the octahedral hexaaquo (free) Co(II) over tetrahedral (bound) high-spin Co(II) complexes (Krizek et al., 1991; Regan & Clarke, 1990).

H64C, H64D, and H64N gp32s. Co(II)-substituted H64C, H64D, and H64N gp32s were reconstituted under anaerobic conditions. The appearance of blue-green color upon addition of $CoCl_2$ to the metal-free protein is indicative of the complex formation between metal and protein molecule. The complete optical absorption spectra of all three Co(II)-substituted His^{64} mutants compared to wild-type gp32 are shown in Figure 4.

Co(II)-substituted H64C gp32 is characterized by a significantly red-shifted d−d visible absorption envelope and

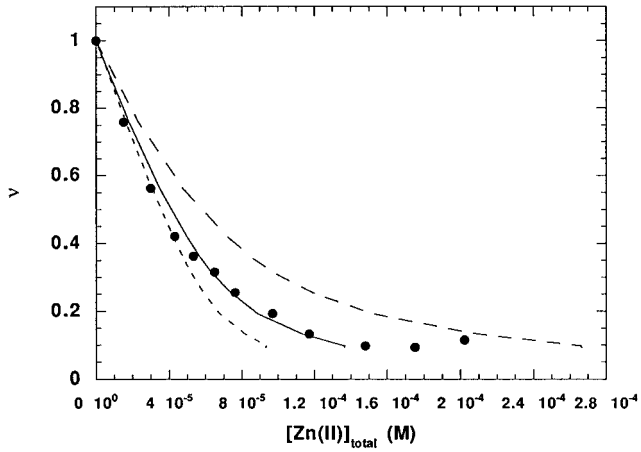

FIGURE 3: Zn(II) titration of Co(II)-substituted wild-type gp32. 2.8 mM total Co(II) was present prior to the first addition of Zn(II). Theoretical isotherms with $K_a = 3.0 \times 10^{11}$ M^{-1} (250-fold higher than $K_{a,Co}$) (—), 1.2×10^{11} M^{-1} (100-fold) (− −), and 6.0×10^{11} M^{-1} (500-fold) (- - -) are plotted to demonstrate the goodness of fit. The concentration of active protein is 85 μM.

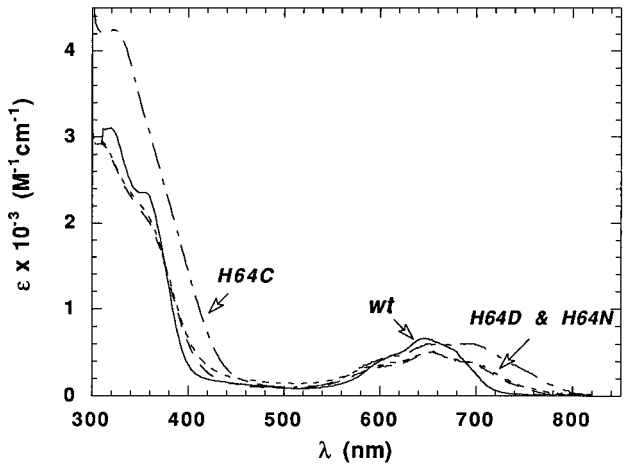

FIGURE 4: The complete corrected optical absorption spectra of Co(II)-substituted H64C (− - −), H64D (− −), and H64N (- - -) gp32s compared with that of wild-type (wt) (—) gp32.

contains additional alternations in the $S^- \rightarrow Co(II)$ ligand-to-metal charge transfer region ($\lambda \leq 450$ nm) relative to the Co(II)-substituted wild-type protein. Although the maximum of the d−d transition envelope occurs at 694 nm, the envelope appears to form a plateau between 650 and 694 nm.[3] 646 nm defines the maximum of d−d absorption envelope for wild-type gp32. The extinction coefficient at 694 nm is estimated to be 640 M^{-1} cm^{-1}, strongly suggestive of tetrahedral or distorted tetrahedral coordination geometry. Two shoulders are also found at 750 and 600 nm. The absorption envelope for H64C gp32 is much broader than that for wild-type protein, especially in the lower energy region. These spectral changes are consistent with the observed spectral consequences of replacement of a non-thiolate with a thiolate ligand in model Co(II) coordination compounds (Corwin et al., 1988). In addition, there are also alterations in the near-UV region, transitions assignable to

[3] Absorption spectra of Co(II)-substituted H64C gp32 previously published (Qiu & Giedroc, 1994; Guo et al., 1995) are now known to be representative of H64C gp32 which is less than fully saturated with Co(II). These spectra are shown here to correspond to a mixture of metal-linked dimeric and monomolecular, tetrathiolate Co(II) complexes.

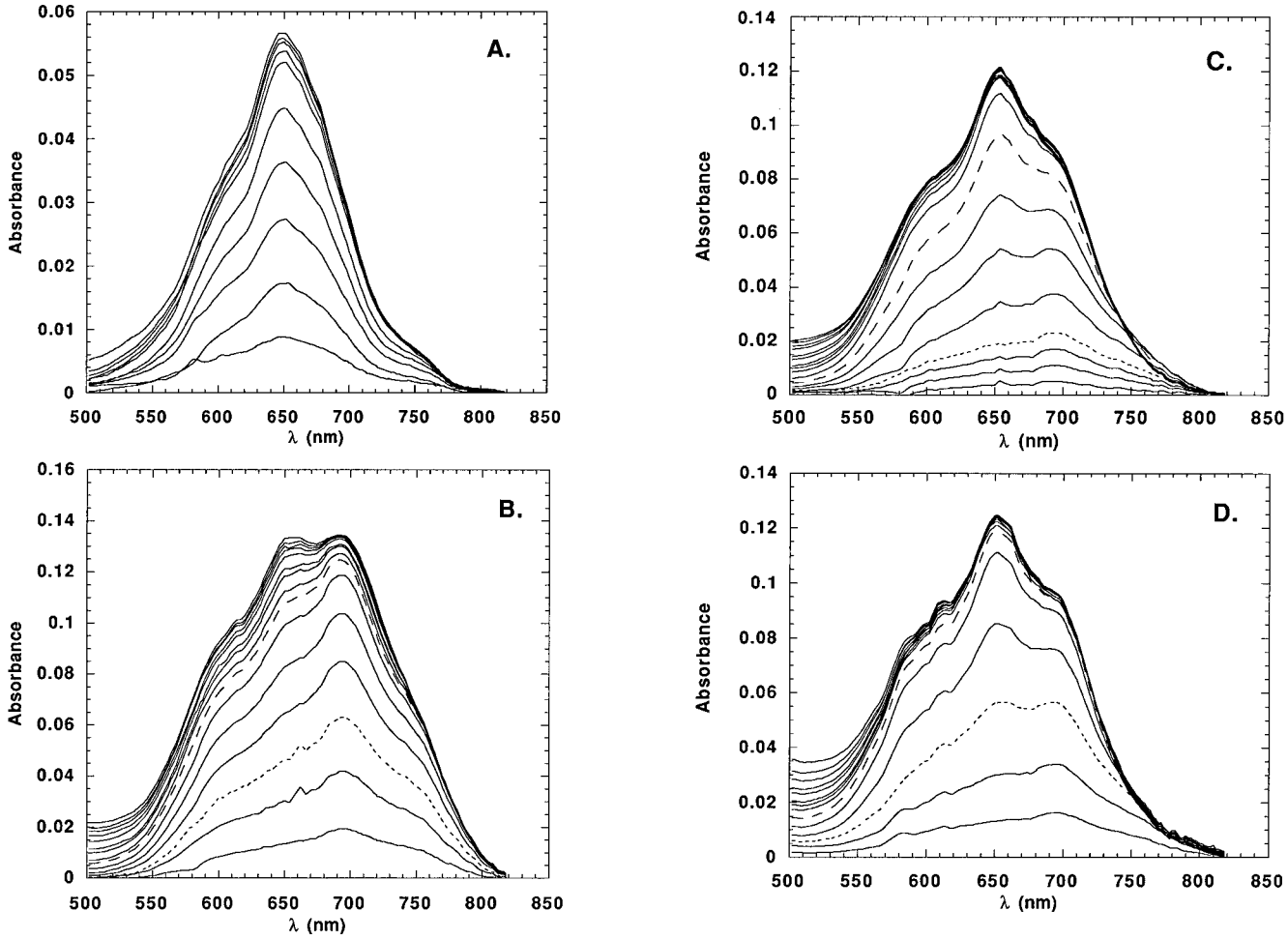

FIGURE 5: The ligand field envelope region of a series of absorption spectra of wild-type (A), H64C (B), H64D (C), and H64N (D) gp32s collected in the presence of increasing $[Co(II)]_{total}$ (0 → 600 μM). The total active protein concentration of wild-type gp32 is 85 μM (A). The total protein concentration is 212 μM, 241 and 249 μM H64C (B), H64D (C), and H64N (D) gp32s, respectively. For panels B−D, the $[Co(II)]_{total}$:$[gp32]_{total}$ ratios follow for individual spectra from lowest to highest absorption. For emphasis, the dotted and dashed spectra correspond to those obtained at 0.5:1 and 1:1 $[Co(II)]_{total}$:$[gp32]_{total}$ ratios, respectively. (B) 0.15, 0.31, 0.46, 0.62, 0.75, 0.91, 1.04, 1.20, 1.50, 1.80, 2.24, 2.70, 2.99, 3.30, 3.76; (C) 0.12, 0.24, 0.38, 0.50, 0.62, 0.74, 0.87, 0.99, 1.13, 1.23, 1.35, 1.59, 1.97, 2.34, 2.71, 3.09, 3.46, 3.70; (D) 0.18, 0.33, 0.51, 0.68, 0.84, 0.99, 1.17, 1.33, 1.50, 1.85, 2.18, 2.52, 3.04, 3.72.

S^-→Co(II) charge transfer bands. For example, the intensity of absorption at ≈330 nm is increased while the shoulder at 350 nm in the wild-type gp32 moves to longer wavelengths and becomes broader in the H64C protein. This observation is also consistent with increasing sulfur coordination (Corwin et al., 1988). X-ray absorption and EXAFS of the Co(II)- and Zn(II)-substituted H64C gp32s are consistent with tetrathiolate, tetrahedral coordination (Guo et al., 1995).

In contrast to H64C gp32, the optical spectra of Co(II)-substituted H64D and H64N gp32s show maximum absorption in the d−d transition envelope at 650 nm, revealing only a small red shift relative to wild-type gp32 (646 nm). This suggests that the number of thiolate ligands in each coordination complex is identical and indistinguishable from wild-type gp32. The ligand-field transition envelopes of these two Co(II)−proteins are again broader than that of the wild-type protein, which indicates a wider energy distribution of the electronic transitions associated with mutant protein−Co(II) complexes. It is both interesting and important that the spectra of these two mutants are essentially identical with $\epsilon_{650} \approx 500\ M^{-1}\ cm^{-1}$. Considering the distinct nature of the side chains of aspartate (a potential ligand) and asparagine (a nonliganding amino acid), these data suggest that neither of these two mutant side chains are involved in the coordination of the metal. Instead, a H_2O molecule or other solvent ligand (e.g., Cl^-) provides the fourth ligand in these complexes.

In contrast to wild-type gp32 (Figure 2A), the absorption at λ_{max} for His^{64} mutant gp32s (e.g., 694 nm for H64C gp32) does not increase linearly with incremental additions of Co(II), but instead shows strongly sigmoidal behavior (data not shown). The ligand field absorption spectra of each of the His^{64} mutants obtained during the course of a titration with Co(II) are shown in Figure 5 (panels B−D) compared with the same data for wild-type gp32 (panel A). As can be seen, the origin of the sigmoidicity is due to fact that the absorption spectrum clearly changes during the course of a titration of each of the His^{64} mutant proteins with Co(II), in contrast to wild-type gp32. In all three cases, a spectrum with relatively red-shifted ligand field bands is observed early in the titration, i.e., at low metal to protein ratios. For example, the absorption maximum of the ligand field envelope for H64C, H64D, and H64N gp32s at $[Co(II)]_{total}$:$[gp32]_{total}$ ratios ≤0.5 appears at ≈700 nm accompanied by a distinct shoulder at ≈750−760 nm and a secondary maximum at ≈650 nm. As the $[Co(II)]_{total}$ increases beyond a $[Co(II)]_{total}$:$[gp32]_{total}$ of 0.5, the ligand field absorption maximum shifts to ≈650 nm for H64D and H64N gp32s,

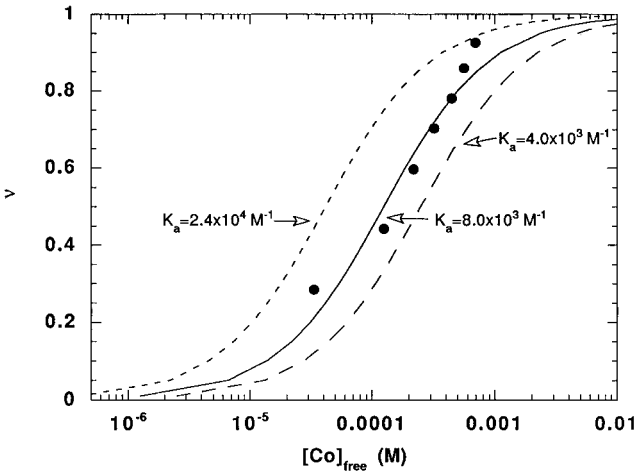

FIGURE 6: Determination of the Co(II) binding constant for H64C gp32 using an EGTA-bleaching experiment (see Materials and Methods for details). The same kind of experiment was carried out with H64D and H64N gp32s to determine K_a (Table 2) (data not shown).

with the 700 nm peak becoming the secondary maximum and the shoulder at 750 nm essentially lost. In the case of H64C gp32 (panel B), the 650 nm band becomes of nearly comparable intensity to that of the 700 nm absorption band, with the entire spectral envelope remaining red-shifted relative to the wild-type, H64D, and H64N gp32s.

Taken together, these spectral changes strongly suggest formation of tetrathiolate complex at low $[Co(II)]_{total}$ (or low $[Co(II)]_f$) which rearranges to form the final end-point complex at high or saturating $[Co(II)]_{total}$. Since H64D and H64N gp32s only have three Cys in the coordination site, formation of a tetrathiolate, tetrahedral coordination complex at $[Co(II)]_{total}$:$[gp32]_{total} \leq 0.5$ requires formation of a 2:1 protein−Co(II) complex via coordination by two pairs of cysteinate ligands from each of two protein monomers. As the concentration of $[Co(II)]_{total}$ increases beyond a $[Co(II)]_{total}$:$[gp32]_{total}$ ratio of 0.5, formation of the final 1:1 monomolecular complex via coordination by the expected ligands and with the anticipated geometry is favored. Michael et al. (1992) have observed a similar partitioning of free Co(II) between monomolecular and cross-linked dimeric forms of non-native TFIIIA-type zinc finger peptides during the course of a Co(II) titration. In fact, the absorption spectra obtained at low $[Co(II)]_{total}$ for H64D and H64N gp32s bear a striking qualitative similarity to the metal-linked dimeric complexes of TFIIIA peptides formed at $[Co(II)]_{total}$:$[gp32]_{total} \leq 0.5$ (Michael et al., 1992). Preliminary modeling of these absorption data with an equilibrium partitioning scheme which takes into account formation of 2:1 and 1:1 gp32−metal complexes suggests that this is indeed an appropriate model.[4]

The formation of distinct complexes during the course of a titration makes it impossible to estimate the binding affinity directly from the Co(II) titration data as was done for wild-type gp32. Therefore, to estimate the stability of 1:1 gp32−Co(II) complex without complications from formation of the 2:1 complex, an EGTA-bleaching experiment was employed. A representative experiment is shown in Figure 6 for Co(II)-substituted H64C gp32. This experiment was performed by challenging the 1:1 gp32:Co(II) complex present at excess

[4] J. Guo and D. P. Giedroc, to be published elsewhere.

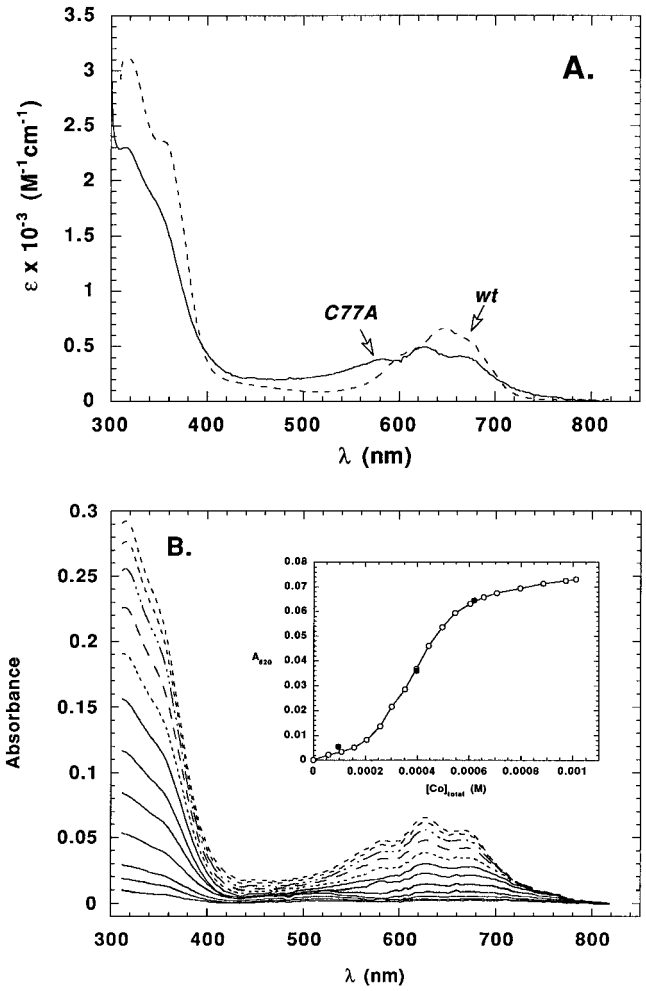

FIGURE 7: (A) The complete corrected optical absorption spectrum of Co(II)-substituted C77A gp32 (—) compared with that of wild-type gp32 (- - -). (B) Absorption spectra obtained for C77A gp32 collected in the presence of increasing $[Co(II)]_{total}$ (0 → 650 μM). The total protein concentration is 280 μM (213 μM apoprotein). Note the presence of significant absorption in the 500 nm region at low $[Co(II)]_{total}$. Inset: Absorbance at 620 nm obtained from the spectra shown plotted as a function of $[Co(II)]_{total}$. The data points given by the filled circles were obtained from individual reaction mixtures which were incubated anaerobically for 24 h at the same [C77A gp32] and the indicated concentration of $[Co(II)]_{total}$. The solid line has no theoretical significance.

Co(II) with increasing concentrations of EGTA. During the EGTA-induced displacement followed optically, $[Co(II)]_f$ decreases. As expected, as $[Co(II)]_f$ decreases, the absorption spectra suggest that the dominant species in solution changes from a 1:1 to 2:1 protein−metal complex, just like that observed in the normal titration (data not shown). Only those spectra which were clearly indicative of the 1:1 complex as the major species were used to estimate the binding affinity of the protein for Co(II) (see Materials and Methods). K_a for H64C, H64D, and H64N gp32s determined in this way are $(8.0 \pm 0.9) \times 10^3$ M^{-1}, $(1.0 \pm 0.2) \times 10^4$ M^{-1}, and $(1.1 \pm 0.2) \times 10^4$ M^{-1}, respectively. In Figure 6, the theoretical binding isotherm defined by K_a obtained from nonlinear least squares fitting is superimposed on the data points along with theoretical isotherms defined by 2-fold smaller to 3-fold larger K_a. The K_a's of these metal complexes are decreased by about 5 orders of magnitude relative to that of wild-type gp32.

C77A gp32. Substitution of Cys^{77} with Ala removes both the charge and side chain of this ligand. The optical

spectrum of Co(II)-substituted C77A gp32 is shown in Figure 7A compared with wild-type Co(II)-gp32. The maximum of the d−d transition envelope is blue-shifted relative to that of the wild-type, which is consistent with a fewer number of thiolate ligands involved in metal coordination in this mutant. There are two shoulders at 580 and 670 nm, with the entire envelope again somewhat broader than that of wild-type gp32.

Figure 7B shows full absorption spectra collected for C77A gp32 (213 μM apoprotein) during the course of a Co(II) titration; the inset shows a plot of A_{620} vs $[Co(II)]_{total}$ derived from the spectra shown in the main body of the figure. The solid spectra correspond to the first seven titration points in the inset, i.e., at $[Co(II)]_{total} \leq 0.35$ mM, with the remaining dashed and dot-dashed spectra obtained at later points in the titration. Several points can be made from these spectra. At low $[Co(II)]_{total}$:$[apo\ C77A\ gp32]_{total}$ ratios (≤ 1.0), there appears to be maximum development of a weak absorption envelope centered about 520 nm, with little development of the absorption in the 620 nm region. At $[Co(II)]_{total}$:$[gp32]_{total}$ ≥ 1.0 and up to ≈ 4 mol equiv, significant absorption of the major 620 nm ligand field envelope is observed, with the 520 nm transition region becoming a shoulder on this envelope. Assuming 100% active apo C77A gp32 molecules, approximately 3−4 mol equiv appear to be required to saturate the apoprotein. Other titrations carried out at different input [apo C77A gp32] give the same qualitative trends in the data (data not shown). To ensure that the titration shown is at equilibrium, the absorption spectra of three individual anaerobic 24-h incubations of protein and Co(II) concentrations identical to conditions obtained during the course of the titration were recorded (filled circles, Figure 7B, inset). The excellent correspondence of these data with the titration obtained in the normal fashion show that the isotherm shown is indeed at equilibrium.

The spectral shape, λ_{max}, and estimated molar absorption obtained at low $[Co(II)]_{total}$:$[gp32]_{total}$ ratios are suggestive of an octahedral Co(II) complex, perhaps with some thiolate ligation character. Only after this site is significantly populated does the expected tetrahedral Cys_2-His-H_2O or pentacoordinate complex form (Corwin et al., 1987).[5] This of course suggests that the Co(II):C77A gp32 stoichiometry is significantly greater than 1, a point that cannot be rigorously established by these data.

An EGTA-bleaching experiment was again employed to estimate the thermodynamic stability of the Co(II)−C77A gp32 complex, subject to the caveat that the determined value may be representative of a mixture of molecular species. Because of the sigmoidal shape of the Co(II) titration, the concentration of active protein cannot be estimated from the direct titration isotherms. Assuming an $\epsilon_{620} \approx 500\ M^{-1}\ cm^{-1}$, which is a reasonable value consistent with the other single Cys substitution mutants (see below), K_a for the Co(II)−C77A gp32 complex is $(6.6 \pm 1.7) \times 10^3\ M^{-1}$.

[5] Tetrahedral and pentacoordinate complexes are most easily distinguished on the basis of molar absorptivity, with the pentacoordinate site of significantly lower absorptivity (Corwin et al., 1987). Since we have no estimate of the concentration of active molecules or the metal:gp32 stoichiometry, the molar absorptivity of Co(II)-C77A gp32 cannot be estimated. This precludes unambiguous discrimination between four- and five-coordinate complexes.

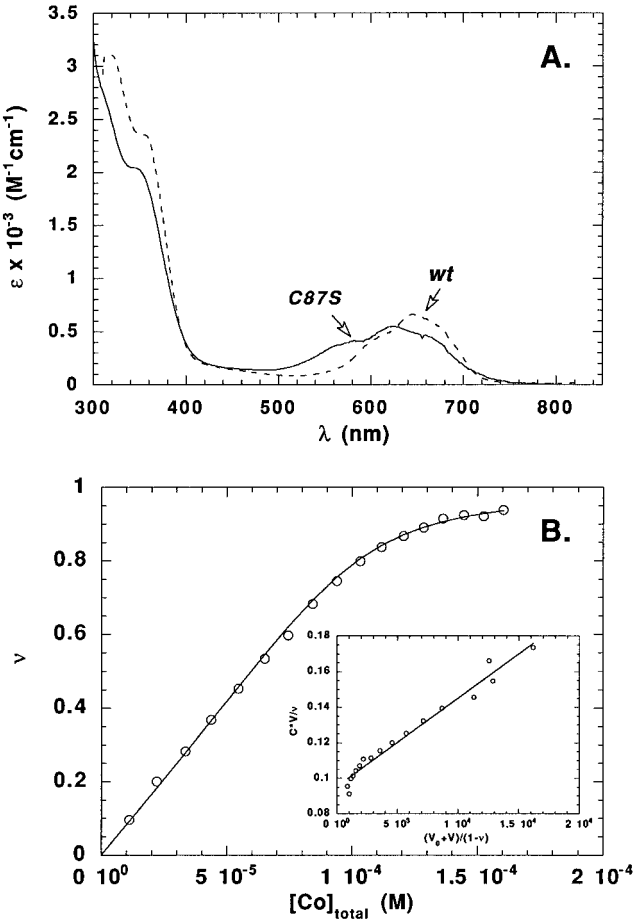

FIGURE 8: (A) The complete corrected optical absorption spectrum of Co(II)-substituted C87S gp32 (—) compared with that of wild-type gp32 (- - -). (B) Co(II) titration of C87S gp32 plotted as ν vs total Co(II) concentration. Open circles are data points with the solid line representing a theoretical isotherm described by K_a for a 1:1 binding model. Inset: Linear fit using the method of Klemba and Regan (1995) (see text for details).

C87S gp32. Like alanine, serine can be considered a nonliganding substitution for Cys[87]. The optical spectrum of Co(II)-substituted C87S gp32 is shown in Figure 8A compared to that of wild-type Co(II)-substituted gp32 (Guo et al., 1995). It is not surprising to see that the maximum absorption of the d−d transition envelope is at 624 nm, which is a shorter wavelength than the peak absorption of the wild-type gp32. Like C77A gp32, the intensity of absorption in both the near-UV and visible regions is lower than that for wild-type gp32, which is expected since the number of the thiolate ligands involved in the Co(II) coordination decreases from three for the wild-type to two for the mutant (Corwin et al., 1988).

A direct Co(II)-binding titration has been performed by adding free Co(II) into reduced apo C87S gp32 under anaerobic conditions. As above, the formation of Co(II)−C87S gp32 complex was followed by collecting the UV−visible spectrum after each addition of Co(II) titrant. These absorption spectra show that the absorption at the maximum wavelength increases linearly with the addition of Co(II) at the beginning of the titration (data not shown). Extrapolation of the initial linear portion of the curve indicates that $\approx 80\%$ of the apoprotein is competent to form a metal complex (data not shown). This permits determination of the molar

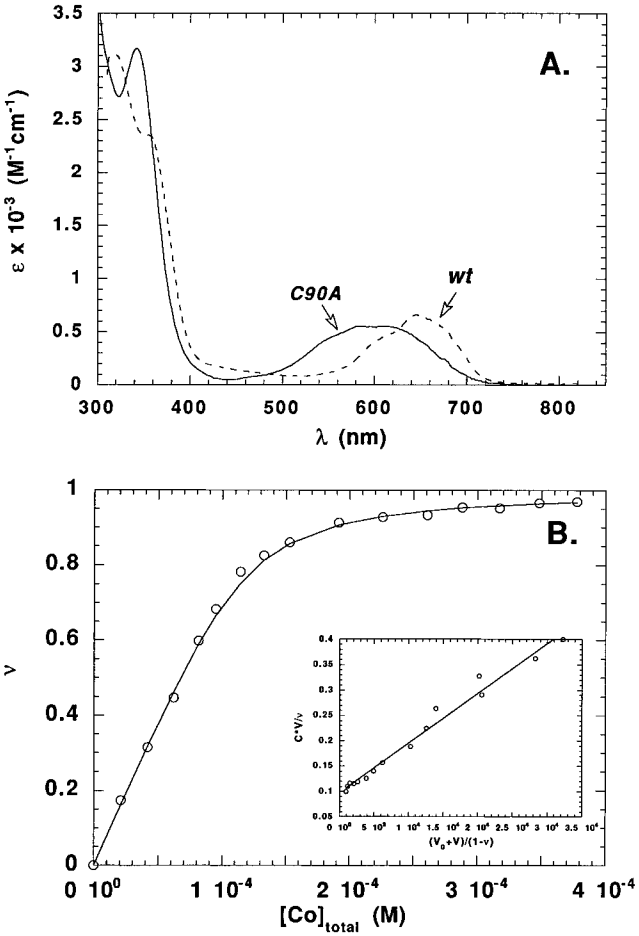

FIGURE 9: (A) The complete corrected optical absorption spectrum of Co(II)-substituted C90A gp32 (—) compared with that of wild-type gp32 (- - -). (B) Co(II) titration of C90A gp32 as in Figure 8.

extinction coefficient for the Co(II) complex as 535 ± 10 M^{-1} cm^{-1}. Subsequent analysis was done using eqs 3 and 4.

A second independent method of data analysis was performed as described in the Materials and Methods (Klemba & Regan, 1995). The maximum absorbance at 624 nm (A_{max}) at saturation was estimated by least squares curve fitting from a double-reciprocal plot ($1/A_i$ vs $1/[Co]_{t,i}$) using eq 5. The determined A_{max} was then used to calculate the fractional saturation, ν_i, following each *i*th addition of Co(II). A linear least squares fit of the plot CV/ν_i vs $(V_0 + V)/(1 - \nu_i)$ (see Materials and Methods) returns a slope of $1/K_a$ (Figure 8B, inset). The main body of Figure 8B shows a superimposition of a theoretical isotherm defined by the best-fit parameters (K_a and ϵ) on the experimental data. Two independent experiments give a K_a of $(1.7 \pm 0.3) \times 10^5$ M^{-1}.

C90A gp32. The corrected optical spectrum of Co(II)-substituted C90A gp32 is shown in Figure 9A. As seen in C87S gp32, the d−d transition envelope broadens to cover a wider range of energies than that observed for the wild-type gp32. The maximum absorption wavelength is 620 nm which is again at the blue side of the wild-type absorption peak. Compared to C87S gp32, the transition envelope of C90A gp32 shifts to slightly higher energies while the intensity of the shoulder in the 550−580 nm region appears higher than that found in the spectra of Co(II)−C87S and Co(II)−C77A gp32s.

Direct Co(II)-binding titrations were carried out exactly as described for C87S gp32. The corrected absorption at 620 nm increases linearly with the addition of Co(II) in the initial part of the titration (data not shown) as found for C87S gp32, indicative of simple 1:1 rectangular hyperbolic binding behavior. The analysis of two independent titrations exactly as described for C87S gp32 gives an extinction coefficient at 620 nm of 550 ± 50 M^{-1} cm^{-1}, and an affinity of C90A gp32 for Co(II) of $(1.1 \pm 0.2) \times 10^5$ M^{-1} (Figure 9B).

The Co(II)-binding properties of wild-type and mutant gp32s are summarized in Table 2.

DISCUSSION

Bacteriophage T4 gene 32 protein is an extensively studied paradigm for a class of accessory proteins in DNA replication and recombination which bind nonspecifically and preferentially to single-stranded nucleic acids. Since gp32 was shown to contain 1 mol of Zn(II) per monomer (Giedroc et al., 1986), gp32 has been studied as a model for the function of inorganic zinc in a nucleic acid binding or zinc finger protein (Giedroc et al., 1987, 1989, 1990, 1992; Qiu & Giedroc, 1994; Qiu et al., 1994; Shamoo et al., 1991, 1993; Nadler et al., 1990).

Removal of zinc from gp32 makes the protein hypersensitive to proteases and markedly decreases its thermostability (Keating et al., 1988), revealing that zinc coordination is important for maintaining the conformational stability of the molecule (Giedroc et al., 1987, 1989). A detailed molecular understanding of metal binding begins with the identification of the metal-liganding residues in the primary structure of the protein. Early work established unambiguously that Cys^{77}, Cys^{87}, and Cys^{90} are involved in metal coordination (Giedroc et al,, 1989), while more recent work has shown that His^{64}, and not His^{81}, donates the non-thiolate ligand (Qiu & Giedroc, 1994). This finding is supported from quantitative analysis of the XAS and EXAFS analysis obtained for wild-type and mutant gp32s (Guo et al., 1995) and is consistent with the crystal structure of gp32 core fragment solved recently (Shamoo et al., 1995). The crystallographic model reveals that His^{64} and Cys^{77} are derived from two sequential β-strands within a three-stranded β-sheet, while Cys^{87} and Cys^{90} may well cooperate in nucleating an α-helix. The coordination bond formed by Cys^{77} is by far the most buried of the four protein ligands which is followed by His^{64}, Cys^{87}, and Cys^{90} in solvent accessibility. The S^{γ} atom of Cys^{90} and the N^{δ} atom of His^{64} form part of the van der Waals surface of the ssDNA binding cleft.

In this paper, we have investigated the structural consequences of independently mutating each of the amino acid ligands to the Zn(II) site of T4 gene 32 protein to potentially liganding and/or nonliganding residues. The optical spectra of all Co(II)-substituted gp32s reveal that tetrahedral Co(II) coordination geometry is retained in all mutants. However, in all cases, the ligand-field absorption envelope is broader than that observed for wild-type gp32, which suggests either a greater distortion from perfect tetrahedral coordination symmetry relative to wild-type gp32, or is the result of conformational averaging (Garmer & Krauss, 1993). The overall intensity of the d−d transition envelope is significantly reduced, which is consistent with a larger range of energies associated with individual electronic transitions associated with the intense $^4A_2 \rightarrow {}^4T_1$ (P) transition envelope (Garmer & Krauss, 1993). There is also a distinct red shift

Table 2: Co(II)-Binding Properties of Wild-Type and Metal Ligand Substitution Mutant gp32s at pH 7.5, 25 °C

gp32	ligand field absorption max (nm)	extinction coeff (ϵ) (M^{-1} cm^{-1})	K_a (M^{-1})	$\Delta G°$ (25 °C) (kcal mol^{-1})
wild-type intact	646	660	$1.3 (\pm 0.1) \times 10^9$ [a]	$-12.4 (\pm 0.1)$
wild-type core	646	660	$0.9 (\pm 0.1) \times 10^9$	$-12.2 (\pm 0.1)$
H64C	694	≈ 640 [b]	$8.0 (\pm 0.9) \times 10^3$	$-5.3 (\pm 0.1)$
H64D	650	≈ 500 [b]	$1.0 (\pm 0.2) \times 10^4$	$-5.5 (\pm 0.1)$
H64N	650	≈ 500 [b]	$1.1 (\pm 0.2) \times 10^4$	$-5.5 (\pm 0.1)$
C77A	620	n.d. [c]	$6.6 (\pm 1.7) \times 10^3$ [d]	$-5.2 (\pm 0.2)$ [d]
C87S	624	535 ± 10	$1.7 (\pm 0.3) \times 10^5$	$-7.1 (\pm 0.1)$
C90A	620	550 ± 50	$1.1 (\pm 0.2) \times 10^5$	$-6.9 (\pm 0.1)$

[a] Obtained by nonlinear least squares curve fitting using a single set of data. [b] Estimated from the Co(II)-binding isotherms. [c] Not determined. [d] Estimated using an $\epsilon_{620} = 500$ M^{-1} cm^{-1}, consistent with other single Cys substitution mutant gp32s.

in the ligand-field transition envelope (500−800 nm) and corresponding increase in the $S^- \rightarrow Co(II)$ LMCT region ($\lambda \leq 450$ nm) as the number of thiolate donor atoms in the complex increases from two (C77A, C87S, and C90A gp32s) to three (wild-type, H64D, and H64N gp32s) to four (H64C gp32). This shows that, at excess Co(II), an amino acid substitution of one of the four ligands does not significantly perturb the ability of each of the other ligands to make a coordination bond to the metal at saturating [Co(II)].

Strikingly, the affinity of the mutants for Co(II), K_a, is reduced by 10^4-10^5-fold relative to wild-type gp32, ranging from 6×10^3 to 2×10^5 M^{-1} for mutant gp32s to 1.2×10^9 M^{-1} for the native protein; this corresponds to a destabilization of 5−7 kcal mol^{-1} depending upon the substitution (Table 2). Although the extent of destabilization is large in every case, each ligand appears to make a distinct contribution to coordination chemistry of the site, and possibly the globular conformation of the molecule.

Substitution of His^{64} with potentially liganding (Cys, Asp) and nonliganding (Asn) amino acids reveals that Cys^{64} forms a new coordination bond to the metal in Co(II)−H64C gp32 as revealed by optical absorption (Figure 4) and X-ray absorption spectroscopies (Qiu & Giedroc, 1994; Guo et al., 1995). However, the new coordination bond formed between Cys^{64} and Co(II) appears not to contribute greatly to the stability of the chelate since K_a measured for H64C, H64D, and H64N gp32s are identical within experimental error (Table 1). Since Asp and Asn have very different capacities to serve as zinc ligands yet give rise to essentially indistinguishable optical spectra in the UV−visible region (Figures 4 and 5), this suggests that neither participates in metal coordination. If this is the case, the optical spectra of the Co(II)-substituted H64D and H64N gp32s require that a water molecule or other solvent ion complete the tetrahedral coordination complex. Even in the case of the H64C mutant, it cannot be ruled out that Cys^{64} is a "partial occupancy" ligand which is readily exchangeable with solvent molecules. However, our attempts to obtain evidence for an open coordination site in H64C or C87S gp32s with exogenous thiolate- and imidazole-based ligands (e.g., benzenethioate, β-mercaptoethanol, and 4-methylimidazole) have not met with success.

The $Cys^{77} \rightarrow$Ala substitution reduces the stability of the chelate by about 6−7 kcal mol^{-1}, essentially indistinguishable from that of the His^{64} mutants. Both the His^{64} mutants and C77A gp32 show sigmoidal Co(II)-binding isotherms, although the origin of the behavior is different in the two cases. A coupled equilibrium scheme in which free Co(II) is partitioned between 1:1 and 2:1 gp32−Co(II) complexes has been proposed to interpret the Co(II) binding properties of the His^{64} mutants, analogous to that previously proposed for mutant zinc finger peptides (Michael et al., 1992) and *de novo* designed zinc sites (Klemba & Regan, 1995) (Figure 5). In C77A gp32, there is no spectral evidence for the formation of tetrathiolate chelate. However, like the His^{64} mutants, the spectral data do suggest that more than one complex is formed during the course of a Co(II) titration; specifically, a higher affinity octahedral complex and a low affinity tetrahedral or pentacoordinate complex.[4] The Co(II)-binding isotherm of C77A gp32 is certainly suggestive of cooperativity of metal binding, which implies that the stoichiometry of the binding of Co(II) to C77A gp32 may well be larger than one.

Since none of the substitutions at position His^{64} retain high chelate stability (including cysteine which ultimately forms a coordination bond), it seems plausible that His^{64} plays a unique role in directing the coordination chemistry of the site. Indeed, the crystal structure of gp32 shows that His^{64} resides in the central β-strand of a three-stranded antiparallel β-sheet which connects subdomain II to subdomain I, shown schematically in Figure 10. As can be seen (Figure 10), both residues 64 and 77 are involved in extensive main-chain hydrogen bonding; in addition, the side chains of His^{64} (N^{δ}) and Cys^{77} (S^{γ}) donate and accept hydrogen bonds to the α-carbonyl oxygen of Val^{109} and the main-chain amide of Ser^{79}, respectively. Since this region of the molecule appears strongly organized, the side chains of residues 64 and 77 might be quite restricted even in the metal-free apoproteins. Therefore, unlike the situation in TFIIIA-type zinc fingers (Krizek et al., 1993), neither of the two liganding substitutions (His→Cys and His→Asp) at residue 64 of gp32 form good coordination bonds with the metal ion. We anticipate that similar results would be obtained for liganding substitutions of Cys^{77} as well.

C87S and C90A gp32s form more stable Co(II) complexes than those formed by substitution mutants of the other two liganding sites, with the difference ≈10−20-fold in K_a (Table 2). In contrast to His^{64} and Cys^{77} mutant proteins, these mutants form the expected 1:1 protein−Co(II) complex throughout the course of a Co(II) titration (Figures 8 and 9). The crystal structure of the wild-type core gp32 reveals that Cys^{87} is at the −1 position relative to the start of an α-helix with Cys^{90} positioned within the first turn of the same α-helix; both S^{γ} atoms of Cys^{87} and Cys^{90} are partially exposed to solvent (Shamoo et al., 1995). The α-helical region which contains Cys^{87} and Cys^{90} may well be relatively disordered in the apoprotein. This flexible or disordered state may make it favorable for these two ligands to establish

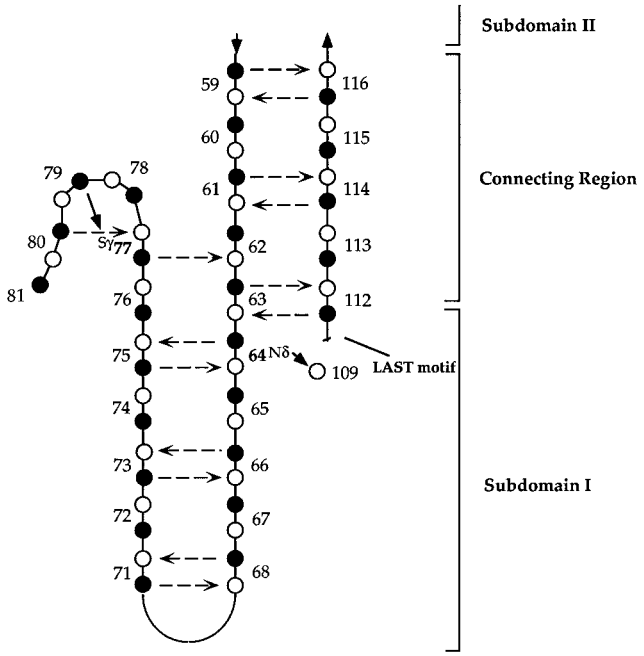

FIGURE 10: Two-dimensional cartoon representation of secondary structure and hydrogen-bonding interactions around the Zn(II) coordination site of the gp32 core fragment. Filled circles represent backbone amides, and open circles, backbone carbonyl oxygens. The broken arrows represent main-chain hydrogen bonds present in the β-sheet structure. Two side chain-to-main chain hydrogen bonds are indicated by the solid arrows: His^{64} $N^{\delta}H\cdots O{=}C$ Val^{109} and Ser^{79} $NH\cdots S^{\gamma}$ Cys^{77}. Residues 88−97, 98−101, and 102−106 not shown form three α-helices which complete the secondary structure of subdomain I. Based on the crystal structure of gp32 core fragment complexed to ssDNA (Shamoo et al., 1995).

optimal interactions with the metal ion both in the native monomolecular state as well as in the proposed bimolecular tetrathiolate complexes formed by H64C, H64D, and H64N gp32s. Cys^{87} and Cys^{90} may therefore function analogously to the metal ligands in TFIIIA-like zinc fingers where metal ion coordination drives the formation of the secondary structure. In gp32, the extent of formation of secondary structure need not be large and could be quite small due to localized changes (Giedroc et al., 1987; Qiu et al., 1994). Preliminary far-UV CD spectra obtained for the metal-free apo forms of the His^{64} substitution mutants show that these molecules are folded to an extent essentially identical to that of apo wild-type gp32, which give rise, upon metal coordination, to only small changes in CD intensity (data not shown), i.e., on the order of that obtained with Zn(II)- or Co(II)-containing wild-type gp32 (Giedroc et al., 1987; Qiu et al., 1994).

We conclude that the local protein secondary structure in gp32 dictates the effect of nonliganding substitutions in a *residue-specific manner*, such that individual ligands of the tetrahedral zinc complex play distinct roles in maintaining the high stability and coordination geometry of the site. Cys^{87} and Cys^{90} contribute about equally in the absence of the other to the stability of the site (≈5 kcal mol^{-1}) while substitution of His^{64} or Cys^{77} is even more destabilizing (≈7 kcal mol^{-1}). This degree of destabilization is of the expected order of magnitude based on previous studies with naturally occurring zinc sites (Kiefer & Fierke, 1994; Mely et al., 1991), but far greater than that observed with *de novo* designed sites (Klemba & Regan, 1995). These findings lead to model where His^{64} and Cys^{77} play critical roles in positioning of the metal ion for the eventual formation of a stable tetrahedral complex. Upon formation of these two coordination bonds, the two more solvent-exposed cysteines, Cys^{87} and Cys^{90}, then complete the tetrahedral complex, the act of which nucleates the α-helix and the folding of globular domain I. Without a functioning ligand at Cys^{87} or Cys^{90}, the site is simply destabilized to an equivalent degree, with the remaining Cys clearly forming a coordination bond in the absence of the other. In contrast, when the native ligand is absent at either His^{64} or Cys^{77}, instead of forming a native-like site, the coordination chemistry strongly drives either the formation of metal-linked bimolecular tetrathiolate, tetrahedral coordination complex (e.g., in the His^{64} mutants) or the formation of complexes of apparently alternative coordination geometries (e.g., C77A gp32). The bimolecular sites in particular must be of relatively high affinity since they very efficiently compete with EGTA for Co(II) in normal metal competition experiments (data not shown), unlike the final monomolecular site.

At excess Co(II), tetrahedral metal coordination geometry appears to be maintained in each of the gp32 variants, with the fourth ligand provided by the substituted amino acid (e.g., H64C gp32), a water molecule, or other solvent-derived ligand (Cl^{-}). However, in every case, the chelate suffers large decrements in thermodynamic stability (5−7 kcal mol^{-1}) depending upon the substitution, even when a new protein-derived coordination bond is formed (e.g., H64C gp32). These findings suggest that a nonoptimal coordination bond is basically as stabilizing as no coordination bond. This is consistent with studies on CAII (Keifer & Fierke, 1994; Ippolito et al., 1995) and contrasts sharply with the TFIIIA-like zinc finger (Krizek et al., 1993). Our results help to explain a common phenomenon encountered by investigators working toward *de novo* design of zinc complexes (for a review, see Regan, 1995). As is becoming apparent, it does not seem very difficult to design or redesign a metal site to which the metal binds to the predicted ligands and desired geometry; however, the binding site is rarely found to be as stable as naturally occurring sites (Klemba et al., 1995; Klemba & Regan, 1995). One notable exception to this is the incorporation of a fourth protein-derived ligand into the native His_3 site of CAII by displacing the zinc-bound hydroxide which increases the zinc binding affinity by ≈200-fold (Ippolito et al., 1995).

Current work is aimed at determining the structures of mutant metal complexes of gp32 and evaluating to what extent other redesigned sites containing potentially liganding substitutions are tolerated. Based on the native gp32 structure, Cys^{87} and Cys^{90} are excellent candidates for further substitution, perhaps in the context of non-native ligands at residue 64. For example, a bulky liganding side chain incorporated at residue 87 and/or 90 may partially overcome the large attenutation in stability exhibited by a mutant containing a shorter side chain at residue 64 (e.g., Cys^{64}), in a way which optimizes the metal−ligand bond distance, which appears to play the primary role in maintaining high chelate stability (Ippolito et al., 1995).

ACKNOWLEDGMENT

We are indebted to Dr. Paul Lindahl, Department of Chemistry, for extensive use his anaerobic facilities and Dr. Paul Fitzpatrick, Department of Biochemistry and Biophys-

ics, for use of his Hewlett-Packard 8452A spectrophotometer. We also acknowledge Mr. Jon Christopher, TAMU, for his assistance in generating the structural models of the gene 32 core fragment from coordinates refined and made available by Dr. Yousif Shamoo, Yale University.

REFERENCES

Alexander, R. S., Kiefer, L. L., Fierke, C. A., & Christianson, D. W. (1993) *Biochemistry 32*, 1510−1518.
Berg, J. M., & Merkle, D. L. (1989) *J. Am. Chem. Soc. 111*, 3759−3761.
Berg, J. M., & Shi, Y. (1996) *Science 271*, 1081−1085.
Coleman, J. E. (1992) *Annu. Rev. Biochem. 61*, 897−946.
Corwin, D. T., Jr., Fikar, R., & Koch, S. A. (1987) *Inorg. Chem. 26*, 3079−3080.
Corwin, D. T., Jr., Gruff, E. S., & Koch, S. A. (1988) *Inorg. Chim. Acta 151*, 5−6.
Garmer, D. R., & Krauss, M. (1993) *J. Am. Chem. Soc. 115*, 10247−10257.
Giedroc, D. P. (1994) in *Encyclopedia of Inorganic Chemistry* (King, R. B., Ed.) pp 4392−4406, John Wiley & Sons, Sussex, England.
Giedroc, D. P., Keating, K. M., Williams, K. R., Konigsberg, W. H., & Coleman, J. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8452−8456.
Giedroc, D. P., Keating, K. M., Williams, K. R., & Coleman, J. E. (1987) *Biochemistry 26*, 5251−5259.
Giedroc, D. P., Johnson, B. A., Armitage, I. M., & Coleman, J. E. (1989) *Biochemistry 28*, 2410−2418.
Giedroc, D. P., Khan, R., & Barnhart, K. (1990) *J. Biol. Chem. 265*, 11444−11455.
Giedroc, D. P., Qiu, H., Khan, R., King, G. C., & Chen, K. (1992) *Biochemistry 31*, 765−774.
Guo, J., Wang, S., Dong, J., Qiu, H., Scott, R. A., & Giedroc, D. P. (1995) *J. Am. Chem. Soc. 117*, 9437−9440.
Ippolito, J. A., & Christianson, D. W. (1994) *Biochemistry 33*, 15241−15249.
Ippolito, J. A., Baird, T. T., Jr., McGee, S. A., Christianson, D. W., & Fierke, C. A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5017−5021.
Kiefer, L. L., & Fierke, C. A. (1994) *Biochemistry 33*, 15233−15240.
Kiefer, L. L., Ippolito, J. A., Fierke, C. A., & Christianson, D. W. (1993a) *J. Am. Chem. Soc. 115*, 12581−12582.
Kiefer, L. L., Krebs, J. F., Paterno, S. A., & Fierke, C. A. (1993b) *Biochemistry 32*, 9896−9900.
Klemba, M., & Regan, L. (1995) *Biochemistry 34*, 10094−10100.
Klemba, M., Gardner, K. H., Marino, S., Clarke, N. D., & Regan, L. (1995) *Nature, Struct. Biol. 2*, 368−373.
Klug, A., & Rhodes, D. (1987) *Trends Biochem. Sci. 12*, 464−469.
Krizek, B. A., & Berg, J. M. (1992) *Inorg. Chem. 31*, 2984−2986.
Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L., & Berg, J. M. (1991) *J. Am. Chem. Soc. 113*, 4518−4523.
Krizek, B. A., Merkle, D. L., & Berg, J. M. (1993) *Inorg. Chem. 32*, 937−940.
Maniatis, T., Frisch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Martell, A. E., & Smith, R. M. (1974) *Critical Stability Constants*, Vol. 1, Plenum Press, New York.
Mely, Y., Cornille, F., Fournié-Zaluski, M.-C., Darlix, J.-L., Roques, B. P., & Gérard, D. (1991) *Biopolymers 31*, 899−906.
Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T., & Berg, J. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4796−4800.
Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609−1614.
Nadler, S. G., Roberts, W. J., Shamoo, Y., & Williams, K. R. (1990) *J. Biol. Chem. 265*, 10389−10394.
Pavletich, N. P., & Pabo, C. O. (1991) *Science 252*, 809−817.
Qiu, H., & Giedroc, D. P. (1994) *Biochemistry 33*, 8139−8148.
Qiu, H., Kodadek, T., & Giedroc, D. P. (1994) *J. Biol. Chem. 269*, 2773−2781.
Rebar, E. J., & Pabo, C. O. (1994) *Science 263*, 671−673.
Regan, L. (1995) *Trends Biochem. Sci. 20*, 280−285.
Regan, L., & Clarke, N. D. (1990) *Biochemistry 29*, 10878−10883.
Schmiedeskamp, M., & Klevit, R. E. (1994) *Curr. Opin. Struct. Biol. 4*, 28−35.
Shamoo, Y., Webster, K. R., Williams, K. R., & Konigsberg, W. H. (1991) *J. Biol. Chem. 266*, 7967−7970.
Shamoo, Y., Tam, A., Konigsberg, W. H., & Williams, K. R. (1993) *J. Mol. Biol. 232*, 89−104.
Shamoo, Y., Friedman, A. J., Parsons, M. R., Konigsberg, W. H., & Steitz, T. A. (1995) *Nature 376*, 362−366.
Villemain, J. L., & Giedroc, D. P. (1993) *Biochemistry 32*, 11235−11246.
Zeng, J., Vallee, B. L., & Kägi, J. H. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9984−9988.

BI9617769