Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor

TAO PAN AND JOSEPH E. COLEMAN

The Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06510

ABSTRACT   The transcription factor GAL4 from *Saccharomyces cerevisiae* contains a "zinc-finger"-like motif, Cys-Xaa$_2$-Cys-Xaa$_6$-Cys-Xaa$_6$-Cys-Xaa$_2$-Cys-Xaa$_6$-Cys, within its DNA-binding domain. A GAL4 fragment consisting of residues 1–147 plus two additional residues from the cloning vector [denoted GAL4(149*)] has been cloned and overexpressed in *Escherichia coli*. This fragment includes the entire DNA-binding domain (residues 1–74). The homogeneous GAL4-(149*) protein contains 1–1.5 moles of Zn(II) per mole of protein. The GAL4(149*) protein binds tightly to the specific 17-base-pair palindromic DNA sequence found at GAL4 binding sites as shown by gel-retention assays using a ^{32}P-labeled 23-mer containing this sequence. Removal of the intrinsic Zn(II) by EDTA at low pH abolishes binding to the 23-mer. The GAL4(149*) apoprotein can be reconstituted with Zn(II), Cd(II), or Co(II) with restoration of specific DNA binding. Titration of GAL4(149*) apoprotein with ^{113}Cd(II) shows two ^{113}Cd(II) binding sites on the molecule, one with δ of 707 ppm, suggesting coordination to four sulfur atoms, and one with δ of 669 ppm, suggesting coordination to three or four sulfur atoms. Because GAL4(149*) protein contains only six cysteine residues within its DNA-binding domain, the precise coordination of the two Cd(II) ions cannot be stated with certainty; one or more shared $-S^-$ ligands could exist. GAL4(149*) protein contains ≈40% α-helix and ≈20% β-sheet, estimated from circular dichroism. Removal of the native Zn(II) ion causes limited unfolding of secondary structure, but less than one turn of α-helix. The binding of Zn(II), Cd(II), and, to a lesser extent, Co(II) to GAL4(149*) apoprotein protects the protein from proteolysis by trypsin, which produces a 13-kDa DNA-binding core.

GAL4 is a transcription factor required for galactose utilization in *Saccharomyces cerevisiae* (1). *In vivo*, GAL4 protein binds to 17-base-pair (bp) palindromic sequences (upstream activation sequence G; UAS$_G$) upstream from galactose-inducible genes (2). Although intact GAL4 protein consists of 881 amino acids, the determinants required for recognition of UAS$_G$ sequences are located in the N-terminal 74 amino acids, as shown by deletion mutagenesis (3). The N-terminal 74 amino acids of GAL4 include six cysteine residues, four of which have been predicted to form a single Cys$_2$-Cys$_2$ "zinc finger" (4), analogous to the two Cys$_2$-Cys$_2$ zinc fingers found in the DNA-binding domains of the steroid hormone receptor proteins (5, 6).

To examine the structural basis of this recognition, we have cloned into T7 overexpression vectors in *Escherichia coli* DNA fragments of GAL4 transcription factor consisting of residues (1–74 + 2), (1–92 + 1), and (1–147 + 2). Overproduction of the entire *GAL4* gene product in *E. coli* has not been successful (7 and unpublished results). On the other hand, induction of the engineered genes for all the above fragments resulted in large overproduction of a protein

product; however, only the domain consisting of (1–147 + 2) produced a soluble protein. We purified GAL4(1–147 + 2) [denoted GAL4(149*)] in high yield to >95% homogeneity. We found that GAL4(149*) protein isolated from *E. coli* incorporates stoichiometric amounts of Zn(II) and requires Zn(II) for binding to its specific DNA sequence. The native Zn(II) can be replaced by Cd(II) and Co(II) with restoration of specific DNA binding. The characteristics of the GAL4(149*) metalloprotein and apoprotein are reported here.

MATERIALS AND METHODS

Cloning and Overproduction of GAL4(149*) Protein. A cDNA encoding GAL4 protein (residues 1–881) was cloned into a *Nde* I/*Hind*III site of a modified pAR3039 (ref. 8; unpublished results) by ligating the *Sph* I/*Hind* III fragment from pBM1123 (supplied by M. Johnston, Washington University, Saint Louis) to a synthetic Nde I/*Sph* I linker. A *Spe* I stop-codon linker was then introduced into the *Cla* I site at amino acid 147 to yield pTPT7G1 with GAL4(147) protein under the control of the T7 RNA polymerase promoter. The *Spe* I linker adds two amino acids (Leu-Asp) onto GAL4(1–147) protein. The natural sequence after amino acid 147 is Ile-148 and Asp-149. Hence, we have termed our construct GAL4(149*). Overproduction of GAL4(149*) construct was achieved by adding isopropyl β-D-thiogalactose (1 mM) to BL21(DE3) cells/pTPT7G1 at mid-logarithmic phase, and the cells were harvested 4 hr later. GAL4(149*) protein represented ≈10% of the total *E. coli* protein.

Purification of GAL4(149*) Construct. The purification procedure reported for GAL4 fragments (7) did not yield high amounts of GAL4(149*). The procedure to be described results in much higher yields, ≈2 mg of GAL4(149*) protein per g (wet weight) of cells. Fifteen grams of cells were suspended in 40 ml of lysis buffer (50 mM Tris·HCl, pH 8.0/200 mM KCl/1 mM EDTA/1 mM dithiothreitol) and sonicated twice for 3 min on ice. Polymin P was then added over a 10-min period to 0.5%, followed by centrifugation at 25,000 × g for 20 min at 4°C. After dialysis of the supernatant against standard column buffer (STD) [10 mM Tris·HCl, pH 8.0/1 mM EDTA/1 mM 2-mercaptoethanol/10% (vol/vol) glycerol] plus 50 mM NaCl, the dialysate was loaded onto a pre-equilibrated Trisacryl-SP column and washed sequentially with STD plus 50 mM NaCl, STD plus 150 mM NaCl, and STD plus 300 mM NaCl. The fractions from the 300 mM NaCl wash were combined and dialyzed against STD plus 50 mM NaCl. The dialysate was then loaded onto a Cibacron Blue 3GA agarose column, washed with STD plus 50 mM NaCl and STD plus 250 mM NaCl, and then eluted with a NaCl gradient, STD plus 250–1000 mM NaCl. GAL4(149*) protein elutes in the 500–1000 mM NaCl fractions and is >95% pure.

Gel-Retardation Assays. Gel-retardation assays to detect GAL4(149*) binding to DNA were performed on 6% polyacrylamide gels as described (9). The buffer was 20 mM Tris·HCl/1 mM dithiothreitol/80 mM NaCl, pH 8.0 (buffer G), and sample volume was 20 μl. *GAL4*-specific DNA was

Table 1. Metal content of GAL4(149*) construct

| GAL4(149*)	Treatment	Mol/mol of protein		
		Zn	Cd	Co
Native	Dialysis (prep. 1)	1.27 ± 0.02	–	–
	Dialysis (prep. 2)	1.49 ± 0.02	–	–
	Dialysis (prep. 3)	1.14 ± 0.03	–	–
Native	Dialysis + 10 mM EDTA, pH 8	1.01 ± 0.04	–	–
Apoprotein	Dialysis + 10 mM EDTA, pH 5	0.05 ± 0.01	–	–
Cd(II)[†]	Dialysis + 1 mM CdCl$_2$,[‡] pH 8	0.30	0.68	–
Co(II)	Dialysis + 100 mM CoCl$_2$,[§] pH 5	0.05	–	1.15[¶]

[†]Cd(II)GAL4(149*) metalloprotein can also be formed by addition of stoichiometric amount of metal ion to the apoprotein.
[‡]Dialysis was done against 1 mM CdCl$_2$ at 4°C for 20 hr, followed by dialysis versus metal free buffer.
[§]Dialysis was carried out twice against 100 mM CoCl$_2$ at 4°C for 20 hr under nitrogen, followed by dialysis versus metal free buffer, pH 8.
[¶]Co(II) can be removed by exhaustive dialysis.

a synthetic ^{32}P-labeled 23-bp oligonucleotide (sequence shown below), incorporating a single *GAL4* 17-bp recognition sequence (3).

```
5'-GATCC CGGAAGACTCTCCTCCG G
        G GCCTTCTGAGAGGAGGC CCTAG
```
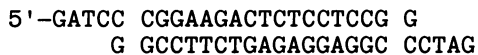
^{113}Cd NMR. The NMR was performed on a Bruker AM-500 spectrometer (110.93 MHz for ^{113}Cd) with a 10-mm broadband probe. Samples were ≈0.54 mM GAL4(149*) protein/50 mM phosphate/250 mM NaCl, pH 8.0 at 25°C; sample volumes were 2.0 ml. A 45° pulse and a recycle time of 2 sec were used.

Zinc, Cadmium, and Cobalt Analyses. These were performed by atomic absorption spectroscopy using an Instrumentation Laboratory (Lexington, MA) IL157 spectrometer.

RESULTS

Metal Content of GAL4(149*) Construct. GAL4(149*) contains from 1 to 1.5 mol of Zn(II) per mole of protein (Table 1). Zn(II) cannot be removed by EDTA at pH 8, but exhaustive dialysis will reduce the Zn(II) content to 1 mol per mol of protein. GAL4(149*) apoprotein can be prepared by dialysis at pH 5 against 10 mM EDTA/50 mM sodium acetate buffer. After removal of the Zn(II)–EDTA complex by extensive

dialysis, the pH can be returned to 8 with the production of a stable apoprotein that contains 0.05 mol of Zn(II) per mol. Zn(II) can be reconstituted to the apoprotein by adding a 1:1 molar ratio of the metal ion in buffer G and incubating for 60 min at 4°C. Both Cd(II) and Co(II) can be substituted for Zn(II) (Table 1).

Binding of GAL4(149*) to DNA. Native GAL4(149*) retains the ^{32}P-labeled 23-mer double-stranded DNA incorporating the 17-bp-specific sequence on the gel-retention assay (Fig. 1A). When the apoprotein is used, there is no retention of the DNA. Two complexes between GAL4(149*) protein and the specific 23-mer, a major and a minor one, which migrate at slightly different rates are seen on the gel-retention assay. The great specificity of the Zn(II)-dependent binding of GAL4 protein to the upstream activation sequence G (UAS$_G$) DNA is best demonstrated by the gel-retention assay, in which there is a large excess of both competing nonspecific DNA as well as GAL4 protein (Fig. 1A).

Readdition of Zn(II) in a 1:1 molar ratio to the apoprotein reestablishes retention of the specific 23-bp fragment (Fig. 1A). There may be a slight increase in the efficiency of retention at the 2:1 and 3:1 molar ratios of Zn(II) to protein. Increasing the Zn(II)-to-protein ratio to 50:1 has no further effect. Cd(II) can be substituted for Zn(II) with full retention of the specific DNA at a 1:1 Cd(II)-to-protein ratio. Co(II) will induce some binding of the specific DNA, but a 5:1
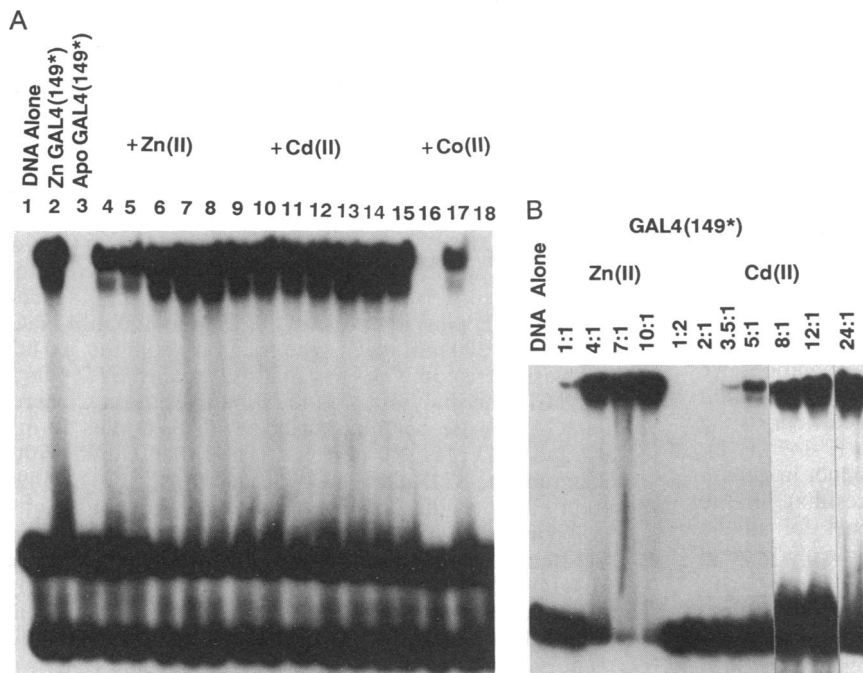

Fig. 1. Gel retardation of DNA by GAL4(149*) protein. (*A*) Retention of *GAL4*-specific ^{32}P-labeled 23-mer (10 nM) by 1 μM of purified GAL4(149*) protein in the presence of 20 μM unlabeled calf thymus DNA. Lanes: 1, DNA alone; 2, Zn(II)GAL4(149*); 3, GAL4(149*) apoprotein; 4–9, GAL4(149*) apoprotein reconstituted with 1:1, 1:2, 1:3, 1:5, 1:10, and 1:50 protein-to-Zn(II) ratios; 10–15, GAL4(149*) apoprotein reconstituted 1:1, 1:2, 1:3, 1:5, 1:10, and 1:50 protein-to-Cd(II) ratios; 16–18, GAL4(149*) apoprotein reconstituted with 1:2, 1:5, and 1:50 protein-to-Co(II) ratios. (*B*) Retention of GAL4-specific ^{32}P-labeled 23-mer (0.1 μM) as a function of the protein-to-DNA ratio of native Zn(II)GAL4(149*) and reconstituted Cd(II)GAL4(149*). The reaction mix contained 1 μM calf thymus DNA. Ratios above each lane indicate the protein-to-DNA ratio.
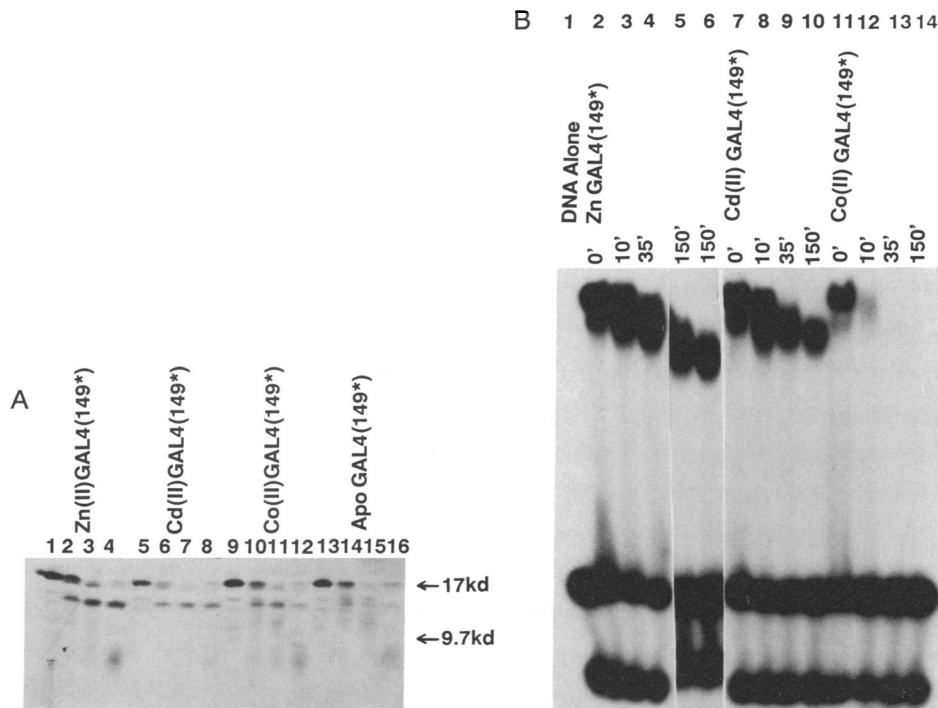
FIG. 2. Limited proteolysis of GAL4-(149*) protein by trypsin. (A) Time course of trypsin proteolysis. Samples were withdrawn at 0, 10, 35, and 150 min. Trypsin cleavage was stopped by boiling in SDS buffer for 5 min. Lanes: 1–4, Zn-(II)GAL4(149*); 5–8, Cd(II)GAL4-(149*); 9–12, Co(II)GAL4(149*); 13–16, GAL4(149*) apoprotein. (B) Gel retardation of ^{32}P-labeled *GAL4*-specific 23-mer by the tryptic fragments of GAL4(149*) protein. Trypsin cleavage was stopped by addition of soybean trypsin inhibitor in molar excess to trypsin before mixing with the ^{32}P-labeled 23-mer. Samples were withdrawn at 0, 10, 35, and 150 min after starting hydrolysis. Lanes: 1, DNA alone; 2–5, Zn(II)GAL4(149*); 6, Zn(II)-GAL4(149*), 150 min at a GAL4(149*)-to-trypsin ratio of 100:1 (wt/wt); 7–10, Cd(II)GAL4(149*); 11–14, Co(II)GAL4-(149*). In both A and B 10 μM GAL4-(149*) protein was mixed with a 400:1 (wt/wt) ratio of protein to trypsin (unless otherwise noted) in buffer G at 25°C. Co(II)GAL4(149*) contained 1.15 mol of Co(II) per mol of protein.

Co(II)-to-protein molar ratio is required before significant binding of the specific DNA is seen. Excess Co(II) destroys the binding to DNA.

The gel-retardation assay does not lend itself to the determination of the minimum protein-to-DNA stoichiometry present in a given complex. While it is sometimes possible to isolate a 1:1 protein–DNA complex by cutting out the complex from the gel, the several systems we have investigated by this technique require an excess of protein at the start to retain most DNA with the complex. This effect appears to derive from the loss of free protein as the sample enters the gel, because free protein migrates either toward the cathode or not at all. To better compare the affinities of native Zn(II)GAL4(149*) and the reconstituted Cd(II)GAL4(149*) for the specific 23-mer, gel-retention assays for both proteins as a function of the initial protein-to-DNA ratio in the presence of a 10:1 excess of calf thymus DNA are shown in Fig. 1B. The native Zn(II) protein requires a 10:1 molar ratio of protein to DNA to retain all labeled DNA. The Cd(II) protein requires approximately twice the concentration of the Zn(II) protein to retain comparable amounts of DNA (Fig. 1B). A moderate drop in DNA binding affinity has previously been seen when Cd(II) replaces Zn(II) in the single-stranded DNA binding protein, gene 32 protein from T4 (10). The GAL4(149*) apoprotein retains no DNA under these conditions. Zn(II)GAL4(149*) demonstrates some binding to a ^{32}P-labeled DNA of random sequence in the absence of calf thymus DNA (data not shown). The random DNA, however, is completely displaced by calf thymus DNA; this nonspecific binding mode is also Zn(II) dependent.

Zn(II) Protects GAL4(149*) Against Proteolysis by Trypsin. When the native Zn(II)GAL4(149*) is treated with trypsin [1:400 (wt/wt) trypsin to GAL4(149*)], rapid but limited proteolysis results in a 13-kDa core that resists further cleavage (Fig. 2A). In contrast, exposure of the GAL4(149*) apoprotein to trypsin rapidly degrades the protein to small peptides. After reconstitution with Cd(II) limited proteolysis is once again seen. Co(II) is less effective than either Zn(II) or Cd(II) in restoring resistance to complete proteolysis, and the 13-kDa Co(II) protein core is rapidly cleaved by trypsin. The 13-kDa core can bind to the specific recognition sequence for GAL4 (Fig. 2B).

Circular Dichroism of Zn(II)GAL4(149*) and GAL4(149*) Apoprotein. The native Zn(II)GAL4(149*) has significant molar ellipticity in the wavelength region of the peptide bond chromophores, $-23.0 \times 10^{5\circ}\cdot cm^2$ per dmol at 208 nm and $-15.8 \times 10^{5\circ}\cdot cm^2$ per dmol at 222 nm (Fig. 3). These values require that considerable amounts of α-helical and β-sheet structure be present in GAL4(149*) protein. Although the CD spectrum cannot be precisely matched by a combination of the three CD curves reported for homopolypeptides in the α, β, and random-coil configurations, a reasonable graphical fit is obtained by a combination of 40% α-helix, 20% β-sheet, and 40% random coil (11). Application of a computer program for calculating the Chou–Fasman prediction of secondary structure including β turns from the amino acid sequence gives 50% α-helix, 20% β-sheet, and 30% random coil for the GAL4(149*) amino acid sequence. Removal of Zn(II) results in a relatively small but significant change in the secondary structure, as shown by a decrease in negative molar ellipticity of $+7.5 \times 10^{4\circ} \cdot cm^2$ per dmol at 222 nm and $+3 \times 10^{5\circ}\cdot cm^2$ per dmol at 208 nm. Reconstituted Cd(II)GAL4(149*) shows the same CD spectrum as the native protein.
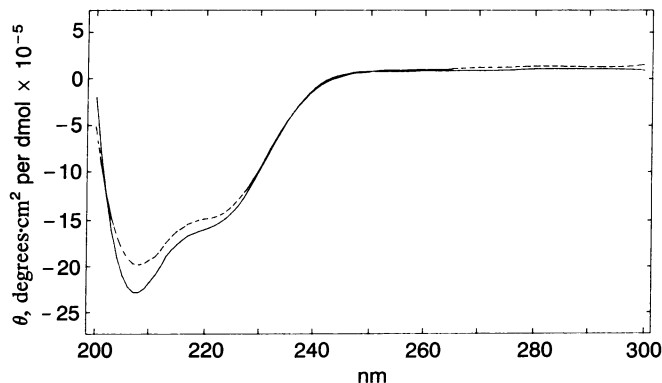

FIG. 3. Circular dichroism of Zn(II)GAL4(149*) and GAL4(149*) apoprotein. CD spectra of both Zn(II)GAL4(149*) (—) and GAL4(149*) apoprotein (----) in 10 mM Tris/150 mM NaCl/1 mM 2-mercaptoethanol/50 μM EDTA/10% glycerol, pH 8.0 at 28°C. Protein concentrations were 5 μM for both species.

Spectroscopic Properties of the Co(II)GAL4(149*) and Cd(II)GAL4(149*). Co(II) optical spectra: Visible d–d transitions of Co(II) can be used to distinguish octahedral from tetrahedral ligand geometry around Co(II), and near-UV sulfur to Co(II) charge-transfer bands can often be used to detect sulfur ligands to the Co(II). Although the Co(II)-substituted GAL4(149*) protein has the intense charge-transfer bands at 340 nm (ε = 2.2 × 10^3 mol^{-1}·cm^{-1}) and 305 (ε = 3 × 10^3 mol^{-1}·cm^{-1}) expected from sulfur ligation, there are no intense d–d transitions in the vicinity of 700 nm typical of tetrahedral Co(II) complexes. A comparison of the absorption spectrum of Co(II)GAL4(149*) protein to the tetrahedral four $-S^-$ ligand site in Co(II) alcohol dehydrogenase (12) is shown in Fig. 4. A Co(II) protein spectrum similar to this, with no tetrahedral Co(II) d–d transitions, is observed for the Co(II)-substituted B site or exchangeable Zn(II) site in *E. coli* RNA polymerase (13).

^{113}Cd NMR of ^{113}Cd(II)GAL4(149*) Protein. When two equivalents of ^{113}Cd(II) are added to GAL4(149*) apoprotein, the ^{113}Cd(II) NMR shows two signals of equal intensity at δ of 707 and 669 ppm (Fig. 5). The chemical shifts of ^{113}Cd(II) bound to protein sites with known ligand composition show that ^{113}Cd chemical shifts can be used to predict the number of sulfur donor atoms coordinating the Cd(II) ion (14–16). The δ of 707 ppm is within the chemical shift range expected for a S_4 donor set, whereas a δ of 669 ppm is compatible with either a S_3X or S_4 set (see *Discussion*). At present we have been unsuccessful in attempting to generate a GAL4(149*) protein with only one of these sites occupied with ^{113}Cd(II), either by addition of one ^{113}Cd(II) or removal of one of the bound ^{113}Cd(II) ions after reconstitution. Both ^{113}Cd(II) ions seem to be removed simultaneously, suggesting there may be some cooperative binding.

DISCUSSION

The zinc analysis on GAL4(149*) protein demonstrates that GAL4 is a zinc metalloprotein (Table 1). Likewise the recognition of the specific DNA sequence to which GAL4 transcription factor binds depends on the presence of Zn(II) (Fig. 1A). The sequence of the GAL4(149*) fragment is given in Fig. 6. Six cysteine residues occur in this fragment, Cys-11, Cys-14, Cys-28, and Cys-31, proposed to be involved in formation of a Zn(II) complex (4), and two additional cysteine
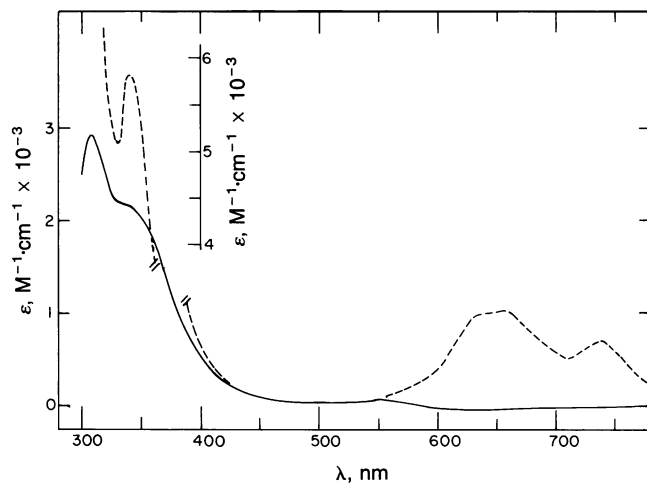
FIG. 5. ^{113}Cd NMR spectrum of GAL4(149*) protein. Spectral width is 15.2 kHz (136 ppm), and the number of transients is 9800. A 50-Hz line broadening was applied for spectral enhancement; acquisition conditions are described in text. Chemical shift is plotted relative to that of 0.1 M ^{113}CdClO₄, δ = 0. PPM, parts per million.

residues, Cys-21 and Cys-38. Mutation studies on GAL4 (18) revealed that two regions within the DNA binding domain (residues 1–74) are important in specific DNA binding. All missense mutations bearing a single amino acid change in the immediate vicinity of the putative zinc-finger (residues 10–38) abolish specific DNA binding, whereas most missense mutations in the region downstream (residues 40–50) are accompanied by partial loss of specific DNA-binding affinity.

Cd(II) effectively substitutes for Zn(II) in restoring the specific DNA-binding function to GAL4(149*) apoprotein (Fig. 1). On the other hand, Co(II) is not as effective in restoring the binding of GAL4(149*) to its specific DNA sequence (Fig. 2B). The native GAL4(149*) protein assumes a significant amount of α-helical and β-sheet secondary structure as shown by the UV CD (Fig. 3). The removal of the Zn(II) must result in only modest unfolding of this secondary structure, because the molar ellipticity becomes less negative by only 7.5 × $10^{4\circ}$·cm^2 per dmol at 222 nm and 3 × $10^{5\circ}$·cm^2 per dmol at 208 nm (Fig. 3). When the change at 222 nm is

FIG. 4.   Visible absorption spectrum of Co(II)GAL4(149*) protein (—). The spectrum is corrected against Zn(II)GAL4(149*) of the same *A* at 280 nm. The concentration of Co(II)GAL4(149*) used for this spectrum is ≈40 μM. (---) Horse liver alcohol dehydrogenase with Co(II) substituted at the structural site, replotted from ref. 12. The extinction coefficient between 300–350 nm for liver alcohol dehydrogenase should be referred to the middle ordinate.
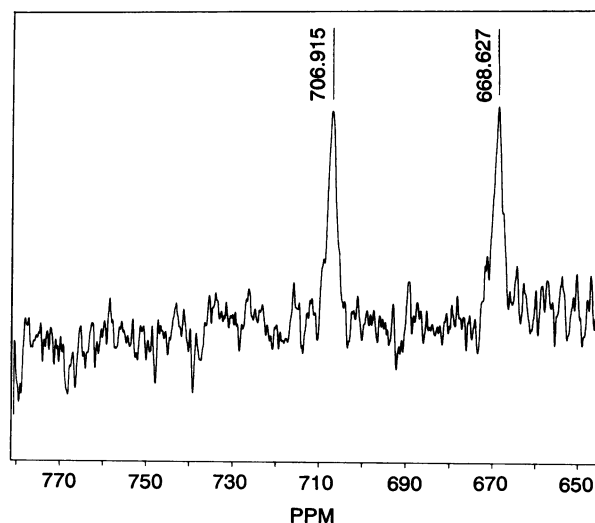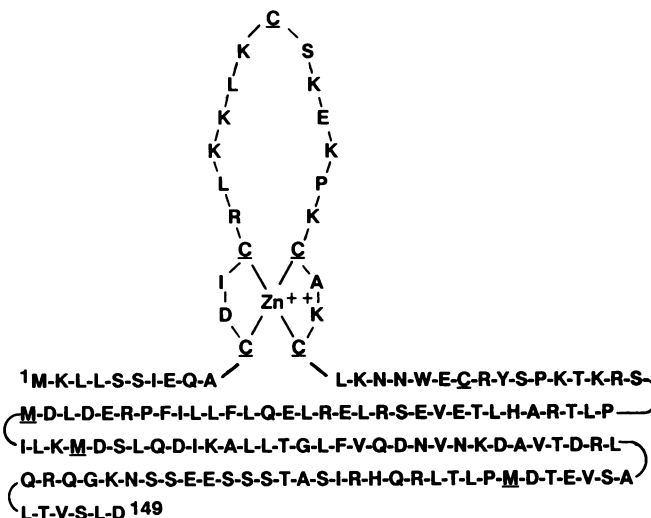
FIG. 6.   Amino acid sequence of the GAL4(149*) protein. The amino acid sequence (one-letter code) of the intact GAL4 transcription factor is taken from ref. 17, and the proposed Zn(II) ligands are as indicated in ref. 4.

Biochemistry: Pan and Coleman

Proc. Natl. Acad. Sci. USA 86 (1989)   3149

assumed to be unfolding of α-helix, then only $\approx$ two residues are induced to form α-helix by Zn(II) binding. The considerably larger change in ellipticity at 208 nm suggests that unfolding of some other type of secondary structure also occurs on Zn(II) removal, quite possibly β-turns. It seems reasonable to conclude that Zn(II) binding is involved in inducing the final conformation of a section of the polypeptide backbone of GAL4(149*) protein involved directly or indirectly in forming the DNA-binding surface. Disorganization of this section on Zn(II) removal must account for the loss of specific DNA recognition (Fig. 1A).

Extensive proteolysis and differential scanning calorimetry studies of the Zn(II) binding domain in the single-stranded DNA-binding protein, gene 32 protein from bacteriophage T4, show that Zn(II) protects a core domain from further proteolysis as well as imparting a significant increase in thermal stability to the protein (10, 19). Likewise for GAL4-(149*) protein, the Zn(II) protects a 13 kDa core from complete proteolysis (Fig. 2). Because this core contains the DNA-binding surface (Fig. 2B), it probably represents the minimal DNA-binding subdomain.

There are two general types of zinc-finger domains in transcription factors: those with the ligand composition, Cys_2-His_2—e.g., transcription factor IIIA from *Xenopus* oocytes (20, 21)—and those with the ligand composition, Cys_2-Cys_2—e.g., the steroid receptor proteins (5, 6). Few physicochemical studies have been reported on representatives of the Cys_2-Cys_2 group, but both the DNA-binding fragment of the glucocorticoid receptor and now the appropriate GAL4 fragment contain the expected Zn(II) content (Table 1) (6). The structure surrounding the metal ion is not clear in proteins belonging to this group. Number and types of amino acid residues between pairs of putative cysteine ligands is highly variable, and the sequence often contains extra cysteines. Hence, precise folding schemes have not been well defined. For GAL4(149*), the present data show that, while the metal ion induces additional folding of the polypeptide chain, the induced folding must be highly localized within a well-defined secondary structure and does not appear to increase the content of α-helix (Fig. 3).

The optical absorption spectrum of the Co(II) derivative of GAL4(149*) protein does not suggest tetrahedral geometry (Fig. 4). Co(II) in this case, however, may not induce the normal configuration of the coordination complex, because the restoration of DNA binding is clearly defective (Fig. 1A). ^{113}Cd NMR of the ^{113}Cd(II) GAL4(149*) protein shows that a metal site with four S$^-$ ligands exists in GAL4(149*) protein, the site with δ of 707 ppm (Fig. 5). Most such sites in proteins have been found to consist of some variant of tetrahedral coordination geometry. The human glucocorticoid receptor has been shown to contain two Zn(II) ions per molecule within the DNA-binding domain of 150 amino acid residues that contains two Cys_2-Cys_2 zinc fingers (5, 6). The presence of the two Zn(II) ions is required for the binding of the 150-residue fragment to the specific DNA sequence recognized by the glucocorticoid receptor. Extended x-ray absorption fine structure of the Zn(II) protein fragment suggests that both Zn(II) ions occupy tetrahedral sites consisting of four S$^-$ ligands (6).

The unexpected finding in the case of GAL4(149*) protein is the presence of a second metal binding site with coordination to at least three sulfur donors (Fig. 5). At present we cannot state whether the occupancy of this site affects DNA binding. Although GAL4(149*) protein, as isolated, consistently contains >1 mol of Zn(II) per mol, the protein can be treated with metal-free solutions, such that the average Zn(II)

content is reduced to one mol per mol without obvious loss of DNA binding. The conditions of the gel-retention assay, however, do not lend themselves to careful correlation of DNA-binding affinity as a function of Zn(II) stoichiometry. In fact, the gel in Fig. 1A could be interpreted to show some difference in efficiency of retention between the 1:1, 2:1, and 3:1 Zn(II) stoichiometries, although this is less evident for the Cd(II) derivative.

The distribution of ligands in GAL4(149*) protein that provide for sites with two such downfield ^{113}Cd NMR signals is not obvious in the absence of more structural information. Although four $-S^-$ coordination can explain one of the ^{113}Cd NMR signals, there are only two additional $-S^-$ ligands possible. Possibly the two Cd(II) ions share one or more $-S^-$ ligands, creating a two-metal cluster. Lack of ^{113}Cd-^{113}Cd coupling might be thought to rule this alternative out, but such coupling (30–50 Hz) may not be easy to resolve. There are also three methionine residues toward the C-terminal region of GAL4(149*) transcription factor, and the sulfur in this ether linkage can be a donor to Cd(II) (22). The methionine residues are not within the sequence involved in DNA binding, but the second Cd(II) site may not directly affect ligand binding. More definitive definition of structure of the metal complexes awaits more precise probes like extended x-ray absorption fine structure and the examination of fluorescence quenching upon nucleotide binding to gain a more exact measure of DNA-binding affinity versus metal ion stoichiometry.

1. Oshima, Y. (1982) in *Molecular Biology of the Yeast Saccharomyces*, eds. Strathern, J., Jones, E. & Broach, J. K. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), Vol. 1, pp. 159–180.
2. Giniger, E., Varnum, S. M. & Ptashne, M. (1985) *Cell* **40**, 767–774.
3. Keegan, L., Gill, G. & Ptashne, M. (1986) *Science* **231**, 699–704.
4. Johnston, M. (1987) *Nature (London)* **328**, 353–355.
5. Severne, Y., Wieland, S., Schaffner, W. & Rusconi, S. (1988) *EMBO J.* **7**, 2503–2508.
6. Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B. & Yamamoto, K. R. (1988) *Nature (London)* **334**, 543–546.
7. Lin, Y.-S., Carey, M. F., Ptashne, M. & Green, M. R. (1988) *Cell* **54**, 659–664.
8. Studier, F. W. & Moffatt, N. (1986) *J. Mol. Biol.* **189**, 113–130.
9. Fried, M. G. & Crothers, D. M. (1981) *Nucleic Acids Res.* **5**, 6505–6525.
10. Giedroc, D. P., Keating, K. M., Williams, K. R. & Coleman, J. E. (1987) *Biochemistry* **26**, 5251–5259.
11. Greenfield, N. & Fasman, G. D. (1969) *Biochemistry* **8**, 4108–4116.
12. Maret, W., Andersson, I., Dietrich, H., Schneider-Bernlohr, H., Einarsson, R. & Zeppezauer, M. (1979) *Eur. J. Biochem.* **98**, 501–509.
13. Giedroc, D. P. & Coleman, J. E. (1986) *Biochemistry* **25**, 4969–4978.
14. Bobsein, B. R. & Myers, R. J. (1980) *J. Am. Chem. Soc.* **102**, 2454–2455.
15. Armitage, I. M. & Otvos, J. D. (1982) in *Biological Magnetic Resonance*, eds. Berliner, L. J. & Reuben, J. (Plenum, New York), Vol. 4, pp. 79–144.
16. Giedroc, D. P., Johnson, B. A., Armitage, I. M. & Coleman, J. E. (1989) *Biochemistry* **28**, 2410–2418.
17. Laughon, A. & Gesteland, R. F. (1984) *Mol. Cell. Biol.* **4**, 260–267.
18. Johnston, M. & Dover, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2401–2405.
19. Keating, K. M., Ghosaini, L. R., Giedroc, D. P., Williams, K. R., Coleman, J. E. & Sturtevant, J. M. (1988) *Biochemistry* **27**, 5240–5245.
20. Miller, J., Mclachlan, A. P. & Klug, A. (1985) *EMBO J.* **4**, 1609–1614.
21. Diakun, G. P., Fairall, C. & Klug, A. (1986) *Nature (London)* **324**, 698–699.
22. Engeseth, H. R., McMillin, D. & Otvos, J. D. (1984) *J. Biol. Chem.* **259**, 4822–4826.