Protein Science (1992), *1*, 1403–1412. Cambridge University Press. Printed in the USA.
Copyright © 1992 The Protein Society

Effects of DNA binding and metal substitution on the dynamics of the GAL4 DNA-binding domain as studied by amide proton exchange

TED MAU, JAMES D. BALEJA,[1] AND GERHARD WAGNER

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

(RECEIVED July 9, 1992; REVISED MANUSCRIPT RECEIVED August 21, 1992)

Abstract

Backbone amide proton exchange rates in the DNA-binding domain of GAL4 have been determined using ^{1}H–^{15}N heteronuclear correlation NMR spectroscopy. Three forms of the protein were studied—the native Zn-containing protein, the Cd-substituted protein, and a Zn–GAL4/DNA complex. Exchange rates in the Zn-containing protein are significantly slower than in the Cd-substituted protein. This shows that Cd-substituted GAL4 is destabilized relative to the native Zn-containing protein. Upon DNA binding, global retardation of amide proton exchange with solvent was observed, indicating that internal fluctuations of the DNA-recognition module are significantly reduced by the presence of DNA. In all forms of the protein, the internal dyad symmetry of the DNA-recognition module of GAL4 is reflected by the backbone amide proton exchange rates.

Keywords: amide exchange; DNA-binding protein; GAL4; metal substitution; NMR; transcriptional activation

GAL4 is a transcriptional activator that binds as a dimer to the galactose upstream activating sequence (UAS_G) in the yeast genome and activates the transcription of genes encoding enzymes of the galactose metabolism pathway (Bram & Kornberg, 1985; Giniger et al., 1985; Johnston, 1987; Carey et al., 1989). Its DNA-binding domain, a bimetal thiolate cluster (Pan & Coleman, 1990a, 1991; Povey et al., 1990; Gadhavi et al, 1991), has been termed a class 3 Zn-containing DNA-binding protein, whereas class 1 consists of the classical zinc fingers, and class 2 is represented by the steroid receptor DNA-binding domains (Harrison, 1991). Two three-dimensional (3D) structures of GAL4 fragments that include the DNA-binding domain have been determined by NMR methods: a Zn-containing peptide comprising residues 7–49 (Kraulis et al., 1992) and a Cd-substituted protein comprising residues 1–65 (Baleja et al., 1992). In addition, the structure of Cd-GAL4(1–65) in complex with DNA has been solved crystallographically (Marmorstein et al., 1992). The structure of the protein–DNA complex reveals a dimeric binding motif. The N-terminal DNA-recognition modules lie in the major groove of the DNA at the far ends of the DNA-recognition site. They are connected by extended chains to a dimerization element that forms a coiled-coil in the GAL4 dimer.

This work investigates the dynamics of the DNA-recognition module of GAL4 both in its free form and in complex with a consensus target DNA site by measuring hydrogen exchange rates with NMR spectroscopy. Hydrogen exchange can be used to study two aspects of protein dynamics. The first involves the qualitative classification of labile protons into a rapidly exchanging group and a slowly exchanging group. Identification of the rapidly exchanging protons establishes a map of the solvent-exposed surface of a protein, whereas identification of the slowly exchanging amides maps the protein interior. This method can be used to identify interfaces between proteins and ligands by comparing exchange rates of the free protein and that of a protein–ligand complex. Retardation of hydrogen–deuterium exchange has been used, for example, to map the antibody binding site in cytochrome *c* (Paterson et al., 1990) and to characterize antigen-induced changes in the structure of an antigen-binding unit of immunoglobulin G (Takahashi et al., 1991).

Quantitative measurement of exchange rates for interior amide protons, the second aspect of an amide exchange

Reprint requests to: Dr. Gerhard Wagner, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., Boston, Massachusetts 02115.

[1] Present address: Department of Biochemistry, Tufts University School of Medicine, Boston, Massachusetts 02111.

study, provides insight into the stability and dynamics of a protein (Hvidt & Nielsen, 1966; Englander et al., 1972; Wagner, 1983; Englander & Kallenbach, 1984). Here we make use of this tool to investigate the internal mobility of GAL4. The rates at which backbone amide protons of the GAL4 DNA-recognition module exchange with solvent deuterons are determined by the direct exchange-out method in which protein samples lyophilized from H_2O are dissolved in D_2O, and the decay of intensities for the amide proton NMR signals are monitored over time as amide protons are replaced by deuterons from the solvent. Compared to earlier studies where homonuclear correlation spectroscopy (COSY) spectra were used to monitor the exchange (Wagner & Wüthrich, 1982), here the more sensitive ^1H-^{15}N correlated spectra were employed. Rates for the free native Zn-containing protein are compared to those for the DNA-bound protein. Two metallo-forms of GAL4 are also studied, one containing the native Zn^{2+} ion and the other containing Cd^{2+}. Cadmium substitution was used in previous NMR studies of GAL4 to take advantage of the NMR-sensitive ^{113}Cd nucleus (Pan & Coleman, 1989, 1990a,b, 1991; Povey et al., 1990; Gadhavi et al., 1991; Gardner et al., 1991; Baleja et al., 1992). The exchange rate comparisons reveal subtle structural and dynamic differences between free and bound forms of the DNA-recognition module and between the two metallo-forms of GAL4.

Results

NMR assignment of ^1H and ^{15}N resonances in three forms of GAL4

Resonance assignments for backbone amide protons and nitrogens of the DNA-recognition module of GAL4 (residues 11–40) were based on previously published ^1H resonance assignments (Gadhavi et al., 1990; Gardner et al., 1991; Pan & Coleman, 1991) and on experiments carried out in our laboratory (Baleja & Wagner, unpubl.) (Table 1). For Cd-GAL4 (Fig. 1A), well-dispersed ^1H-^{15}N cross peaks in heteronuclear single quantum coherence (HSQC) spectra were assigned from the known chemical shifts of the attached protons, whereas overlapping resonances were assigned by examining tiers of unique ^{15}N frequency from a 3D ^{15}N-dispersed nuclear Overhauser effect spectroscopy (NOESY) spectrum of Cd-GAL4 and matching the NOE cross peaks with those previously identified in two-dimensional (2D) homonuclear NOESY spectra. In this way, all backbone ^1H-^{15}N resonances from the recognition module were assigned. Cross-peaks in the HSQC spectrum of Zn-GAL4 (Fig. 1B) were assigned from previously published ^1H chemical shifts (Gadhavi et al., 1990) and by comparison to the ^1H-^{15}N HSQC spectrum of Cd-GAL4. Twenty-three backbone amide protons of the recognition module could be clearly identified on this basis. The remaining backbone amide protons were unassigned due to overlapping ^1H chemical shifts and the lack of ^{15}N-dispersed 3D NMR data.

The protein–DNA complex was formed by adding aliquots of DNA to a 0.7 mM protein solution. Formation of the complex was monitored by one-dimensional ^1H spectra and by 2D ^1H-^{15}N HSQC spectra. Some amide resonances remained relatively sharp throughout the course of the titration; others broadened beyond detection due to some intermediate exchange process (or processes) between the free and bound forms of the protein.

As a nucleus of the protein is in an equilibrium between its chemical environment in the free molecule and that in the complex, its monitored resonance undergoes characteristic changes that reflect the kinetics of the chemical exchange (Harris, 1983). If the exchange lifetime is short compared to the NMR time scale, a single resonance will be observed at the chemical shift frequency corresponding to the weighted average of the chemical shifts of the two environments (fast exchange). If the exchange lifetime is much longer than the NMR time scale, the observed spectrum is the superposition of the spectra of the two species (slow exchange). Intermediate exchange results in broadened line shape and a spectrum with intermediate appearances. In this respect, the NMR time scale is primarily a function of the chemical shift difference for the exchanging proton in the two chemical environments and is roughly between 1 and 1×10^{-5} s in this study.

Assignment of resonances in the HSQC spectrum of the Zn-GAL4–DNA complex (Fig. 1C) was more difficult than in the free form of the protein due to the intermediate exchange behavior of many resonances. As a consequence of multiexchange phenomena (see Discussion), changing sample conditions (ionic strength, temperature, and protein:DNA ratio) was of limited utility for obtaining better spectra. Some assignments, for resonances that were in fast exchange, could be made simply by following the assigned resonances of the free protein during the course of the titration. Other assignments, involving amide protons with good chemical shift dispersion, were based on assigned cross-peaks in the 2D homonuclear NOESY spectrum of the complex and by comparison to the ^1H-^{15}N HSQC spectrum of free Zn-GAL4. ^1H-^{15}N cross-peaks in the spectrum could be assigned for 18 backbone amides, as well as for the side-chain amides of four residues (Table 1).

Amide proton exchange rates in Cd-GAL4, Zn-GAL4, and the Zn-GAL4–DNA complex

Rate constants for amide exchange are presented in Table 2. For Cd-GAL4 at 25 °C and pH 6.67, nine cross-peaks were seen in the first HSQC spectrum acquired after exchange-out was initiated (Fig. 2A; Kinemage 1). A lower limit of 0.194 min^{-1} is calculated for the rate constants of other amide protons that have exchanged completely within the first 13 min. For Zn-GAL4, at 35 °C

Table 1. *^1H and ^{15}N NMR assignments of amide protons for the DNA recognition module of Cd-GAL4, Zn-GAL4, and the Zn-GAL4–DNA complex at 25 °C*[a]

	Cd-GAL4 (pH 7.18)		Zn-GAL4 (pH 6.21)		Zn-GAL4–DNA[b] (pH 5.42)	
	^1H	^{15}N	^1H	^{15}N	^1H	^{15}N
Backbone						
C11	8.96	123.87	9.35	125.10	9.51	124.66
D12	9.07	120.57	9.13	120.87	9.12	119.98
I13	7.86	122.03	[c]	[c]	[c]	[c]
C14	7.77	124.81	7.86	124.56	7.99	125.99
R15	8.54	119.97	8.49	119.32	8.48	119.09
L16	7.91	122.66	[c]	[c]	[c]	[c]
K17	8.02	117.67	8.20	117.61	[c]	[c]
K18	7.71	118.52	7.71	118.70	[c]	[c]
L19	8.53	120.65	8.66	120.61	[c]	[c]
K20	8.37	122.65	[c]	[c]	[c]	[c]
C21	8.82	133.83	8.80	134.54	[c]	[c]
S22	8.94	127.23	9.05	127.32	9.08	127.96
K23	9.64	120.85	9.50	120.42	9.60	121.03
E24	7.99	119.75	7.32	118.92	[c]	[c]
K25	8.35	119.93	[c]	[c]	[c]	[c]
P26	–	–	–	–	–	–
K27	7.77	120.58	7.80	120.02	7.83	119.89
C28	9.43	126.19	10.04	127.86	10.04	128.08
A29	8.17	123.27	[c]	[c]	[c]	[c]
K30	8.15	121.36	[c]	[c]	[c]	[c]
C31	7.88	124.62	7.97	124.56	7.98	124.39
L32	8.45	121.44	8.65	121.28	8.64	121.26
K33	7.80	119.94	7.82	120.01	7.86	119.97
N34	7.00	114.47	7.02	114.14	7.08	114.27
N35	7.54	117.98	7.50	117.83	7.49	117.74
W36	8.34	120.65	8.49	120.47	8.54	120.59
E37	8.68	123.74	8.74	124.01	8.80	124.24
C38	8.17	132.46	8.10	133.13	7.93	132.53
R39	8.00	131.65	7.99	130.92	7.81	130.69
Y40	8.97	125.28	8.94	124.90	8.93	125.27
S41	9.19	123.18	9.11	123.25	9.11	124.66
P42	–	–	–	–	–	–
K43	8.55	123.70	8.56	123.59	8.57	124.73
Side chain						
Q9 ϵNH_2	7.61, 6.95	114.39	7.63, 6.93	114.37	7.87, 7.08	114.40
N34 δNH_2	6.96, 6.20	115.79	7.00, 6.35	115.93	6.97, 6.25	115.90
N35 δNH_2	7.53, 6.80	113.72	7.52, 6.79	113.61	7.52, 6.78	113.70
W36 NϵH	10.09	130.04	10.06	129.97	10.08	130.09

[a] Chemical shifts are given in ppm. ^1H chemical shifts are referenced to trimethylsilyl-propionate (TSP), and the uncertainty in measurement is 0.01 ppm. ^{15}N chemical shifts are referenced to external $^{15}NH_4Cl$ (2.9 M in 1 M HCl at 20 °C) at 24.93 ppm, with an uncertainty of 0.05 ppm.

[b] The DNA:protein (monomer) ratio was approximately 1.5:1.

[c] Amide resonance is unassigned because of peak overlap or weak signal intensity due to large line-width.

and pH 7.02, 10 backbone amide protons exchanged sufficiently slowly with solvent to be observed in the first HSQC spectrum acquired after initiation of exchange-out (Fig. 2B). In addition to the nine slowly exchanging amide protons identified in Cd-GAL4, that of D12 was also observed. The exchange rates of six backbone amide protons, C11, C14, R15, C28, C31, and L32, were determined quantitatively. The resonances of four backbone amide protons, D12, L19, E24, and W36, were seen only in the first HSQC spectrum and have decayed to the level of the noise in the second spectrum. Their rate constants were given upper and lower limits of 0.252 and 0.101 min^{-1}, respectively. The former value represents a lower limit of the exchange rates for amide protons that exchanged too fast to be seen in this experiment.

In the presence of DNA, and at 35 °C and pH 6.81, 12 amide protons were observed in the first HSQC spectrum, 8 of which were assigned (Fig. 2C). Exchange rates for the

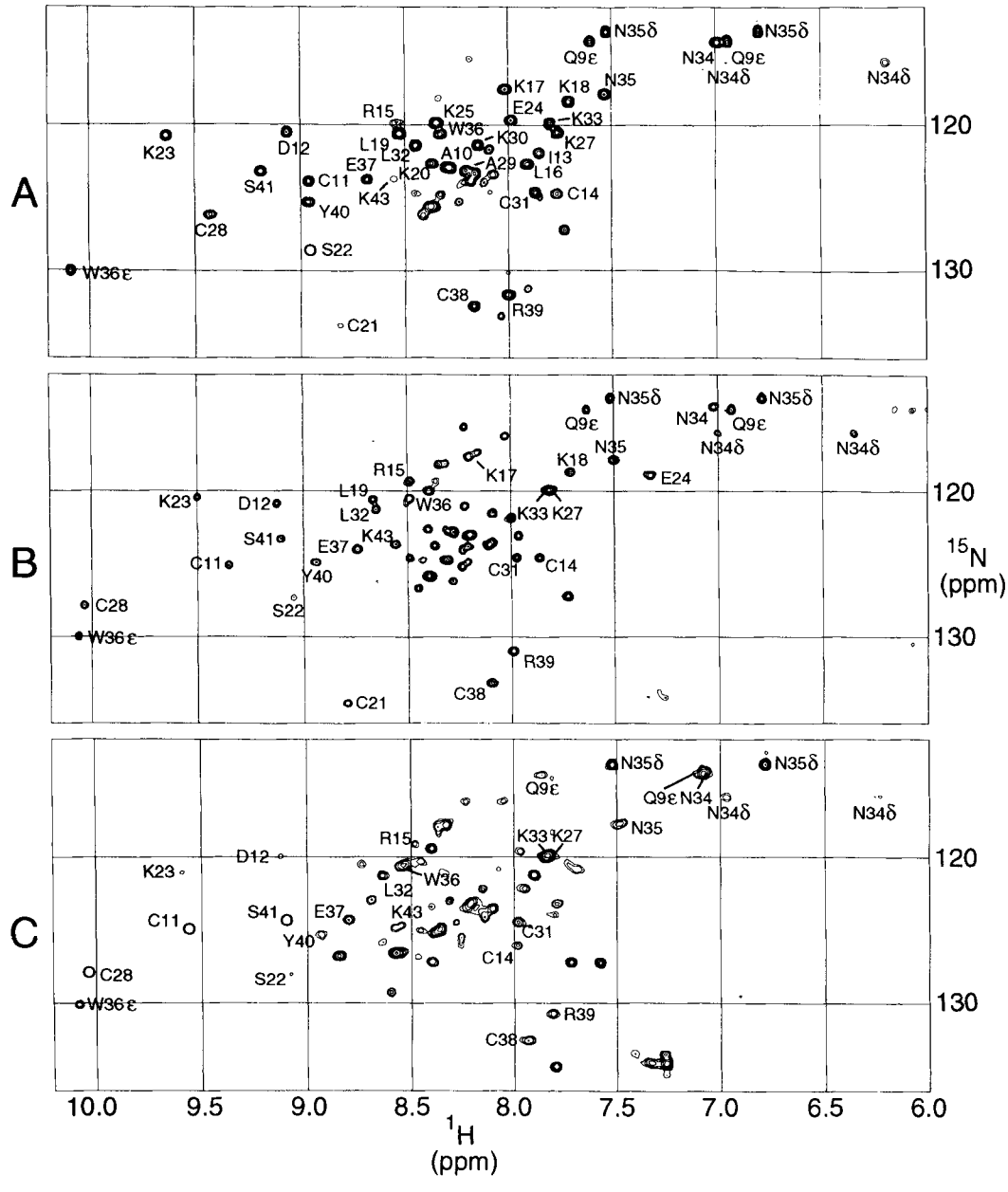

Fig. 1. ^{1}H-^{15}N HSQC spectra at 25 °C in H_2O. **A:** Cd-GAL4 – 1.5 mM, pH 7.18. **B:** Zn-GAL4 – 0.6 mM, pH 6.21. **C:** Zn-GAL4-DNA-complex at a protein (monomer):DNA ratio of approximately 1.5:1, 0.6 mM, pH 5.42. Cross-peaks indicated by open circles are visible when the spectra are presented at lower threshold settings. However, these settings would also obscure cross-peaks in the crowded spectral regions centered at $^{1}H = 8.3/^{15}N = 124$ ppm.

NH of C11, D12, C14, R15, C28, C31, and L32 were obtained, and that of W36 was estimated (see Table 2). The rate constants for Zn-GAL4-DNA were determined with rather large uncertainties due to the low signal-to-noise of the spectra, although useful comparisons to rate constants for the free protein could still be made.

Discussion

Structure of the recognition module

A drawing for the DNA-recognition module of GAL4 (Baleja et al., 1992) is shown in Figure 3 and Kinemage 1. The module consists of an α-helix from residues 11 to 18 followed by an extended loop from residues 19 to 23, leading to a tight turn at cis-proline 26. Residues 28–35 comprise a second α-helix, and residues 36–40 form an extended strand. The second helix-extended strand element is nearly a structural duplication of the first. The internal dyad symmetry is revealed when the backbone of residues 10–22 is superimposed onto that of residues 27–39 after a 180° rotation (Baleja et al., 1992; Kraulis et al., 1992; Marmorstein et al., 1992). Each metal ion is liganded by three cysteines from one helix-strand element and a fourth from the other. Two of the six cysteines thus ligand both

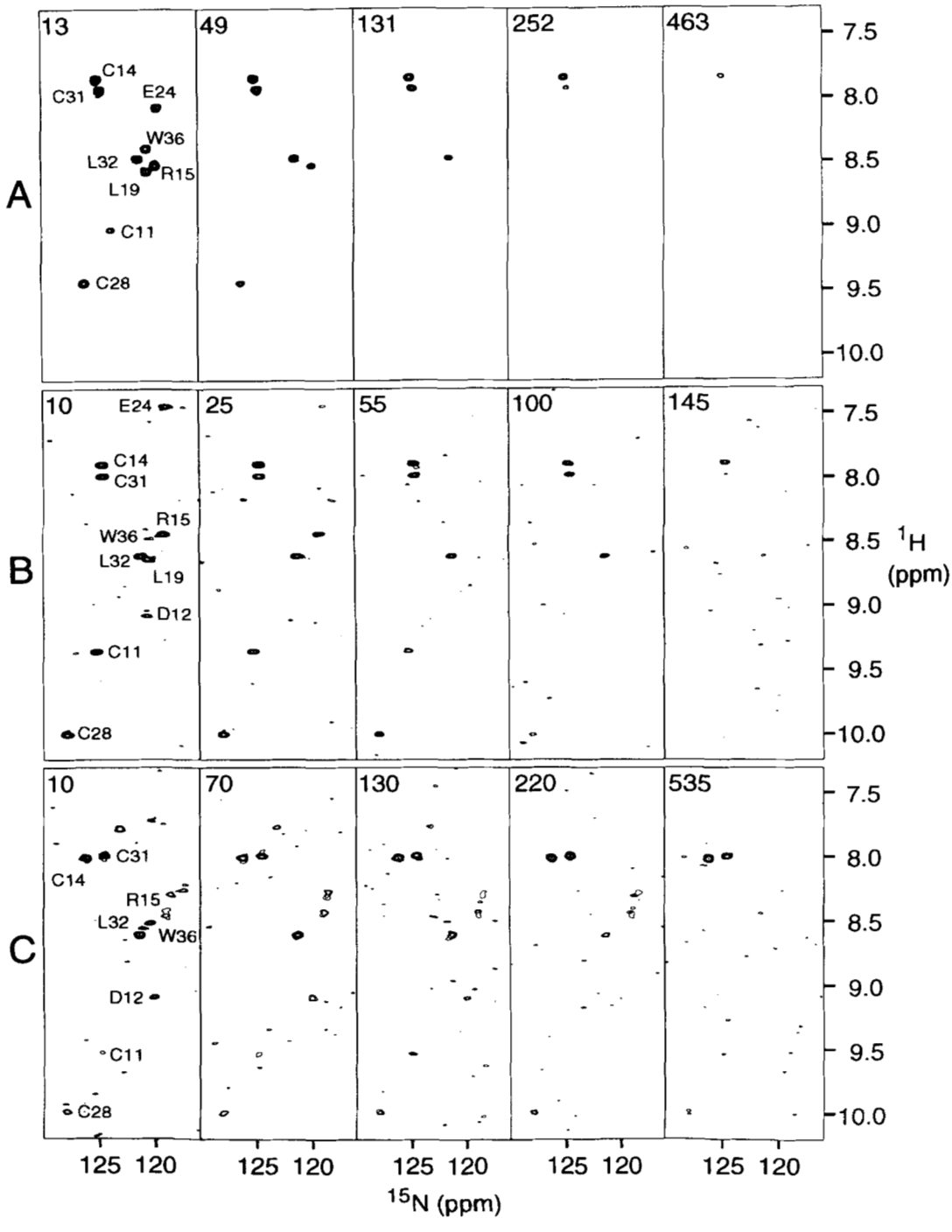

Fig. 2. Amide proton exchange-out of (**A**) Cd-GAL4 at 25 °C, pH 6.67; (**B**) Zn-GAL4 at 35 °C, pH 7.02; and (**C**) Zn-GAL4–DNA complex at a protein (monomer):DNA ratio of 1.5:1 at 35 °C, pH 6.81. The number at the upper left corner of each spectrum indicates the start of acquisition in minutes after initiation of exchange-out.

metal ions (Pan & Coleman, 1990a, 1991; Povey et al., 1990; Gadhavi et al., 1991). Other fungal transcription factors, that have a high sequence homology with GAL4 (Johnston, 1987), would presumably share the same twofold symmetric pattern of coordination of the two metal ions by the six cysteines.

The internal twofold symmetry of the module is reflected by the exchange rates of the backbone amide protons (Table 2). The two slowest-exchanging amide protons are those of C14 and C31, two cysteines related by dyad symmetry. The only other two cysteines with slow-exchanging amide protons are the bridging cysteines C11 and C28, also related by dyad symmetry. The two remaining conserved cysteines, the dyad-related C21 and C38, have relatively fast-exchanging amide protons that are not observed in Cd-, Zn-, or DNA-bound GAL4. The three

Table 2. *Backbone amide proton exchange rate constants*[a]

	Cd-GAL4 (25 °C, pH 6.67)	Zn-GAL4 (35 °C, pH 7.02)	Zn-GAL4-DNA[b] (35 °C, pH 6.81)
C11	78.7 ± 4.4	25.6 ± 1.2	1.1 ± 0.4
D12	>194[c]	176 ± 75[d]	3.4 ± 0.6
C14	5.9 ± 0.1	13.2 ± 0.7	0.4 ± 0.1
R15	36.0 ± 1.5	46.2 ± 3.1	1.4 ± 0.3
L19	78.1 ± 3.7	176 ± 75[d]	[e]
E24	64.0 ± 4.9	176 ± 75[d]	[e]
C28	29.4 ± 1.1	24.2 ± 2.1	1.2 ± 0.4
C31	13.1 ± 0.2	17.1 ± 0.8	0.4 ± 0.1
L32	16.3 ± 0.3	22.6 ± 1.7	2.8 ± 0.3
W36	92.6 ± 2.5	176 ± 75[d]	176 ± 75[d]

[a] See Materials and methods for details. The observed rate constants (k_{obs}) are given in 10^{-3} min^{-1}.

[b] Exchange rates are measured at a DNA:protein (monomer) ratio of 0.64:1. They represent upper limits of the actual rates at which amide protons in DNA-bound GAL4 exchange with solvent, because the rates are measured during exchange between the free and the bound proteins.

[c] The amide proton exchanged completely before acquisition of the first spectrum.

[d] The resonance was seen only in the first spectrum collected after initiation of exchange-out, and the given rate constant reflects estimated upper and lower limits of the exchange rate, as described in Materials and methods.

[e] The amide proton was unassigned in the protein-DNA complex.

subsets of exchange rates also correspond well with the relative stability of the environments of the amide protons. Not surprisingly, the slowest pair of amide protons is found in the middle of the α-helices (C14, C31). The fastest pair of amide protons belongs to residues in the extended strands (C21, C38), and the intermediate group belongs to cysteine residues at the N-termini of the α-helices (C11, C28). The validity of this grouping in all three species of GAL4 examined also argue against drastic changes in the structure of the bimetal-thiolate cluster when Cd(II) is substituted for Zn(II) or when GAL4 binds to DNA.

The hydrogen-bonding pattern in the bimetal thiolate cluster partially provides a rationale for the difference in exchange rates among the three subsets of cysteine amide protons. The backbone amide protons for four of the six conserved cysteines are thought to hydrogen bond to sulfur atoms of other cysteines in the bimetal thiolate cluster (Kraulis et al., 1992). These involve C11(NH)–C21(S), C14(NH)–C11(S), C28(NH)–C38(S), and C31(NH)–C28(S), the latter pair related to the first by dyad symmetry.[2] The amide protons of C14 and C31, the slowest-exchanging group, are each involved in two hydrogen bonds, one to the thiolate of a cysteine in the cluster, and the other to the carbonyl of the $(i - 3)$th residue to the first amino acid of the α-helix. The amide protons of C11 and C28, the second group, are each involved in a single hydrogen bond to the thiolate of a cysteine in the cluster. Possible H-bond acceptors for the amide protons of the fastest exchanging group, C21 and C38, are not obvious from the structural model. Because hydrogen exchange is slow at neutral pH for some cysteines, these amide protons must be involved in intramolecular hydrogen bonds. The slowest exchanging amides of C14 and C31 appear to have two H-bond acceptors (a carbonyl and the sulfur), whereas the faster exchanging amides of C11 and C28 have only one reasonable H-bond acceptor (the thiolate). Alternatively, other aspects of the GAL4 structure such as hydrophobic packing and solvent accessibility could contribute to the observed grouping of these slowly exchanging amides.

In comparison to the extensively studied amide proton exchange rates of the basic pancreatic trypsin inhibitor (Wagner, 1983), it may at first appear that the amides of the DNA-binding domain of GAL4 have unusually fast exchange rates. However, measurements here have to be performed at neutral pH because metal ligation requires cysteines to be ionized. The exchange rates are comparable to those measured in other metal-binding proteins, such as metallothionein (Messerle et al., 1992), or other large proteins studied at neutral pH. The amide exchange rates of basis pancreatic trypsin inhibitor (BPTI) are several orders of magnitude slower than those of GAL4 even

[2] No H-bond-mediated scalar coupling to Cd larger than about 8 Hz was discerned from our analysis of amide proton multiplet patterns, as was recently observed for a Cd-substituted rubredoxin (Blake et al., 1992).

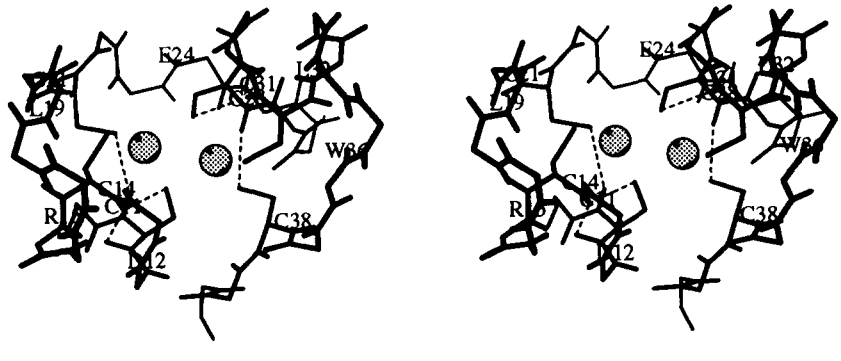

Fig. 3. Schematic stereo drawing of the DNA-recognition module of GAL4. The numbers indicate the six cysteine residues and additional residues that have slowly exchanging amide protons. Putative hydrogen bonds to slowly exchanging cysteinyl amide protons are indicated by dashed lines.

at neutral pH, but this reflects more the unusual stability properties of the trypsin inhibitor rather than remarkably rapid amide exchange for the DNA-binding domain of GAL4.

From the crystal structure of the 1–65 residue fragment of GAL4 with specific target DNA (Marmorstein et al., 1992), one might expect that some of the amide protons in the observed α-helices, specifically residues 11–18, 28–35, and 50–64, would exchange slowly with bulk solvent. In this study, all of the slowly exchanging amides for the uncomplexed Zn and Cd forms of GAL4 are consistent with being in the recognition module. No slowly exchanging amides are observed for the C-terminal helix (residues 50–64). This region does not seem to form a stable α-helix in the absence of bound DNA. It should be mentioned that at neutral pH intrinsic NH exchange is rather fast – absence of slowly exchanging amides under these conditions does not necessarily dictate the absence of a helix. However, additional evidence against this helix exists from our earlier NMR structural studies of this fragment of GAL4 (Baleja et al., 1992), where no strong NOEs were observed for amino acid residues 52–64, indicating that this region of the protein is at least partially disordered in solution and in the absence of DNA. On the contrary, several additional slowly exchanging amides are observed for the GAL4–DNA complex (Fig. 2C), which are as yet unassigned. Consistent with a GAL4 dimer bound to DNA, these may arise from the α-helices of a coiled-coil dimerization element, which is present through the C-terminal part of the protein (Marmorstein et al., 1992).

The exchange rates for the amides of the Zn-, Cd-, and DNA-bound forms of GAL4 were measured under slightly different conditions of pH and temperature. The small differences in pH could have been corrected by adjustment with small amounts of acid or base prior to NMR data collection. However, the valuable information inherent in the rates for the fastest members of the slow-exchange group would be lost in the increased lag time between sample resolubilization and spectral acquisition. Fortunately, small differences in pH may be easily corrected for in the base-catalyzed regime of amide exchange by assuming a 10-fold increase in rate for every unit increase in pH (Englander et al., 1972). The actual rate dependence on pH may be somewhat less, because the pH dependence of the equilibrium constant of the exchange reaction often decreases slightly from pH 4 to pH 8 in cases where pH-dependent protein stabilization occurs (Richarz et al., 1979). The magnitude of this effect in this work is taken to be negligible.

^{1}H-^{15}N HSQC spectra for the Zn-GAL4–DNA complex have the highest resolution and sensitivity at a temperature of 35 °C. Therefore amide exchange studies for the protein–DNA complex and for Zn-GAL4 were undertaken at this temperature. At 35 °C, however, the exchange for many amide protons of Cd-GAL4 is too rapid to be measured accurately by the exchange-out method. Amide exchange was instead measured at 25 °C for Cd-GAL4, and temperature effects corrected for by assuming a threefold decrease in rate for the 10 °C decrease in temperature, as previously observed for amide exchange in peptides (Englander et al., 1972). The threefold decrease may be a slight undercompensation because proteins are often structurally sensitive to temperature. In any case, any inaccuracies in these assumptions do not alter the conclusions drawn from the comparisons in amide exchange rates as discussed below.

Cd- vs. Zn-GAL4

Table 3 compares the six amide protons whose rate constants were determined in both metallo-forms of GAL4, using data corrected for differences in experimental conditions. The amide protons exchange 3–21 times faster in Cd-GAL4 than in Zn-GAL4, indicating lower stability over much of the recognition module. Specifically, four of the six amide protons participate in hydrogen bonding within the helices. Their exchange rates are roughly fivefold higher in Cd-GAL4 than in Zn-GAL4. Assuming the exchange rate difference can be attributed solely to the energy difference of the hydrogen bond in which the amide proton is involved, a fivefold difference in exchange rate corresponds to an energy difference of approximately 1 kcal/mol. The weakening of hydrogen bonds translates into a slight decrease in the stability of the α-helices and of the entire recognition module. The reduction in stability by cadmium substitution into a zinc metalloprotein has been observed for the bacteriophage T4 gene 32 protein (g32P) (Keating et al., 1988). The non-native Cd-g32P has a slightly lower melting temperature and free energy of unfolding than Zn-g32P. Cadmium substitution is likely to cause similar reduction of stability for GAL4, despite the higher intrinsic affinity of Cd(II) for the thiolate ligands than Zn(II) (Keating et al., 1988). This may be a consequence of the fact that the optimal Cd-S bond

Table 3. *Comparison of amide proton exchange rates in Cd- and Zn-GAL4*

	Cd-GAL4[a] (35 °C, pH 7.02)	Zn-GAL4[b] (35 °C, pH 7.02)	$\frac{k_{obs}(\text{Cd-GAL4})}{k_{obs}(\text{Zn-GAL4})}$
C11	529 ± 4.4	25.6 ± 1.2	20.8 ± 1.2
C14	39.6 ± 0.1	13.2 ± 0.7	3.0 ± 0.2
R15	242 ± 1.5	46.2 ± 3.1	5.3 ± 0.4
C28	197 ± 1.1	24.2 ± 2.1	8.2 ± 0.8
C31	88.0 ± 0.2	17.1 ± 0.8	5.2 ± 0.3
L32	109 ± 0.3	22.6 ± 1.7	4.9 ± 0.4

[a] Exchange rates (in 10^{-3} min^{-1}) from the original data shown in Table 2 are corrected to 35 °C and pH 7.02 by multiplication of a constant, 6.716 (see text).

[b] Exchange rate constants (in 10^{-3} min^{-1}) taken from Table 2.

lengths are slightly different from Zn-S bonds, thus forcing some strain in the structure evolutionarily optimized for Zn binding. Alternatively, because the radius of Cd is larger than that of Zn (by 25%), the faster exchange observed for the Cd-bound protein could be due to enhanced breathing dynamics induced by the larger metal (see below).

Free vs. DNA-bound GAL4

Comparison of the 3D structures of free GAL4 determined by NMR methods (Baleja et al., 1992; Kraulis et al., 1992) and bound GAL4 determined by X-ray crystallography (Marmorstein et al., 1992) shows the structure of the recognition module to be almost identical in both (Baleja et al., 1992). This study, however, reveals significant changes in the dynamics of the module upon DNA binding. Table 4 compares the exchange rates of six amide protons in free and DNA-bound Zn-GAL4 and shows the rates to be reduced by 5–24-fold in the latter. Interestingly, solvent exchange for the backbone NH of C28 and C31, two conserved cysteines located away from the DNA-binding surface, is retarded by DNA binding to the same degree as is that for C11 and C14, their symmetry-related counterparts on the DNA-contacting side of the recognition module (Fig. 3). This observation implies that the retardation of exchange is primarily a consequence of the restrained internal motions of the module, a global effect of DNA binding, rather than the protection from solvent offered by bound DNA, which would result in local retardation of amide proton exchange. In the "breathing model" of amide proton exchange (Hvidt & Nielsen, 1966; Englander et al., 1972), a protein exists in an equilibrium between a closed state (C), in which the intramolecular hydrogen bonds involving an amide proton are intact, and a transient open state (O), in which the bonds are broken and the amide proton can exchange with a solvent deuteron. The results here suggest that the presence of specifically bound DNA alters the equilibrium constant between the open and closed states in the following pathway (Hvidt & Nielsen, 1966) and retards amide proton exchange as a result (Eq. 1).

$$\mathrm{C(H)} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{O(H)} \xrightarrow{k_3} \mathrm{O(D)} \underset{k_1}{\overset{k_2}{\rightleftharpoons}} \mathrm{C(D)}. \qquad (1)$$

Application of Equation 1 to amide exchange kinetics assumes in the Zn-GAL4–DNA complex that the protein under consideration is a dimer specifically bound to one DNA molecule. The breathing model represented by Equation 1, however, does not take into account additional dynamic processes that include, at least, equilibria between free and bound forms and between monomeric and dimeric states of the protein. The kinetics therefore may be considerably more complex.

Despite the incomplete assignment of the DNA-recognition module backbone amide proton resonances in the HSQC spectra of Zn-GAL4 and the Zn-GAL4–DNA complex, meaningful information is still obtained in this study, because comparison of exchange rates is based on slowly exchanging amide protons assigned in both Cd- and Zn-GAL4, and in both free and DNA-bound GAL4. Assignment of the four unidentified slow-exchanging amide protons in the Zn-GAL4–DNA complex will enable further characterization of the effect of DNA binding on the dynamics of GAL4. Future efforts will be directed to the assignment of these and other unassigned amide protons in the ^1H-^{15}N HSQC spectra.

Table 4. *Comparison of amide proton exchange rates in Zn-GAL4 and Zn-GAL4–DNA*

	Zn-GAL4[a] (35 °C, pH 7.02)	Zn-GAL4-DNA[b] (35 °C, pH 7.02)	k_{obs}(Zn-GAL4) / k_{obs}(Zn-GAL4-DNA)
C11	25.6 ± 1.2	1.81 ± 0.4	15.0 ± 4.0
C14	13.2 ± 0.7	0.70 ± 0.1	19.4 ± 3.8
R15	46.2 ± 3.1	2.21 ± 0.3	21.5 ± 4.3
C28	24.2 ± 2.1	1.93 ± 0.4	13.3 ± 3.9
C31	17.1 ± 0.8	0.73 ± 0.1	24.0 ± 4.4
L32	22.6 ± 1.7	4.56 ± 0.3	5.0 ± 0.7

[a] Exchange rate constants (in 10^{-3} min^{-1}) taken from Table 2.
[b] Exchange rates (in 10^{-3} min^{-1}) from Table 2 are corrected to pH 7.02 by multiplication of a constant, 1.622 (see text).

Materials and methods

Protein and DNA

The DNA-binding domain of GAL4 was overexpressed in *Escherichia coli* strain XA-90 from a plasmid vector carrying the gene conferring ampicillin resistance and the coding region of the N-terminal 65 residues of GAL4 under the control of the *tac* promoter. This construct contains an additional phenylalanine at the 66th position as a cloning artifact (originally described by Carey et al., 1989). Clones used in large-scale preparations were first tested for induction. Plasmid-bearing cells were streaked out on LB agar plates, and resultant colonies were spot-tested on plates with and without 4 mM isopropylthiogalactoside (IPTG). Cells producing GAL4(65) are expected to devote most of their protein synthesis machinery when induced with IPTG and to divide at a reduced rate. Colonies that grew weaker on IPTG-containing plates were therefore selected. Inducible GAL4 protein production was confirmed by growing these clones in small volumes of rich medium (16 g/L tryptone, 10 g/L yeast extract, 10 g/L NaCl, pH adjusted to 7.0 with NaOH, and 100 mg/L ampicillin), inducing with IPTG, and examining the cell lysate by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE).

Large-scale preparations typically used 2 L of culture. Cells were grown at 37 °C in modified M9 medium (1.2%

Na_2HPO_4, 0.3% KH_2PO_4, 0.05% NaCl, 0.1% NH_4Cl, 0.5% glucose, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.5 mM $FeCl_3$, 5 mg/mL thiamine monophosphate chloride) to mid-log phase and induced with 0.3 mM IPTG, shortly after the addition of 0.2 mM $CdCl_2$ or 0.5 mM $ZnSO_4$. The culture was allowed to grow for 4–5 h after induction, at which point the cells were then harvested by centrifugation to typically yield 14–16 g of cells (wet pellet weight). Cells were washed with 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7.5, 0.2 M NaCl, centrifuged, and resuspended in lysis buffer (20 mM HEPES, pH 7.5, 0.2 M NaCl, 10% sucrose, and 10 mM $CdCl_2$ or 100 mM $ZnSO_4$). The suspension was frozen in liquid nitrogen for storage at −20 °C and subsequently thawed prior to cell lysis. Subsequent steps were carried out at 4 °C. Protease inhibitors (0.05 mg/mL phenylmethylsulfonylfluoride, 0.5 mg/mL pepstatin A, 1 mg/mL leupeptin) and 14 mM β-mercaptoethanol were added to the thawed suspension. Lysozyme was not used for cell lysis because it has similar chromatographic properties as GAL4(65) and presented a difficulty in purification when used in trial preparations. Cells were lysed instead by sonication with 30-s pulses until the OD_{600} reached 10–20% of the original reading. The suspension was centrifuged at 14,000 rpm for 30 min, and 5% polyethylenimine was added dropwise to the supernatant to a final concentration of 0.5% over a period of 20 min to precipitate DNA. The suspension was stirred for an additional 20 min and centrifuged at 14,000 rpm for 30 min. Ammonium sulfate was added to the supernatant to a concentration of 40%, and the suspension was stirred for 20 min. After centrifugation at 14,000 rpm for 20 min, the pellet was resuspended in our standard buffer (SB: 20 mM sodium phosphate buffer, pH 7.0, 0.2 M NaCl, 10 mM $CdCl_2$ or 100 mM $ZnSO_4$, 10% glycerol, 14 mM β-mercaptoethanol, and protease inhibitors). The protein solution was dialyzed against SB for 4 h with two changes of buffer, diluted 2× with SB, and loaded onto an S Sepharose Fast Flow (Pharmacia) column preequilibrated in SB. GAL4(65) was eluted with a linear 0.2–0.6 M NaCl gradient. GAL4 fractions were identified by SDS-PAGE, pooled, diluted 2–3× with SB (without NaCl), and loaded onto a heparin agarose (Sigma Type II) column preequilibrated in SB. A linear gradient of 0.2–0.55 M NaCl was applied, and GAL4(65) eluted in a sharp peak at a NaCl concentration of 0.4 M. The fractions were identified by SDS-PAGE, pooled, and concentrated to a volume of 1–2 mL using the Amicon Micro-Ultrafiltration System (Model 8MC) with a YM3 DIAFLO membrane. Gel filtration (Pharmacia Sephadex G-50) was employed to purify the protein sample further and to exchange with 20 mM sodium phosphate buffer, pH 6.0, 0.1 M NaCl, 10 mM $CdCl_2$ or 100 mM $ZnSO_4$, 0.05% NaN_3. GAL4(65) fractions were identified by SDS-PAGE, pooled, and concentrated by ultrafiltration to millimolar concentration. The protein was greater than 97% pure as judged from SDS-PAGE, and the yield of several preparations ranged from 5 to 20 mg/L as determined by a Coomassie blue dye-binding assay (Bio-Rad Laboratories).

The DNA used was a 19-bp oligonucleotide with 5′ single-base overhangs, originally designed for crystallographic studies of the protein–DNA complex. It contains the 17-bp consensus GAL4 binding site (Marmorstein et al., 1992) and has the sequence:

```
5′  A C C G G A G G A C A G T C C T C C G G
3′    G G C C T C C T G T C A G G A G G C C T.
```

NMR spectroscopy

Experiments were performed with 0.5–1.5 mM of protein in 20 mM sodium phosphate buffer at neutral pH, 0.1 M NaCl, and either 10 μM $CdCl_2$ or 100 μM $ZnSO_4$. Protein samples in H_2O buffer were adjusted to pH 6.3 ± 0.2 and lyophilized. Exchange-out was initiated by dissolving the sample in D_2O. Generally 10–15 min elapsed before the acquisition of the first NMR spectrum to allow for sample preparation, temperature equilibration, stabilization of the lock signal, and quick shimming. A series of 2D ^1H–^{15}N correlation spectra were then collected for up to 25 h with the HSQC pulse sequence of Bodenhausen and Ruben (1980) on Bruker AMX-500 or AMX-600 spectrometers. The water resonance was suppressed by a proton trim pulse applied prior to magnetization transfer to ^{15}N (Messerle et al., 1989). Each spectrum was acquired with 96 increments of 4 scans of 2K data points with a total acquisition time of 9 min for Cd-GAL4, and with 64 increments of 12 scans of 2K data points with a total acquisition time of 15 min for Zn-GAL4 and the Zn-GAL4–DNA complex. pH values were measured after exchange was completed. Amide proton exchange in the presence of DNA was measured with a sample containing 0.58 mM protein (concentration as a monomer) and 0.37 mM DNA to yield a protein:DNA ratio so that the protein should be saturated with DNA.

Calculation of amide proton exchange rate constants

Peak volumes in HSQC spectra were integrated with the 2D integration routine in the NMR data processing software FELIX (Hare Research, Woodinville, Washington) using a manually picked rectangular footprint. The peak intensity in the spectrum gives the degree of protonation at the start of the experiment, i.e., for the first t_1 increment. The decay due to exchange during the recording of the 2D spectrum will broaden the signal along t_1, but not affect the intensity. Therefore, one should use the time at the point at which one-half of the scans during the first t_1 increment have accumulated for most accurate timing of the exchange-out procedure. However, for rate con-

stants determined with four or more data points, and with our signal-to-noise ratios, measurement times of about 12 min, and amide life-times of at least 10 min, in practice the use of the average time during acquisition of an entire 2D spectrum gave the same result. Peak intensities for each amide proton were fitted to a single exponential decay using a nonlinear least-squares fitting routine in the data manipulation package PLOT (New Unit, Inc.). For amide protons with four or more data points, an uncertainty in the rate constant is calculated. This uncertainty is the pessimist's error, defined as the amount the rate constant must change in order to alter χ^2 by 1, assuming other parameters in the fit take on the worst possible values at their limits of certainty. The calculated value thus reflects the quality of the fit.

For amide protons seen only in the first spectrum taken after initiation of exchange-out, a few assumptions were made in arriving at an estimate for the exchange rates. The signals for resonances of several amide protons with accurately determined rate constants are observed usually to have decayed to noise level when their intensity $I(t)$ reaches $0.08 \times I_0$, where I_0 is the extrapolated full intensity when exchange-out was initiated. Based on this approximation, upper and lower limits of the rate constant were estimated by taking the time of the beginning and the end of the first experiment, respectively, to be t in $I(t) = 0.08 \times I_0$. The upper limit so obtained also represents the lower limit of exchange rates for those amide protons not observed in the first spectrum, the majority in this study.

Acknowledgments

We thank Dr. Ronen Marmorstein and Professor Stephen Harrison for the gift of DNA oligonucleotide, the GAL4 overproducing plasmid and bacterial strain, and advice on protein purification. This work was supported by NIH grant GM47467 (G.W.), the Medical Research Council of Canada (J.D.B.), and the Alberta Heritage Foundation for Medical Research (J.D.B.).

References

Baleja, J.D., Marmorstein, R., Harrison, S.C., & Wagner, G. (1992). Solution structure of the DNA-binding domain of Cd_2-GAL4 from *S. cerevisiae*. *Nature 356*, 450–453.

Blake, P.R., Park, J.B., Adams, M.W.W., & Summers, M.F. (1992). Novel observation of NH···S(Cys) hydrogen-bond-mediated scalar coupling in ^{113}Cd-substituted rubredoxin from *Pyrococcus furiosus*. *J. Am. Chem. Soc. 114*, 4931–4933.

Bodenhausen, G. & Ruben, D.J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett. 69*, 185–189.

Bram, R. & Kornberg, R. (1985). Specific protein binding to far upstream activating sequences in polymerase II promoters. *Proc. Natl. Acad. Sci. USA 82*, 43–47.

Carey, M., Kakidani, H., Leatherwood, J., Mostashari, F., & Ptashne, M. (1989). An amino-terminal fragment of GAL4 binds DNA as a dimer. *J. Mol. Biol. 209*, 423–432.

Englander, S.W., Downer, N.W., & Teitelbaum, H. (1972). Hydrogen exchange. *Annu. Rev. Biochem. 41*, 903–924.

Englander, S.W. & Kallenbach, N.R. (1984). Structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys. 16*, 521–655.

Gadhavi, P.L., Davis, A.L., Povey, J.F., Keeler, J., & Laue, E.D. (1991). Polypeptide–metal cluster connectivities in Cd(II) GAL4. *FEBS Lett. 281*, 223–226.

Gadhavi, P.L., Raine, A.R.C., Alefounder, P.R., & Laue, E.D. (1990). Complete assignment of the ^{1}H NMR spectrum and secondary structure of the DNA binding domain of GAL4. *FEBS Lett. 276*, 49–53.

Gardner, K.H., Pan, T., Narula, S., Rivera, E., & Coleman, J.E. (1991). Structure of the binuclear metal-binding site in the GAL4 transcription factor. *Biochemistry 30*, 11292–11302.

Giniger, E., Varnum, S.M., & Ptashne, M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell 40*, 767–774.

Harris, R.K. (1983). *Nuclear Magnetic Resonance Spectroscopy*. Pitman, London.

Harrison, S.C. (1991). A structural taxonomy of DNA-binding domains. *Nature 353*, 715–719.

Hvidt, A. & Nielsen, S.O. (1966). Hydrogen exchange in proteins. *Adv. Protein Chem. 21*, 287–386.

Johnston, M. (1987). A model fungal gene regulatory mechanism: The *GAL* genes of *Saccharomyces cerevisiae*. *Microbiol. Rev. 51*, 458–476.

Keating, K.M., Ghosaini, L.R., Giedroc, D.P., Williams, K.R., Coleman, J.E., & Sturtevant, J.M. (1988). Thermal denaturation of T4 gene 32 protein: Effects of zinc removal and substitution. *Biochemistry 27*, 5240–5245.

Kraulis, P.J., Raine, A.R.C., Gadhavi, P.L., & Laue, E.D. (1992). Structure of the DNA-binding domain of zinc GAL4. *Nature 356*, 448–450.

Marmorstein, R., Carey, M., Ptashne, M., & Harrison, S.C. (1992). DNA recognition by GAL4: Structure of a protein–DNA complex. *Nature 356*, 408–414.

Messerle, B.A., Schäffer, A., Váošak, M., Kägi, J.H.R., & Wüthrich, K. (1992). Comparison of the solution conformations of human [Zn_7]-metallothionein-2 and [Cd_7]-metallothionein-2 using nuclear magnetic resonance spectroscopy. *J. Mol. Biol. 225*, 433–443.

Messerle, B.A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989). Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with H_2O solutions. *J. Magn. Reson. 85*, 608–613.

Pan, T. & Coleman, J.E. (1989). Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor. *Proc. Natl. Acad. Sci. USA 86*, 3145–3149.

Pan, T. & Coleman, J.E. (1990a). GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster. *Proc. Natl. Acad. Sci. USA 87*, 2077–2081.

Pan, T. & Coleman, J.E. (1990b). The DNA binding domain of GAL4 forms a binuclear metal ion complex. *Biochemistry 29*, 3023–3029.

Pan, T. & Coleman, J.E. (1991). Sequential assignments of the ^{1}H NMR resonances of $Zn(II)_2$ and $^{113}Cd(II)_2$ derivatives of the DNA-binding domain of the GAL4 transcription factor reveal a novel structural motif for specific DNA recognition. *Biochemistry 30*, 4212–4222.

Paterson, Y., Englander, S.W., & Roder, H. (1990). An antibody binding site on cytochrome c defined by hydrogen exchange and two-dimensional NMR. *Science 249*, 755–759.

Povey, J.F., Diakun, G.P., Garner, C.D., Wilson, S.P., & Laue, E.D. (1990). Metal ion coordination in the DNA binding domain of the yeast transcriptional activator GAL4. *FEBS Lett. 266*, 142–146.

Richarz, R., Sehr, P., Wagner, G., & Wüthrich, K. (1979). Kinetics of the exchange of individual amide protons in the basic pancreatic trypsin inhibitor. *J. Mol. Biol. 130*, 19–30.

Takahashi, H., Odaka, A., Kawaminami, S., Matsunaga, C., Kato, K., Shimada, I., & Arata, Y. (1991). Multinuclear NMR study of the structure of the Fv fragment of anti-dansyl mouse IgG2a antibody. *Biochemistry 30*, 6611–6619.

Wagner, G. (1983). Characterization of the distribution of internal motions in the basic pancreatic trypsin inhibitor using a large number of internal NMR probes. *Q. Rev. Biophys. 16*, 1–57.

Wagner, G. & Wüthrich, K. (1982). Amide proton exchange and surface conformation of the basic pancreatic trypsin inhibitor (BPTI) in solution: Studies with two-dimensional nuclear magnetic resonance. *J. Mol. Biol. 160*, 343–361.