Volume 266, number 1,2, 142-146 FEBS 08525 June 1990

Metal ion co-ordination in the DNA binding domain of the yeast transcriptional activator GAL4

Jane F. Povey[1], Gregory P. Diakun[2], C. David Garner[3], Stephen P. Wilson[3] and Ernest D. Laue[1]

[1]*Department of Biochemistry, Tennis Court Road, Cambridge, CB2 1QW,* [2]*MRC/SERC Biology Support Laboratory, Daresbury Laboratory, Daresbury, Warrington, WA4 4AD* and [3]*Department of Chemistry, University of Manchester, M13 9PL, UK*

Received 19 March 1990

The structure of the DNA binding domain of the yeast transcriptional activator GAL4 was investigated by extended X-ray fine structure (e.x.a.f.s.). Two samples of GAL4 were studied, one containing cadmium as a structural probe (Cd(II)GAL4) and the other containing the 'native' zinc (Zn(II)-GAL4). The results suggest that the structure of the DNA binding domain of GAL4 contains a two metal ion cluster distinguishing it from the 'zinc finger' proteins typified by the *Xenopus laevis* transcription factor TFIIIA.

GAL4; Structure; DNA binding; Extended X-ray fine structure; Zinc finger; *Saccharomyces cerevisiae*

1. INTRODUCTION

GAL4 activates transcription of the GAL1 and GAL10 genes in the yeast *Saccharomyces cerevisiae* by binding to DNA sequences within the upstream activating sequence for galactose (UASg) [1]. The amino terminal DNA binding domain of the protein (residues 1-147) contains a sequence (Cys-X_2-Cys-X_6-Cys-X_6-Cys-X_2-Cys-X_6-Cys) which is highly conserved amongst several other fungal gene regulatory proteins [2,3 (Fig. 1)]. It has been proposed [4,5] that these cysteines may co-ordinate zinc, creating a structure analogous to the 'zinc fingers' of TFIIIA [6]. More recently a ^{113}Cd NMR study suggested that the protein contained two binding sites in which each cadmium was co-ordinated by 3 or 4 sulphur atoms [7]. In this paper we report the results of extended X-ray fine structure (e.x.a.f.s.) studies on both Cd(II) and Zn(II) GAL4 purified from an over-expressing strain of *E.coli*. We have studied the Cd(II) form of the protein to avoid problems of background contamination and loss of metal during the purification; cadmium binds much more tightly and the background levels are considerably lower.

2. MATERIALS AND METHODS

2.1. *Protein purification*

The protein were purified as described by Lin et al. [8], with the following differences; (i) zinc ions were omitted from all buffers to prevent background contamination, (ii) a DNA-cellulose column was used instead of the DEAE-cellulose column, and (iii) in order to produce the Cd(II) form of the protein, cadmium chloride was added to the growing cells to 0.3 mM at the same time as the isopropyl β-D-thiogalactopyranoside (IPTG). Acid-washed glassware and EDTA washed plasticware were used in the purification of samples for atomic absorption and e.x.a.f.s. spectroscopy [6]. Protein concentrations were measured by amino acid analysis.

2.2. *Gel retardation assays*

Gel retardation assays were carried out on a 4% acrylamide gel as described by Fried & Crothers [9]. The protein was incubated with an *Eco*RI-*Hin*dIII fragment of the pUC18 polylinker into which the 17 base pair consensus GAL4 DNA binding site had been inserted in the *Bam*HI site. All incubations were carried out in a buffer containing 20 mM Hepes pH 7, 50 mM potassium chloride, 5 mM magnesium chloride, 10 μM zinc acetate/cadmium chloride, 0.2 mg/ml acetylated bovine serum albumin, 50% glycerol and poly(dI-dC).

2.3. *Atomic absorption spectroscopy*

Zinc and cadmium concentrations were determined using a flame method on a Perkin-Elmer Model 380 atomic absorption spectrometer.

2.4. *Cadmium substitution experiments*

Molar equivalents of cadmium chloride were added to samples of the Zn(II) protein before recording UV spectra on a Perkin Elmer 555 UV-VIS spectrophotometer every 15 min until there were no further changes (approximately 2 h). The final spectra obtained were compared.

2.5. *E.x.a.f.s. studies*

E.x.a.f.s. spectra were recorded at the Zn and Cd K-edge as fluorescence excitation spectra. The data were recorded on the Wiggler beamline using station 9.2 of the Synchrotron Radiation Source at Daresbury Laboratory. The electron storage ring operated at 2 GeV with an average current of 150 mA. A silicon (220) double crystal monochromator was employed to avoid harmonic contamination. A step size of approximately 0.04k was used to collect the data. An array of 12 germanium solid state detectors with individual electronics [10] monitored the X-ray fluorescence excitation spectra, with simultaneous measurement of a zinc or cadmium foil. The samples

Correspondence address: E.D. Lane, Department of Biochemistry, Tennis Court Road, Cambridge, CB2 1QW, UK

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

were dialysed against 5 mM Hepes pH 7, assayed for DNA binding activity and lyophilised before being packed into teflon cells (sample window dimension of 12 mm width by 4 mm height by 5 mm depth). A total of 8 1-h scans were collected for each sample of GAL4 at approximately 80K using a liquid nitrogen cryostat. Prior to averaging the data, each spectrum from the individual germanium detector was inspected and weighted in proportion to its edge height. Any spectrum showing irregularities was omitted from the averaging process. A check for radiation damage around the immediate area of the metal ion was carried out by inspecting each scan to see if changes in the edge and near edge structure occurred; non were visible.

Data analysis was accomplished via the single scattering curved-wave method of e.x.a.f.s. calculation, and phase-shifts were derived from 'ab initio' calculations as described previously [11-13].

2.6. *Reactions with para-hydroxy mercuribenzoate*

Reactions were carried out as described by Riordan and Vallee [14] in 20 mM Hepes/200 mM sodium chloride pH 7, monitoring the absorbance at 250 nm.

3. RESULTS

When the DNA binding domain of GAL4 (GAL4 (1-147)) was purified from *E.coli*, atomic absorption spectroscopy showed that it contained 2.43 (±0.39) mol of zinc per mol of protein. Monitoring the absorbance at 245 nm (characteristic of Cd-S charge transfer) in competition assays showed that the protein has a much higher affinity for cadmium than it does for zinc. Titration of the protein with cadmium showed no further spectral change after addition of two equivalents of cadmium. A GAL4 fragment, purified from cells grown in the presence of cadmium, was also found to contain an average of 2.07 (±0.28) mol of cadmium per mol of protein. The Cd(II) form bound to the GAL4 DNA binding site with a similar affinity to that of the native zinc form (see Fig. 1).

E.x.a.f.s. spectroscopy was used to investigate the immediate environment of cadmium and zinc in GAL4. This technique provides information on the number of ligands, their type and distance from the atom being probed. The X-ray absorption near edge structure (x.a.n.e.s.) data are a sub-set of the e.x.a.f.s. data and are indicative of ligand geometry. The x.a.n.e.s. data of both Cd(II) and Zn(II) GAL4 are very similar to those of the following species: Cd(II) and Zn(II) metallothionein [15,15a], the DNA binding domain of the glucocorticoid receptor [16] and $[NEt_4]_2[Zn(SPh)_4]$, a model compound in which the zinc is tetrahedrally co-ordinated by 4 sulphur atoms (data not shown). This suggests a tetrahedral co-ordination about each metal ion in GAL4.

The e.x.a.f.s. data, and their corresponding Fourier transform, of Zn(II) GAL4 (see Fig. 2) can best be simulated with a model consisting of three sulphur ligands at 0.230 nm (2.30 Å), a single light atom (either oxygen or nitrogen) at 0.196 nm (1.96 Å) and a sulphur atom at 0.334 nm (3.34 Å) (see Table I). Assuming that GAL4 (1-147) has two metal binding sites this result could be interpreted in one of two ways. Firstly, each metal ion might be co-ordinated by three sulphur atoms and a single oxygen (or nitrogen) ligand. Alternatively, one metal ion might be co-ordinated by four sulphur ligands and the other by two sulphur ligands and two oxygen (or nitrogen) ligands. Using e.x.a.f.s. it is impossible to distinguish between these possibilities as the data result from an average of the two zinc sites. For the e.x.a.f.s. data, and their corresponding Fourier transform, of Cd(II) GAL4 the best simulation is achieved with 4 sulphur ligands at a bond distance of 0.247 nm (2.47 Å) (see Table I).

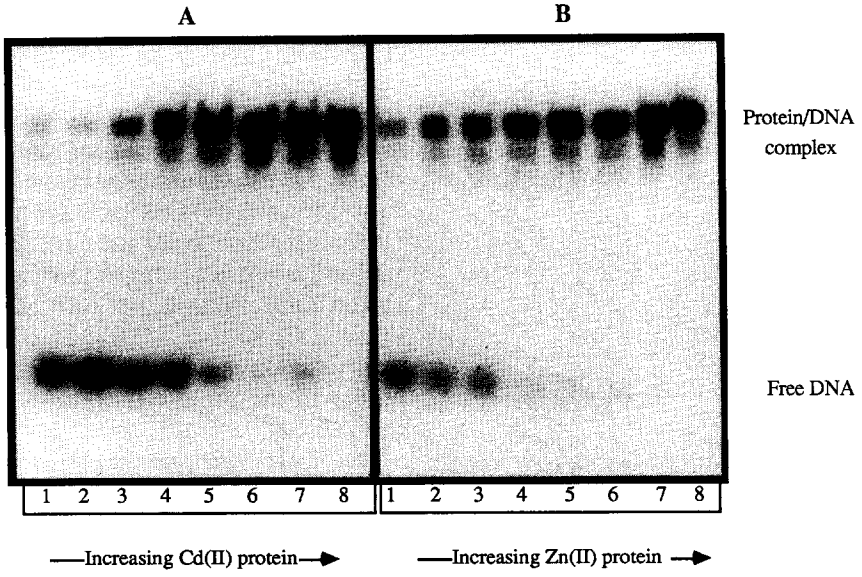


Fig. 1. A gel retardation assay showing the relative DNA binding affinities of the Zn(II) and Cd(II) forms of the GAL4 (1-147) fragment.

Fig. 2. Cd and Zn K-edge e.x.a.f.s. data, and their associated Fourier transforms. The continuous lines represent the experimental spectra, the diamonds the simulations. (A) Zn(II) GAL4 fitted with a model consisting of 3 sulphur ligands and one oxygen ligand with a further sulphur atom at approximately 3.3 Å, and (B) Cd(II) GAL4 fitted with a model consisting of four sulphur ligands. The e.x.a.f.s. data are presented as $\chi(k)k^3$ plotted against k where k is in $Å^{-1}$. $\chi(k)$ is the function describing the oscillations of the e.x.a.f.s. spectrum; it is multiplied by k^3 to enhance features at high k. k is the momentum of the photoelectron given by $k = [(E-E_o)2m/\hbar^2]^{1/2}$ where m is its mass, E its energy and E_o the energy of the absorption edge.

It is surprising, given the similar DNA binding affinities (see Fig. 1), that the co-ordination spheres of the two different metals are not identical. We were able to check the phase shifts, and the calculation procedure, using three different model compounds in which the zinc ion was co-ordinated by either 4 sulphur, 3 sulphur and one oxygen or two sulphur and two oxygen atoms. The results of the e.x.a.f.s. simulations compared very well with the crystal structures giving us confidence in the refinement procedure. We also found that para-hydroxy mercuribenzoate [14] reacts with 2 sulphydryl groups per protein molecule in Zn(II) GAL4, but none in Cd(II) GAL4. These results further emphasise the differences in metal ion co-ordination between the Zn(II) and the Cd(II) forms of the protein.

4. DISCUSSION

Using a combination of atomic absorption spectroscopy and amino acid analysis, we have been able to show that the DNA binding domain of the yeast transcriptional activator GAL4 contains either two cadmium or two zinc ions per protein molecule, depending on the growth conditions. Results obtained from a titration of the Zn(II) protein with cadmium support this conclusion. As the Cd(II) form of the protein bound to DNA with a similar affinity to the native Zn(II) form we were able to use cadmium as a structural probe in the e.x.a.f.s. studies.

The results of the e.x.a.f.s. studies suggest that the metal ions in GAL4 are either co-ordinated by 4 sulphur atoms in the case of Cd(II) GAL4, or by a mixture of sulphur and oxygen atoms in the case of Zn(II) GAL4.

Table I

Parameters used to simulate the e.x.a.f.s. data

Protein	Ligands	σ^2 (nm^2)	R(nm)
Zn(II) GAL4	3S	0.000065	0.230
	1O	0.000065	0.195
	1S	0.00004	0.334
Cd(II) GAL4	4S	0.000075	0.247

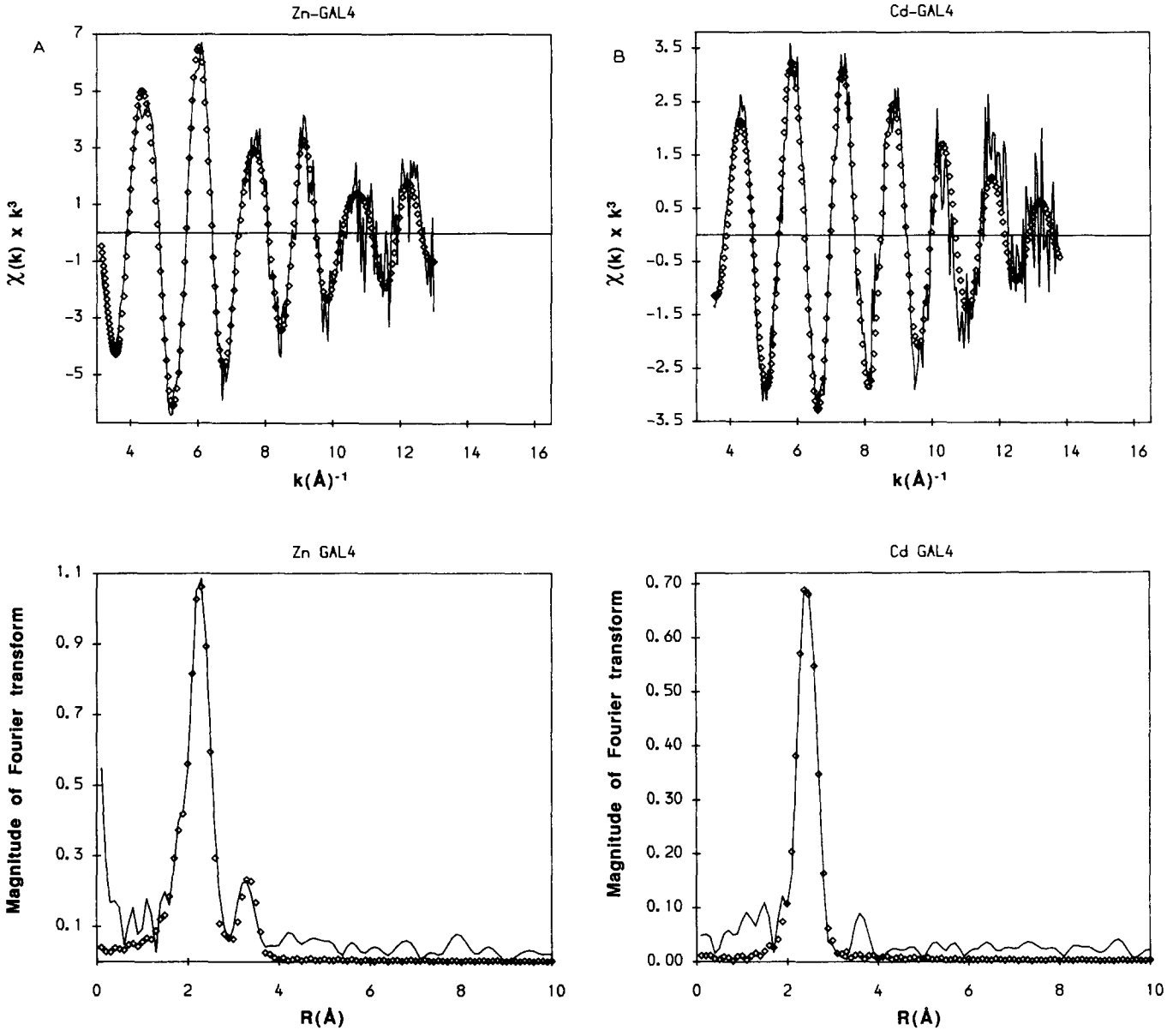
σ^2 is a Debye-Waller like factor and is equivalent to ΔR rms the mean square deviation of an atom from its mean position. R is the distance of atoms from the metal ion as determined from the e.x.a.f.s. data. The errors in the fitting procedure are approximately ± 0.002 nm.

The sulphur atoms co-ordinated to the zinc or cadmium could in principle be provided by either cysteine or methionine residues. As only the cysteine residues are strongly conserved in both position and number in the fungal gene regulatory proteins to which GAL4 is closely homologous [3], it is probably that they are the co-ordinating ligands. In GAL4, these 6 cysteines co-ordinate the two metal ions each of which, in the Cd(II) protein, are co-ordinated by 4 sulphur atoms. It is likely, therefore, that the structure of the DNA binding domain must differ from the structure of the zinc finger proposed for TFIIIA.

Given the similar DNA binding affinity, we expected that the structure of Zn(II) GAL4 would be the same as that of the Cd(II) protein. We speculate that both the oxygen ligand and the sulphur atom at 3.34 Å are due to the co-ordination of the zinc ions by the sulphonic acid group from the Hepes buffer. This could mean that the buffer is displacing one or more of the natural ligands. Consistent with this hypothesis para-hydroxy mercuribenzoate reacts with two sulphydryl groups per protein molecule in Zn(II) GAL4, but with none in Cd(II) GAL4, suggesting that in Zn(II) GAL4 two of the cysteine residues are not involved in zinc co-ordination. It remains a possibility, however, that the oxygen that we see in the e.x.a.f.s. results from excess zinc bound to the protein. Because of this ambiguity we also studied the e.x.a.f.s. of the Cd(II) protein.

For Cd(II) GAL4 we propose that each monomer of the protein contains a two metal ion cluster where each cadmium ion is tetrahedrally co-ordinated by four cysteine residues, two of which act as bridging ligands. It is known that the GAL4 protein binds to DNA as a dimer [17] raising the possibility that it is metal-linked as suggested for the HIV Tat protein [18]. Although the results obtained so far cannot rule out this possibility, we have no evidence for the formation of such a metal-linked dimer. Indeed, our own unpublished results and those of Carey et al. [17], show that the main dimerisation function is not contained within the cysteine-rich region of the protein suggesting that the dimer is not metal linked.

Despite the fact that the Zn(II) protein has a different metal ion co-ordination it can still bind with similar affinity to DNA. Further work is in progress to determine whether Zn(II) GAL4 has a substantially different structure or whether the solvent or buffer is simply displacing one or more of the natural ligands.



Acknowledgements: We thank the Director of Daresbury Laboratory for provision of facilities, Mrs Juliette Jacoby for amino acid analysis, Michael Carey, Melvyn Hollis, Hitoshi Kakidani and Mark Ptashne for *E. coli* strains, help and advice. This work was supported by the SERC and its Cambridge Centre for Molecular Recognition.

REFERENCES

[1] Giniger, E., Varnum, S.M. and Ptashne, M. (1985) Cell 40, 767-774.
[2] Keegan, L., Gill, G. and Ptashne, M. (1986) Science 231, 699-704.
[3] Pfeifer, K., Kim, K.S., Kogan, S. and Guarente, L. (1989) Cell 56, 291-301.
[4] Johnston, M. (1987) Nature 328, 353-355.
[5] Klug, A. and Rhodes, D. (1987) Trends Biochem. Sci. 12, 464-469.
[6] Miller, J., McLachlan, A.D. and Klug, A. (1985) EMBO J. 4, 1609-1614.
[7] Pan, T. and Coleman, J.E. (1989) Proc. Natl. Acad. Sci. USA 86, 3145-3149.
[8] Lin, Y.S., Carey, M.F., Ptashne, M. and Green, M.R. (1988) Cell 54, 659-664.
[9] Fried, M. and Crothers, D. (1981) Nucleic Acids Res. 9, 6505-6525.
[10] Cramor, S.P., Tench, O., Yocum, M. and George, G.W. (1988) Nuclear Instr. Methods A266, 586-591.
[11] Lee, P.A. and Pendry, J.B. (1975) Phys. Rev. B11, 2795-2811.
[12] Perutz, M.F., Hasnain, S.S., Duke, P.J., Sessler, J.L. and Hahn, J.F. (1982) Nature 295, 535-538.
[13] Gurman, S.J., Binsted, N. and Ross, I. (1984) J. Phys. Chem. 17, 143-151.
[14] Riordan, J.F. and Vallee, B.L. (1972) in: Methods in Enzymology, Vol. 25. (Hirs, C.H.W. and Timasheff, S.N., eds.), Academic Press, New York, pp. 449-456.

[15] Abrahams, I.L., Garner, C.D., Bremner, I., Diakun, G.P. and Hasnain, S.S. (1985) J. Am. Chem. Soc. 107, 4596–4597.
[15a] Abrahams, I.L., Bremner, I., Diakun, G.P., Garner, C.D., Hasnain, S.S., Ross, I. and Vasák, M. (1986) Biochem. J. 236, 585–589.
[16] Freedman, L.P., Luisi, B.F., Korszum, Z.R., Basavappa, R., Sigler, P.B. and Yamamoto, K.R. (1988) Nature 334, 543–546.
[17] Carey, M.F., Kakidani, H., Leatherwood, J., Mostashari, F. and Ptashne, M. (1989) J. Mol. Biol. 209, 423–432.
[18] Frankel, A.D., Bredt, D.S. and Pabo, C.O. (1988) Science 240, 70–73.