Volume 281, number 1,2, 223–226 FEBS 09581 April 1991

ADONIS 001457939100291 2

Polypeptide–metal cluster connectivities in Cd(II) GAL4

Paresh L. Gadhavi[1], Adrian L. Davis[1], Jane F. Povey[1], James Keeler[2] and Ernest D. Laue[1]

[1] *Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK* and [2] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK*

Received 5 February 1991

Two-dimensional ^{1}H-^{113}Cd correlation NMR spectra have been used to determine the polypeptide/metal cluster connectivities in Cd(II) GAL4. The results show that the protein contains a two metal ion cluster where Cys-11 and Cys-28 are the bridging ligands.

GAL4; 2D ^{1}H-^{113}Cd NMR; Metal ion co-ordination; DNA binding; Metal cluster; Zinc finger

1. INTRODUCTION

The DNA binding domain of the yeast transcriptional activator GAL4 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 [1]. Using EXAFS and NMR spectroscopy respectively, we [2] and others [3] have suggested that the protein contains a two metal ion cluster distinguishing it from the 'zinc finger' proteins typified by the *Xenopus laevis* transcription factor (TF)IIIA [4]. Further NMR studies of the structure of the zinc-containing DNA binding domain have been carried out on a purified 43 amino acid fragment (GAL4(7–49)) recently obtained by us from *E. coli* [5]. We have reported the complete assignment of the ^{1}H NMR spectrum and identification of secondary structure in the Zn(II) protein [6].

The polypeptide chain is folded about two metal ions, and as a result rather fewer NOE enhancements are present in the NOESY spectrum than might be expected for a protein of this molecular weight; a similar observation has also been made in the spectra of metallothioneins [7]. In order to obtain a greater number of distance constraints on the structure and to establish firmly that GAL4 does indeed contain a two metal ion cluster, we have determined the polypeptide/metal cluster connectivities in Cd(II) GAL4 by using ^{1}H-^{113}Cd two-dimensional correlation experiments. These results allow us to propose a model for the $Cd(II)_2$ Cys_6 cluster within the DNA binding domain of GAL4.

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

2. MATERIALS AND METHODS

2.1. *Protein purification*

The Cd(II) GAL4(7–49) protein was prepared by adding a 5-fold excess of ^{113}Cd to the Zn(II) protein purified as described previously [6]. Samples for NMR spectroscopy were concentrated to 1–2 mM and adjusted to pH 5.4 in the presence of 10% $^{2}H_2O$ and 5–10 mM ^{113}cadmium perchlorate.

2.2. *NMR spectroscopy*

The spectrometer used to record and process the NMR spectra (a Bruker AMX 600) operated at 600 MHz for protons. All spectra were recorded from samples held at 283 K. The ^{1}H-^{113}Cd two-dimensional correlation experiments were recorded using the standard sequence which involves evolution of heteronuclear multiple quantum coherence (HMQC) during the t_1 period [18]. The HMQC pulse sequence used was $\pi/2(^{1}H)$-τ-$\pi/2$ (^{113}Cd)-$t_1/2$-$\pi(^{1}H)$-$t_1/2$-$\pi/2(^{113}Cd)$-τ-acquire (^{1}H)/decouple (^{113}Cd). The cross-peaks in the resulting spectrum show the chemical shifts of ^{113}Cd and ^{1}H atoms ($C^{\beta}H$ in cysteine residues) which are directly coupled. In some experiments a further $\pi/2(^{1}H)$ pulse is added just prior to data acquisition. The resulting spectrum shows the same cross-peaks as in an HMQC as well as further 'relay' peaks which indicate a correlation between ^{113}Cd and $C^{\alpha}H$ in cysteine residues, even if $C^{\alpha}H$ is not directly coupled to ^{113}Cd. The spectral width in the f_2 dimension was 7246 Hz and the acquisition time in t_2 was 0.283 s. In f_1 the spectral width was 14049 Hz and t_1 was incremented in 64 equal steps to a maximum of 4.01 ms. The TPPI method of frequency discrimination was used [9,10] as was presaturation of the H_2O resonance during the relaxation delay.

3. RESULTS

Double quantum filtered COSY [11] and X-filtered COSY [12] experiments were initially recorded to confirm that all six cysteine residues were coordinated to a cadmium ion (data not shown). The cysteine residues in Cd(II) GAL4(7–49) were initially assigned by comparison with the $C^{\alpha}H$ and $C^{\beta}H$ chemical shifts in the Zn(II) protein [6]; a partial sequential assignment was used to confirm these assignments. A series of HMQC and HMQC-relay spectra were then recorded, using different multiple quantum preparation periods, τ, of between 10 and 60 ms. Recording several spectra in this

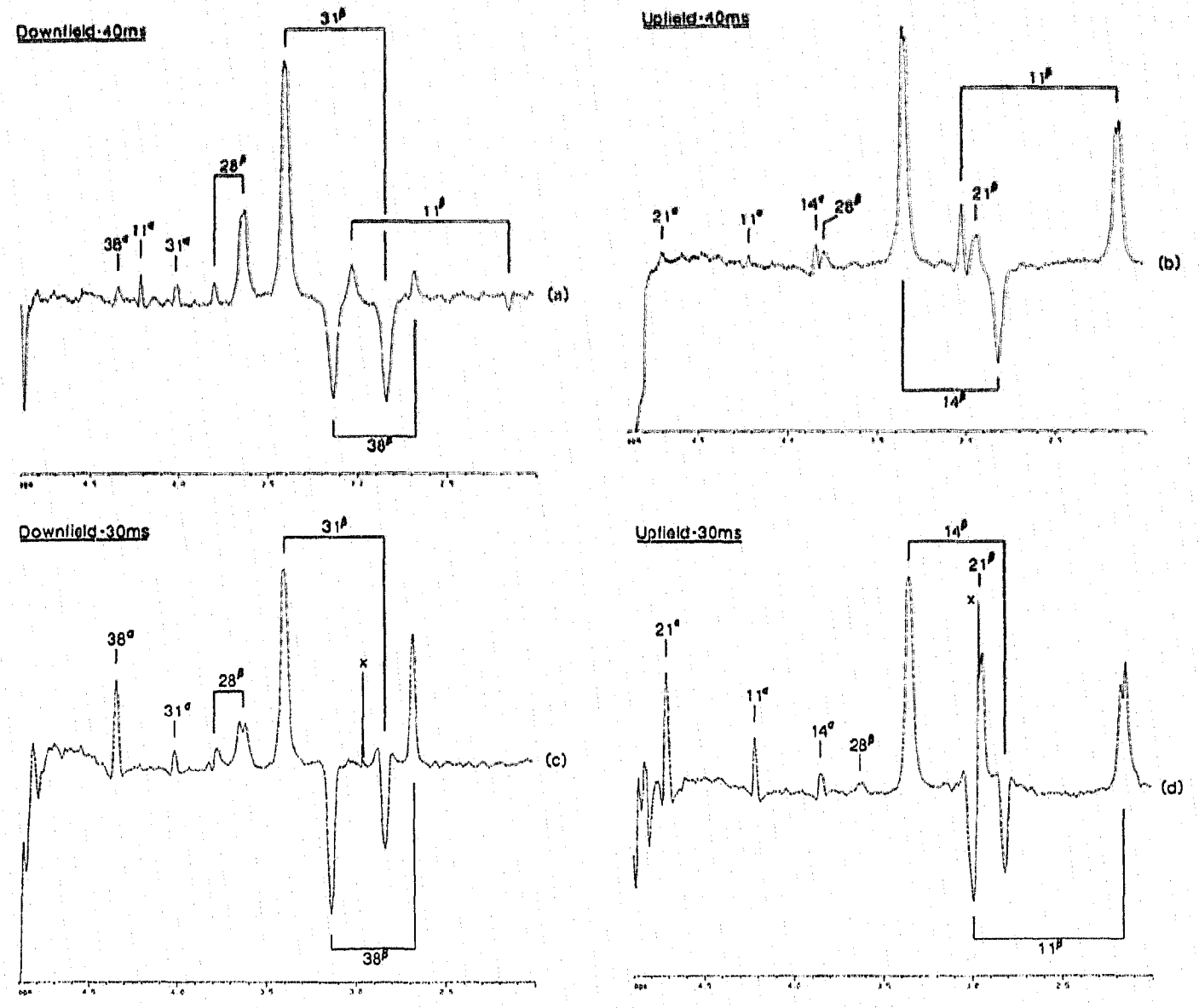

Fig. 1. Cross-sections, taken parallel to f_2, from the ^1H-^{113}Cd two-dimensional relayed correlation spectrum (HMQC-relay) of Cd(II) GAL4(7–49). Sections (a) and (c) are taken at the f_1 frequency corresponding to the downfield Cd resonance, and sections (b) and (d) likewise refer to the upfield resonance. The $C^{\alpha}H$ resonances of cysteine residues appear between 3.85 and 4.70 ppm and $C^{\beta}H$ resonances of cysteine residues appear between 2.15 and 3.80 ppm. The signals are identified by the position of the cysteine residue in the protein sequence. The evolution of the magnetisation under the ^1H, ^1H couplings during the delay, τ, leads to the varying amplitudes and phases of the cross-peaks. Some peaks also show splittings due to ^1H, ^1H coupling. Cross-sections (a) and (b) are taken from a spectrum recorded with the delay τ set to 40 ms, while for (c) and (d) the delay was 30 ms. In the latter X marks an instrumental artefact.

way is essential as the cross-peak intensities and phases have a complex dependence on τ and it is possible that some cross-peaks will be missing just due to an unfortunate choice of τ. All spectra showed just two cadmium resonances having very similar chemical shifts to those previously reported by Pan and Coleman for GAL4(62*) [13]. When taken together with the assignment of the cysteine residues in Cd(II) GAL4(7–49), these spectra clearly establish the polypeptide-metal cluster connectivities in the Cd(II) protein. For example, in the HMQC-relay spectrum recorded with τ = 40 ms the downfield cadmium resonance shows correlations to the $C^{\beta}Hs$ of cysteine residues 11, 28, 31 and 38 and relay peaks to the $C^{\alpha}Hs$ of cysteine residues 11, 31 and 38 (Fig. 1a). Some of the correlations to the $C^{\alpha}Hs$ are more clearly detected in a similar spectrum recorded with τ = 30 ms (Fig. 1c); the $C^{\alpha}H$ resonance of Cys-28 is buried under the residual H_2O signal. The upfield cadmium resonance shows correlations to both the $C^{\alpha}H$ and the $C^{\beta}Hs$ of cysteine residues 11, 14 and 21 (Fig.

1b). Again some of the correlations to the $C^{\alpha}Hs$ are more clearly seen in the 30 ms spectrum (Fig. 1d). In addition, correlations to one of the $C^{\beta}Hs$ of Cys-28 (at 3.80 ppm) and to the other $C^{\beta}H$ of Cys-28 (at 3.63 ppm) were observed in the 40 ms and 30 ms spectrum, respectively (Fig. 1b and 1d).

These results were all confirmed by repeating the experiments on a freshly prepared sample of the protein.

4. DISCUSSION

Although the assignment of the proton spectrum of the Cd(II) protein is not yet complete we are reasonably confident that the chemical shifts of the $C^{\alpha}H$ and $C^{\beta}Hs$ of the six cysteine residues are very similar to those found in the Zn(II) protein (Table I), perhaps implying that the metal ion coordination is similar. The HMQC spectra clearly show that in the Cd(II) protein one metal ion is coordinated by cysteine residues 11, 14, 21 and 28, whilst the other is coordinated by cysteine residues 11, 28, 31 and 38. It can therefore be concluded that the Cd(II) form of the GAL4(7-49) protein does indeed contain a two metal ion cluster where Cys-11 and Cys-28 act as bridging ligands (see Fig. 2).

As a result of our EXAFS studies of GAL4(1-147) [2], and the NMR studies of GAL4(62*) by Pan and Coleman [3], we had expected such a cluster. However, we find that the bridging ligands are different in GAL4(7-49) from those suggested by the NMR studies of GAL4(62*) by Pan and Coleman [3]. The chemical shifts of the $C^{\alpha}H$ and the $C^{\beta}Hs$ of the cysteine residues in GAL4(62*) are reproduced (from [3]) in Table II; they are ordered so as to aid comparison with our own results for GAL4(7-49). With one exception (the $C^{\beta}Hs$ of Cys 11 and spin system 4) the chemical shifts in GAL4(62*) appear to be very similar to those found in GAL4(7-49). On the basis of an analysis of the ^{1}H-^{113}Cd heteronuclear couplings, Pan and Coleman found that in GAL4(62*) their spin systems 5 and 6 were the bridging cysteine ligands. In GAL4(7-49) we can find no evidence, either from an analysis of couplings or from the ^{1}H-^{113}Cd correlation experiments, that these two spin systems represent the bridging cysteine ligands. Instead we find that Cys-11 and Cys-28 (their spin systems 4 and 1) are the bridging ligands (see above). In addition their tentative sequential assignment assigns spin systems 5 and 6 to Cys-38 and Cys-21, whilst ours would assign them to Cys-31 and Cys-14, respectively.

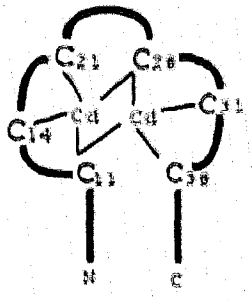

Fig. 2. A model of the $Cd(II)_2Cys_6$ metal ion cluster in GAL4(7-49).

The method that we have used to identify the bridging cysteines is based on the occurrence of cross-peaks at certain frequency coordinates. As the occurrence of such cross-peaks indicates *unambiguously* the presence of a coupling or chain of couplings, the interpretation of the spectra is straightforward. Pan and Coleman used *quantitative* measurements of splittings from cross-peak multiplets to identify the bridging cysteines. Such an approach is undoubtedly both valid and elegant, but is subject to uncertainties caused by the poor definition of multiplets inevitable in protein NMR spectra. Having identified which spin systems are from the bridging cysteines, the assignment of these to particular cysteine residues in the sequence relies on the validity of the assignment of the spectrum. Our assignment [6] appears to be consistent with the available NMR data but at this stage may be regarded as tentative.

It is possible that the Cd(II) form of GAL4(7-49) has different metal ion coordination from GAL4(62*) or from the Zn(II) form of GAL4(7-49). It is also possible

Table I

^{1}H NMR chemical shifts of the cysteine residues in Cd(II) and Zn(II) GAL4(7-49) at pH 5.4 and 283K (relative to internal TSP)

	Protein					
	Cd(II) GAL4(7-49)			Zn(II) GAL4(7-49)		
	Proton			Proton		
	$C^{\alpha}H$	$C^{\beta}H(1)$	$C^{\beta}H(2)$	$C^{\alpha}H$	$C^{\beta}H(1)$	$C^{\beta}H(2)$
Cys^{11}	4.21	2.15	3.01	4.17	2.21	2.93
Cys^{14}	3.85	2.82	3.35	3.81	2.82	3.27
Cys^{21}	4.70	2.94	-	4.72	2.92	-
Cys^{28}	4.85	3.63	3.80	4.67	3.58	3.70
Cys^{31}	4.01	2.84	3.40	4.02	2.80	3.32
Cys^{38}	4.34	2.68	3.14	4.41	2.58	3.18

Table II

^{1}H NMR chemical shifts of the cysteine residues in Cd(II) and Zn(II) GAL4(62*) at 308K (from [3])

	Protein					
	Cd(II) GAL4(62*) at pH 8.0			Zn(II) GAL4(62*) at pH 5.4		
	Proton			Proton		
	$C^{\alpha}H$	$C^{\beta}H(1)$	$C^{\beta}H(2)$	$C^{\alpha}H$	$C^{\beta}H(1)$	$C^{\beta}H(2)$
Spin system 4	4.22	2.99	3.14	4.12	2.88	-
Spin system 6	3.75	2.77	3.26	3.73	2.73	3.17
Spin system 2	4.64	2.95	-	4.65	2.84	-
Spin system 1	4.94	3.60	3.73	4.67	3.50	3.60
Spin system 5	3.98	2.82	3.35	3.98	2.75	3.25
Spin system 3	4.33	2.67	3.04	4.29	2.55	3.04

that the metal coordination is variable. In an attempt to show that the preparation of the Cd(II) form of the GAL4(7-49) protein was reproducible, we prepared a fresh sample and repeated all the NMR experiments. The results were identical. In addition, the fact that some of the ring current shifted resonances (e.g. those of the $C^{\beta}Hs$ of Asn-34 and Cys-28) are so similar in the 'native' Zn(II) form of the protein [6] and the Cd(II) form (data not shown) also argues against this possibility. Moreover, the general similarity of the chemical shifts of the $C^{\alpha}H$ and the $C^{\beta}Hs$ of the six cysteine residues in GAL4(7-49) and GAL4(62*) would suggest that the structures of the two proteins are not too dissimilar, but further speculation must await the completion of both three-dimensional structure determinations.

Acknowledgements: We thank Peter Alefounder for the plasmid expressing GAL4(1-49) and Hitoshi Kakidani and Mark Ptashne for expression plasmids and *E. coli* strains. This work was supported by SERC and its Cambridge Centre for Molecular Recognition.

REFERENCES

[1] see e.g. Fig. 1 from Pfeifer, K., Kim, K.S., Kogan, S. and Guarente, L. (1989) Cell 56, 291-301.
[2] Povey, J.F., Diakun, G.P., Garner, C.D., Wilson, S.P. and Laue, E.D. (1990) FEBS Lett. 266, 142-146.
[3] Pan, T. and Coleman, J.E. (1990) Proc. Natl. Acad. Sci. USA 87, 2077-2081.
[4] Miller, J., McLachlan, A.D. and Klug, A. (1985) EMBO J. 4, 1609-1614.
[5] Gadhavi, P.L., Alefounder, P.R., Green, B. and Laue, E.D. (in preparation).
[6] Gadhavi, P.L., Raine, A.R.C., Alefounder, P.R. and Laue, E.D. (1990) FEBS Lett. 276,49-53.
[7] Braun, W., Wagner, G., Wörgötter, E., Vasák, M., Kägi, J.H.R. and Wüthrich, K. (1986) J. Mol. Biol. 187, 125-129.
[8] Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) J. Am. Chem. Soc. 105, 7188-7190; and Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) J. Magn. Reson. 55, 301-315.
[9] Bodenhausen, G., Vold, R.L. and Vold, R.R. (1980) J. Magn. Reson. 37, 93-106.
[10] Marion, D. and Wüthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967-974.
[11] Piantini, U., Sørensen, O.W. and Ernst, R.R. (1982) J. Am. Chem. Soc. 104, 6800-6801.
[12] Wörgötter, E., Wagner, G. and Wüthrich, K. (1986) J. Am. Chem. Soc. 108, 6162-6167.
[13] Pan, T. and Coleman, J.E. (1990) Biochemistry 29, 3023-3029.