NMR chemical shift perturbation mapping of DNA binding by a zinc-finger domain from the yeast transcription factor ADR1

MIA SCHMIEDESKAMP, PONNI RAJAGOPAL, AND RACHEL E. KLEVIT

Biomolecular Structure Center and Department of Biochemistry, University of Washington, Seattle, Washington 98195

(RECEIVED March 3, 1997; ACCEPTED May 16, 1997)

Abstract

Mutagenesis studies have revealed that the minimal DNA-binding domain of the yeast transcription factor ADR1 consists of two Cys₂-His₂ zinc fingers plus an additional 20 residues proximal and N-terminal to the fingers. We have assigned NMR ¹H, ¹⁵N, and ¹³C chemical shifts for the entire minimal DNA-binding domain of ADR1 both free and bound to specific DNA. ¹H chemical shift values suggest little structural difference between the zinc fingers in this construct and in single-finger constructs, and ¹³C_α chemical shift index analysis indicates little change in finger structure upon DNA binding. ¹H chemical shift perturbations upon DNA binding are observed, however, and these are mapped to define the protein–DNA interface. The two zinc fingers appear to bind DNA with different orientations, as the entire helix of finger 1 is perturbed, while only the extreme N-terminus of the finger 2 helix is affected. Furthermore, residues N-terminal to the first finger undergo large chemical shift changes upon DNA binding suggesting a role at the protein–DNA interface. A striking correspondence is observed between the protein–DNA interface mapped by chemical shift changes and that previously mapped by mutagenesis.

Keywords: DNA-binding domain; protein–DNA interaction; protein NMR; transcription factor; zinc finger

The classic Cys₂-His₂ zinc finger is a ubiquitous eukaryotic DNA-binding motif first identified in the *Xenopus* transcription factor TFIIIA (Brown et al., 1985; Miller et al., 1985). Hundreds of potential Cys₂-His₂ family members have since been identified by homology to the consensus sequence (Y,F)-X-C-X_{2,4}-C-X₃-F-X₅-L-X₂-H-X_{3,5}-H-X_{2,6} (Jacobs, 1992). Structural studies of zinc fingers free in solution and bound in tandem arrays to DNA demonstrate a common ββα fold (reviewed in Kaptein, 1992; Berg, 1993; Schmiedeskamp & Klevit, 1994). The conserved cysteine and histidine residues act as ligands for tetrahedral coordination of a zinc ion and the remaining conserved residues form a hydrophobic core. Co-crystal structures reveal that residues at positions –1, 2, 3, and 6 of the α-helix often participate in specific DNA binding (Pavletich & Pabo, 1991, 1993; Fairall et al., 1993; Kim & Berg, 1996). Further evidence of the identity of critical DNA-contacting residues in Cys₂-His₂ zinc fingers has come from mutagenesis studies, in which substitution at these positions leads to diminished affinity or altered specificity of DNA binding (reviewed in Berg, 1992).

While both structural and mutational studies have elucidated details of zinc finger–DNA binding, the techniques have largely

been applied to separate systems, allowing little opportunity for direct comparison of the two complementary types of data. An exception is the ADR1 transcription factor from the yeast *Saccharomyces cerevisiae*, which regulates expression of the glucose-repressible alcohol dehydrogenase, ADH2, and various peroxisomal proteins through an upstream activating sequence known as UAS1 (Denis & Young, 1983; Shuster et al., 1986; Simon et al., 1991). Comprehensive mutagenesis of the ADR1 DNA-binding domain, which contains two Cys₂-His₂ zinc fingers, has been performed (Hartshorne et al., 1986). These studies include deletion analysis (Thukral et al., 1989, 1991a); random mutagenesis followed by loss- or gain-of-function screens (Blumberg et al., 1987; Camier et al., 1992; Cook et al., 1994); and alanine-scanning (Thukral et al., 1991b) and change-of-specificity mutagenesis (Thukral et al., 1992; Cheng & Young, 1995). Studies have also been conducted on UAS1 DNA to identify critical base pairs in the ADR1 binding site (Cheng et al., 1994). Complementing these genetic and biochemical data, individual ADR1 zinc fingers have been characterized by solution NMR studies, including detailed structure determinations for the wild-type N-terminal finger and a mutant version of the C-terminal finger (Parraga et al., 1988; Klevit et al., 1990; Xu et al., 1991; Hoffman et al., 1993; Bernstein et al., 1994).

ADR1 is of particular interest because it requires a region outside the zinc-finger domains for high-affinity DNA binding (Thukral et al., 1989, 1991a). Most zinc-finger motifs studied thus far occur

Reprint requests to: Rachel E. Klevit, Biomolecular Structure Center, Box 357742, University of Washington, Seattle, Washington 98195-7742; e-mail: klevit@u.washington.edu.

in arrays of three or more and require no flanking sequences for DNA binding. Three notable exceptions have been characterized structurally. In both the *Drosophila* Tramtrack (Fairall et al., 1993) and yeast SWI5 proteins (Neuhaus et al., 1992), stable folding of the N-terminal zinc finger requires an additional, third strand in the β -sheet. The *Drosophila* GAGA factor contains a single zinc finger and requires approximately 30 residues adjacent and N-terminal to the finger to bind DNA (Pedone et al., 1996; Omichinski et al., 1997). ADR1 presents an additional exception, for while its two zinc fingers fold in the absence of any additional sequence (Párraga et al., 1988; Klevit et al., 1990), high-affinity DNA-binding requires approximately 20 residues N-terminal to the first finger (Thukral et al., 1991a). In addition, a mutation has been identified in this flanking sequence that can suppress defects in DNA binding arising from mutations in either zinc finger (Camier et al., 1992).

Structural studies of ADR1 have thus far been limited to 30-residue synthetic peptides encompassing individual zinc fingers (Xu et al., 1991; Hoffman et al., 1993; Bernstein et al., 1994). Analysis of a larger construct in complex with DNA may provide an understanding of the role of the N-terminal flanking region and the inability of the zinc fingers alone to bind DNA with high affinity. This paper presents the first NMR studies of a construct containing the entire minimal DNA-binding domain of ADR1, including both zinc fingers and the N-terminal flanking region. NMR chemical shift assignments are presented for the construct both free and bound to DNA. These assignments are compared to data for the isolated fingers, and chemical shift perturbations upon DNA binding are mapped. The resulting picture of protein–DNA interaction complements data from existing mutagenesis studies.

Results and discussion

Design of constructs

The construct used in this study, ADR1z, spans residues 75–161 of ADR1, encompassing both zinc-finger motifs as well as a region N-terminal to the first finger (Fig. 1a). This sequence was chosen based on deletion mutagenesis results that defined the minimal DNA-binding domain of ADR1. While deletion of residues C-terminal to the zinc fingers has no effect on DNA binding, deletions N-terminal to the fingers diminish binding affinity in electrophoretic mobility shift assays (EMSA) (Thukral et al., 1989). N-terminal truncation at residue 84 of ADR1 results in a minor decrease in binding affinity. Further deletion to residue 88, however, results in loss of all detectable binding (Thukral et al., 1991a). The choice of Met 75 as a starting point for the ADR1z construct allowed inclusion of the pivotal region of residues 84 to 88 while minimizing the size of the polypeptide for NMR analysis. Additionally, it allowed utilization of a native AUG start codon, yielding expressed polypeptide of entirely native sequence. EMSA analysis indicates an apparent dissociation constant on the order of 12 nM for ADR1z binding to UAS1, the 22-base-pair palindromic recognition site for ADR1 (data not shown).

Mutagenesis shows that all base pairs important for specific ADR1 binding fall within UAS1 (Cheng et al., 1994). While two ADR1 molecules must bind to UAS1 to activate *ADH2* expression, all available biochemical evidence indicates that ADR1 binds without positive cooperativity (Thukral et al., 1991a). ADR1 binds as strongly to DNA bearing a single UAS1 half-site as to each half of the full UAS1 palindrome (Thukral et al., 1991a). In addition, ethylation interference (Cheng et al., 1994) and methylation inter-

Fig. 1. Sequence of constructs used for NMR studies. **A:** The sequence of ADR1z, the minimal ADR1 DNA-binding domain construct. Brackets indicate the extent of the shortest deletion mutant found to bind DNA with high affinity and of the longest deletion mutant found not to bind DNA. The circle indicates the position of a suppressor mutation, R91K. Positions –1, 3, and 6 of the finger 2 α -helix are labeled. **B:** The sequence of UAS1 and of the 14mer DNA used in complex with ADR1z. This DNA contains one half-site of UAS1.

ference (Thukral et al., 1991a) footprints of ADR1 bound to a lone UAS1 halfsite fall entirely within the outer 10 base pairs of the half-site. Interference patterns identical to these are observed for each half of a full UAS1 palindrome bound by ADR1 (Thukral et al., 1991a; Cheng et al., 1994). Based on these results, only a single half-site of UAS1 was included in the DNA sequence used for this study to minimize the size of the final protein–DNA complex for NMR analysis.

DNA oligonucleotides of different lengths were evaluated; they contained at least the outer 10 base pairs of a UAS1 half-site in varying contexts. NMR spectra of ADR1z bound to each of these sequences yielded essentially identical peak positions. The chosen 14mer was the minimal length to retain good solubility characteristics in complex with ADR1z (Fig. 1b). This 14mer contains native UAS1 sequence except for two A/T base pairs from the center of the palindrome, which have been replaced by G/C base pairs to increase duplex stability. Previous mutagenesis studies showed that ADR1 exhibits no base preference at these positions (Cheng et al., 1994).

ADR1z forms a 1:1 complex with the 14mer DNA that is soluble at 2 mM under NMR conditions. This complex displays slow exchange behavior on the NMR time scale (data not shown). HSQC spectra of ^{15}N -labeled ADR1z both free and bound to 14mer are

shown in Figure 2. Each spectrum displays a single resonance for each amide except for several residues at the N- and C-termini. The N-terminal doubling is likely due to heterogeneous processing of the initiating methionine, as detected by electrospray mass spectrometry (data not shown). Doubling at the two C-terminal residues is unexplained; however, such behavior has been reported for another zinc-finger peptide (Dutnall et al., 1996). Due to the large chemical shift perturbations for many resonances in ADR1z upon DNA binding and the slow exchange behavior of the complex, assignments were obtained independently for the free and bound polypeptide.

Resonance assignment strategy

The cornerstone of the assignment process for both free and DNA-bound ADR1z was the HNCA, HN(CO)CA pair of triple-resonance experiments (Fig. 3). These experiments provide connectivities between sequential alpha carbons, resolved by the well-dispersed backbone nitrogen and amide proton resonances. Due to considerable redundancy of $^{13}\text{C}_\alpha$ chemical shift, however, additional experiments were required to complete backbone assignments.

In the case of free ADR1z, ambiguities in the backbone assignments were resolved using isotopic labeling of specific amino acids (Muchmore et al., 1989). Samples labeled in *Escherichia coli* with ^{15}N -leucine or ^{15}N -glutamate were analyzed to classify backbone amide peaks in ^1H - ^{15}N HSQC experiments. Due to differential scrambling of the isotope label to L, I, V, and A in the leucine experiment, and to D, E, Q, F, Y, N, L, V, I, S, and A in the glutamate experiment, peaks were assigned to the following classes: D, E, Q, or F; L; V or I; A; N, S, or Y; and C, H, K, G, R, or T (unlabeled). Further clarification of amino acid type was obtained

from samples grown on ^{15}N -labeled media to which ^{14}N -lysine or arginine was added (Matsuo et al., 1991). Because metabolic scrambling of the amino nitrogen is inefficient for these amino acids, HSQC spectra of these samples lack peaks for lysine and arginine, respectively, allowing their identification.

Additional sequential information was obtained from specific ^{13}C labeling (Matsuo et al., 1991). Leucine, arginine, or lysine labeled with ^{13}C at the backbone carbonyl position was incorporated into uniformly ^{15}N -labeled ADR1z. In HSQC spectra of these samples, backbone amides sequential to the incorporated ^{13}C label yield crosspeaks split in the ^{15}N dimension, allowing identification of the $i+1$ neighbors to L, K, and R residues.

Assignment of the backbone ^{15}N , NH, and $^{13}\text{C}_\alpha$ resonances of free ADR1z was completed using these specific labeling experiments, the HNCA, HN(CO)CA triple-resonance experiments, and ^{15}N -edited TOCSY and NOESY experiments, which yielded traditional NOESY-based connectivities (Wüthrich, 1986).

An identical assignment strategy was applied to DNA-bound ADR1z. Due to poorer signal from the ^{15}N -glutamate sample, however, the differential labeling results were not as definitive as those for the free protein, resulting in a number of remaining ambiguities in backbone assignments. These were resolved with HNCACB and CBCA(CO)NH triple resonance spectra, which provide connectivities between sequential $^{13}\text{C}_\beta$ resonances as well as between sequential $^{13}\text{C}_\alpha$ resonances.

Side-chain ^{13}C and ^1H assignments for ADR1z and ADR1z-DNA were obtained from 3D HCCH-TOCSY experiments on uniformly ^{13}C , ^{15}N -labeled samples in D_2O . Due to relatively weak signals from the arginine side-chain carbons in the ADR1z-DNA spectrum, a sample specifically labeled with $^{13}\text{C}_6$ -arginine was

Fig. 2. ^{15}N -HSQC spectra of ADR1z and ADR1z-DNA. The HSQC spectrum of ADR1z uniformly labeled with ^{15}N is shown in red, that of DNA-bound ADR1z in blue. The free spectrum was acquired at pH 5.4, 30°C; the bound spectrum was acquired at pH 7.0, 37°C. These differences in temperature and pH alone do not perturb the spectrum of free ADR1z significantly (data not shown).

Fig. 3. 3D triple-resonance HNCA experiments. Strips are taken at the amide proton and nitrogen positions of various residues. Observed peaks are alpha carbon resonances of residues *i* and *i*-1. **A:** HNCA strips for ADR1z. **B:** HNCA strips for DNA-bound ADR1z.

prepared. 2D HCCH-TOCSY and HCCH-COSY experiments on this sample provided assignments for these residues. Proton assignments for aromatic spin systems in free ADR1z were obtained from homonuclear ^1H D $_2\text{O}$ TOCSY and NOESY spectra. Side-chain NH_2 assignments were obtained from 3D ^{15}N -edited NOESY, CBCA(CO)NH, and HN(CO)CA spectra.

Nearly complete ^{13}C , ^{15}N , and ^1H assignments are presented for ADR1z and ADR1z-DNA in Table I and Table II of the Supplementary Material. Missing assignments for ADR1z-DNA are generally due to unavoidable spectral overlap, particularly for the side-chain spin systems of K100, L101, and K130, which lie adjacent to very intense random coil spin systems.

Comparison to single zinc-finger peptides: finger-finger interactions

The two zinc fingers of ADR1 have been studied individually by NMR spectroscopy in the context of 30-residue synthetic peptides spanning the consensus zinc finger sequences. ^1H resonance assignments have been published for the N-terminal finger (ADR1b; residues 102–130) (Hoffman et al., 1993), and for variations of the

C-terminal finger, including a wild-type peptide (ADR1a; 130–159) (Xu et al., 1991) and a proline-to-alanine double mutant (PAPA) (Bernstein et al., 1994). A comparison of NMR chemical shift assignments for the fingers of ADR1z with wild-type single-finger assignments is shown in Figure 4.

The overall pattern of chemical shift differences indicates that there is little, if any, perturbation of either finger in the context of the two-finger construct. Few resonances exhibit chemical shift differences larger than the apparent ^1H linewidth of about 0.06 ppm, despite a 5 °C difference in temperature of acquisition.

Fig. 4. Comparison of ADR1z chemical shifts to the chemical shifts of single zinc-finger peptides. Chemical shifts for the two zinc fingers of ADR1z are compared to shifts for the analogous ADR1b and ADR1a single finger peptides. In the case of ADR1a, the shifts used correspond to the minor conformational form (Xu et al., 1991). **A:** The difference in NH chemical shift; **B:** The difference in H_α chemical shift; **C:** The difference in side-chain proton chemical shift. Positions where Zif268 (X) and MBP-1 (*) fingers interact with one another are marked in panel C. Horizontal lines in each panel indicate the apparent ^1H linewidth in ADR1z spectra (0.06 ppm).

Of those resonances that do exhibit perturbations, the majority belong to residues that lie near the termini of the single-finger peptides and are thus liable to end effects. For the remaining residues, no backbone or side-chain chemical shift difference greater than 0.12 ppm is observed, except for β and γ proton resonances of Q116 in the N-terminal finger.

Limited interactions between adjacent zinc fingers in tandem arrays have been observed in several cases. In the Zif268 zinc finger-DNA co-crystal structure, the side chain of an arginine equivalent to ADR1 R124 hydrogen bonds to the backbone carbonyl of the equivalent of T142 in each finger-finger interaction (Pavletich & Pabo, 1991). The two zinc fingers of human enhancer binding protein MBP-1 appear to interact in solution even in the absence of DNA. While the resulting relative finger orientation differs from that in Zif268, the participating residues fall at similar positions. A threonine at the equivalent of ADR1 S125 forms a putative hydrogen bond with the highly constrained side chain of a lysine at the equivalent of ADR1 T142 (Omichinski et al., 1992). In contrast, there appears to be no interaction between adjacent fingers in yeast SWI5 double-finger constructs in solution, based on chemical shifts and NOE patterns (Nakaseko et al., 1992).

ADR1z presents a second example in which no interfinger interaction occurs in the absence of DNA. Discounting end effects, only the side chain of Q116 is significantly perturbed in the N-terminal finger,¹ while no side-chain resonances are significantly perturbed in the C-terminal finger of the construct. Specifically, no side-chain perturbations are observed near the locations of residues shown to interact in Zif268 or MBP-1 (Fig. 4c). While it is possible that these residues participate in finger-finger interactions when bound to DNA, it appears that they do not in the free polypeptide.

Comparison to single zinc-finger peptides: Finger structure

The coincidence of ADR1z and ADR1b chemical shifts implies that the NMR structure determined for the single-finger ADR1b peptide (Hoffman et al., 1993) is a good model for finger 1 in the longer ADR1z construct. In particular, the N-terminus does not interact with the zinc fingers in the free ADR1z polypeptide. In SWI5 and Tramtrack, the first zinc finger folds in the presence of additional N-terminal residues which form a third strand to the β -sheet (Neuhaus et al., 1992; Fairall et al., 1993). In ADR1z, the lack of chemical shift perturbations of the N-terminal finger in the presence of flanking residues suggests that no such interaction occurs in this construct.

The C-terminal finger of ADR1z also appears to be nearly identical to the corresponding single-finger ADR1a peptide. A detailed structure has not been calculated for this peptide due to the existence of a second conformational form (Xu et al., 1991). A solution structure has been reported for the P131A,P133A double mutant, PAPA, which suffers less from peak doubling (Bernstein et al., 1994). These mutations alter DNA-binding affinity, however (Thukral et al., 1991b). Structural studies of ADR1z currently underway will provide our first opportunity to characterize fully the structure of the wild-type C-terminal finger.

¹The chemical shift differences for Q116 are unexplained. It is possible that the differences arise due to deamidation of glutamine to glutamate in the ADR1b peptide, which was heated during folding (Hoffman et al., 1993). Unfortunately, mass determinations that might have detected a 1 Da change in mass were not performed on the ADR1b NMR samples.

¹³C chemical shifts and secondary structure of ADR1z: Zinc fingers

A simple measure of secondary structure in proteins is the ¹³C chemical shift index (CSI) (Spera & Bax, 1991). Systematic downfield chemical shift differences from random coil values are observed for C α resonances in stretches of α -helix, while analogous upfield shifts are found in extended structures, including β -strands.

The difference of ¹³C α chemical shifts from random coil for free ADR1z is shown in Figure 5a. Strong downfield secondary shifts indicative of α -helix are observed for residues Q116 to S125 of finger 1, and for the analogous R144 to K153 of finger 2. The calculated structure of ADR1b contains a helix extending from Q116 to H126 (Hoffman et al., 1993). In addition, R115 is in an extended conformation and A114 is in right-handed helical space in ADR1b, in agreement with the respective upfield and downfield alpha carbon chemical shift perturbations observed for these positions in each finger of ADR1z.

¹³C chemical shift assignments and a CSI analysis have been published for three zinc fingers of TFIIIA (Liao et al., 1994). The pattern of chemical shifts in the C-terminal half of each finger of ADR1z is quite similar to that seen for the fingers of TFIIIA. Each helix begins at the identical position, preceded by a markedly

Fig. 5. ¹³C α chemical-shift index for ADR1z. **A:** Plot of the difference of ADR1z ¹³C α chemical shifts from random coil values. **B:** Plot of the difference of DNA-bound ADR1z ¹³C α chemical shifts from random coil values. A downfield shift from random coil yields a positive value in these graphs.

upfield-shifted alpha carbon. In each case, the helical pattern of strong downfield shifts appears to break just prior to the last histidine ligand.²

The cysteine ligands of ADR1z show pronounced downfield shifts from random coil C_{α} caused by zinc coordination, as was observed in the fingers of TFIIIA (Liao et al., 1994). Unfortunately, further comparisons in this region are inappropriate, as the TFIIIA fingers have four residues between cysteines while ADR1 fingers possess only two intervening residues. In the calculated structure of ADR1b, residues F104, V105, A112, and F113 comprise a β -sheet (Hoffman et al., 1993). Consistent with this result, in ADR1z, the C_{α} of each residue experiences an upfield perturbation from random coil indicative of extended structure. The analogous residues in finger 2 also exhibit upfield shifts, although that of Y132 is difficult to interpret as this residue precedes a proline; residues immediately N-terminal to prolines exhibit upfield chemical shift perturbations even in a random coil setting (Wishart et al., 1995).

In the structure of ADR1b, residues C106, T110, and R111 participate in a β -bulge, while residues E107, V108, and C109 form a reverse turn (Hoffman et al., 1993). In ADR1z finger 1, each of these residues, including the cysteines mentioned previously, exhibits a $^{13}C_{\alpha}$ shift downfield from random coil. The $^{13}C_{\alpha}$ chemical shift pattern of ADR1z finger 2, however, suggests a very different structure in this turn region. The residues between and just after the finger 2 cysteines (G135–R139) are shifted upfield of the analogous residues in finger 1 by an average of 2 ppm. This region is preceded by two prolines, P131 and P133, which are not present in finger 1. Mutation of these prolines to alanine in the single-finger PAPA peptide results in a structure very similar to that of ADR1b, including the β -bulge and reverse turn (Bernstein et al., 1994). It is thus possible that one or both of the prolines is responsible for a different turn structure in the wild-type finger 2.

P133 is more likely than P131 to cause a substantial change in turn structure. Two of the fingers of Zif268 have proline at the position analogous to P131, and these fingers fold into structures much like those of ADR1b and PAPA (Pavletich & Pabo, 1991; Hoffman et al., 1993; Bernstein et al., 1994). In contrast, the residue analogous to P133 adopts a backbone conformation inaccessible to proline in the CX₂C loops of the Zif268 fingers, as well as in PAPA (Bernstein et al., 1994). It was therefore concluded that backbone rearrangements would be required to accommodate the wild-type proline residue at position 133 of ADR1 (Bernstein et al., 1994); the results of the carbon chemical shift analysis of ADR1z indicate that the G135–R139 region of finger 2 does indeed differ in structure from the analogous E107–R111 region of finger 1.

TFIIIA and ADR1 have similar linker sequences connecting adjacent zinc fingers. In the TFIIIA three-finger construct, the $^{13}C_{\alpha}$ chemical shifts of the linker residues lay upfield of random coil, suggesting an extended structure even in the absence of DNA (Liao et al., 1994). This supposition was supported by the pattern of NOEs observed as well. No such regular pattern of secondary shifts is observed for the linker of ADR1z; these residues exhibit $^{13}C_{\alpha}$ chemical shifts near random coil values. It has already been established that the zinc fingers themselves do not interact in ADR1z.

It is also unlikely that they are maintained in any fixed orientation by the linker.

¹³C chemical shifts and secondary structure of ADR1z: N-terminus

As described above, the essential N-terminal flanking residues do not appear to interact with the fingers of ADR1z. It is further clear from the $^{13}C_{\alpha}$ chemical shift profile of the N-terminus that this region of the polypeptide is essentially random coil in the absence of DNA (Fig. 5a). The only significant perturbations from random coil chemical shifts in this region are induced by prolines, which cause upfield shifts of about 2 ppm in the residues preceding them, even in an otherwise random coil context (Wishart et al., 1995). ^{15}N linewidth measurements also support the view that the N-terminus of ADR1z is more mobile than the zinc fingers, as this region displays linewidths an average of 3 Hz smaller than those of the rest of the polypeptide (data not shown). The random coil nature of the N-terminal region of ADR1z must be interpreted with caution, however. It is not only possible that this region acquires structure in the presence of DNA, but it may also have a defined structure in the full-length protein that is not stable in the minimal ADR1z construct.

¹³C chemical shifts of bound ADR1z: Zinc fingers

When ADR1z binds to the 14mer DNA, a majority of residues experience large perturbations of backbone amide 1H and ^{15}N chemical shift, as seen in Figure 2. The $^{13}C_{\alpha}$ chemical shift profile of the zinc fingers remains fairly constant, however, suggesting that their structure is not changed much upon DNA binding (Fig. 5b). The beginning and ending points and overall profile of downfield shifts defining each α -helix is retained upon DNA binding, as is the extended nature of the residue immediately preceding the helix. The pattern of downfield shifts in the Cys–Cys loop of finger 1 is reiterated in the DNA-bound form, and the alternate pattern observed in the same region for finger 2 is retained as well. Both P131 and P133 appear to remain *trans* on the basis of beta and gamma carbon shifts.

That the zinc fingers do not alter structure upon DNA binding was expected from comparison of previously determined structures of free and bound fingers. For example, the solution NMR structure calculated for single-finger ADR1b displays close similarity to the fingers of Zif268 bound to DNA, with main-chain RMSDs of 0.61 and 0.66 Å between ADR1b and Zif268 fingers 2 and 3, respectively. Similar RMSD values of 0.66 and 0.64 Å are observed between the Zif268 fingers and the PAPA structure (Bernstein et al., 1994). These values are comparable to RMSDs observed among the Zif268 fingers themselves. The $^{13}C_{\alpha}$ secondary shift profile of the ADR1z zinc fingers both free and bound to DNA supports the notion that the fingers are minimally perturbed upon DNA binding.

¹³C chemical shifts of bound ADR1z: N-terminus

The greatest $^{13}C_{\alpha}$ chemical shift perturbations upon DNA binding are in the N-terminal flanking sequence of ADR1z (Fig. 5b). Residues from E88 through R95 each experience a 1 to 2 ppm perturbation. In contrast, residues N-terminal to K85 retain random coil $^{13}C_{\alpha}$ values. These are exactly the residues found to be unnecessary for high-affinity binding in mutagenesis experiments (Thukral et al., 1991a).

²As the random coil $^{13}C_{\alpha}$ chemical shift of histidine coordinated to a zinc ion is unknown, the value used in this study is that of a protonated histidine, to match the published analysis of TFIIIA.

No pattern suggestive of entirely regular secondary structure is apparent in the $^{13}\text{C}_\alpha$ shifts of the perturbed N-terminal region. Although it appears that the region is no longer random coil in the presence of DNA, the exact nature of its structure is difficult to predict from chemical shift, as this measure may be sensitive not only to secondary structure, but also to proximity to the shift-inducing DNA bases themselves. A definitive description of the secondary structure of the N-terminus in DNA-bound ADR1z must therefore await a detailed analysis of NOE and coupling constant information that is currently underway.

Chemical shift perturbations upon DNA binding: Zinc fingers

Although the backbone structure of the zinc fingers does not appear to change significantly upon DNA binding, contact with DNA is expected to induce perturbations in ^1H chemical shifts at the intermolecular interface. In fact, a large number of resonances are affected in the presence of DNA; perturbations of NH, αH , and βH are shown in Figure 6. Although most chemical shift perturbation mapping to date has been performed using backbone amide resonances, which are easily measured, side-chain resonance perturbations should be the best indicator of intermolecular interactions. Figure 7 presents a summary of side-chain perturbations, with each residue placed in a category either representing no significant perturbation (blue), with no side chain or alpha proton resonance experiencing a change of greater than 0.1 ppm; severe perturbation (red), with an average side-chain perturbation of at least 0.2 ppm or an individual resonance perturbation of at least 0.3 ppm; or moderate perturbation (green), for all cases in-between.

Among the most severely perturbed side chains in the zinc fingers are those of R115, H118, R121, and R143. Each of these positions has been implicated in direct contact to DNA bases. In several X-ray crystal structures of zinc finger–DNA complexes available, the majority of base-specific contacts are provided by the side chains of residues at positions –1, 3, and 6 of the α -helix (Pavletich & Pabo, 1991; Fairall et al., 1993; Kim & Berg, 1996). In ADR1, these positions correspond to residues 115, 118, and 121 of finger 1 and residues 143, 146, and 149 of finger 2. Mutation of ADR1 residues 115, 118, 121, or 143 to alanine results in dramatically reduced or undetectable DNA binding in EMSA studies and loss of transactivation *in vivo* (Thukral et al., 1991b). Furthermore, mutations at positions 115, 118, and 121 can alter the binding-site preference of ADR1, suggesting that these residues provide direct base contacts (Thukral et al., 1992). The ^1H chemical shift perturbation map of ADR1z lends further support to the idea that the side chains of R115, H118, R121, and R143 make close approach to the DNA at the intermolecular interface.

Additional evidence for the behavior of H118 is available from ^{15}N -HMQC experiments, which allow observation of the protonation state and chemical shift of the N^δ and N^ϵ positions of imidazole rings (Pelton et al., 1993). In these spectra H118 is easily distinguished from the other histidines in ADR1z: H122, H126, H150, and H155 are used to coordinate zinc and thus have unusual spectral properties (Fig. 8). Each of these ligands carries a single proton on the N^δ position; this tautomeric state is uncommon but it is consistent with the use of N^ϵ as a zinc ligand. Moreover, the N^ϵ chemical shifts of the ligand residues are atypical, ranging from 212 to 220 ppm; similar values of around 211 ppm have been observed in stromelysin for histidines that bind zinc (Gooley et al., 1993). In contrast, H118 bears a single proton at its N^ϵ position

Fig. 6. Chemical shift perturbations of ADR1z upon binding to DNA. **A:** Differences for NH resonances; **B:** differences for H_α resonances; **C:** differences for H_β resonances. A downfield shift upon DNA binding yields a positive value in these graphs. Positions where specific contacts to DNA are expected are marked by asterisks.

and the chemical shifts of the protonated and unprotonated nitrogens are a more typical 170 and 241 ppm, respectively, in free ADR1z.

Upon DNA binding, the nitrogen chemical shifts of the H118 side chain are severely perturbed. The unprotonated N^δ resonance shifts upfield by 14 ppm, while the protonated N^ϵ shifts downfield by 11 ppm. These large chemical shift changes are consistent with a role for H118 in contacting DNA. In particular, nitrogens that serve as hydrogen-bond donors have been shown to experience downfield perturbations of around 10 ppm (Bachovchin, 1986); this is similar to the change seen for the protonated N^ϵ of H118 upon DNA binding, suggesting that this position may act as a hydrogen-bond donor.

Fig. 7. Perturbations to ADR1z ^1H chemical shifts upon DNA binding. Residues are classified as severely perturbed (red), moderately perturbed (green), or unperturbed (blue). See text for definitions of categories. **A:** Classifications are shown as colored circles under the sequence of ADR1z. Brackets indicate the extent of the shortest sequence that binds DNA with high affinity and the longest sequence that fails to bind DNA. The site of the suppressor mutation, R91K, is circled. Conserved residues are colored orange in the sequence. Arrows mark non-conserved residues where mutations deleterious to DNA binding have been identified. Those arrows that are red mark the sites of suspected base-specific contacts. **B:** The same data are shown on backbone traces of a zinc finger, with a comparison to direct DNA contacts made by Zif268. **C, D:** The same perturbation data are displayed on a three-dimensional model of the fingers of ADR1 bound to DNA. The model was generated by substituting fingers 2 and 3 of the Zif268 co-crystal structure with the analogous sequence of ADR1. DNA is colored cyan in panels C and D. Structures are rendered using the MOLSCRIPT program (Kraulis, 1991).

The imidazole ring of H122 is also perturbed upon DNA binding, although not as severely as that of H118 (Fig. 8).³ The H122 N^δ , N^ϵ , C^δH , and $\text{C}^\epsilon\text{H}$ resonances experience perturbations of 3.3, 1.0, 0.35, and 0.03 ppm, respectively. A role for this residue in DNA binding is likely; in many zinc finger–DNA interactions observed in co-crystal structures, ligand histidines analogous to H122 contact the DNA backbone through the imidazole N^δ (Pavletich & Pabo, 1991, 1993; Fairall et al., 1993; Kim & Berg, 1996). The only other conserved DNA backbone contact in these structures is made by a basic residue analogous to R111. Figure 7B shows that

both H122 and R111 are among the ADR1z finger 1 residues whose side chains are severely perturbed upon DNA binding.

Missing canonical contacts

Residues L146 and R149 of ADR1 fall at canonical base-contacting positions 3 and 6 of the helix of finger 2. The resonances of L146 are only moderately perturbed and those of R149 are unaffected by DNA binding, however, unlike analogous positions in finger 1 which are severely perturbed (Fig. 7). Mutation of either of these residues to alanine caused little or no change in the DNA-binding affinity of ADR1 (Thukral et al., 1991b). Thus, NMR data support mutagenesis results suggesting that finger 2 uses only its -1 position as a specific DNA contact. Our NMR data also suggest a lack of canonical DNA backbone contacts by finger 2. Chemical shifts of the H150 and R139 residues are unperturbed, although analogous residues provide DNA backbone contacts in other zinc finger–DNA structures (Fig. 7, 8).

³ Among the zinc ligands, only the tautomer pattern for H122 appears to be severely perturbed upon DNA binding. While the patterns for bound ADR1z are not formally assigned, it seems clear that the imidazole rings of H150 and H155 experience no significant chemical shift perturbation, whereas that of H126 is only slightly perturbed from the free values (see Fig. 8). This leaves one tautomer pattern unaccounted for in bound ADR1z, which must belong to H122.

Fig. 8. ADR1z ^{15}N -HMQC spectra. Panel A is a schematic of tautomer patterns expected for two protonation states of a typical histidine imidazole ring. Only the $^2J_{\text{NH}}$ couplings are large enough to yield strong crosspeaks. Panel B and panel C are ^{15}N -HMQC spectra of free ADR1z (pH 5.5) and DNA-bound ADR1z (pH 7.5), respectively. Panel D is an overlap emphasizing the tautomer patterns of the zinc ligands of ADR1z. The spectrum of free ADR1z at pH 7.2 is shown in pink and that of DNA-bound ADR1z at pH 7.5 in blue. The peaks for DNA-bound H118 are very weak; the spectrum in panel C is plotted at a very low level to display this pattern.

Differences in the protein–DNA interfaces of finger 1 and finger 2 are apparent throughout the C-terminal portions of the fingers (Fig. 6, 7). In finger 1, there are severe side-chain perturbations at residues 115, 118, 121, and 122, and moderate perturbations at most other positions in the helix, all the way through the last histidine ligand and into the linker. This would seem to indicate full involvement of the helical region of finger 1 with DNA, in line with use of all three canonical positions to make specific base contacts. In contrast, finger 2 experiences severe chemical shift perturbation only for the predicted base contact R143, and moderate perturbations only from R144 through I148. No perturbations are seen for the remainder of the finger, including the unused canonical +6 contact position, R149, and the unused canonical backbone contact position, H150. In contrast to finger 1, it seems that only a portion of the helix of finger 2 is in proximity to the DNA, and that close approach may be made only by the extreme N-terminus of the helix. Thus, unlike the similar dispositions of the three fingers of Zif268, each of which makes two or more specific contacts to DNA (Pavletich & Pabo, 1991; Elrod-Erickson et al., 1996), the chemical shift data presented here suggest that the fingers of ADR1 approach the DNA with different orientations.

Unperturbed portions of fingers

In the structures of zinc finger–DNA complexes solved thus far, contacts to the DNA arise mostly from the face of the α -helix, while the Cys–Cys loop lies on the “back” of the helix, away from the DNA (Pavletich & Pabo, 1991, 1993; Fairall et al., 1993). The chemical shift perturbation map displays this trend for ADR1, with no perturbation to the side chains of the Cys–Cys loop of finger 2 and the reverse turn of finger 1 (Fig. 7).

The severe perturbations that do occur in the β -sheet of finger 1 are likely fully explained by a limited extension of interactions

across the Cys–Cys loop. Preliminary analysis of NOE data indicates that S103 is juxtaposed to A114 in DNA-bound ADR1z, as NOE crosspeaks occur between S103 and A114, between S103 and F113, and between R102 and A114 (data not shown). In the free protein, S103 is likely near A112 instead; this is where it was found in the single-finger ADR1b peptide, which has chemical shifts nearly identical to free ADR1z (Hoffman et al., 1993). A change in S103 position upon DNA binding may thus be responsible for the severe chemical shift perturbations observed for R102, S103, A112, F113, and A114, including the extremely large downfield perturbation of the A114 NH resonance. Furthermore, juxtaposition of S103 to A114 may be important for the proper orientation of the N-terminus in the protein–DNA complex. Indeed, the mutation A114V is deleterious to ADR1 activity (Blumberg et al., 1987); this effect may be due to the necessity of an S103/A114 interaction.

Otherwise, the lack of chemical shift perturbation in exposed portions of the fingers suggests that the severely perturbed N-terminus does not interact with these regions of the protein. It is formally possible that the effect of the N-terminus on DNA-binding affinity is mediated through interaction with the fingers themselves. Such an interaction might allow the fingers to fold properly, as was the case with the third strand of β -sheet in SWI5 and Tramtrack (Neuhauser et al., 1992; Fairall et al., 1993), or it might increase the affinity of the fingers for DNA. These possibilities seemed unlikely, given that each finger folds in the absence of flanking sequences to a structure very much like the active binding form observed in zinc finger–DNA complexes (Hoffman et al., 1993; Bernstein et al., 1994), and given that, in the unbound form, the random coil N-terminal flanking sequence does not interact with the fingers. The ^1H chemical shift map of bound ADR1z also tends to rule out an interaction of the N-terminus with the

exposed portions of the fingers in the context of the complex. Thus, the most likely possibility is that the role of the N-terminus is to interact with the DNA itself and thereby increase binding affinity.

N-terminus

The ^1H chemical shift perturbation map reflects the importance of the N-terminus from E88 onward, as most residues following this critical point are moderately or severely shifted, while residues preceding this juncture are not perturbed (Fig. 6, 7). Large perturbations to NH and αH chemical shifts are found throughout the N-terminus after residue E88; these could result from the acquisition of secondary structure from random coil and/or proximity to DNA. In this region, side-chain resonances are also severely perturbed. As side chains are less liable to perturbation due to acquisition of secondary structure than are backbone resonances, these changes likely indicate involvement in the binding interface.

In addition to the abundant ^1H chemical shift changes in the N-terminus, the side-chain $^{15}\text{N}^\delta$ of N89 is perturbed out of the usual range of chemical shifts for side-chain NH_2s , by a downfield shift of ~ 2 ppm. The perturbed peaks are less sharp than those of other side-chain NH_2s . Examination of HSQC spectra collected at high pH also indicates that one of the side-chain amide protons is less liable to exchange with solvent than the other, suggesting protection from solvent.

N89 and the set of highly perturbed side chains extending from L92 to T96 fall in exactly the region of the N-terminus where mutations have been observed to affect DNA binding. Mutation of P87 or P97 to leucine causes a reduction in DNA-binding affinity (Thukral et al., 1991a). The mutation R91K suppresses defects in DNA binding arising from mutations in either zinc finger and increases DNA binding by two- to three-fold in a wild-type background (Camier et al., 1992). The R91 side chain does not experience a significant chemical shift perturbation upon DNA binding, however, suggesting that this residue is not involved in DNA contact in the wild-type construct. Mutation to lysine likely allows the formation of a novel contact that contributes to the overall binding energy. Residues N-terminal to D84, which exhibit random coil $^{13}\text{C}_\alpha$ shifts and unperturbed ^1H chemical shifts, have been identified as non-essential to DNA binding by deletion mutagenesis (Thukral et al., 1991a). Thus chemical shift perturbation maps and mutagenesis data identify the region from P87 to P97 as critical to the DNA-binding contribution of the N-terminus.

Perturbations to the DNA site upon ADR1z binding

There is often a one-to-one correspondence of contacts between the -1, 3, and 6 positions of the α -helix of a zinc finger and a three-base-pair subsite in its target DNA (Pavletich & Pabo, 1991; Fairall et al., 1993; Kim & Berg, 1996). Thus ADR1 finger 1 is predicted to bind 3'-GAG-5' in the UAS1 half-site, using R115, H118, and R121. Finger 2 is likewise predicted to bind 3'-GTT-5' with R143, L146, and R149 (Klevit, 1991). Mutagenesis of the DNA site has confirmed the importance of the 3'-GAGG-5' sequence (Cheng et al., 1994), in agreement with the demonstrated importance of residues 115, 118, 121, and 143. In addition, changes to the preceding and following A/T base pairs affect DNA binding to a lesser extent (see Fig. 9).

To explore effects of ADR1z binding on DNA, imino proton resonances were assigned for the 14mer DNA both free and in

Fig. 9. Perturbations to 14mer DNA upon ADR1z binding. Perturbations of the chemical shifts of imino protons are plotted. The sequence of the 14mer DNA is shown underneath, with the expected binding sites for finger 1 and finger 2 underlined. Expected side chain-base contacts are indicated. Asterisks mark positions where ADR1 exhibits a base preference; small asterisks indicate a weak preference (Cheng et al., 1994).

complex with ADR1z, using conventional ^1H imino-imino NOESY connectivities. Chemical shift perturbations to the imino protons upon ADR1z binding are mapped in Figure 9. Only a few imino protons of the DNA site show significant chemical shift changes upon ADR1z binding. These could result from contacts by ADR1z, or from perturbation of the DNA structure induced by protein binding.

Consistent with the notion of extensive contacts from finger 1 and very limited contacts from finger 2, the major changes in the DNA upon binding occur in the 3'-GAG-5' binding site predicted for finger 1. Additionally, perturbation of the upstream T confirms that this base pair, while not predicted to make contact with the finger based on canonical rules, nevertheless plays a role in the site, as suggested by mutagenesis (Cheng et al., 1994) and methylation interference footprinting (Thukral et al., 1991a). The lack of perturbation for the 5' T of the predicted 3'-GTT-5' subsite for finger 2 is consistent with the earlier conclusion that R149 does not contact the DNA. Finally, the largest chemical shift perturbation, a downfield shift of 0.6 ppm, occurs at the T in the A/T base pair predicted to be contacted by H118. This relatively large shift may be due to a ring current effect of the aromatic histidine side chain. Were the histidine ring to contact the A/T base pair edge-on, the imino proton would fall in the downfield shift profile of the imidazole ring. Thus the overall imino chemical shift pattern is consistent with the picture that emerges from protein chemical shifts in which finger 1 is more intimately involved in contacting the DNA than is finger 2.

Conclusions

The chemical shift assignments presented here represent the first NMR study of the minimal ADR1 DNA-binding domain, including both zinc fingers and an additional 20 residues required for high-affinity DNA binding, as well as the second set of assignments reported for a zinc finger-DNA complex. The results clearly demonstrate that the zinc fingers fold to the $\beta\beta\alpha$ fold found in

structures of other zinc fingers and that the fingers do not interact with each other or with the random-coil N-terminal portion of the protein in the absence of DNA. The NMR data presented here suggest that the zinc fingers of ADR1 do not change structure upon DNA binding, with the exception of a limited extension of interactions across the Cys-Cys loop of finger 1. In contrast, the N-terminus becomes structured in the presence of DNA. Its $^{13}\text{C}_\alpha$ chemical shifts are not clearly consistent with the formation of regular secondary structure, however, including α -helix. This view is supported by a preliminary analysis of NOE data, which reveals no regular patterns of short-range sequential NOEs in the N-terminus. The behavior of the ADR1z N-terminus may thus contrast with that of GAGA and SWI5, in which residues adjacent to the N-terminal zinc fingers adopt a helical fold (Dutnall et al., 1996; Omichinski et al., 1997).

The dramatic chemical shift perturbations of the N-terminus upon DNA binding are remarkably consistent with mutagenesis data that previously defined the essential N-terminal region. ^1H and ^{13}C chemical shift perturbations become significant starting at residue 86, which falls at the border of the sequence required for high-affinity binding (Thukral et al., 1991a). Furthermore, the residues most perturbed lie between P87 and P97, flanking the site of the affinity-enhancing suppressor mutation R91K (Camier et al., 1992). The severely perturbed N-terminus does not appear to interact with the exposed, unperturbed portions of the zinc fingers in the protein-DNA complex. These data taken together suggest that the region P87-P97 plays a role in contacting DNA and thereby increases the binding affinity of the construct.

There are other proteins known to utilize an additional region outside of a well-defined motif to supplement DNA binding. These include λ repressor (Jordan & Pabo, 1988), the homeodomain proteins (Kissinger et al., 1990), GATA-1 (Omichinski et al., 1993), winged helix-turn-helix proteins (Clark et al., 1993), and the LEF-1 HMG-domain protein (Love et al., 1995). In each of these cases, accessory residues make direct contact with DNA to increase binding affinity. The binding affinity of the SWI5 Cys₂-His₂ zinc finger protein is also increased by additional residues N-terminal to the three-strand β -sheet of finger 1 (Dutnall et al., 1996). The N-terminal sequence adjacent to the *Drosophila* GAGA zinc finger makes specific contacts to the major and minor grooves of the DNA, extending the site recognized by the protein (Omichinski et al., 1997). It seems probable that more zinc finger proteins will be found that use non-consensus sequences to increase DNA binding, in analogy to other recognized motifs.

The pattern of ADR1z side-chain chemical shift perturbations presented here is consistent with the picture from mutagenesis studies that finger 1 uses all three canonical DNA-contacting positions of its helix to bind UAS1, while finger 2 uses only the first (-1) position. The perturbation map further suggests that the two fingers approach DNA in different orientations, with the α -helix of finger 1 fully involved in contacting the binding site and only the extreme N-terminus of the helix of finger 2 approaching the DNA. This pattern is different from that of the Zif268 structure, in which all three fingers approach DNA in a similar fashion (Pavletich & Pabo, 1991). Furthermore, fingers in Zif268, GLI, and Tramtrack employ analogues of ADR1 residues R111/R139 and H122/H150 to contact the phosphodiester backbone (Pavletich & Pabo, 1991, 1993; Fairall et al., 1993). While ADR1z finger 1 residues R111 and H122 both experience chemical shift perturbations upon DNA binding, neither of the analogous finger 2 residues are perturbed significantly. Thus, in addition to minimal contacts to DNA bases,

our data suggest that finger 2 does not participate in these canonical phosphate contacts, further reducing its contribution to DNA binding.

L146, at position 3 of the α -helix, may be responsible for the unusual pattern of contact by ADR1 finger 2 (Bernstein et al., 1994; Cheng & Young, 1995). While this residue does not make an important DNA contact in the wild-type protein (Thukral et al., 1991b), the mutation L146H leads to a strong, new base preference at the UAS1 position predicted to be contacted by helical position 3. This mutation also leads to a newfound preference at the base predicted to be contacted by helical position 6, R149 (Cheng & Young, 1995). The wild-type leucine at position 3 of the helix apparently prevents interaction of R149 with the DNA; this interference can be removed by altering L146.

The importance of L146 was anticipated by a modeling study in which the structure of the ADR1 finger 2 PAPA mutant was superimposed on finger 3 of the Zif268 cocrystal structure (Bernstein et al., 1994). The resulting model predicted a steric clash between L146 and the DNA, suggesting that the helix of finger 2 does not approach the major groove in the Zif268 orientation. This model is strongly supported by the ^1H chemical shift perturbation map of ADR1z-DNA.

The idea that some zinc fingers can make use of their entire helix to contact DNA while others cannot has been put forward in a general way by Suzuki et al. (Suzuki et al., 1994). They categorized Cys₂-His₂ fingers as either "good" (likely to use the whole helix for contacts) or "bad" (unable to do so), based on the nature of residues near the N-terminal end of the helix. It was suggested that an abundance of short residues in this area would make for a "good" finger, while longer or bulky residues in this area would create a "bad" finger that would make suboptimal contacts. By this criterion, L146 may render finger 2 of ADR1 a "bad" finger; data from two separate techniques, mutagenesis and NMR spectroscopy, independently demonstrate the minimal contact made by finger 2.

Most Cys₂-His₂ zinc fingers studied thus far use at least two base-specific contacts per finger to bind DNA, and most zinc finger proteins contain arrays of three or more fingers. In ADR1, a suboptimal one-contact finger falls in the context of a rare two-finger array. This anomaly likely explains the importance of finger 1 making use of three canonical DNA contacts. Mutagenesis has shown that loss of any of those contacts seriously compromises DNA binding. In addition, the suboptimal finger in a two-finger context likely explains the requirement for N-terminal DNA-contacting sequences. In this sense, ADR1z is similar to the GAGA DNA-binding domain. In that system only a single zinc finger is present; DNA contacts from the N-terminus compensate for the lack of additional fingers, and the lone finger makes specific contacts using three canonical positions (Omichinski et al., 1997).

We now have evidence from the striking agreement of chemical shift maps and mutagenesis data that one way ADR1 overcomes the handicap of very limited DNA contact by finger 2 is through the use of three contacts from finger 1 and an unusual N-terminal DNA-contacting sequence. Further structural studies underway should reveal the precise role and disposition of the N-terminus, as well as the precise orientation of each of the fingers.

Materials and methods

Subcloning

The ADR1z construct extends from codon 75 to 161 of ADR1. An engineered *Nde*I site took advantage of the native ATG at codon

75, with a change at the preceding codon from AAT to CAT. An in-frame TAA stop codon was introduced immediately following codon 161, followed by a *Bam*HI site. Overexpression in *E. coli* was achieved using the T7 polymerase system (Studier et al., 1990). The ADR1z construct was introduced into plasmid pET11a (Novagen) between the unique *Nde*I and *Bam*HI sites, and propagated in *E. coli* strain BL21(DE3). In this system, the translation start codon is the ATG included in the *Nde*I site, and therefore the resulting polypeptide contains native sequence only.

Overexpression and purification

Cell cultures were grown at 37 °C on minimal medium containing 1 g/L ammonium chloride and 4 g/L glucose as the sole nitrogen and carbon sources, respectively. In cases of uniform ¹⁵N and/or ¹³C labeling, isotopically labeled material (Isotec, Cambridge Isotope Labs) was substituted. ¹³C-glucose was used at half the normal level. Specific ¹⁵N-labeling experiments used 80 mg/L ¹⁵N-leucine or ¹⁵N-glutamate (CIL) just prior to induction, along with all other unlabeled amino acids added initially to twice the level suggested (Davis et al., 1980). Specific ¹³C-labeling experiments used 100 mg/L 1-¹³C-leucine, 1-¹³C-lysine, or ¹³C₆-arginine (CIL), added just prior to induction to a medium uniformly labeled with ¹⁵N (Matsuo et al., 1991).

Zinc chloride was added to the cultures at an OD₆₀₀ of 0.25 to a final concentration of 200 μM. IPTG was added to 1 mM at an OD₆₀₀ of 0.5. Induction continued for 10 h, at which point the cells were spun from the medium. Cell extracts obtained by French Press were treated with DNAase, spun at 10 K, and applied to a CM Sephadex C-25 (Pharmacia) cation exchange column at 4 °C in a buffer containing 50 mM HEPES pH 7.6, 5 mM DTT, and 0.1 mM EGTA. ADR1z was eluted with a 0.2–2 M NaCl gradient. The polypeptide was dialyzed against 0.1% acetic acid, lyophilized, reduced with 30 mM DTT, and then applied to a C4 reverse phase HPLC column (Waters) in 29% acetonitrile/0.1% TFA. A 29–35% acetonitrile gradient was used to elute ADR1z, at approximately 98–99% purity as judged by SDS-PAGE.

The molecular mass of purified ¹⁵N-labeled polypeptide was measured as 10,335 and 10,203 Da by electrospray mass spectrometry, as compared to the expected mass of 10,334.7 Da, indicating partial post-translational removal of the N-terminal methionine. Final yields averaged 0.5 μmol ADR1z per L culture.

Oligonucleotide preparation

DNA oligonucleotides were synthesized by the solid-state phosphoramidite method. Single strands were purified by size exclusion chromatography on Sephadex G50 superfine resin (Pharmacia) and then annealed in 400 mM NaCl by slow-cooling from 82 °C. Single-stranded DNA was separated from double-stranded products of the annealing reaction on Biogel HTP hydroxylapatite (Biorad).

NMR samples

ADR1z samples were typically 2 mM protein in 25 mM deuterated tris/acetate buffer containing 5 mM DTT and 40 μM EDTA. Samples prepared in H₂O contained 10% D₂O. The samples were reconstituted in the presence of two equivalents of zinc chloride by raising the pH above the pK_a of the ligand histidines. Free protein samples for NMR study were typically prepared at pH 5.4 to

prevent oxidation of the sole non-ligand cysteine while avoiding protonation of the histidine zinc ligands. This pH also reduces exchange with solvent. Samples intended for addition to DNA were prepared at pH 7 due to the improved solubility characteristics of the resulting complex at this pH.

ADR1z-DNA samples were prepared by slow addition of a free ADR1z sample prepared as above to 1.1 equivalents of 14mer DNA in the same buffer system at room temperature. All protein-DNA samples were maintained at or near pH 7, due to poor solubility at lower pH. The resulting samples ranged from 1.5 to 2 mM in final ADR1z-DNA concentration.

NMR experiments

NMR experiments were performed on a Bruker DMX500 spectrometer equipped with a 5 mM triple-axis pulsed field gradient three-channel Bruker probe. Experiments on the free protein were generally conducted at 30 °C, while those on the protein-DNA complex were done at 37 °C. The higher temperature was chosen to sharpen linewidths for the 20 kD complex; temperatures for the free protein were lowered to minimize oxidation of the non-ligand cysteine at residue 140.

¹H chemical shifts are reported relative to TSP at 0 ppm for each of the buffer/temperature combinations used. ¹⁵N chemical shifts are referenced relative to the ¹H zero ppm frequency (Live et al., 1984). ¹³C chemical shifts are referenced to dioxane at 67.8 ppm for comparison with reported random coil values (Richarz & Wüthrich, 1978).

2D ¹H-¹⁵N HSQC correlation spectra were acquired using gradients for coherence selection and artifact suppression (Kay et al., 1992). Sixteen transients were collected per FID. Typically, the spectral width in the acquired dimension was 6410 Hz, and 450 increments were collected in the ¹⁵N dimension, with a spectral width of 2000 Hz. In these and other experiments the recycle delay was generally 1.2 s, and the ¹⁵N carrier was set to 116 ppm.

Constant time HNCA and HN(CO)CA triple-resonance experiments (Grzesiek & Bax, 1992) were performed on uniformly ¹⁵N,¹³C-labeled ADR1z, free and bound to DNA, using the WATERGATE sequence for water suppression. Quadrature detection in t₁ was obtained by TPPI. Spectral widths and the number of complex points acquired were 3521 Hz and 24 points in ¹³C, 2000 Hz and 40 points in ¹⁵N, and 4006 Hz and 512 points in ¹H. Thirty-two transients were collected per FID for free ADR1z and 64 per FID for DNA-bound ADR1z. The ¹H carrier was set to 8.6 ppm and the ¹³C carrier to 52 ppm. Triple-resonance CBCA(CO)NH and HNCACB experiments (Muhandiram & Kay, 1994) were also performed on DNA-bound ADR1z. Gradients were used for coherence selection and artifact suppression. Quadrature detection in the indirectly detected dimensions was achieved by the States method. In these experiments a spectral width of 7000 Hz with 30 complex points was used for the ¹³C dimension, and the ¹³C carrier was set to 43 ppm. Thirty-two transients per FID were collected.

HCCH-TOCSY experiments (Bax et al., 1990; Kay et al., 1993) were acquired in D₂O on uniformly ¹⁵N,¹³C-labeled ADR1z samples. Spectra of DNA-bound ADR1z were collected using isotropic mixing times of 17.9 ms and 25 ms; the spectrum of free ADR1z was obtained with a 19.6 ms mixing time. Sixteen transients were collected per FID. Gradients were used for artifact suppression, and quadrature in the indirectly detected dimensions was obtained using the States method. The spectral widths and the number of

complex points acquired were 3501 Hz and 118 points in ^1H , 2641 Hz and 29 points in ^{13}C , and 5000 Hz and 512 points in the acquisition dimension. The carrier was placed at 43 ppm in the ^{13}C dimension and at 3 ppm in both ^1H dimensions.

Three-dimensional ^{15}N -edited NOESY-HSQC experiments were acquired with mixing times of 150 ms for free ADR1z and 75 ms for the ADR1z–DNA complex, using the WATERGATE sequence for suppression of the water signal (Sklenár et al., 1993). Quadrature detection in t_1 was obtained by the States method. Standard homonuclear ^1H NOESY spectra in H_2O were also collected on both free and ADR1z-bound 14mer DNA for assignment of the DNA imino protons, with a spectral width of 10500 Hz. 2D and 3D ^{15}N -edited HSQC-TOCSY spectra were recorded with isotropic mixing times of 76 ms for free ADR1z and 31 ms for the ADR1z–DNA complex. In addition, a homonuclear ^1H TOCSY experiment with a 72-ms mixing time was recorded in D_2O for the assignment of aromatic side chains in free ADR1z.

Two-dimensional magnitude ^{15}N -HMQC experiments using gradients for coherence selection were used to determine the tautomeric forms of histidine side chains (Pelton et al., 1993). The ^{15}N spectral width was 7645 Hz, and 96 points in the ^{15}N dimension were collected. The ^{15}N carrier was placed at 200 ppm. Spectra of free ADR1z were acquired with 64 transients; the spectrum of DNA-bound ADR1z was acquired with 1024 transients. The value of $1/2J_{\text{HN}}$ was taken as 22 ms.

NMR experiments were processed using FELIX95 (Biosym Technologies). Data in the indirectly detected dimensions were typically extended 50% by linear prediction. FIDs were weighted by a squared sine-bell window shifted by $\pi/2$ prior to Fourier transform.

Electronic supplementary material

^1H , ^{15}N , and ^{13}C NMR assignments for ADR1z both free and bound to DNA are provided as electronic supplementary material in the files titled “free ADR1z assignment table” and “bound ADR1z assignment table.”

Acknowledgments

We thank Dr. Elton T. Young and his colleagues for providing information on ADR1 mutants, for providing starting materials for the ADR1z subclone, and for performing a DNA-binding assay on ADR1z. We are indebted to Ms. Beverly Castner for preparing a number of the protein samples used. We thank Dr. Lewis Kay and Dr. Walter Chazin for providing some of the pulse programs. We thank Dr. Jon Herriott for his long-standing interest in this project and Dr. Elinor Adman for her thoughtful comments on this work. This work was supported by NIH Program Project P01 GM32681, by NIH RO1 GM55369, and by the Murdock Charitable Trust. M.S. was supported by an NSF Predoctoral Fellowship and by an NIH predoctoral traineeship in Molecular Biophysics, T32 GM08268.

References

- Bachovchin WW. 1986. ^{15}N NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: Evidence for a moving histidine mechanism. *Biochemistry* 25:7751–7759.
- Bax A, Clore GM, Gronenborn AM. 1990. ^1H – ^1H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C -enriched proteins. *J Magn Reson* 88:425–431.
- Berg JM. 1992. Sp1 and the subfamily of zinc finger proteins with guanine-rich binding sites. *Proc Natl Acad Sci USA* 89:11109–11110.
- Berg JM. 1993. Zinc-finger proteins. *Curr Opin Struct Biol* 3:11–16.
- Bernstein BE, Hoffman RC, Horvath S, Herriott JR, Klevit RE. 1994. Structure of a histidine–X₄–histidine zinc finger domain: Insights into ADR1–UAS1 protein–DNA recognition. *Biochemistry* 33:4460–4470.
- Blumberg H, Eisen A, Sledziewski A, Bader D, Young ET. 1987. Two zinc fingers of a yeast regulatory protein shown by genetic evidence to be essential for its function. *Nature* 328:443–445.
- Brown RS, Sander C, Argos P. 1985. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. *FEBS Lett* 186:271–274.
- Camier S, Kacharovsky N, Young ET. 1992. A mutation outside the two zinc fingers of ADR1 can suppress defects in either finger. *Mol Cell Biol* 12:5758–5767.
- Cheng C, Kacharovsky N, Dombek KM, Camier S, Thukral SK, Rhim E, Young ET. 1994. Identification of potential target genes for ADR1p through characterization of essential nucleotides in UAS1. *Mol Cell Biol* 14:3842–3852.
- Cheng C, Young ET. 1995. A single amino acid substitution in zinc finger 2 of ADR1p changes its binding specificity at two positions in UAS1. *J Mol Biol* 251:1–8.
- Clark KL, Halay ED, Lai E, Burley SK. 1993. Co-crystal structure of the HNF-3/forkhead DNA-recognition motif resembles histone H5. *Nature* 364:412–420.
- Cook WJ, Mosley SP, Audino DC, Mullaney DL, Rovelli A, Stewart G, Denis CL. 1994. Mutations in the zinc-finger region of the yeast regulatory protein ADR1 affect both DNA binding and transcriptional activation. *J Biol Chem* 269:9374–9379.
- Davis RW, Botstein D, Roth JR. 1980. *A manual for genetic engineering: Advanced bacterial genetics*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.
- Denis CL, Young ET. 1983. Isolation and characterization of the positive regulatory gene ADR1 from *Saccharomyces cerevisiae*. *Mol Cell Biol* 3:360–370.
- Dutnall RN, Neuhaus D, Rhodes D. 1996. The solution structure of the first zinc finger domain of SWI5: A novel structural extension to a common fold. *Structure* 4:599–611.
- Eisen A, Taylor WE, Blumberg H, Young ET. 1988. The yeast regulatory protein ADR1 binds in a zinc-dependent manner to the upstream activating sequence of ADH2. *Mol Cell Biol* 8:4552–4556.
- Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO. 1996. Zif268 protein–DNA complex refined at 1.6 Å: A model system for understanding zinc finger–DNA interactions. *Structure* 4:1171–1180.
- Fairall L, Schwabe JWR, Chapman L, Finch JT, Rhodes D. 1993. The crystal structure of a two-zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature* 366:483–487.
- Gooley PR, Johnson BA, Marcy AI, Cuca GC, Salowe SP, Hagmann WK, Esser CK, Springer JP. 1993. Secondary structure and zinc ligation of human recombinant short-form stromelysin by multidimensional heteronuclear NMR. *Biochemistry* 32:13098–13108.
- Grzesiek S, Bax A. 1992. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* 96:432–440.
- Hartshorne TA, Blumberg H, Young ET. 1986. Sequence homology of the yeast regulatory protein ADR1 with *Xenopus* transcription factor TFIIIA. *Nature* 320:283–287.
- Hoffman RC, Horvath SJ, Klevit RE. 1993. Structures of DNA-binding mutant zinc finger domains: Implications for DNA binding. *Protein Sci* 2:951–965.
- Jacobs GH. 1992. Determination of the base recognition positions of zinc fingers from sequence analysis. *EMBO J* 11:4507–4517.
- Jordan SR, Pabo CO. 1988. Structure of the lambda complex at 2.5 Å resolution: Details of the repressor–operator interactions. *Science* 242:893–899.
- Kaptein R. 1992. Zinc-finger structures. *Curr Opin Struct Biol* 2:109–115.
- Kay LE, Keifer P, Saarinen T. 1992. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* 114:10663–10665.
- Kay LE, Xu G-Y, Singer AU, Muhandiram DR, Forman-Kay JD. 1993. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^1H and ^{13}C correlations in H_2O samples of proteins. *J Magn Reson Ser B* 101:333–337.
- Kim CA, Berg JM. 1996. A 2.2 Å resolution crystal structure of a designed zinc finger protein bound to DNA. *Nature Struct Biol* 3:940–945.
- Kissinger CR, Liu B, Martin-Blanco E, Kornberg TB, Pabo CO. 1990. Crystal structure of an engrailed homeodomain–DNA complex at 2.8 Å resolution: A framework for understanding homeodomain–DNA interactions. *Cell* 63:579–590.
- Klevit RE. 1991. Recognition of DNA by Cys₂His₂ zinc fingers. *Science* 253:1367.
- Klevit RE, Herriott JR, Horvath SJ. 1990. Solution structure of a zinc-finger domain of yeast ADR1. *Proteins: Struct Funct Genet* 7:215–226.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Liao X, Clemens K, Cavanagh J, Tennant L, Wright PE. 1994. ^1H , ^{15}N and ^{13}C resonance assignments for the first three zinc fingers of transcription factor IIIA. *J Biomol NMR* 4:433–454.
- Live DH, Deaver DG, Agosta WC, Cowburn D. 1984. Long range hydrogen bond

- mediated effects in peptides: ^{15}N NMR study of gramicidin S in water and organic solvents. *J Am Chem Soc* 106:1939–1941.
- Love JJ, Li X, Case DA, Giese K, Grosschedl R, Wright PE. 1995. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* 376:791–795.
- Matsuo H, Shirakawa M, Ohkubo T, Yamazaki T, Kyogoku Y. 1991. Assignments of ^1H - ^{15}N magnetic resonances and identification of secondary structure elements of the λ -cro repressor. *J Biomol NMR* 1:191–204.
- Miller J, McLachlan AD, Klug A. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* 4:1609–1614.
- Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW. 1989. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol* 177:44–73.
- Muhandiram DR, Kay LE. 1994. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J Magn Reson Ser B* 103:203–216.
- Nakaseko Y, Neuhaus D, Klug A, Rhodes D. 1992. Adjacent zinc-finger motifs in multiple zinc-finger peptides from SW15 form structurally independent, flexibly linked domains. *J Mol Biol* 228:619–636.
- Neuhaus D, Nakaseko Y, Schwabe JWR, Klug A. 1992. Solution structures of two zinc-finger domains from SW15 obtained using two-dimensional ^1H nuclear magnetic resonance spectroscopy. A zinc-finger structure with a third strand of β -sheet. *J Mol Biol* 228:637–651.
- Omichinski JG, Clore GM, Robien M, Sakaguchi K, Appella E, Gronenborn AM. 1992. High-resolution solution structure of the double Cys₂His₂ zinc finger from the human enhancer binding protein MBP-1. *Biochemistry* 31:3907–3917.
- Omichinski JG, Clore GM, Schaad O, Felsenfeld G, Trainor C, Appella E, Stahl SJ, Gronenborn AM. 1993. NMR structure of a specific DNA complex of the Zn-containing DNA-binding domain of GATA-1. *Science* 261:438–446.
- Omichinski JG, Pedone PV, Felsenfeld G, Gronenborn AM, Clore GM. 1997. The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode. *Nature Struct Biol* 4:122–132.
- Párraga G, Horvath SJ, Eisen A, Taylor WE, Hood L, Young ET, Klevit RE. 1988. Zinc-dependent structure of a single-finger domain of yeast ADR1. *Science* 241:1489–1492.
- Pavletich NP, Pabo CO. 1991. Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252:809–817.
- Pavletich NP, Pabo CO. 1993. Crystal structure of a five-finger GLI-DNA complex: New perspectives on zinc fingers. *Science* 261:1701–1707.
- Pedone PV, Ghirlando R, Clore GM, Gronenborn AM, Felsenfeld G, Omichinski JG. 1996. The single Cys₂-His₂ zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding. *Proc Natl Acad Sci USA* 93:2822–2826.
- Pelton JG, Torchia DA, Meadow ND, Roseman S. 1993. Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated III^{Glc}, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Sci* 2:543–558.
- Richarz R, Wüthrich K. 1978. Carbon-13 NMR chemical shifts of the common amino acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers* 17:2133–2141.
- Schmiedeskamp M, Klevit RE. 1994. Zinc finger diversity. *Curr Opin Struct Biol* 4:28–35.
- Shuster J, Yu J, Cox D, Chan RVL, Smith M, Young E. 1986. ADR1-mediated regulation of ADH2 requires an inverted repeat sequence. *Mol Cell Biol* 6:1894–1902.
- Simon M, Adam G, Rapatz W, Spevak W, Ruis H. 1991. The *Saccharomyces cerevisiae* ADR1 gene is a positive regulator of transcription of genes encoding peroxisomal enzymes. *Mol Cell Biol* 11:699–704.
- Sklenář V, Piotto M, Leppik R, Saudek V. 1993. Gradient-tailored water suppression for ^1H - ^{15}N HSQC experiments optimized to retain full sensitivity. *J Magn Reson Ser A* 102:241–245.
- Spera S, Bax A. 1991. Empirical correlation between protein backbone conformation and α and β ^{13}C nuclear magnetic resonance chemical shifts. *J Am Chem Soc* 113:5490–5492.
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185:60–89.
- Suzuki M, Gerstein M, Yagi N. 1994. Stereochemical basis of DNA recognition by Zn fingers. *Nucleic Acids Res* 22:3397–3405.
- Thukral SK, Eisen A, Young ET. 1991a. Two monomers of yeast transcription factor ADR1 bind a palindromic sequence symmetrically to activate ADH2 expression. *Mol Cell Biol* 11:1566–1577.
- Thukral SK, Morrison ML, Young ET. 1991b. Alanine scanning site-directed mutagenesis of the zinc fingers of transcription factor ADR1: Residues that contact DNA and that transactivate. *Proc Natl Acad Sci USA* 88:9188–9192.
- Thukral SK, Morrison ML, Young ET. 1992. Mutations in the zinc fingers of ADR1 that change the specificity of DNA binding and transactivation. *Mol Cell Biol* 12:2784–2792.
- Thukral SK, Taviani MA, Blumberg H, Young ET. 1989. Localization of a minimal binding domain and activation regions in yeast regulatory protein ADR1. *Mol Cell Biol* 9:2360–2369.
- Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. 1995. ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. Investigations of nearest-neighbor effects. *J Biomol NMR* 5:67–81.
- Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: John Wiley & Sons, Inc.
- Xu RX, Horvath SJ, Klevit RE. 1991. ADR1a, a zinc finger peptide, exists in two folded conformations. *Biochemistry* 30:3365–3371.