Paramagnetic Cobalt as a Probe of the Orientation of an Accessory DNA-Binding Region of the Yeast ADR1 Zinc-Finger Protein†

Mia Schmiedeskamp and Rachel E. Klevit*

Biomolecular Structure Center and Department of Biochemistry, Box 357742, University of Washington, Seattle, Washington 98195-7742

Received June 9, 1997; Revised Manuscript Received August 13, 1997⊗

ABSTRACT: The minimal DNA-binding domain of the yeast ADR1 transcription factor consists of two Cys_2−His_2 zinc fingers and an additional 20 residues N-terminal and proximal to the fingers. The accessory sequence likely plays a role in contacting DNA. Paramagnetic cobalt was incorporated into the fingers of an ADR1 DNA-binding construct (ADR1z) to serve as a probe of the proximity of the accessory sequence to the zinc fingers. NMR signals from the accessory region are not perturbed by cobalt incorporation. Previous studies showed that this region is random coil in the ADR1z construct in the absence of DNA; it does not adopt a fixed orientation with respect to the cobalt sites. In contrast, many residues of the accessory region are perturbed by cobalt in the DNA-bound form of the protein, suggesting this region becomes constrained. This observation agrees with previous results showing a disorder-to-order transition for the accessory region upon DNA binding. Furthermore, these results indicate that the accessory region lies close to the fingers in the protein−DNA complex. This region thus does not extend along the DNA away from the zinc fingers; it more likely binds the same stretch of DNA contacted by the zinc fingers. Comparison to the behavior of other zinc-finger proteins that utilize an accessory DNA-binding sequence suggested that the region of ADR1 proximal to the zinc fingers might form an α-helix. Analysis of sequential NOEs in the accessory region of DNA-bound ADR1z reveals no helical structure.

The Cys_2−His_2 zinc finger is the most ubiquitous eukaryotic DNA binding motif known (Berg & Shi, 1996). The consensus sequence for this motif is (Y,F)-X-C-$X_{2,4}$-C-X_3-F-X_5-L-X_2-H-X_{3-5}-H-X_{2-6}, in which the conserved cysteine and histidine ligands coordinate a single zinc ion tetrahedrally, and the remaining conserved residues participate in a small hydrophobic core. Each such zinc finger sequence folds into an independent structural domain, generally consisting of a short two-strand β-sheet packed against a three-turn α-helix [reviewed in Kaptein (1992), Berg (1993), and Schmiedeskamp and Klevit (1994)]. Cys_2−His_2 zinc fingers typically occur in tandem arrays; while some two-finger arrays have been shown to bind DNA with high affinity, arrays of three or more zinc fingers are much more common.

Several Cys_2−His_2 zinc-finger proteins have been shown to require sequences adjacent to the finger arrays for tight DNA binding, including the yeast SWI5 and ADR1 proteins and the *Drosophila* Tramtrack and GAGA proteins. In some cases, the contribution of the accessory sequence is merely the stabilization of the fold of one of the fingers. In SWI5 and in Tramtrack, for example, residues immediately adjacent to the most N-terminal finger form a third strand of β-sheet for that finger (Neuhaus et al., 1992; Fairall et al., 1993). In SWI5, Neuhaus et al. (1992) showed that the N-terminal finger fails to fold without this extra strand; in Tramtrack, the same is likely true, and a cocrystal structure indicates that the additional strand does not make any contacts with DNA (Fairall et al., 1992, 1993).

GAGA is a rare single-finger protein; it will not bind DNA without approximately 20 residues N-terminal to its finger (Pedone et al., 1996). An NMR structure of the GAGA−DNA complex has shown that this accessory sequence interacts extensively with DNA, making contacts in the major groove, across the DNA backbone, and in the minor groove (Omichinski et al., 1997). In SWI5, a region N-terminal to the extra zinc finger β-strand also increases the DNA binding affinity (Dutnall et al., 1996). This region folds into a helix that is not intimately involved with the zinc finger; it may play its role by making direct contact with the DNA.

The transcription factor ADR1, which regulates expression of glucose-repressible alcohol dehydrogenase in the yeast *Saccharomyces cerevisiae*, possesses a two-finger array that is insufficient for high-affinity DNA binding (Denis & Young, 1983; Hartshorne et al., 1986; Thukral et al., 1989). Tight binding requires approximately 20 residues N-terminal and proximal to the zinc-finger array (Thukral et al., 1991a). This accessory region contains a large number of basic residues. The fingers fold to canonical structures in the absence of flanking sequences (Hoffman et al., 1993; Bernstein et al., 1994). Previous NMR work suggested that the fingers do not interact extensively with the accessory region even in the presence of DNA, suggesting that the role of the nonfinger sequence is the direct contact of DNA (Schmiedeskamp et al., 1997). The importance of the N-terminus may be explained in part by the limited DNA contact made by one of the two zinc fingers of ADR1. Mutagenesis and NMR studies suggest that the C-terminal finger of ADR1 makes only a single specific zinc finger−DNA contact, and that otherwise it interacts minimally with

† This work was supported by NIH Grants P01 GM32681 and R01 GM55369 and by the Murdock Charitable Trust. M.S. was supported by an NSF Predoctoral Fellowship and by an NIH training grant in Molecular Biophysics (T32 GM08268).

* Corresponding author: Biomolecular Structure Center, Box 357742, University of Washington, Seattle, WA 98195-7742. E-mail: klevit@u.washington.edu. Fax: (206) 543-8394.

⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1997.

S0006-2960(97)01364-0 CCC: $14.00

DNA (Thukral et al., 1991b; Schmiedeskamp et al., 1997). Such suboptimal contact by one of the two zinc fingers in ADR1 may explain the need for contacts from an accessory sequence.

There are several questions about the required nonfinger region of ADR1 that can be posed in light of known accessory DNA-binding sequence behaviors. The first is the issue of where the accessory arm binds. In some proteins that use DNA-binding arms, the arm extends away from the DNA site contacted by a conserved DNA binding motif, interacting with adjacent bases. This is the case with the accessory arms of λ repressor (Jordan & Pabo, 1988) and the homeodomains (Kissinger et al., 1990). In other cases, the accessory arm binds the opposite face of the DNA site contacted by a conserved DNA binding motif, as is true for the accessory arm of chicken GATA-1 (Omichinski et al., 1993). The DNA-contacting accessory arm of the GAGA zinc finger protein displays both types of behavior. The portion of the GAGA arm immediately adjacent to the zinc finger contacts bases adjacent to the zinc finger-binding site, while the extreme N-terminus of the GAGA arm binds to the back of the zinc-finger site (Omichinski et al., 1997).

A second question concerns the secondary structure of the accessory arm. Most accessory arms studied to date adopt irregular but fairly extended structures, including those of λ repressor, the homeodomains, and GATA-1. The accessory sequences from GAGA and SWI5, however, exhibit helical structure in the region immediately adjacent to the zinc finger (Dutnall et al., 1996; Omichinski et al., 1997). It is possible that the ADR1 arm adopts a similar fold.

We addressed these issues using a polypeptide, ADR1z, that spans the minimal DNA-binding domain of ADR1, including both zinc fingers and the N-terminal accessory sequence. NMR assignments are available for ADR1z both free in solution and bound to DNA (Schmiedeskamp et al., 1997). We first used a paramagnetic shift agent, cobalt, as a probe of the proximity of the accessory sequence to the zinc fingers. We next assessed secondary structure by observation of NOE connectivities in the N-terminal accessory sequence of DNA-bound ADR1z.

MATERIALS AND METHODS

Samples. Pure ADR1z and DNA for NMR studies were obtained as reported previously (Schmiedeskamp et al., 1997). ADR1z samples for ^{15}N-edited spectra were prepared by uniform labeling with $^{15}NH_4Cl$ and those for ^{13}C-edited spectra by double labeling with [^{13}C]glucose and $^{15}NH_4Cl$, as described previously (Schmiedeskamp et al., 1997). Samples for ^{15}N-edited three-dimensional NOESY and ^{13}C-edited three-dimensional NOESY experiments were folded with zinc and bound to DNA as described (Schmiedeskamp et al., 1997). These samples were prepared at pH 7.0.

Samples containing cobalt were prepared in a similar manner, with the following exceptions. The use of reducing agents was avoided in samples that contained cobalt. Therefore, all steps in the folding of ADR1z in the presence of cobalt and subsequent addition to DNA were carried out anaerobically, in degassed solutions under an atmosphere of argon. $CoCl_2$ was used in the folding reactions in place of $ZnCl_2$, at a level of 0.9 equiv of cobalt per finger. Quantitation of the cobalt stock solution was achieved with UV−visible spectroscopy of cobalt chelated by 2,2′,2″-terpyridine (Storm & Dunn, 1985). The incorporation of cobalt into the zinc fingers was confirmed by following the visible spectrum of the protein sample at 635 nm (Párraga et al., 1990). Preparation of the ADR1z−DNA complex proceeded as described previously (Schmiedeskamp et al., 1997), except that these steps were also conducted under argon in the absence of reducing agents. DNA-free ADR1z samples were maintained at pH 5.4 as before; ADR1z−DNA samples were prepared at pH 7.0.

NMR Spectra. 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 conducted at 37 °C, as described previously (Schmiedeskamp et al., 1997).

Two-dimensional ^{1}H−^{15}N-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 ^{15}N dimension, with a spectral width of 2000 Hz. In experiments involving cobalt, the ^{1}H spectral width was expanded to 12 500 Hz. In these and other experiments, the recycle delay was generally 1.2 s, and the ^{15}N carrier was set to 116 ppm.

The three-dimensional ^{15}N-edited NOESY−HSQC experiment was carried out with a mixing time of 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 achieved with the States method. Thirty-two transients per FID were collected. The spectral widths and the number of complex points acquired were 6410 Hz and 128 points in ^{1}H, 1562 Hz and 17 points in ^{15}N, and 4006 Hz and 512 points in the acquisition dimension. The ^{1}H carrier was placed at water in t_1 and at 8.5 ppm in the acquired dimension.

The ^{13}C-edited three-dimensional NOESY experiment was carried out with a mixing time of 75 ms for the ADR1z−DNA complex in D_2O (Majumdar & Zuiderweg, 1993). Gradients were used for artifact suppression. Quadrature detection in t_1 was achieved with the States method. Sixteen transients per FID were collected. The spectral widths and the number of complex points acquired were 5000 Hz and 110 points in ^{1}H, 2641 Hz and 27 points in ^{13}C, and 5121 Hz and 512 points in the acquisition dimension. The ^{13}C carrier was set to 43.4 ppm.

NMR experimental data 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 transformation.

Paramagnetic Shift Calculations. Calculation of predicted paramagnetic shifts due to cobalt utilized the magnetic susceptibility tensor empirically determined for a Cys_2−His_2 zinc finger by Berg and colleagues (Harper et al., 1993). Because the characteristics of the tensor depend primarily on the metal and on the nature and the geometry of the ligand atoms (La Mar et al., 1973), this tensor should be generally applicable to Cys_2−His_2 zinc fingers, which share nearly identical tetrahedral geometries at their conserved metal centers.

First, we calculated the surface at which a 0.2 ppm perturbation by cobalt is expected. This calculation was performed for Berg's consensus zinc-finger peptide, CP1, using the tensor mentioned above (Harper et al., 1993). The resulting isoshift surface has three lobes, centered on the cobalt. The major axis of the tensor bisects one of the sulfur−cobalt−nitrogen angles.

Next, we generated a model of ADR1 fingers bound to DNA by overlapping ADR1 finger structures obtained with NMR (Hoffman et al., 1993; Bernstein et al., 1994) with fingers 2 and 3 of the Zif268 cocrystal structure (Pavletich & Pabo, 1991). The ADR1b single-finger structure (ADR1 finger 1) superimposes on Zif268 finger 2 with a main chain rmsd of 0.67 Å . The PAPA single-finger structure (ADR1 finger 2) superimposes on finger 3 of Zif268 with a main chain rmsd of 0.65 Å (Bernstein et al., 1994).

The final model was prepared by superimposing the 0.2 ppm isoshift surface calculated for CP1 on each of the fingers in the ADR1−DNA model. This was accomplished by superposition of CP1 (and thus its paramagnetic tensor) on each finger using the metal center and the metal ligands only. This was judged to be the most relevant superposition, because the nature and orientation of the tensor are determined by the metal−ligand atom geometry (La Mar et al., 1973).

RESULTS

Orientation of the N-Terminus of ADR1z. Cobalt can be readily substituted for zinc in Cys_2−His_2 zinc fingers (Párraga et al., 1990; Krizek et al., 1993); in fact, the seminal crystal structures of Zif268 and GLI fingers bound to DNA were obtained in whole or in part from cobalt-bound forms (Pavletich & Pabo, 1991, 1993). Cobalt binds approximately 10000-fold less tightly than does zinc to a Cys_2−His_2 consensus peptide, but with a dissociation constant still on the order of 6×10^{-8} M (Krizek et al., 1993). Cobalt(II) is a paramagnetic metal that causes broadening of NMR resonances and changes in chemical shifts in a predictable way through space (La Mar et al., 1973). It thus can be used as a probe of the proximity of various parts of a protein to a metal-binding site. In the case of ADR1, cobalt incorporated into the metal centers of the zinc fingers should serve as a probe of the proximity of the essential N-terminal accessory region to the fingers. The specific question to be answered is this: does the N-terminus of ADR1 extend along DNA away from the zinc fingers, or does it bind close to them?

In answering this question, we used a protein construct designed previously for NMR studies of an ADR1−DNA-binding domain/DNA complex (Schmiedeskamp et al., 1997). This construct, ADR1z (ADR1 residues 75−161), contains both zinc fingers of ADR1 and the minimal adjacent sequence required for high-affinity DNA binding as defined by deletion mutagenesis (Figure 1A). Truncation of the N-terminus of ADR1 at residue 84 results in a small loss in DNA-binding affinity, but further truncation to residue 88 results in complete loss of detectable binding (Thukral et al., 1991a). ADR1z binds UAS1, the palindromic ADR1 target DNA sequence, with a dissociation constant of 12 nM, measured with a gel electrophoretic mobility shift assay (EMSA). The DNA construct used for NMR studies is based on UAS1. It spans 14 base pairs and includes a single half-site for ADR1 binding (Figure 1B). Numerous biochemical and mutagenesis studies have shown that a single half-site of UAS1 is sufficient for high-affinity ADR1 binding (Thukral et al., 1991a; Cheng et al., 1994). NMR assignments for ADR1z and ADR1z bound to 14mer have been published (Schmiedeskamp et al., 1997).

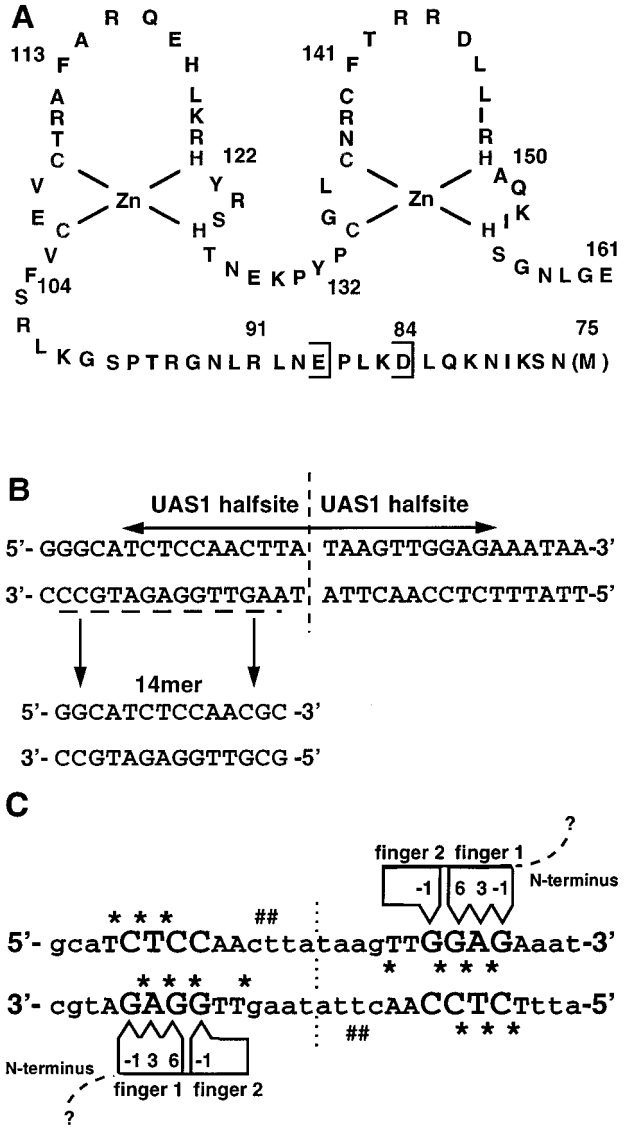

FIGURE 1: ADR1 DNA-binding domain and its specific target DNA site, UAS1. The sequence of the ADR1z construct used for NMR studies is shown in panel A. Brackets mark the extent of the shortest N-terminal deletion mutant that binds DNA with high affinity and the longest that fails to bind. Panel B shows the palindromic UAS1 DNA site and the related 14mer DNA used in this study. Panel C is a schematic of symmetric ADR1 binding to UAS1, in which specific contacts are made from positions −1, 3, and 6 of the helix of finger 1 and from position −1 of the helix of finger 2 (Thukral et al., 1992). Bases for which ADR1 exhibits a preference are shown in capitals; a strong preference is indicated by large capitals and a weak preference by small capitals (Cheng et al., 1994). Sites at which ethylation interferes with binding of ADR1(75−229) are marked by black asterisks. Additional ethylation sites that interfere with binding of ADR1(17−229) are shown by pound sign (#) (Cheng et al., 1994).

Figure 2A shows the ^{1}H-^{15}N-HSQC spectrum of (paramagnetic) cobalt-bound ADR1z overlapped with that of (diamagnetic) zinc-bound ADR1z. The majority of resonances are broadened and experience an upfield chemical shift perturbation in the presence of cobalt. Berg and col-

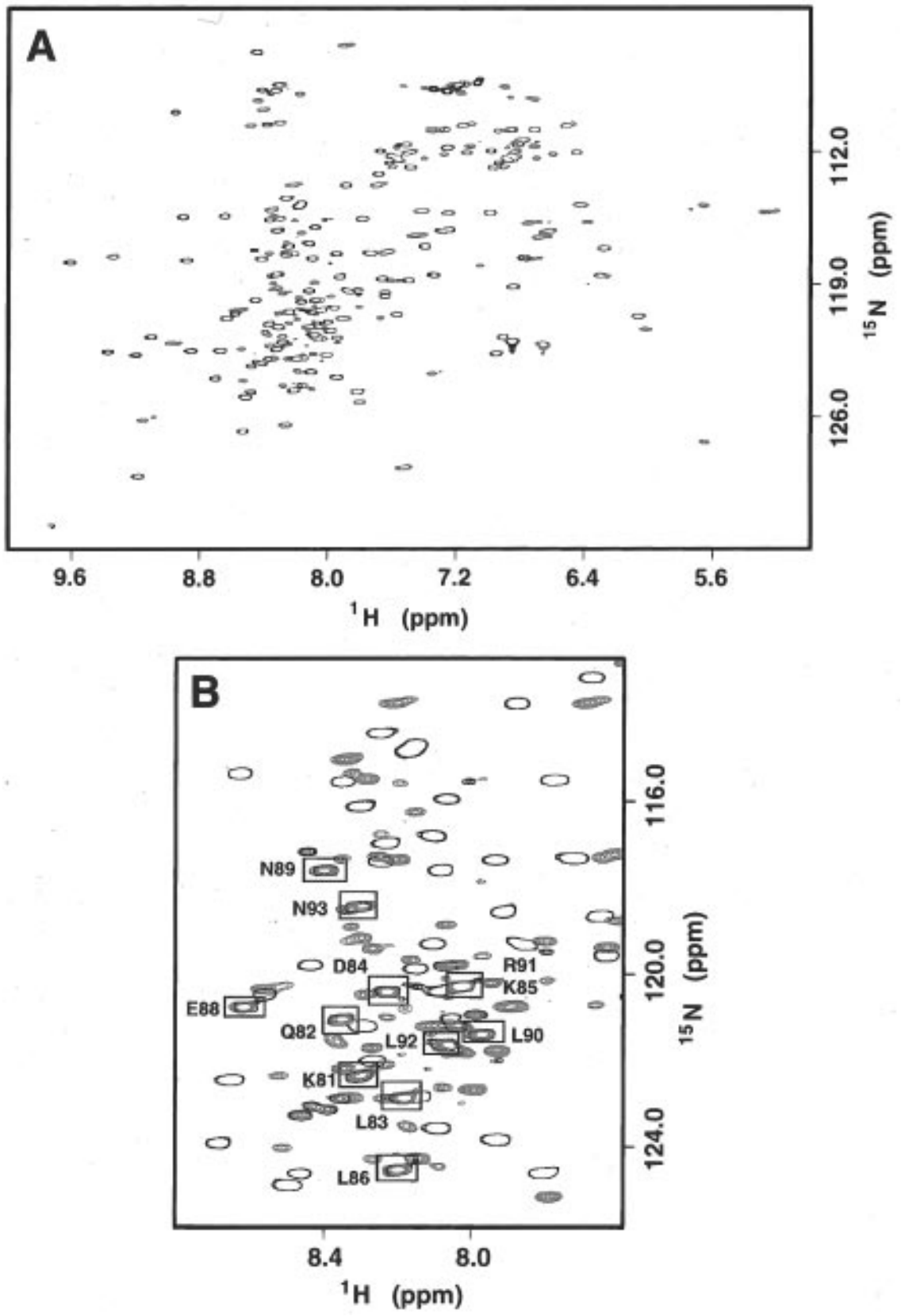

FIGURE 2: Effects of paramagnetic cobalt on ADR1z spectra. ^{1}H-^{15}N-HSQC spectra of zinc−ADR1z (blue) and cobalt−ADR1z (red) are overlapped. Panel A demonstrates that many peaks shift upfield in the presence of cobalt. None of the peaks from the N-terminal accessory sequence (78−100) are shifted or broadened; a sampling of these peaks is shown in panel B.

leagues (Harper et al., 1993) have empirically determined the magnetic susceptibility tensor for a cobalt-bound Cys_2−His_2 zinc-finger consensus peptide, CP1. The results for ADR1z are consistent with their observation for CP1 where the zinc finger lies almost wholly in the lobe of the tensor causing upfield shifts (Harper et al., 1993). In contrast, none of the N-terminal accessory residues of ADR1z appear to be affected by the presence of cobalt in the fingers (Figure 2B). This result is consistent with the previous observation in which the N-terminus is largely random coil in this

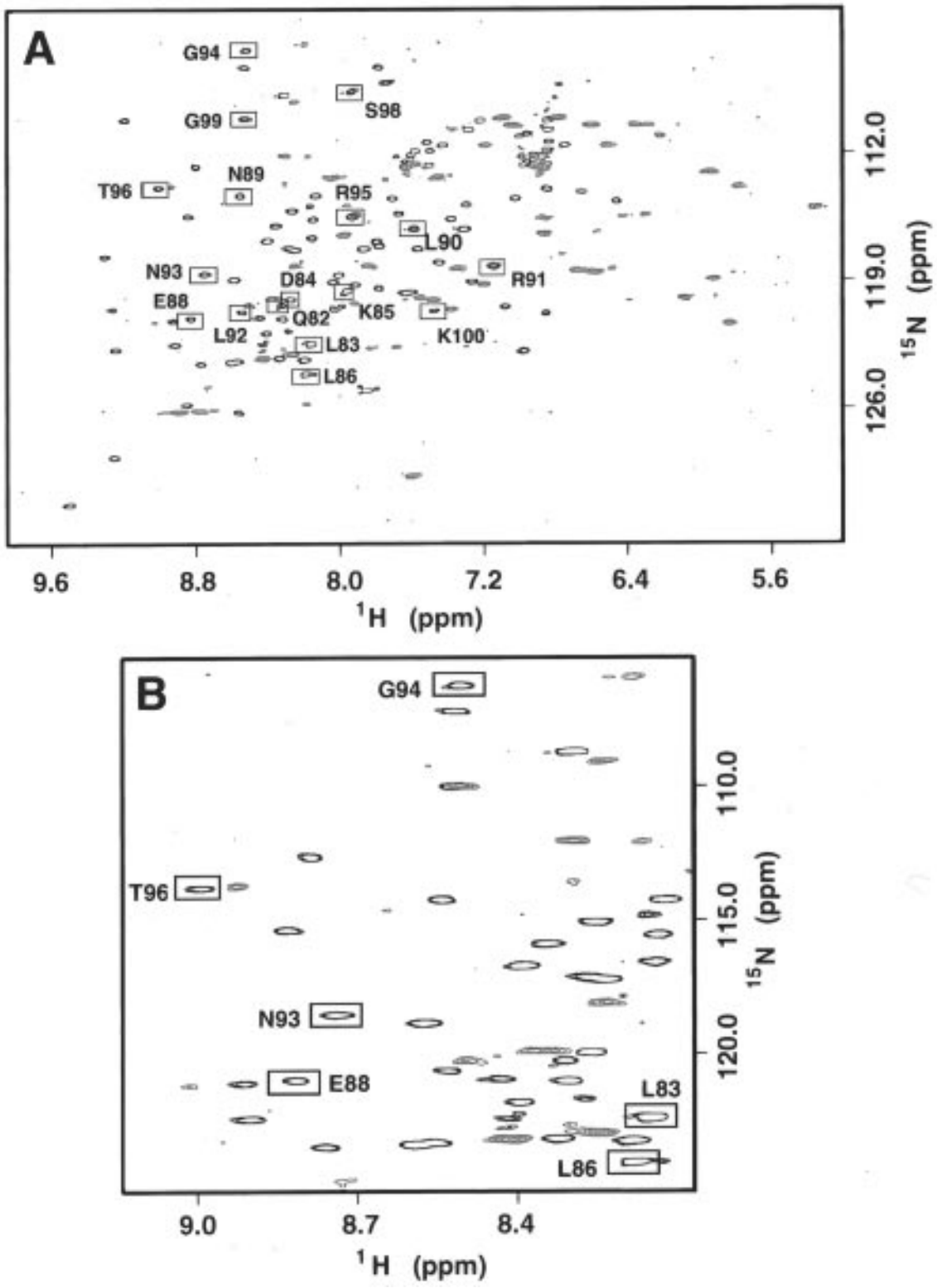

FIGURE 3: Effects of paramagnetic cobalt on spectra of DNA-bound ADR1z. ^{1}H-^{15}N-HSQC spectra of zinc−ADR1z/DNA (blue) and cobalt−ADR1z/DNA (red) are overlapped. Panel A demonstrates that many peaks shift upfield in the presence of cobalt, as in free ADR1z; in the presence of DNA, however, many peaks from the N-terminal accessory sequence shift as well. Panel B emphasizes a number of peaks from the accessory sequence; several are perturbed by at least 0.2 ppm.

construct in the absence of DNA (Schmiedeskamp et al., 1997). This region is thus not expected to have any fixed orientation with respect to the cobalt, and it likely spends a majority of time in conformations distant in space from the metal centers.

Figure 3A shows a similar HSQC spectrum of DNA-bound cobalt−ADR1z overlapped with the spectrum of DNA-bound zinc−ADR1z. Once again, the effect of cobalt on a majority of the peaks is an upfield shift, as expected for the zinc fingers. In addition, a number of peaks from the N-terminus are now shifted away from their zinc-bound positions (Figure 3B). This result implies that the accessory region is not mobile and random coil in the presence of DNA as it is in the free ADR1z construct (compare to Figure 2B). Previous

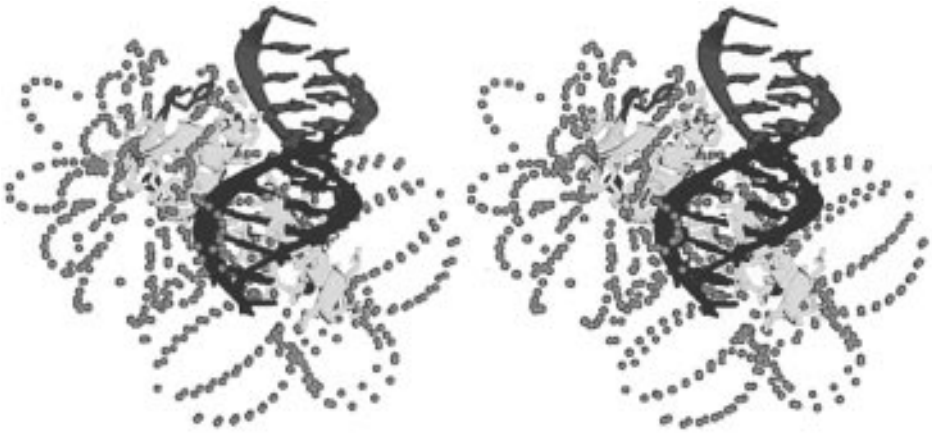

FIGURE 4: Stereoview of predicted 0.2 ppm cobalt isoshift surfaces for a model of ADR1 zinc fingers bound to DNA. Nuclei that experience at least a 0.2 ppm chemical shift perturbation upon cobalt binding are predicted to lie on or within the isoshift surfaces delineated by the pink spheres. Each zinc finger (yellow ribbon) is predicted to lie in the equatorial lobe of its own tensor. Cobalt is represented by black spheres. The N-terminal accessory sequence will arise from finger 1 at the position colored in green. Because much of the N-terminus lies within the 0.2 ppm isoshift surface, it must lie near the fingers and the DNA site they bind (dark blue). The N-terminus does not extend away from the fingers along the DNA colored in cyan, as it would then quickly leave the 0.2 ppm isoshift lobes.

observations of backbone carbon chemical shifts, proton chemical shift perturbations, line widths, and heteronuclear NOEs indeed indicate that the accessory region undergoes a disorder-to-order transition upon ADR1z−DNA binding (Schmiedeskamp et al., 1997; D. Hyre and R. Klevit, manuscript in preparation). The effect of cobalt on the N-terminal accessory residues further implies that this region lies near the zinc fingers when ADR1z binds DNA.

Amide resonances throughout the N-terminus are perturbed by cobalt, including those from residues Q82−N89, L92−T96, and K100 (Figure 3A). Several of these resonances are shifted by at least 0.2 ppm. While no explicit assignments of the cobalt−ADR1z/DNA complex have been made,[1] it is clear that no peaks in the cobalt spectrum lie within 0.2 ppm of the initial zinc-bound positions of residues L83, K85, L86, E88, N93, and G94 from the N-terminal accessory region (Figure 3B). Other N-terminal resonances may be shifted this much or more as well; this cannot be unambiguously determined without assignments in the presence of cobalt.

Because several accessory residues exhibit unambiguous paramagnetic shift perturbations of at least 0.2 ppm, we sought to understand what the 0.2 ppm paramagnetic isoshift surfaces might look like for DNA-bound ADR1 zinc fingers. We constructed a model in which the fingers of ADR1 are bound to DNA in the orientation observed in the Zif268 cocrystal structure (Pavletich & Pabo, 1991). On each finger in this model, we superimposed 0.2 ppm isoshift surfaces which were calculated on the basis of the empirical magnetic susceptibility tensor determined for the cobalt-bound Cys_2−His_2 finger peptide, CP1 (Harper et al., 1993). The resulting model is presented in Figure 4. Any nucleus that experiences a paramagnetic shift of 0.2 ppm or more is predicted to lie on or within one of the isoshift surfaces drawn.

The model presented in Figure 4 shows that nuclei which are shifted by 0.2 ppm or more lie close to the zinc fingers. It is clear from this model that the N-terminus, portions of which must lie within the 0.2 ppm isoshift surfaces, cannot fulfill its role in DNA binding by extending along the DNA away from the zinc fingers. Instead, it must lie nearer to the UAS1 DNA site (depicted in dark blue), close to the zinc fingers. Indeed, residue N93 from the N terminus must lie significantly closer to a metal-binding site than even the 0.2 ppm isoshift surfaces drawn. Among the peaks in the cobalt spectrum that lie near the zinc-bound position of N93, the closest is 0.5 ppm away in ^{1}H chemical shift and 0.35 ppm away in ^{15}N chemical shift; the next closest peak is only 0.39 ppm away in ^{1}H, but it is fully 1.4 ppm away in ^{15}N (Figure 3B).

Portions of the DNA bound by the zinc fingers (dark blue) pass through the 0.2 ppm isoshift surface depicted in the model in Figure 4. Of particular interest is the strand contacted by the fingers themselves, which lies nearest the viewer in the stereo representation. The backbone of this strand passes through the 0.2 ppm isoshift surface, while the backbone of the opposite strand does not. Additionally, at several points, an isoshift lobe juts out over the minor groove opposite the zinc finger-binding site. These may be likely sites for interaction of the N-terminal accessory sequence with DNA.

The HSQC spectra shown in Figure 3A suggest that the backbone amide resonances of L90, R91, S98, and G99 are largely unperturbed by cobalt in ADR1z−DNA. Although this conclusion cannot be made with certainty, as it is possible that cobalt has shifted peaks from elsewhere in the spectrum to overlap with the zinc-bound positions of these four residues, the fact that the putatively unperturbed resonances belong to residues that are proximal in sequence argues against coincidence. If these peaks are indeed unperturbed, this may suggest that these residues lie near the convergence of the equatorial and axial lobes of the isoshift surfaces. This is a region where large differences in cobalt-induced chemical shift perturbation can occur over very short distances, perhaps explaining why some residues are perturbed by cobalt while their neighbors appear to be unaffected.

Secondary Structure of the ADR1z N Terminus. Like ADR1, the GAGA and SWI5 zinc-finger proteins use regions adjacent to the zinc-finger array to increase DNA-binding affinity. In these proteins, the accessory arms fold into a

[1] Assignment of the cobalt−ADR1z/DNA complex is hampered by the significant line broadening induced by the paramagnetic centers.

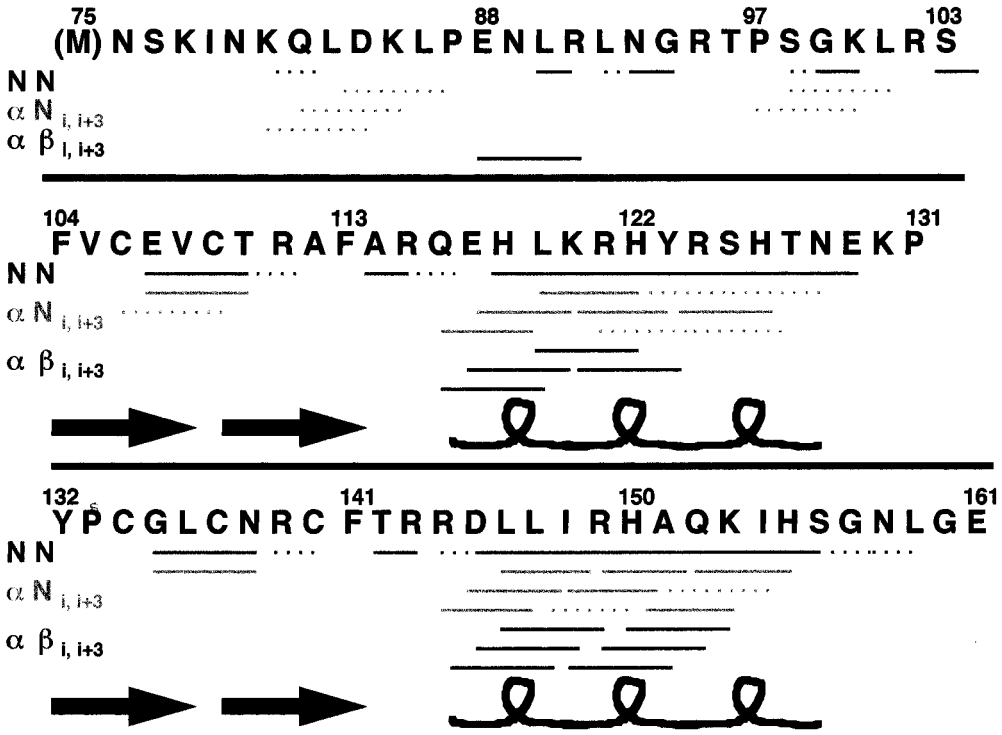

FIGURE 5: Sequential connectivity map for DNA-bound ADR1z. Several types of sequential NOEs diagnostic for α-helix are mapped versus the sequence for ADR1z−DNA, including NN, $\alpha N_{i,i+3}$, and $\alpha\beta_{i,i+3}$ NOEs. Strong patterns of helical NOEs are apparent in the zinc fingers, but not in the N-terminal accessory region (75−102). Dashed lines indicate observed NOEs that cannot be unambiguously assigned, due to chemical shift overlap.

helical structure in the region immediately preceding the β-sheet of the N-terminal zinc finger (Dutnall et al., 1996; Omichinski et al., 1997). In both cases, this helix plays a role in increasing DNA binding. We wanted to know whether a similar structure could be detected in the N-terminal region of ADR1.

As a measure of helical secondary structure, we looked for patterns of sequential NOE connectivities diagnostic of helix (Wüthrich, 1986). Our search included NN, $\alpha N_{i,i+3}$, and $\alpha\beta_{i,i+3}$ NOEs, using ^{15}N-edited and ^{13}C-edited three-dimensional NOESY spectra. Figure 5 shows the results of this NOE analysis for ADR1z bound to DNA. Portions of the zinc fingers that are expected to be helical on the basis of canonical finger structures include residues Q116−H126 in finger 1 and residues R144−H155 in finger 2. These regions have in fact been shown to be helical by other measures, including carbon chemical shift (Schmiedeskamp et al., 1997), and they clearly possess patterns of NOE connectivities diagnostic for helical structure, including strong sequential NN, $\alpha N_{i,i+3}$, and $\alpha\beta_{i,i+3}$ NOEs.

In contrast, no portion of the ADR1z N-terminus has any hallmarks of helical structure, including the region adjacent to the β-sheet of the first finger where helices were found in both SWI5 and GAGA. Thus, ADR1z appears to be a case different from either of these proteins. On the basis of the NOE analysis and on backbone carbon chemical shift profiles (Schmiedeskamp et al., 1997), there is no evidence for completely regular secondary structure in the N-terminal region. This region is probably fairly extended.

DISCUSSION

The cobalt paramagnetic shift data presented here suggest that the N-terminal accessory region of ADR1z lies close to the zinc fingers when bound to DNA, likely binding the same DNA site to which the zinc fingers bind. This conclusion is consistent with earlier data from ethylation interference footprints of ADR1 binding (Figure 1C). These experiments showed that ethylation of phosphates within UAS1 perturbed ADR1 binding, but ethylation of phosphates outside the site did not (Cheng et al., 1994).

Mutagenesis studies have shown that the zinc fingers of ADR1 contact bases near the edges of the palindromic UAS1 sequence, as diagrammed in Figure 1C (Thukral et al., 1992; Cheng et al., 1994). Ethylation of phosphates at the edges of UAS1 interferes with ADR1 binding, as expected for the zinc fingers (Cheng et al., 1994). Ethylation of phosphates near the center of UAS1, however, interferes with ADR1 binding as well. In particular, ethylation of phosphates at the center of the UAS1 palindrome interferes with binding when the entire N-terminus of ADR1 is included in the protein construct (ADR1 17−229), but modification of these central phosphates does not interfere with binding by a shorter, minimal construct (ADR1 75−229) (Cheng et al., 1994). This result implies that, when the entire N-terminus is present, the N-terminus turns back at some point from the edge of the palindrome, where it connects to finger 1, to extend toward the center of the palindrome. In doing so, the region of the N terminus adjacent to the zinc fingers might lie close to the fingers and to UAS1. Our NMR studies show that, even when the full N-terminus is not present, the minimal accessory sequence does indeed turn back from the edge of the palindrome to lie close to the zinc fingers.

Ethylation of a phosphate four base pairs from the center of the palindrome on the strand contacted by the zinc fingers interferes with binding by a construct with even a minimal length N-terminus (Figure 1C) (Cheng et al., 1994). While this phosphate falls at the extreme edge of the triplet

DNA subsite predicted by canonical rules to be bound by finger 2, the finger is unlikely to be the source of the contact. The only residues of finger 2 that experience chemical shift perturbations upon DNA binding are those that bind to the opposite end of the finger-2 subsite (Schmiedeskamp et al., 1997). In fact, very little of finger 2 is perturbed, suggesting quite limited contact with the DNA. In some zinc fingers, a phosphate analogous to the one in question is contacted by the first histidine zinc ligand (Kim & Berg, 1996). While the analogous histidine in ADR1 finger 1 is perturbed upon DNA binding, the equivalent histidine in finger 2 remains absolutely unperturbed (Schmiedeskamp et al., 1997).

Thus, it is unlikely that finger 2 is responsible for the phosphate contact suggested by ethylation interference at the edge of the finger-2 subsite. Furthermore, the backbone phosphate contact in question falls on the same strand of DNA that the zinc fingers bind. As discussed in the Results, this strand (but not its opposite) passes through the 0.2 ppm isoshift surface in our model, as must the N-terminal accessory sequence. It thus seems plausible that the minimal N-terminus might be responsible for this contact with the DNA backbone. If the N-terminus turned toward the center just after leaving finger 1 and adopted an extended conformation, the minimal essential sequence could just span the distance to the center of the palindrome.

Indeed, we find no evidence for helical structure like that observed in other accessory sequences used by the zinc-finger proteins GAGA and SWI5 (Dutnall et al., 1996; Omichinski et al., 1997). On the basis of the sequential NOEs presented here and backbone carbon chemical shift data presented previously (Schmiedeskamp et al., 1997), the N-terminus of ADR1 is probably largely extended. It might thus be able to cover the distance to nearly the center of the UAS1 palindrome.

Prediction of the exact course of the ADR1z N-terminus is difficult; previous experience has shown that the structure and function of accessory regions are diverse. For example, in the GAGA N-terminus alone, there are a helix, extended regions, and a tight hairpin turn, and this region makes major groove contacts and minor groove contacts and crosses the DNA backbone (Omichinski et al., 1997). In the case of ADR1z, we already know that it will differ somewhat from the cases of GAGA and SWI5, on the basis of the lack of helical structure in the N-terminus. Further complicating any prediction is the apparent lack of regular secondary structure in the N-terminal region of DNA-bound ADR1z. Description of the precise course of the N-terminal accessory arm must await a full structure determination of the ADR1z−DNA complex.

One point that is clear, however, is that the minimal N-terminus does not make extensive specific contacts in the major groove. The only portion of the UAS1 sequence in which specific bases are strongly required is that part where the zinc fingers bind in the major groove (Cheng et al., 1994). One base pair immediately 3′ to the finger-1 subsite does show a weak preference for specific bases (Cheng et al., 1994). The reason for this is unclear. This base pair may be contacted by finger 1 itself; in many other cases, the +2 helical position has been implicated in binding to an analogous base neighboring the primary finger subsite (Pavletich & Pabo, 1991; Fairall et al., 1993; Elrod-Erickson et al., 1996; Kim & Berg, 1996). Alternatively, this base pair might play a role in maintaining the appropriate DNA structure. Finally, it is possible that this base pair is contacted by the ADR1 N-terminal accessory sequence as it exits finger 1. No unambiguous evidence for cobalt paramagnetic shift perturbations is available for residues between R95 and the start of the finger at R102 (Figure 3A).[2] It is thus possible that these residues lie outside the 0.2 ppm isoshift surfaces drawn in Figure 4, and interact with the base pair neighboring the finger-1 subsite. If this is the case, they must do so in a different manner than that of the equivalent region of GAGA, which uses a helix to interact with the equivalent base pair (Omichinski et al., 1997).

Extensive interactions between the highly basic N-terminus and the DNA backbone are quite likely. The complex is extremely sensitive to ionic strength; the HSQC spectrum of bound ADR1z is replaced by that of free ADR1z when only 400 mM NaCl has been added to an NMR sample (data not shown). EMSA studies have also shown that salt easily disrupts the ADR1−DNA interaction *in vitro* (Camier et al., 1992). These results suggest that charged interactions constitute a significant portion of the binding energy of ADR1 for its DNA site.

In contrast, cocrystals of zinc-finger arrays and DNA have been obtained from salt solutions that exceed 400 mM by as much as 2-fold (Pavletich & Pabo, 1991). Such conditions would certainly disrupt an ADR1−DNA interaction. There is no evidence that the zinc fingers of ADR1 are involved in any interactions with DNA dramatically different from those of fingers that have been crystallized; both mutagenesis and NMR studies point to the use of canonical types of contacts (Thukral et al., 1991b, 1992; Schmiedeskamp et al., 1997). Thus, the difference in salt sensitivity almost certainly is due to binding contributions of the ADR1z N-terminus. This result suggests that extensive charge−charge interactions occur between the basic N-terminus and the DNA backbone.

All available evidence points to binding of the N-terminal accessory region of ADR1 to the same DNA site bound by the zinc fingers. Binding of accessory sequences to the opposite face of DNA bound by a well-defined motif has been observed before, for proteins such as GATA-1 (Omichinski et al., 1993), the homeodomains (Kissinger et al., 1990), and the zinc-finger protein, GAGA (Omichinski et al., 1997). In each of these cases, contacts to the minor groove occur. Studies to determine whether the N-terminus of ADR1 also makes contacts with the minor groove of the UAS1 site are underway.

[2] Without assignments for the Co(II)−ADR1z/DNA spectrum, we cannot say how much these resonances have shifted. As shown in Figure 3A, these peaks in the zinc spectrum are overlapped or nearly overlapped by peaks in the Co(II) spectrum. It is formally possible that these resonances have indeed shifted, and that other peaks in the Co(II) spectrum have come to rest near their initial positions. However, the simplest interpretation is that these resonances are not shifted, and thus lie outside the 0.2 ppm isoshift surface.

ACKNOWLEDGMENT

We thank Dr. Jeremy Berg for kindly providing the coordinates of the CP1 NMR structure. We thank Dr. Ponni Rajagopal for invaluable advice on NMR spectroscopy and Dr. David Hyre for helpful comments on this work.

REFERENCES

Berg, J. M. (1993) *Curr. Opin. Struct. Biol. 3*, 11−16.

Berg, J. M., & Shi, Y. (1996) *Science 271*, 1081−1085.

Bernstein, B. E., Hoffman, R. C., Horvath, S., Herriott, J. R., & Klevit, R. E. (1994) *Biochemistry 33*, 4460−4470.

Camier, S., Kacherovsky, N., & Young, E. T. (1992) *Mol. Cell. Biol. 12*, 5758−5767.

Cheng, C., Kacherovsky, N., Dombek, K. M., Camier, S., Thukral, S. K., Rhim, E., & Young, E. T. (1994) *Mol. Cell. Biol. 14*, 3842−3852.

Denis, C. L., & Young, E. T. (1983) *Mol. Cell. Biol. 3*, 360−370.

Dutnall, R. N., Neuhaus, D., & Rhodes, D. (1996) *Structure 4*, 599−611.

Elrod-Erickson, M., Rould, M. A., Nekludova, L., and Pabo, C. O. (1996) *Structure 4*, 1171−1180.

Fairall, L., Harrison, S. D., Travers, A. A., & Rhodes, D. (1992) *J. Mol. Biol. 226*, 349−366.

Fairall, L., Schwabe, J. W. R., Chapman, L., Finch, J. T., & Rhodes, D. (1993) *Nature 366*, 483−487.

Harper, L. V., Amann, B. T., Vinson, V. K., & Berg, J. M. (1993) *J. Am. Chem. Soc. 115*, 2577.

Hartshorne, T. A., Blumberg, H., & Young, E. T. (1986) *Nature 320*, 283−287.

Hoffman, R. C., Horvath, S. J., & Klevit, R. E. (1993) *Protein Sci. 2*, 951−965.

Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893−899.

Kaptein, R. (1992) *Curr. Opin. Struct. Biol. 2*, 109−115.

Kay, L. E., Keifer, P., & Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.

Kim, C. A., & Berg, J. M. (1996) *Nat. Struct. Biol. 3*, 940−945.

Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B., & Pabo, C. O. (1990) *Cell 63*, 579−590.

Krizek, B. A., Merkle, D. L., & Berg, J. M. (1993) *Inorg. Chem. 32*, 937−940.

La Mar, G. N., Horrocks, W. D., Jr., & Holm, R. H., Eds. (1973) *NMR of Paramagnetic Molecules Principles and Applications*, Academic Press, New York.

Majumdar, A., & Zuiderweg, E. R. P. (1993) *J. Magn. Reson., Ser. B 102*, 242−244.

Neuhaus, D., Nakaseko, Y., Schwabe, J. W. R., & Klug, A. (1992) *J. Mol. Biol. 228*, 637−651.

Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., & Gronenborn, A. M. (1993) *Science 261*, 438−446.

Omichinski, J. G., Pedone, P. V., Felsenfeld, G., Gronenborn, A. M., & Clore, G. M. (1997) *Nat. Struct. Biol. 4*, 122−132.

Párraga, G., Horvath, S., Hood, L., Young, E. T., & Klevit, R. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 137−141.

Pavletich, N. P., & Pabo, C. O. (1991) *Science 252*, 809−817.

Pavletich, N. P., & Pabo, C. O. (1993) *Science 261*, 1701−1707.

Pedone, P. V., Ghirlando, R., Clore, G. M., Gronenborn, A. M., Felsenfeld, G., & Omichinski, J. G. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2822−2826.

Schmiedeskamp, M., & Klevit, R. E. (1994) *Curr. Opin. Struct. Biol. 4*, 28−35.

Schmiedeskamp, M., Rajagopal, P., & Klevit, R. E. (1997) *Protein Sci. 6*, 1835−1848.

Sklenár, V., Piotto, M., Leppik, R., & Saudek, V. (1993) *J. Magn. Reson., Ser. A 102*, 241−245.

Storm, M. C., & Dunn, M. F. (1985) *Biochemistry 24*, 1749−1756.

Thukral, S. K., Tavianini, M. A., Blumberg, H., & Young, E. T. (1989) *Mol. Cell. Biol. 9*, 2360−2369.

Thukral, S. K., Eisen, A., & Young, E. T. (1991a) *Mol. Cell. Biol. 11*, 1566−1577.

Thukral, S. K., Morrison, M. L., & Young, E. T. (1991b) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9188−9192.

Thukral, S. K., Morrison, M. L., & Young, E. T. (1992) *Mol. Cell. Biol. 12*, 2784−2792.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York.

BI971364F