Articles

Structure of a Histidine-X_4-Histidine Zinc Finger Domain: Insights into ADR1–UAS1 Protein–DNA Recognition†,‡

Bradley E. Bernstein,§ Ross C. Hoffman,§,∥ Suzanne Horvath,⊥ Jon R. Herriott,§ and Rachel E. Klevit*,§

Department of Biochemistry, SJ-70, University of Washington, Seattle, Washington 98195, and Division of Biology, California Institute of Technology, Pasadena, California 91125

Received September 21, 1993; Revised Manuscript Received December 6, 1993®

ABSTRACT: The solution structure for a mutant zinc finger peptide based on the sequence of the C-terminal ADR1 finger has been determined by two-dimensional NMR spectroscopy. The mutant peptide, called PAPA, has both proline residues from the wild-type sequence replaced with alanines. A nonessential cysteine was also replaced with alanine. The behavior of PAPA in solution implicates the prolines in the conformational heterogeneity reported earlier for the wild-type peptide [Xu, R. X., Horvath, S. J., & Klevit, R. E. (1991) *Biochemistry 30*, 3365–3371]. The solution structure of PAPA reveals several interesting features of the zinc finger motif. The residue immediately following the second cysteine ligand adopts a positive ϕ angle, which we propose is a common feature of this class of zinc fingers, regardless of whether this residue is a glycine. The NMR spectrum and resulting solution structure of PAPA suggest that a side-chain to side-chain hydrogen bond involving an arginine and an aspartic acid analogous to one observed in the Zif268 protein–DNA cocrystal structure exists in solution in the absence of DNA [Pavletich, N. P., & Pabo, C. O. (1991) *Science 252*, 809–817]. A model for the interaction between the two ADR1 zinc fingers and their DNA binding sites was built by superpositioning the refined solution structures of PAPA and ADR1b onto the Zif268 structure. This model offers structural explanations for a variety of mutations to the ADR1 zinc finger domains that have been shown to affect DNA-binding affinity or specificity.

Zinc fingers constitute an important class of DNA-binding motifs in eukaryotes [see El-Baradi and Peiler (1991) for a review]. The motif was originally recognized as a repeating element in the sequence of the *Xenopus* transcription factor TFIIIA (Brown *et al.*, 1985; Miller *et al.*, 1985). The critical elements of the 30-residue structural domain include two cysteine and two histidine residues that serve as ligands for a Zn^{2+} ion, giving the motifs their (Cys_2, His_2) nomenclature, and three conserved hydrophobic residues that interact in a small hydrophobic core. These stable domains consist of a Cys-Cys loop (two antiparallel strands connected by a reverse turn) and an α-helix containing the two histidine ligands, both free in solution, as defined by 2D-NMR[1] (Párraga *et al.*, 1988; Lee *et al.*, 1989b; Klevit *et al.*, 1990; Omichinski *et al.*, 1990; Kochoyan *et al.*, 1991) and when bound to DNA in a cocrystal (Pavletich & Pabo, 1991).

ADR1 is a transcription factor that is required for expression of the gene for the glucose-repressible alcohol dehydrogenase isozyme ADH2 in the yeast *Saccharomyces cerevisiae* (Denis & Young, 1983). The minimal sequence of ADR1 that retains high-affinity sequence-specific DNA-binding activity contains two (Cys_2, His_2) zinc fingers. Mutagenesis of this region has revealed that the sequence-specific interaction with DNA is mediated through the two fingers (Thukral *et al.*, 1991b, 1992). ADR1 binds to a 22-base-pair palindrome upstream activator sequence (UAS1) in the promoter for the *ADH2* gene. Two monomers of ADR1 bind independently to the two identical half-sites within UAS1 (Thukral *et al.*, 1991a). Both ADR1 and UAS1 have been subjected to mutagenesis experiments that have identified the amino acid residues and the base pairs that are critical to the protein–DNA interaction (Thukral *et al.*, 1991b, 1992). Our studies have been aimed at determining the three-dimensional structures of the ADR1 zinc fingers with the goal of providing a structural foundation on which to understand the functional properties of the protein.

Our structural studies have focused on each of the two finger domains from ADR1 as independent domains, defining their structures in solution using 2D-NMR spectroscopy. Although a single (Cys_2, His_2) zinc finger does not show detectable site-specific DNA binding, single fingers fold stably into a structure consistent with those reported for functional multifinger domains. Previous NMR analysis of the C-terminal finger from ADR1, ADR1a, demonstrated that the peptide exists in multiple conformations (Xu *et al.*, 1991). The conformational heterogeneity manifests itself as peak doubling in spectra of ADR1a. This peak doubling has made a quantitative determination of the wild-type ADR1a solution

† These studies were supported by NIH Grant P01 GM-32681. B.E.B. was supported by an NIH Medical Scientist Training Program (T32 GM-07266). R.C.H. was supported by an NIH Training Grant in Molecular Biophysics (T32 GM-08268). R.E.K. was supported by an American Heart Association Established Investigatorship.

‡ The top 10 structures and the average structure have been deposited in the Brookhaven Protein Data Bank under the filename 1PAA.

* To whom correspondence should be addressed.

§ University of Washington.

∥ Current address: Department of Protein Chemistry, Immunex Research and Development Corp., 51 University St., Seattle, WA 98101.

⊥ California Institute of Technology.

® Abstract published in *Advance ACS Abstracts*, February 1, 1994.

[1] Abbreviations: 2D-NMR, two-dimensional nuclear magnetic resonance spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; P.E.COSY, primitive exclusive correlation spectroscopy; H-bonds, hydrogen bonds; Tris, tris(hydroxymethyl)-aminomethane; COSY, correlation spectroscopy; RELAY, relayed coherence transfer spectroscopy; RMSD, root-mean-square deviation; TOCSY, total correlation spectroscopy.

0006-2960/94/0433-4460$04.50/0

structure difficult, and we have therefore confined ourselves to a description of the overall folding topology of this domain. In contrast, the N-terminal ADR1 finger peptide, ADR1b, yields a single set of lines in its NMR spectrum, and its three-dimensional structure has been determined and refined (Klevit *et al.*, 1990; Hoffman *et al.*, 1993a). Although both ADR1 zinc fingers conform to the minimal zinc finger sequence (Michael *et al.*, 1992), there are two notable differences in their sequences (refer to Figure 5 for sequences). First, ADR1a contains prolines at positions 131 and 133 (numbering used is for the ADR1 protein), whereas ADR1b contains no prolines. Second, ADR1b is a His-X_3-His finger (i.e., it has a three-residue spacing between the zinc-coordinating histidines), while ADR1a is a His-X_4-His finger. It was hypothesized that one of the aforementioned properties is responsible for the conformational heterogeneity observed in the ADR1a peptide (Xu *et al.*, 1991).

In order to test the contribution of the proline residues to the observed ADR1a peak doubling, and in hopes of obtaining clear structural data for the ADR1a domain, a mutant in which both prolines were substituted with alanine (called PAPA) was synthesized. This peptide also contained a Cys140-to-alanine substitution. Since this cysteine is not a zinc ligand and has been shown by alanine-scanning mutagenesis to be noncritical for DNA binding (Thukral *et al.*, 1991), alanine was substituted at position 140 to yield a peptide that is less susceptible to oxidation. The results of two-dimensional ^{1}H NMR studies of the PAPA peptide are reported here. The refined solution structure of this domain is described and related to the known functional properties of this finger domain in ADR1.

MATERIALS AND METHODS

Sample Preparation. A peptide corresponding to the C-terminal zinc finger of ADR1, but with the modifications listed below, was synthesized and purified as previously described (Párraga *et al.*, 1990). The peptide sequence consists of ADR1 residues 130–159 with alanine substitutions at residues corresponding to Pro131, Pro133, and Cys140 in wild-type ADR1. The sequence of the resulting mutant, PAPA, numbered as in ADR1 protein, with the three non-wild-type residues underlined is

$$NH_3^+\text{-}K^{130}\underline{A}Y\underline{A}CGLCNR\underline{A}FTRRDLLIRHAQKIHSGNK^{159}\text{-}COO^-$$

The peptide was analyzed for purity and correctness by analytical HPLC, mass spectrometry, and Edman degradation sequence analysis. NMR samples were prepared in 50 mM deuterated Tris and 25 mM deuterated acetate, pH 5.29 (meter reading, uncorrected for isotope effect), as described previously (Hoffman *et al.*, 1993a).

NMR Methods. NMR experiments were performed on a Bruker AM-500 spectrometer. The sample temperature was held at 25 °C, and the spectral width was 10 000 Hz. Presaturation of the water resonance during the 2.0-s relaxation delay suppressed the solvent signal. 2D spectra were collected as previously described (Hoffman *et al.*, 1993a,b). Felix II (Hare Research, Woodinville, WA) was used for data processing.

Proton resonances were assigned from COSY, TOCSY, and NOESY experiments by standard practice (Wüthrich, 1986). P.E.COSY spectra collected on a D_2O sample were used to determine $^3J_{\alpha\beta}$ coupling constants (Mueller, 1987). These were used in conjunction with NOESY intensities to obtain stereospecific assignments for prochiral C^β methylene protons and leucine C^δ methyl protons where possible as described previously (Hoffman *et al.*, 1993a).

Relative chemical shifts (Δ) used to compare PAPA to other zinc fingers were calculated by subtracting the random-coil chemical shift for the relevant amino acid type (Wüthrich, 1986) from the observed (or reported) chemical shift. Comparisons were performed using residues from the first conserved aromatic residue to the first conserved histidine residue in each finger sequence.

Structure Calculation and Refinement. Two identical NOESY sets collected with 100-ms mixing times were used to generate distance constraints. The spectra were processed as described previously (Hoffman *et al.*, 1993b). The spectra were added together, and the resulting intensities were measured using volume integrals from the additive spectrum. A diagonal-suppressed NOESY spectrum was used to identify cross peaks that were close to the diagonal (Denk *et al.*, 1985). Dihedral angle constraints were included in the structure calculation when unambiguous assignments could be made. Sixteen χ_1 rotamer assignments were made on the basis of coupling constants and/or relative intraresidue NOESY intensities. In addition, χ_2 constraints for Leu146 and Leu147 could be made on the basis of the stereospecific assignments of the leucine C^β protons and the relative intraresidue NOESY cross peaks. No H-bond constraints or tetrahedral zinc constraints were used in the structure calculations. Lists of the constraints used as input for the structure determination were made available to the reviewers and are available as supplementary material.

Solution structures for PAPA were determined using the refinement method called REPENT (Hoffman *et al.*, 1993b), in combination with DIANA (version 1.0) distance geometry software (obtained from K. Wüthrich, ETH, Zurich) and the complete relaxation matrix algorithm CORMA (obtained from T. James, UCSD, San Francisco), as previously described (Hoffman *et al.*, 1993a). In this approach, iterative rounds of refinement are performed to obtain structural models that are more consistent with the experimentally measured NOE intensities, as measured by $R\langle 1/6\rangle$, where $R\langle 1/6\rangle = \sum\{|(I_{obs})^{1/6} - (I_{calc})^{1/6}|\}/\sum(I_{obs})^{1/6}$.

Structural Analysis. Statistics reported in Table 3 were calculated as described previously (Hoffman *et al.*, 1993a). RMSDs were calculated over residues 132–155 and hydrogen bonds predicted for the 10 best structures generated (i.e., the structures with the lowest $R\langle 1/6\rangle$). An average of the 10 structures was calculated and used for subsequent analysis. The top 10 structures and the average structure have been deposited at the Brookhaven Protein Data Bank. Structures were visualized and compared using QUANTA (Polygen, Corp., Watham, MA) and INSIGHT II (Biosym Technologies Inc., San Diego) on a Silicon Graphics workstation. Individual fingers were superpositioned and RMSDs between them calculated using the backbone atoms for residues spanning from the first conserved hydrophobic (analogous to Tyr132 in PAPA) to the first coordinating histidine residue. This method afforded the closest correspondence between His-X_3-His and His-X_4-His fingers. In addition, models of the presumed interaction between the ADR1 zinc fingers and their DNA subsites were created. These models, composed of a single ADR1 finger docked to its respective DNA-binding subsite, were created by superpositioning either the PAPA or the ADR1b average structure on finger 3 of the Zif268 protein–DNA coordinates and by replacing the relevant DNA bases in the Zif268 binding site with their UAS1 counterparts. These superpositions were achieved using the backbone atoms only,

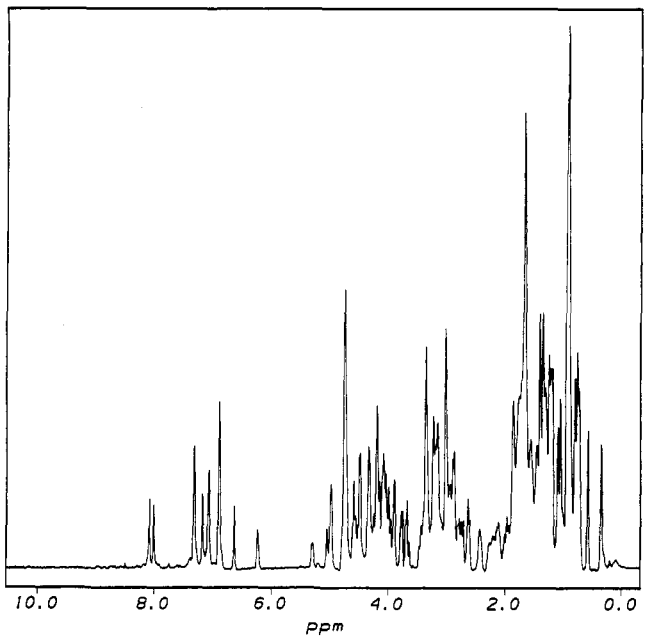

FIGURE 1: ^{1}H NMR spectrum of PAPA. One-dimensional spectrum of PAPA peptide dissolved in 50 mM deuterated Tris and 25 mM deuterated acetate, pH 5.29 (meter reading, uncorrected for isotope effect), in D_2O. The peptide was reconstituted in the presence of 1.1 mol equiv of Zn^{2+}.

starting with the first aromatic residue and ending with the first coordinating histidine residue.

RESULTS AND DISCUSSION

Single-Finger Peptides Assume Multiple Conformations. Previous studies of ADR1a demonstrated that the peptide exists in more than one conformation that interconverts slowly on the NMR time scale, as evidenced by an observed peak doubling in its NMR spectra (Xu *et al.*, 1991). Analysis of the spectra indicated the presence of two folded conformations of ADR1a, a major form present in ~70% and a minor form present in ~30% of the population. Since the presence of proline residues has been shown to be responsible for slowly interconverting species in other proteins, the mutant peptide PAPA was designed in which both proline residues of ADR1a were replaced by alanine. Unlike the spectrum of ADR1a, the 1D spectrum of PAPA (Figure 1) does not show evidence of peak doubling, as is clear from the single sharp resonances of the two histidine C^ϵ protons at ~8.0 ppm. However, when 2D spectra of PAPA with extremely high signal to noise were analyzed, a small amount of peak doubling was detected. Similar handling of spectra obtained for the other finger sequence from ADR1, ADR1b, did not reveal peak doubling. We must therefore conclude that this property is peculiar to the ADR1a sequence, rather than due to some sort of artifact of peptide synthesis or purification.[2] The PAPA spectrum indicates that removal of the two prolines from the ADR1a sequence has a significant effect on the equilibrium between the two forms but does not completely abolish either conformation. Clearly, further mutants of ADR1a must be designed in order to understand this phenomenon more fully. However, the spectral properties of PAPA are good enough to allow a detailed structural analysis of this mutant peptide.

[2] As a control, the single mutant ADR1a:C140A was synthesized and its NMR spectrum analyzed. Like wild-type ADR1a, the spectrum of this peptide displays significant peak multiplicity (D. English and R. Klevit, unpublished results). This rules out the nonligand cysteine at position 140 as the source of spectral heterogeneity originally reported for ADR1a.

The Zinc Finger Signature Fold. Complete assignments were made for PAPA using conventional ^{1}H homonuclear NMR techniques and are listed in Table 1 (see Materials and Methods). A strong correlation has been noted between NH and $C^\alpha H$ resonance frequencies of analogous residues of (Cys_2, His_2) zinc finger peptides (Michael *et al.*, 1992; Lee *et al.*, 1992). Such a correlation suggests that it is mainly the backbone conformation at each residue, rather than the identity of the side chain, that defines its resonance frequency with respect to random-coil values. Thus, the backbone proton resonance frequency shifts can serve as a signature for the zinc finger fold. With this in mind, the NH and $C^\alpha H$ shift values of PAPA were compared residue by residue with those of several other (Cys_2, His_2) zinc finger peptides. As shown in Figure 2, there is remarkable correspondence between PAPA and the seven fingers to which it was compared. The correlation coefficients for pairwise comparisons are listed in Table 2.

Not surprisingly, the strongest correlations are between PAPA and ADR1a, its parent sequence. There is also significant correlation when the resonance frequencies of PAPA are compared with the other ADR1 finger peptide, ADR1b (Hoffman *et al.*, 1993a) and with the peptide based on one of the finger domains from the *Xenopus* protein, Xfin (Lee *et al.*, 1989a). Both ADR1b and Xfin are classic "odd" fingers as dictated by their His-X_3-His spacing and the position of the second conserved aromatic (Kochoyan *et al.*, 1991). In contrast, ZFY-6 is an "even" finger with a His-X_4-His spacing and its conserved aromatic two residues N-terminal from the odd position. PAPA is a hybrid, with the aromatic position of an "odd" finger and the His-X_4-His spacing of an "even" finger. On the basis of its NH shift values, PAPA corresponds more closely to the odd fingers analyzed. Moreover, although the additional residue between the histidines in PAPA causes the NH fingerprint between the coordinating histidines to diverge, the histidine resonances themselves conform closely to those in conventional odd fingers.

Interestingly, PAPA corresponds more closely to the *minor* form of ADR1a on the basis of resonance shifts. The largest chemical shift difference between the two forms of wild-type ADR1a is observed for the amide proton of Tyr132 (~0.8 ppm), the residue between the two prolines (Xu *et al.*, 1991). The chemical shift for the NH of Tyr132 in PAPA corresponds almost exactly to its value in the minor form of ADR1a. As illustrated in Figure 2, the amide proton of the first aromatic residue is shifted significantly upfield relative to the random-coil position in most of the zinc fingers that have been analyzed. The minor form of ADR1a and PAPA also show this large upfield shift, leaving only the major form of ADR1a and MBP-1b as exceptions. Thus, if the backbone frequencies are indeed signatures for the backbone conformation of the zinc fingers, these correlations suggest that PAPA mimics the minor form structure of ADR1a in solution and therefore that replacement of one or both of the prolines causes the equilibrium to favor the so-called minor form. The ADR1a finger resonances in the context of a two-finger synthetic peptide also correspond more closely to the minor form, suggesting that PAPA contains structural elements of the multifinger DNA-binding ADR1c construct (R. X. Xu and R. E. Klevit, unpublished observation).

Analysis of the well-resolved and well-defined NMR spectra of PAPA helped to clarify difficult assignments in ADR1a wild-type and T142I peptides (Xu *et al.*, 1991). Specifically, the $C^\alpha H$ assignments for Cys140 ($\Delta = -0.10$ ppm with respect to random coil) and Arg144 ($\Delta = 0.10$ ppm) in the original report are not consistent with the analogous assignments in

Table 1: Proton Chemical Shifts (in ppm)[a] for PAPA at 25 °C and pH 5.5

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	$C^{\epsilon}H$	$C^{\zeta}H$
Lys130		3.84	1.75	1.21	1.60		
Ala131	8.31	4.44	1.18				
Tyr132	8.74	4.52	2.74,[b] 2.92		6.99	6.81	
Ala133	8.64	4.90	1.29				
Cys134	8.88	4.40	2.79,[b] 3.33				
Gly135	9.28	4.16, 3.90					
Leu136	8.95	4.41	0.60,[b] 1.04	1.37	0.68,[c] 0.50		
Cys137	7.99	4.97	3.14,[b] 3.31				
Asn138	8.11	4.68	2.95		7.51, 6.79[d]		
Arg139	8.01	4.11	1.38,[b] 1.24	1.68, 1.45	3.10, 3.02	7.14[e]	
Ala140	7.95	4.89	1.15				
Phe141	8.48	4.91	2.80,[b] 3.28		7.24	6.82	6.16
Thr142	9.02	4.55	4.49	1.33			
Arg143	7.06	4.71	1.57,[b] 2.13	1.72	3.28	8.43[e]	
Arg144	8.54	3.09	1.51, 1.12	1.37, 1.27	3.08	7.28[e]	
Asp145	8.67	4.10	2.55,[b] 2.65				
Leu146	6.88	3.98	1.93,[b] 1.49	1.61	0.86,[c] 1.11		
Leu147	6.92	3.28	1.30,[b] 2.06	1.58	1.01,[c] 0.99		
Ile148	8.05	3.61	1.79	1.51, 1.22, 0.83 (m)	0.73		
Arg149	7.67	4.00	1.79	1.65	3.16	7.25[e]	
His150	7.52	4.27	2.83,[b] 3.26		7.10	8.01	
Ala151	8.79	3.81	1.59				
Gln152	8.11	3.81	2.19,[b] 2.02	2.55, 2.36		7.38, 6.74[e]	
Lys153	7.85	4.11	1.88, 1.79	1.46, 1.37	1.68	2.95	
Ile154	8.23	4.08	0.93	1.29, 0.28 (m)	0.66		
His155	7.25	5.23	3.24,[b] 3.37		6.54	7.93	
Ser156	8.06	4.27	4.02, 3.95				
Gly157	8.47	4.05, 3.58					
Asn158	8.01	4.66	2.70,[b] 2.80		7.48, 6.95[d]		
Lys159	7.71	4.25	1.59	1.22	0.87		

[a] All chemical shifts are reported relative to sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$ (TSP). [b] Stereospecific assignment, first chemical shift assigned to HB2. [c] Stereospecific assignment, first chemical shift assigned to CD1. [d] $N^{\delta}H$ resonances. [e] $N^{\epsilon}H$ resonances.

PAPA. These residues were also noted as being inconsistent with chemical shifts reported for other zinc fingers (Lee *et al.*, 1992). However, the analogous values in PAPA (A140, Δ = 0.54 ppm; R144, Δ = −1.29 ppm) are consistent with these reports. Since we suspected that the cysteine residue at position 140 might be in part responsible for the poorer quality of the spectra of ADR1a and T142A, a new "wild-type" ADR1a peptide was synthesized in which Cys140 was replaced by alanine. Although the spectra of this peptide still exhibit strong peak doublings, the quality of the spectra is indeed superior, making assignments somewhat clearer (D. English and R. Klevit, unpublished data). The assignments for residues 140 and 144 in C140A (A140, Δ = 0.36 ppm; R144, Δ = −1.28 ppm) are consistent with those in PAPA and therefore with other zinc finger peptides. Even with these assignments in hand, it is difficult to find the corresponding resonances in the original ADR1a spectrum. Since the existence of a cysteine at residue 140 in ADR1a and T142I may perturb the $C^{\alpha}H$ shift relative to the other peptides, it remains unclear whether this assignment is correct as originally reported. On the basis of this analysis, however, we conclude that the original assignment for the $C^{\alpha}H$ of residue 144 in ADR1a was most likely incorrect.

PAPA Structure. The solution structure of the PAPA peptide was calculated from two-dimensional ^{1}H NMR data, as described in Material and Methods. A region of an H_2O NOESY spectrum used to generate distance constraints was made available to reviewers and is available as supplementary material. The 10 structures with the lowest residual error function $R\langle 1/6\rangle$ for observed vs calculated NOE intensities were used for structural analysis (see Materials and Methods); their statistical parameters are presented in Table 3 and superpositions are shown in Figure 3. The backbone structure of the family is very well defined by the NMR data, as illustrated in Figure 3A and indicated by the average RMSD of 0.21 Å (calculated over residues 132–155). Many of the side-chain conformations are also well-defined in the PAPA solution structure (Figure 3B,C). Only 6 of the 30 side chains have average RMSDs greater than 1.0 Å and all of these are hydrophilic, surface residues: Asn138, Arg139, Arg143, Arg144, Arg149, and Lys153.

The solution structure of PAPA reiterates all the structural elements previously described for (Cys_2, His_2) zinc fingers: a Cys-Cys loop consisting of two antiparallel strands connected by a reverse turn, the "fingertip" in which the backbone makes a 90° turn, and a helical region that includes the histidine ligands. The PAPA backbone and zinc coordination geometry closely resemble the refined models for ADR1b (Figure 4), as well as for other experimentally determined zinc finger domains (Hoffman *et al.*, 1993a; Lee *et al.*, 1989b; Omichinski *et al.*, 1990; Kochoyan *et al.*, 1991). The backbone RMSD between the average structures for ADR1b and PAPA is 0.46 Å (calculated as described in Materials and Methods). The remarkable similarity of PAPA to ADR1b (as well as to Zif268, see below) confirms that the sequence of ADR1a is capable of taking on a canonical zinc finger structure when there are alanines at positions 131 and 133. Although the structure of PAPA *per se* cannot directly address the role of the naturally occurring prolines, it can address other important aspects of the ADR1a structure, in particular, the effect of the alternate His-X_4-His spacing. Since a detailed description of the refined solution structure of ADR1b as calculated by an identical protocol has been published recently (Hoffman *et al.*, 1993a), we will focus mainly on a comparison of the solution structures of PAPA and ADR1b. In addition, structural characteristics will be compared to those in the X-ray structure of the three-finger Zif268 protein (Pavletich & Pabo, 1991).

Cys-Cys Loop. The N-terminal residues of PAPA form the Cys-Cys loop, consisting of two antiparallel strands and

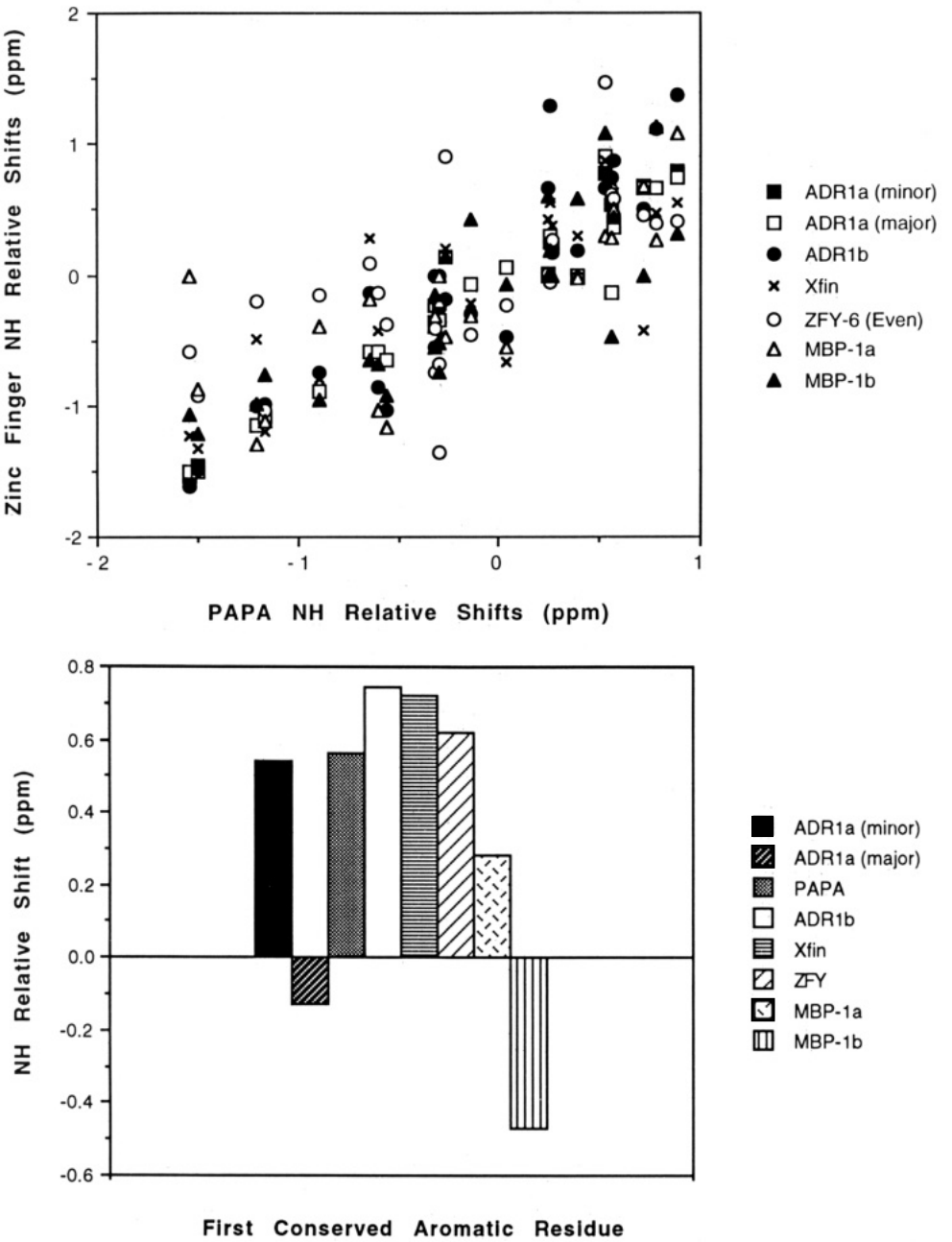

FIGURE 2: Relative chemical shifts for NH resonances of a number of zinc finger sequences. The relative chemical shifts were calculated as described in Materials and Methods. (A, top) The correspondence between the shifts for PAPA and seven other zinc finger sequences are plotted for all residues between and including the first conserved aromatic residue and first conserved histidine residue. (B, bottom) The relative chemical shifts for the NH of the first conserved aromatic residue are shown. For references regarding the chemical shifts, see legend to Table 2.

a reverse turn. The two positions where proline side chains are found in the wild-type sequence have backbone dihedral angles that are consistent with a β-strand conformation: the average (ϕ, ψ) values for Ala131 and Ala133 are ($-90 \pm 50°$, $140 \pm 15°$) and ($137 \pm 6°$, $119 \pm 12°$), respectively. (The wide range of ϕ values in the solution structure for Ala131 is likely due to its proximity to the N-terminus of the peptide.) Proline is limited to ϕ values of $-60 \pm 20°$ and prefers ψ values of $\sim -55°$ or $\sim 145°$. Thus, of the two mutated residues in PAPA, the solution structure clearly indicates that Ala133 exists in a conformation that is not accessible to the wild-type proline. Two of the fingers in Zif268 contain a proline at the position analogous to Pro131 in ADR1a, both of which are in an allowed (ϕ, ψ) conformation of ($\sim -60°$, $\sim -10°$). None of the Zif268 fingers contain proline at the position analogous to Pro133. In all three fingers, the residue in this position takes on (ϕ, ψ) values of ($\sim -110°$, $145°$), consistent with a β-strand conformation and with the value observed for Ala133 in PAPA. Taken together, these results indicate that, in the absence of proline residues, this region of the zinc finger domain exists in an extended β-strand conformation, which may be deformed slightly to accommodate a proline in the position analogous to Pro131 in ADR1a. The consequence of having a proline at position 133 must await a thorough structural analysis of a finger with such a proline residue in place. However, it appears that adjustments of the backbone will be required to accommodate a proline at this position.

Similar to ADR1b, cross-strand H-bonds are predicted between the carbonyl of Tyr132 and the amide of Phe141 and NH–sulfur H-bonds analogous to those reported for ADR1b are predicted (NH136 and NH137 to the S^{γ} of Cys134 and NH139 to Cys137). Also analogous to the solution structure

Table 2: Chemical Shift Correlations between Various Zinc Finger Domains[a]

	ADR1a major[b]	ADR1a minor[b]	ADR1b[c]	MBP-1b[d]	Xfin 31[e]	ZFY-6[f]
PAPA	0.91	0.96	0.84	0.68	0.65	0.41
ADR1a major		0.96	0.78	0.74	0.62	0.47
ADR1a minor			0.82	0.67	0.67	0.49

[a] Relative chemical shifts were calculated as described in Materials and Methods for the NH resonances for residues from the first conserved aromatic through the first conserved histidine residue. Pairwise comparisons between zinc fingers were performed by plotting the relative chemical shifts for each corresponding residue and fitting the resulting plot to a linear function. The numbers in this table are the correlation coefficients for the pairs of data [see, for example, Kurtz (1963)]. [b] Chemical shifts used are for the ADR1a peptide in which a nonessential cysteine was replaced with alanine, C140A (D. English and R. Klevit, unpublished results). [c] As reported in Hoffman *et al.* (1993). [d] As reported in Omichinski *et al.* (1992). [e] As reported in Lee *et al.* (1989a). [f] As reported in Kochoyan *et al.* (1991).

Table 3: Structural Parameters for PAPA

parameter	PAPA
av[a] DIANA target function,[b] Å^2	4.1 ± 0.4
RMSD[c] (N, Cα, C, O), Å	0.21 ± 0.06
RMSD[c] (all heavy atoms), Å	1.13 ± 0.18
av[a] $R(1/6)$, %	9.9 ± 0.1
no. of interproton distance constraints[d]	
intraresidue	159
$i, i+1$	83
$i, i+2$	15
$i, i+3$	19
$i, i+4$	10
long range	50
total	346
no. of dihedral angle constraints[e]	
χ_1	16
χ_2	2

[a] Averages are reported for the 10 refined structures in the family. [b] The target function is a residual error function based on upper and lower bound, dihedral angle constraint, and van der Waals violations for each calculated structure. The target functions reported here were calculated with weighting factors set at 1.0 for each of upper distance bounds, lower distance bounds, and van der Waals and set at 5.0 for dihedral angle constraints. [c] All RMSDs are reported as the average pairwise RMSDs among the 10 refined structures for the residue range 132–155. [d] The upper and lower bound files used in the structure calculations are available as supplementary material. [e] The dihedral angle constraints used in the structure calculations are available as supplementary material.

of ADR1b, the residue following the second zinc-coordinating cysteine residue (Asn138 in PAPA; Thr110 in ADR1b) has a positive ϕ angle. The ϕ and ψ angles adopted by Asn138 (89° and −14°, respectively) are consistent with a G1 β-bulge in which a similarly positioned glycine has (ϕ, ψ) values of (85°, 0°) (Richardson, 1981). As in ADR1b, additional H-bonds predicted between the amide of Cys134 and the carbonyl of Arg139 and between the carbonyl of Cys134 and the amide of Asn138 may help to stabilize this geometry. Although the positive ϕ angle has not been explicitly mentioned in reports of other zinc finger structures, we wondered whether it is a common feature. In (Cys_2, His_2) zinc fingers, the residue that follows the second cysteine ligand is most often glycine (Michael *et al.*, 1992). In addition to the positive ϕ angles in the two ADR1 fingers, an analogous non-glycine residue in Zif268 finger 2 (Met141) also has a positive ϕ (Pavletich & Pabo, 1991). Moreover, we note that, in all rubredoxin structures solved to high resolution by X-ray diffraction, the residue that follows the second cysteine ligand in its Cys-Cys loop (which is always glycine) assumes a positive ϕ angle (Adman *et al.*, 1991.) It is therefore likely that the positive

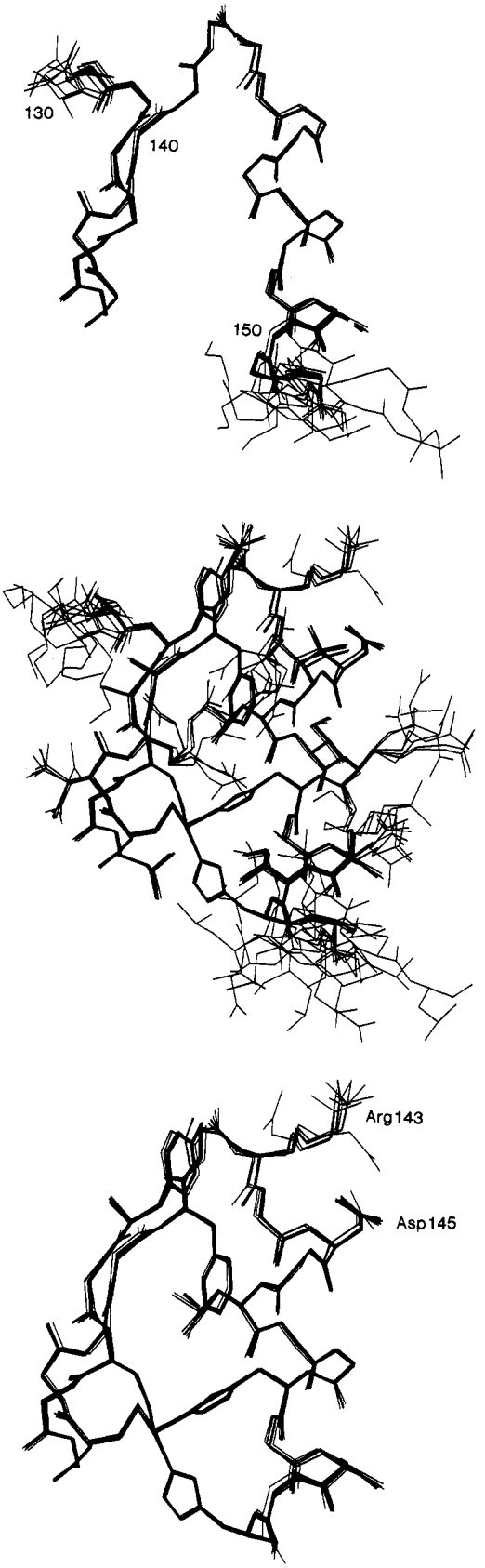

FIGURE 3: Solution structure of PAPA. (A, top) Superposition of the top 10 structures, showing backbone atoms only. All residues of PAPA (i.e., 130–159) are included. (B, middle) Superposition of the top 10 structures, with all heavy atoms included. For clarity, only residues 132–155 are shown in this view. (C, bottom) Superposition of the top 10 structures, showing side-chain heavy atoms for the seven conserved zinc finger residues (four zinc ligands and three hydrophobics). In addition, the side chains of Arg143 and Asp145 are shown.

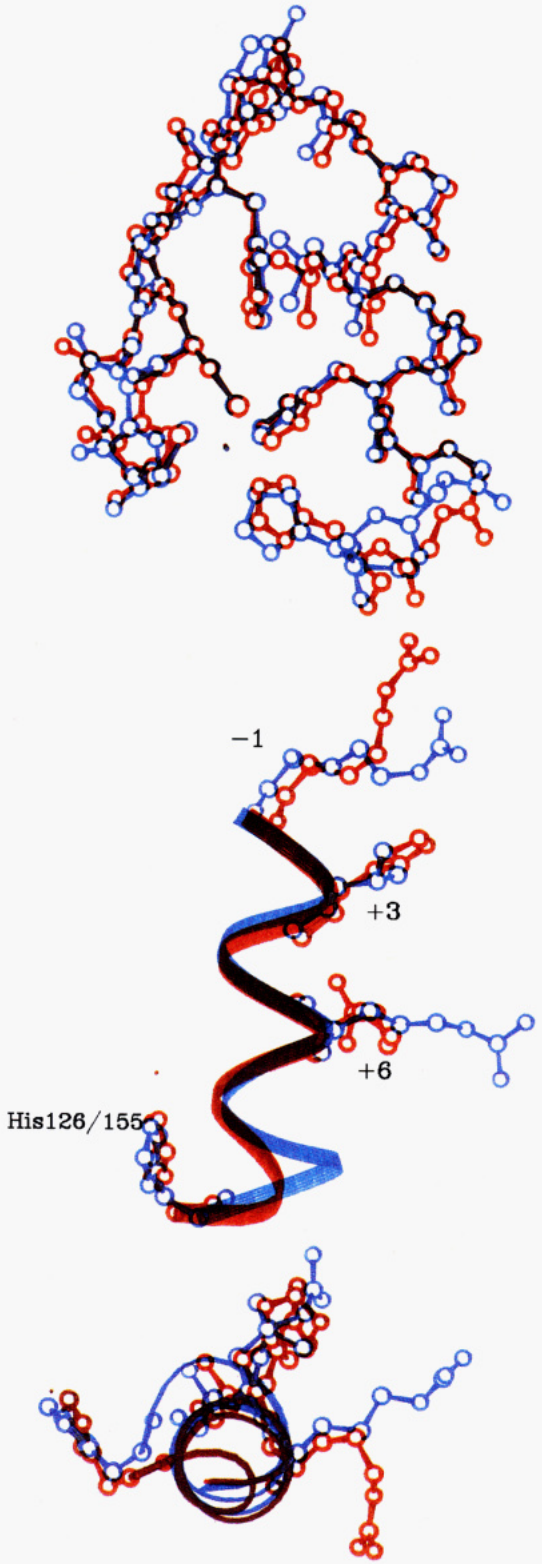

FIGURE 4: Comparison of the solution structures of PAPA (blue) and ADR1b (red). The average structures calculated from the solution structures for each peptide are shown, superimposed as described in Materials and Methods. (A, top) Side-chain atoms for the seven conserved zinc finger residues are included. Note that the main difference between the two structures occurs in the backbone between the two histidine ligands (lower right region of this view). (B, middle) The α-helices of ADR1b and PAPA superimposed. The side chains in the three putative DNA-binding positions (−1, +3, and +6) plus the second histidine residue are shown. Note that the long, hydrophilic side chains at positions −1 and +6 are not well-defined in solution. (C, bottom) Same as (B) but viewed down the helical axis. This view makes clear the tight loop between the histidines in ADR1b and the loose loop in PAPA.

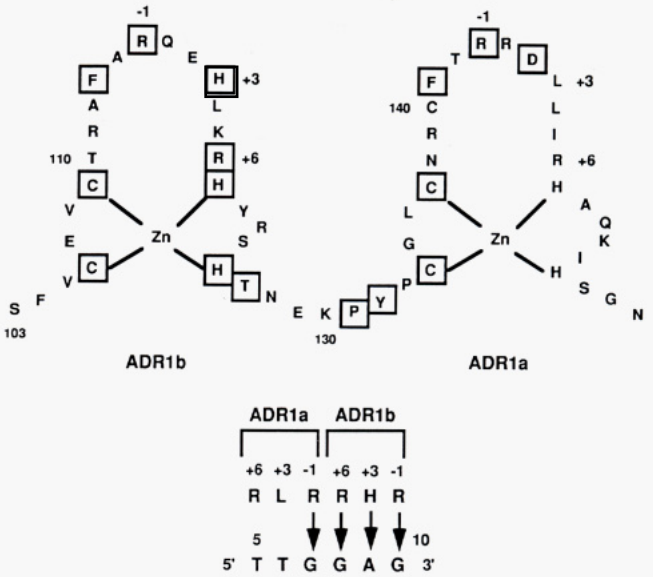

FIGURE 5: Summary of the results of mutagenesis experiments on ADR1. (A, top) The residues where substitution with an alanine resulted in a >10-fold decrease in DNA-binding affinity are boxed in the sequence of the two zinc fingers of ADR1 (Thukral *et al.*, 1991b). (B, bottom) Summary of the protein–DNA contacts revealed from change-of-specificity mutagenesis (Thukral *et al.*, 1992).

ϕ angle assumed by the residue immediately following the second Cys is a feature common to many if not all of the Cys-X_2-Cys zinc finger structures. This arrangement brings the two cysteine side chains into proximity to bind the zinc ion, and the stabilization afforded by zinc binding likely serves to offset the unfavorable dihedral angle at this position.

α-Helix. The α-helical regions of PAPA and ADR1b are shown in Figure 4. Alanine mutagenesis studies revealed that residues at analogous positions in the helical regions of the two ADR1 fingers do not exhibit analogous functional properties (Thukral *et al.*, 1991b, and Figure 5). It is therefore of particular interest that the PAPA structure differs most markedly from ADR1b in this region. A superposition of the PAPA and ADR1b structures, together with dihedral angle information, demonstrates that the helix in each of the molecules begins at an analogous residue (Arg144 in PAPA and Gln116 in ADR1b) and ends at the second histidine. The backbones diverge only between the two histidine ligands: in order to approximate the coordinating histidines to the zinc ion, the three interceding residues in ADR1b tighten to form a 3_{10} helix, while the four residues between the histidines in PAPA loosen such that the dihedrals for Gln152 and Lys153 are in altogether nonhelical geometries. These results appear to be consistent with those observed in NMR studies of another His-X_4-His zinc finger peptide, ZFY-6 (Kochoyan *et al.*, 1991).

A Hydrogen Bond between the Side Chains of Arg143 and Asp145. The N^{ϵ} proton of Arg143 is shifted ~1.2 ppm downfield (to 8.43 ppm) from the other arginine N^{ϵ} protons in the PAPA spectrum, which resonate close to the random-coil position of 7.2 ppm (see Table 1). A similar downfield shift is observed in the wild-type ADR1a spectrum and in the T142I mutant spectrum (Xu *et al.*, 1991). A downfield shift of a nitrogen-bound proton can be indicative of involvement in a H-bond. Interestingly, in the crystal structure of Zif268 the N^{ϵ} protons of the analogous arginines in fingers 1 and 3 are each involved in a H-bond to a carboxylate oxygen of the aspartic acid residue that is two residues to its C-terminal side, forming a H-bond network involving the two side chains and a guanine in the cocrystal (Pavletich & Pabo, 1991). The

analogous residue in PAPA is Asp145, which exists in a defined side-chain conformation in solution, as evidenced by its inequivalent $C^{\beta}H$ resonances, its $^3J_{\alpha\beta}$ couplings, and its NOESY pattern. Thus, the spectral properties of Arg143 and Asp145 suggest that a similar H-bond interaction occurs in PAPA and in ADR1a, even in the absence of DNA.

The PAPA solution structure is consistent with this conclusion in that an Asp145 carboxylate oxygen and the N^{ϵ} proton of Arg143 are in proximity to make such an interaction in a majority of the structures (see Figure 3C). It is important to note that the PAPA solution structures were calculated using only distance, dihedral angle, and van der Waals constraints. Thus, the proximity observed is a consequence of the experimental constraints placed on these two side chains, rather than on an electrostatic potential enforced during the structure calculation. The analogous residues in ADR1b are Arg115 and Glu117, which may be capable of making a comparable interaction. However, the arginine N^{ϵ} proton shift was not observed for this peptide, and the glutamic acid side chain does not take on a preferred orientation in solution (Hoffman *et al.*, 1993a). Either the longer glutamic acid side chain is sterically unable to make an analogous H-bond or the additional loss of entropy that would accompany its formation must override the small gain in enthalpy from the ionic H-bond.

Structure–Function Relationships: Modeling of ADR1–DNA Interactions. An ultimate goal of structural comparisons between ADR1 fingers is to yield insight into the function of the molecule. The residues in the two zinc fingers of ADR1 for which substitution led to a significant diminution of DNA-binding affinity are illustrated in Figure 5. Consistent with the hypothesis that both zinc liganding and the interaction of conserved hydrophobic residues are responsible for stable tertiary structure in a zinc finger motif, substitution of the conserved residues resulted in decreased DNA binding. The three-dimensional structures of the ADR1 fingers, and all other fingers of this class that have been reported, confirm the contribution of the seven conserved residues to the tertiary structure. Therefore, we will focus on the mutations in nonconserved positions that affect DNA binding.

When we conducted these studies, there was only one experimentally determined structure for a zinc finger–DNA complex, that of the Zif268 three-finger fragment bound to an 11-base oligonucleotide (Pavletich & Pabo, 1991). The Zif268 structure provides a detailed model of how zinc fingers can interact sequence-specifically with DNA. For this structure to serve as a guide for understanding ADR1, two criteria should be met: (1) the structures for the individual ADR1 finger domains should be similar to the finger structures in the Zif268 cocrystal, and (2) there should be evidence to suggest that the mode of DNA recognition used by ADR1 is similar to that used by Zif268.

To ascertain whether the first criterion is met, main-chain RMSDs for the 19 residues spanning the first conserved aromatic to the first coordinating histidine were calculated between each ADR1 finger and each Zif268 finger, as listed in Table 4. There is remarkable similarity between the ADR1 and Zif268 fingers, even though the structures were determined by different techniques, using individual domains with no DNA present in the former case and multiple domains bound to DNA in the latter. Therefore, this criterion is satisfied.

In the Zif268–DNA structure, three positions in the zinc finger make direct contact with DNA bases. For consistency, these positions are called –1, +3, and +6 with respect to their position relative to the α-helix. Each of the three fingers in Zif268 uses two of these three possible contacts. As illustrated in the schematic in Figure 5, alanine mutagenesis of ADR1 indicates that all three positions are critical for DNA binding in ADR1b while only the –1 position is critical in ADR1a (Thukral *et al.*, 1991b). DNA mutagenesis studies indicated that each ADR1 finger contacts a 3-bp "subsite" within the 11-bp UAS1 half-site: ADR1b contacts bases 8–10 (GAG) and ADR1a contacts bases 5–7 (TTG). This arrangement, including the polarity of the complex, is also consistent with that observed in the Zif268–DNA complex. Finally, change-of-specificity mutagenesis of ADR1 identified specific side-chain–base interactions, as shown schematically in Figure 5. The above analyses suggest that ADR1 follows the general theme described by the Zif268 finger–DNA interactions and therefore the second criterion is satisfied.

Table 4: Main-Chain RMSDs between ADR1 and Zif268 Zinc Finger Domains[a]

	Zif268 f2	Zif268 f3	ADR1b	PAPA
Zif268 f2		0.418	0.613	0.655
Zif268 f3			0.657	0.644
ADR1b				0.462
PAPA				

[a] RMSDs were calculated for the backbone atoms of residues from the first conserved aromatic through the first conserved histidine residue, as described in Materials and Methods.

As just demonstrated, the zinc fingers of ADR1 appear to behave generally similar to those in Zif268. However, the mode of DNA binding for ADR1 appears to be different in detail. In particular, the +6 DNA-contact position in ADR1a does not mimic analogous residues in either the Zif268 fingers or in ADR1b. The +6 positions in ADR1b and Zif268 fingers 1 and 3 are all arginines that each contact guanine. Although ADR1a also contains an arginine at the +6 position, the relevant base in its UAS1 binding site is thymidine (Thukral *et al.*, 1992). While replacement of the +6 arginine in ADR1b (Arg121) resulted in >100-fold decrease in DNA binding affinity of a two-finger ADR1 construct, the analogous replacement in ADR1a had no detectable effect (Thukral *et al.*, 1991b).

A comparison of the helical regions and the two +6 arginines of ADR1b and PAPA is shown in Figure 4. Although the positions of the ends of these surface-exposed residues are not defined in solution, the superposition shows clearly that the C^{α} positions of Arg121 and Arg149 are in identical positions relative to the rest of the helix so that the side chains extend from identical positions on the helices of ADR1b and PAPA. This suggests that the cause of the difference in the functional properties of Arg121 and Arg149 is not due to structural differences at the site in question and must therefore be due to a more distal effect.

In hopes of gaining insight into the functional differences between the two ADR1 fingers, a model for ADR1 zinc finger–DNA interactions was generated. The average refined structure of each ADR1 finger was simply superpositioned onto Zif268 finger 3 to give the lowest main-chain RMSD, and each 3-bp subsite was "mutated" into the relevant sequence from UAS1 without altering the DNA conformation (see Materials and Methods). As is clear from Figure 3B, the exact positions and conformations of the side chains of surface residues (especially long hydrophilic residues such as arginines and lysines) are not well-defined in solution and are therefore not likely to be representative of their position or conformation when bound to DNA. However, the backbone positions of these residues are defined in the NMR structures and can be considered indicative of their placement in the DNA-bound

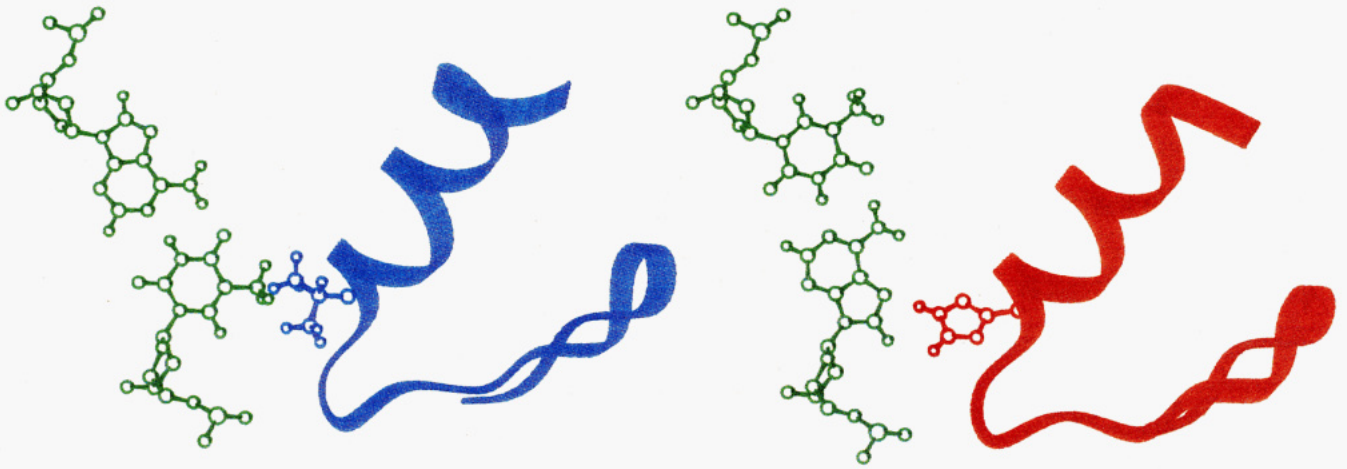

FIGURE 6: Models for the interaction between ADR1 and UAS1. Models were generated by superpositioning the solution structures of PAPA (blue) and ADR1b (red) on the structure of Zif268, as described in Materials and Methods. The putative contacts for position +3 side chains are shown for each finger. (Left) Superposition model for the ADR1a finger. Note the steric clash between the methyl group of Leu146 and the thymidine methyl. (Right) Superposition model for the ADR1b finger. His118 is in excellent distance and geometry to participate in an H-bond with the adenine.

protein. The ADR1–DNA models place the three residues at the −1, +3, and +6 positions of the α-helix facing into the major groove of the DNA, poised to make the interactions predicted from the ADR1 mutational analyses.

In addition to positioning the putative base contacting side chains reasonably with respect to the bases themselves, the ADR1–DNA models also retain the intriguing contact between the N^{δ} of the first zinc-coordinating histidine to a backbone phosphate observed in all three Zif268 fingers. This contact was hypothesized to serve to properly orient finger domains with respect to their DNA sites (Pavletich & Pabo, 1991). While it is not a surprising result for ADR1b, the PAPA–DNA model confirms that such an interaction can also anchor a His-X_4-His finger to its binding site in an orientation similar to the more canonical His-X_3-His fingers. In addition, although the histidine ligand is in proximity to the DNA backbone, the PAPA–UAS1 superposition model does not indicate any clashes between the large His-X_4-His loop and the DNA.

The fact that both ADR1 finger–DNA models affirm central themes predicted from other structural and functional data led us to look for novel findings. We therefore used these models to try to understand why, in contrast to the N-terminal finger, the +3 and +6 position DNA-contact residues in the C-terminal finger of ADR1 do not appear to be critical for DNA binding (see Figure 5). The well-defined nature of the side chains at the +3 position in both ADR1 finger structures makes this residue in the ADR1–DNA models particularly interesting. Superposition models highlighting this position in both ADR1 fingers are shown in Figure 6. Superpositioning places His118, the +3 residue in ADR1b, proximal to the middle base of the subsite (A9). Without any adjustments to the side-chain conformation of the solution structure, the imidazole ring is in excellent H-bonding distance and geometry with the N7 atom of the adenine (adenine nitrogen-to-imidazole proton distance of 1.7 Å). An almost identical interaction was previously observed for the histidine in the +3 position of Zif268 finger 2 and the N7 of a guanine.

The +3 position in ADR1a is Leu146, and the middle base in its DNA subsite is a thymidine (T6). The superposition model for PAPA places Leu146 in proximity to T6, as shown in Figure 6. However, if this finger is in the same relative orientation to its site as the Zif268 fingers, a steric clash is indicated between one of the methyl groups of Leu146 and the thymidine methyl group, with leucine methyl proton to thymidine methyl proton distances of 1.1 Å (compare to the van der Waals radius of a proton of 0.9 Å). The clash could not be relieved by altering the χ_2 for the leucine side chain. Therefore, superpositioning suggests that the ADR1a finger must be adjusted within the major groove at this position, relative to Zif268. Possible results of such an adjustment would be favorable van der Waals interactions and/or hydrophobic interactions between the methyl groups of Leu146 and the methyl group of T6. Substitution of a purine at this position in the PAPA–DNA model would remove these favorable interactions, consistent with mutagenesis results that showed that a cytosine is accommodated at this position, while substitution with either purine base led to a loss of DNA binding (Thukral *et al.*, 1992).

The lack of effect of substituting the +6 position Arg149 residue with alanine in the ADR1a finger is perhaps the most curious result of the alanine mutagenesis study. The ADR1b finger also contains an arginine at this position which change-of-specificity mutation indicates is contacting guanine (Thukral *et al.*, 1992). Substitution of Arg121 with alanine was one of the most detrimental alanine mutations made, suggesting that this interaction contributes significantly to the overall binding free energy. On the other hand, the equivalent ADR1a residue, Arg149, is *not* crucial for DNA affinity (Thukral *et al.*, 1991b). Moreover, unlike the subsites for Zif268 fingers 1 and 3 and ADR1b, the ADR1a binding site contains a thymidine at the site where R149 is expected to interact, rather than a guanine. Wild-type ADR1 will accommodate any other base at this position, strongly suggesting that Arg149 does not directly contact the base (Thukral *et al.*, 1992). Since there is no detectable difference in the structures of ADR1b and PAPA at the +6 positions, we looked for a distal effect. As described above, the superposition model of PAPA–DNA indicates that the ADR1a finger must have its α-helix adjusted in the major groove of the DNA due to the bulky side chain at its +3 position. We suggest that this, in turn, affects the ability of the +6 position Arg149 to make a base contact. Such a neighboring effect of one putative DNA-contact residue on another is consistent with mutagenesis experiments on another zinc finger protein, SP1 (Desjarlais & Berg, 1992). This hypothesis is testable experimentally: substitution of a smaller residue for Leu146 should result in an increased interaction for Arg149. Recent mutagenesis experiments in

which Leu146 was substituted with an alanine side chain are indeed consistent with this hypothesis. Unlike wild-type ADR1, the mutant protein has a preference for a guanine at position 5, the position that Arg149 is predicted to contact (C. Cheng and E. T. Young, personal communication).

CONCLUSIONS

Our goal has been to understand the functional aspects of ADR1 DNA binding in terms of its structure. Toward this end, we have determined and refined the solution structures of zinc finger domains from ADR1 individually. The relevance of this approach is borne out in the finding that the refined structures obtained in the absence of DNA are remarkably similar to the structures of zinc fingers bound to DNA determined by X-ray diffraction. Thus, the solution structures are excellent models from which to derive structure–function information about the DNA-binding domain of ADR1 in particular and of (Cys_2, His_2) zinc finger proteins in general.

The results for the PAPA mutant indicate that the ADR1a sequence is capable of existing in a canonical zinc finger structure that is very similar to those described for ADR1b and other (Cys_2, His_2) fingers. However, the observation that Ala133 in PAPA exists in a conformation that is disallowed for proline indicates that the wild-type domain must have an altered conformation in its N-terminal region. This is consistent with our original report on wild-type ADR1a in which it was stated that the NOEs expected for a classical β-sheet were not observed (Párraga *et al.*, 1988).

Of the two prolines in ADR1a, Pro131 has been implicated as critical for DNA binding since its replacement with alanine resulted in a 100-fold decrease in DNA-binding affinity in a functional two-finger ADR1 construct (Thukral *et al.*, 1991b). Proline is found with high frequency at this position in multifinger proteins, suggesting that this result may be generalizable. Solution NMR studies of two different multiple zinc finger domains, neither of which contains proline in its linkers, have revealed that the dispositions of individual fingers in these constructs are not well-defined in the absence of DNA (Omichinski *et al.*, 1992; Nakaseko *et al.*, 1992). The presence of proline in a linker may serve to limit the range of orientations of two fingers with respect to each other in the free protein. There are prolines in fingers 2 and 3 of the Zif268, both of which are in the *trans* conformation, with (ϕ, ψ) values of ($\sim$–60, –10) in the protein–DNA complex. This conformation is one of the two low-energy conformations for prolines (Cantor & Schimmel, 1980), so it is likely that the prolines are in a similar conformation in the free protein. Although the conformation observed for the two Zif268 prolines is accessible to an alanine residue, an alanine is much less restricted in its allowable backbone conformations. Thus, the substitution of Pro131 with alanine may increase the conformational entropy of the free protein, and the $\sim$3 kcal/mol loss of binding free energy measured for the P131A ADR1 mutant may reflect an increased loss of entropy when the contiguous zinc fingers bind to DNA. A second example of loss of DNA affinity on replacing an analogous proline in the linker of fingers 1 and 2 of TFIIIA has recently been reported (Choo & Klug, 1993).

Several other conclusions that can be drawn from the PAPA solution structure serve to explain earlier observations. First, the residue that follows the second cystine ligand has a positive ϕ angle, similar to the refined ADR1b structure (Hoffman *et al.*, 1993a). This may explain the high degree of conservation of glycine at this position and has implications to the design of novel zinc finger domains, as a glycine at this position should better favor the correct folded structure. Second, the extra residue between the histidine ligands in ADR1a has only a local effect on the conformation of the domain, consistent with the reported structure of another His-X_4-His finger domain (Kochoyan *et al.*, 1991). Comparison of the structure of PAPA and ADR1b in this region shows that the overriding force is the geometry of the zinc ligands, which is nearly the same in the two structures, regardless of the different number of residues between them. Thus, it is the backbone conformation, not the metal–ligand conformation, that adjusts in response to different loop lengths. Third, a side-chain-to-side-chain H-bond is strongly indicated between two surface-accessible, hydrophilic residues in PAPA, namely, Arg143 and Asp145. An analogous H-bond is observed in the DNA-bound form of Zif268, but it was somewhat unexpected to find such a H-bond existing in a domain free in solution. The fact that a similar H-bond is not indicated by the data for ADR1b which contains an analogous arginine but a glutamic acid in the place of Asp145 suggests that there is a delicate balance between the gain in enthalpy and the loss of entropy for this H-bond such that it does not form spontaneously when the longer acidic side-chain is present.

The model created for the interaction between ADR1 finger domains and their respective DNA subsites simply by superpositioning the refined peptide solution structures onto the Zif268–DNA cocrystal structure predicts many of the known DNA-binding properties of the ADR1 fingers and may therefore serve to reveal new insights as well. In particular, the model offers an explanation for the functional properties of the +3 and +6 positions in ADR1a, which mutagenesis experiments had indicated were different from their ADR1b counterparts. Specifically, the model suggests that the ADR1a finger–DNA geometry must be somewhat different from that seen for the three Zif268 fingers, since the analogous geometry results in a steric clash between methyl groups of the +3 Leu146 residue with the methyl group of T6. A slight adjustment of the finger relative to the major groove will alleviate steric overlap and could allow for a favorable interaction between the two methyl groups. This predicted adjustment of ADR1a relative to the DNA may also offer an explanation for the observation that the arginine at the +6 position (Arg149) does not appear to have any interactions with DNA. Taken together, the model explains the mutagenesis results and points to what is certain to become a recurring theme of zinc finger–DNA interaction, namely, the non-independence of neighboring DNA-contacting residues (Desjarlais & Berg, 1992; Bernstein *et al.*, 1993).

In conclusion, the refined structures of the two zinc finger domains from ADR1 have afforded insights into the structural features of this class of DNA-binding motifs. In conjunction with the extensive information available from mutagenesis and functional studies of ADR1, these structural models also offer insights into the structure–function relationships that are central to the sequence-specific DNA recognition between ADR1 and the UAS1 site. The hypotheses offered here on the basis of the ADR1–UAS1 model are currently being tested experimentally, both by further mutagenesis studies and by a direct structural determination of the complex formed between the DNA-binding domain of ADR1 and the UAS1 DNA sequence. This combined approach will undoubtedly yield further insights in the future.

ACKNOWLEDGMENT

The authors thank Mia Schmiedeskamp and Ted Young for helpful discussions and critical reading of the manuscript, David English for sharing his results on the C140A peptide,

Cheng Cheng for sharing his results on Leu146 mutants, and Nikola Pavletich and Carl Pabo for the Zif268–DNA crystal structure coordinates.

SUPPLEMENTARY MATERIAL AVAILABLE

One figure showing the region of the H_2O NOESY spectrum of PAPA and two tables giving distance constraints used to calculate the solution structure of PAPA using DIANA and angle constraints used to generate the structure (9 pages). Ordering information is given on any current masthead page.

REFERENCES

Adman, E. T., Sieker, L. C., & Jensen, L. H. (1991) *J. Mol. Biol. 217*, 337–352.

Bernstein, B. E., Hoffman, R. C., Horvath, S. J., Herriott, J. R., & Klevit, R. E. (1993) *Ann. N.Y. Acad. Sci.* (in press).

Brown, R. S., Sander, C., & Argos, P. (1985) *FEBS Lett. 186*, 271–274.

Cantor, C. R., & Schimmel, P. S. (1980) *Biophysical Chemistry*, Part 1, W. H. Freeman and Co., San Francisco.

Choo, Y., & Klug, A. (1993) *Nucleic Acids Res. 21*, 3341–3346.

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

Denk, W., Wagner, G., Rance, M., & Wüthrich, K. (1985) *J. Magn. Reson. 62*, 350–355.

Desjarlais, J. R., & Berg, J. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7345–7349.

El-Baradi, T., & Peiler, T. (1991) *Mech. Dev. 35*, 155–169.

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

Hoffman, R. C., Xu, R., Klevit, R. E., & Herriott, J. R. (1993b) *J. Magn. Reson., Ser. B 102*, 61–72.

Klevit, R. E., Herriott, J. R., & Horvath, S. J. (1990) *Proteins 7*, 215–226.

Kochoyan, M., Havel, T. F., Nguyen, D. T., Dahl, C. E., Keutmann, H. T., & Weiss, M. A. (1991) *Biochemistry 30*, 3371–3386.

Kurtz, T. E. (1963) *Basic Statistics*, Prentice-Hall, Inc., Englewood CLiffs, NJ.

Lee, M. S., Cavanagh, J., & Wright, P. E. (1989a) *FEBS Lett. 254*, 159–164.

Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., & Wright, P. E. (1989b) *Science 245*, 635–637.

Lee, M. S., Mortishire-Smith, R. J., & Wright, P. E. (1992) *FEBS Lett. 309*, 29–32.

Michael, S. F., Kilfoil, V. J., Schmidt, M. H., Amann, B. T., & Berg, J. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4796–4800.

Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609–1614.

Mueller, L. (1987) *J. Magn. Reson. 72*, 191–197.

Nakaseko, Y., Neuhaus, D., Klug, A., & Rhodes, D. (1992) *J. Mol. Biol. 228*, 619–636.

Omichinski, J. G., Clore, G. M., Appella, E., Sakaguchi, K., & Gronenborn, A. M. (1990) *Biochemistry 29*, 9324–9334.

Omichinski, J. G., Clore, G. M., Robien, M., Sakaguchi, K., Appella, E., & Gronenborn, A. M. (1992) *Biochemistry 31*, 3907–3917.

Párraga, G. E., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., & Klevit, R. E. (1988) *Science 241*, 1489–1492.

Párraga, G. E., Horvath, S. J., 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.

Richardson, J. S. (1981) The Anatomy and Taxonomy of Protein Structure, in *Advances in Protein Chemistry* (Anfinsen, C. B., Edsall, J. T., & Richards, F. M., Eds.) Vol. 34, pp 167–339, Academic Press, Inc., New York, NY.

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, NY.

Xu, R. X., Horvath, S. J., & Klevit, R. E. (1991) *Biochemistry 30*, 3365–3371.