ADR1a, a Zinc Finger Peptide, Exists in Two Folded Conformations†

Robert X. Xu,‡ Suzanna J. Horvath,§ and Rachel E. Klevit*,‡

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

Received November 7, 1990; Revised Manuscript Received January 9, 1991

ABSTRACT: Two-dimensional NMR (2DNMR) studies of several different zinc finger peptides have yielded a picture of the three-dimensional structure of this small DNA-binding motif. Details of the differences among fingers with different sequences may provide some insight into how these domains interact with DNA. Toward this end, we have reanalyzed the 2DNMR spectra of the C-terminal zinc finger sequence from the yeast transcriptional factor ADR1. Although this was the sequence on which our original report describing the overall fold of zinc fingers was based, complete spectral assignments (reported here) were needed to compare this sequence in detail with that of ADR1b, for which we have reported an atomic level structure. In the process of analyzing the spectra of ADR1a and a mutant of ADR1a, it was noted that the peptides give two sets of NMR lines, indicating that this sequence, unlike the other ADR1 zinc finger sequence, exists in two slowly interconverting folded conformations in solution. Residues that exhibit peak doubling are located in the Cys loop, the α-helix, and the extreme C-terminus of the peptide. Differences in NOEs observed for the two forms indicate that there are detectable conformational differences in the Zn^{2+} cluster and in the fingertip region. This conformational flexibility, which has not been observed for other zinc finger peptides, may stem from the presence of an additional residue between the histidine ligands (His-X_4-His versus His-X_3-His).

The DNA-binding motifs known as zinc fingers can be described by the "consensus" sequence Ar-X-Cys-X_{2-4}-Cys-X_3-Ar-X_5-Leu-X_2-His-X_{3-4}-His (where "Ar" stands for an aromatic residue). On the basis of several two-dimensional nuclear magnetic resonance (2DNMR)[1] studies of synthetic zinc finger peptides, the structural determinants of these small domains have been shown to involve a tetrahedral Zn^{2+} cluster that includes conserved Cys and His residues and a small hydrophobic "core" that includes the conserved hydrophobic residues (Parraga et al., 1988; Lee et al., 1989; Klevit et al., 1990; Omichinski et al., 1990). It seems likely that the differences in DNA-binding functions of different zinc finger domains will arise from differences in the nonconserved residues, possibly resulting in small but important differences in the structures of each domain.

In order to compare two naturally occurring zinc finger sequences, the two zinc finger domains from the yeast transcriptional regulator of the ADH2 gene, ADR1, have been synthesized and studied by two-dimensional NMR. Results obtained for the N-terminal domain, ADR1b, have been reported (Klevit et al., 1990). Complete spectral assignments have now been obtained for the C-terminal sequence, ADR1a, and are reported here. As well, the spectrum of a mutant peptide, called A142I (where "A" denotes ADR1a, "142" denotes the mutated residue, and "I" denotes the new amino acid), has been assigned. This mutation is of some interest as the substitution of Thr142 by an isoleucine in the context of the ADR1 protein has been shown to decrease the affinity of ADR1 for its cognate DNA (Blumberg et al., 1987).

In the process of analyzing the spectra of ADR1a and A142I, it was noted that both spectra actually contained two sets of resonances. The results reported here indicate that ADR1a exists as two slowly interconverting folded states. The mutant peptide A142I also exhibits a dynamic equilibrium, but the populations of the two species are different from those observed for the wild-type sequence. Thus, differences in the dynamics of individual domains, as well as structural differences, may be of functional importance.

Materials and Methods

Peptide Synthesis. The peptides were synthesized by the stepwise solid-phase technique (Bruist et al., 1987), purified by reverse-phase HPLC, and characterized by analytical HPLC, amino acid analysis, sequence analysis, and time-of-flight mass spectrometry. Within the detection limits of each of the analyses, the peptide appeared to be completely homogeneous.

Sample Preparation. Peptides were dissolved in 0.5 mL of 50 mM deuterio-Tris and 25 mM deuterioacetic acid in 90% H_2O/10% D_2O or D_2O, with a stoichiometric amount of $ZnCl_2$. Peptide concentrations were determined by UV absorbance (ϵ_{278} = 1300 M^{-1} cm^{-1}) and were usually about 4 mM, pH 5.45, in H_2O and pH 7.45 in D_2O with no correction for the isotope effect.

NMR Measurements. NMR spectra were acquired at 25 °C on a Bruker AM 500 spectrometer. A set of 2D spectra were acquired by using time-proportional phase incrementation (TPPI) in t_1 to achieve F_1 quadrature detection (Marion & Wuthrich, 1983). NOESY (Bodenhausen et al., 1984) spectra were acquired with a mixing time of 150 ms. The TOCSY (Bax & Davis, 1985) experiment used a spin lock field of 10 kHz with a mixing time of 70 ms. DQF-COSY (Rance et al., 1983) used SCUBA (Brown et al., 1988) pulse sequence with t_p = 60 ms immediately after H_2O presaturation in order to recover the signals of NH–αCH correlation. In each case, the

† This work was supported by NIH Grant PO1 GM32681. R.E.K. was supported by an American Heart Association Established Investigatorship.

* To whom correspondence should be addressed.

‡ University of Washington.

§ California Institute of Technology.

[1] Abbreviations: HPLC, high-performance liquid chromatography; 2DNMR, two-dimensional nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy.

0006-2960/91/0430-3365$02.50/0

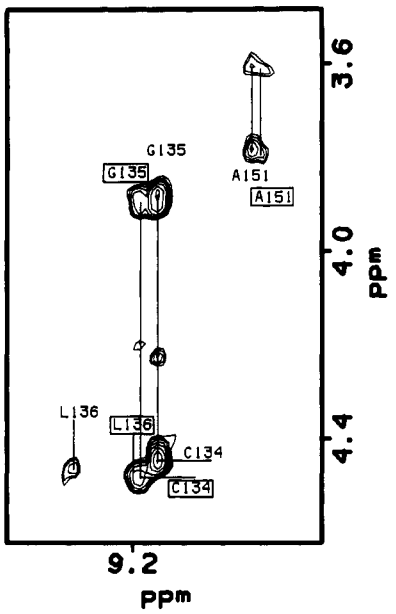

FIGURE 1: Portion of the $d_{\alpha N}$ region of a NOESY spectrum of ADR1a in H_2O (25 °C, pH 5.45), illustrating the doubling of peaks. The minor form is shown in boxes.

spectral width was 6410 Hz, collected into 2K data points in t_2, with 700–800 t_1 experiments acquired. Data processing was carried out by using the software package FTNMR (Hare Research, Woodinville, WA). Chemical shift values are referenced to the H_2O resonance (4.80 ppm at 298 K).

Results and Discussion

Assignments and Secondary Structure of ADR1a. As shown in an earlier report (Parraga et al., 1988), the dispersion of resonances in the spectrum of the C-terminus zinc finger peptide, ADR1a, when reconstituted in the presence of Zn^{2+}, is consistent with a folded structure. Although barely discernible in our original (lower signal-to-noise ratio) spectra of the peptide, some of the resonances in the spectrum are doubled, indicating that the peptide exists in two slowly interconverting species (Figure 1). We have already reported that under certain conditions (i.e., low pH or low Zn^{2+} content) zinc finger peptides interconvert slowly between folded and unfolded states (Parraga et al., 1990). However, the chemical shifts of the extra lines observed in ADR1a in the presence of a stoichiometric amount of Zn^{2+} do not coincide with those for the unfolded peptide. Thus, it appears that the peptide can exist in two folded states. Estimates based on signal intensities show that the major species represents ~70% of the population while the minor species represents ~30%. This distribution was the same at both pH 5.45 and 7.45. It is unlikely that the minor species represents a chemically different population of peptide since a mass spectrometric analysis was consistent with a single species. The smallest observable difference in chemical shift for a doubled resonance in this spectrum is 0.02 ppm (1 Hz), suggesting that the exchange rate between the two forms must be significantly slower than 1 s^{-1}. As well, even at the longest mixing time used (150 ms), no evidence of exchange cross-peaks was observed in a NOESY spectrum, also indicating that the exchange is quite slow.

Resonance assignments were determined by using the conventional approach for homonuclear ^1H 2DNMR. Spin systems were identified from DQF-COSY and TOCSY spectra, and NOESY spectra were then used to obtain sequence-specific resonance assignments. The assignments were completely straightforward and will therefore not be described in detail. It was possible to assign the lines due to the minor form as well as the major form, and complete resonance assignments for both species of ADR1a are given in Table I. Except for specific cases discussed below, identical connectivity patterns (Figure 2) were observed for the two species of ADR1a. A qualitative analysis of the pattern of NOEs observed is consistent with our earlier proposal for the secondary structure of ADR1a. The only regular secondary structure in the domain is a helix starting at residue Asp145, as supported by continuous d_{NN} connectivities from Arg144 to Asn158, $d_{\beta N}$ connectivities from Leu147 to His155, and medium-range NOEs from Asp145 through Gln152.

The presence of two sets of cross-peaks in the ADR1a spectrum will make a quantitative analysis and structure determination quite difficult. By comparing the patterns of NOEs observed for ADR1a with ADR1b [for which a structure has been determined (Klevit et al., 1990)], we can gain insights about the similarities and differences between the two zinc finger domains. The sequential connectivity maps for the two sequences are shown in Figure 2. Not surprisingly, the overall patterns are very similar, confirming that the two sequences fold into similar structures. Residues 102–106 in ADR1b, although in a fairly extended conformation, do not give a classical β-strand connectivity pattern. The corresponding sequence in ADR1a (residues 130–134) contains two proline residues, which cannot adopt the dihedral angles required for a β-strand.

Although the short-range NOEs observed for the helical segments of ADR1a and ADR1b are quite similar, there are differences observed in medium-range NOEs [$d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$, $d_{\beta N}(i, i+3)$, and $d_{\alpha\beta}(i, i+3)$] that suggest that there are significant differences between the α-helices of the two finger domains (Figure 2). In ADR1b, the presence of $d_{\alpha N}(i, i+3)$, $d_{\beta N}(i, i+3)$, and $d_{\alpha\beta}(i, i+3)$ from residue Arg115 to Thr127 indicates a regular α-helix that contains both histidine ligands (Figure 2), and the calculated structure confirms this. The corresponding sequence in ADR1a would be from Asp145 through Ser156. α-Helical medium-range NOEs [i.e., $d_{\alpha N}(i, i+3)$ and $d_{\alpha\beta}(i, i+3)$] are observed only for residues Asp145 through Gln152, although the sequential d_{NN} and $d_{\beta N}$ connectivities continue through Asn158. An obvious difference between ADR1b and ADR1a is that there are three residues between the two Zn^{2+}-binding His residues in ADR1b and four residues between the two His residues in ADR1a. The His-X_3-His spacing allows the imidazole side chains to be oriented for tetrahedral binding in the context of a fairly regular α-helix. On the other hand, if a His-X_4-His sequence were to adopt a regular α-helical conformation, the imidazoles would not both be oriented toward the Zn^{2+}. Additional support for a deviation from an α-helical structure in ADR1a is given by the observation of a strong $d_{\alpha\beta}(i, i+4)$ connectivity between Ala151 and His155.

Two Conformations in ADR1a. A majority of protons in ADR1a yield a single resonance in the NMR spectrum. There are, however, a significant number of protons that give rise to two resonances and, therefore, two sets of cross-peaks in the 2D spectra. While this complicates the analysis of the spectra and also leads to cross-peaks with rather low intensity, all the resonances could be assigned. Figure 3 shows a histogram of the chemical shift differences observed for the two species. The affected residues are not scattered randomly throughout the sequence but are found in three clusters: residues 130–137, residues 145–151, and residue 159. On the basis of the pattern of sequential connectivities described above,

Table I: Spectral Assignments for the Two Forms of ADR1a[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	other
Lys130		4.27	1.84, 1.41	1.34, 1.28	1.61, 1.56	$C^{\epsilon}H$ 2.90
		4.31	1.90	1.46	1.68	$C^{\epsilon}H$ 2.98
Pro131		4.55	2.11	1.92	3.62, 3.48	
		4.47	2.26	1.98	3.68, 3.56	
Tyr132	8.05	5.23	3.00, 2.90		6.95	$C^{\epsilon}H$ 6.87
	8.72	5.23	2.94, 2.76		7.04	$C^{\epsilon}H$ 6.85
Pro133		4.62		2.10, 1.98	3.96, 3.40	
		4.62		2.10, 1.98	3.96, 3.40	
Cys134	8.67	4.44	3.37, 2.72			
	8.72	4.48	3.39, 2.76			
Gly135	9.14	4.23, 3.89				
	9.18	4.21, 3.89				
Leu136	9.32	4.47	1.01, 0.39	1.32	0.72, 0.55	
	9.20	4.47	1.01, 0.45	1.32	0.72, 0.55	
Cys137	7.92	4.96	3.33, 3.30			
	7.95	4.96	3.33, 3.30			
Asn138	8.16	4.84	2.98, 2.89		7.55, 6.85	
	8.16	4.84	2.98, 2.89		7.55, 6.85	
Arg139	8.40	4.12	1.31, 1.24	1.76, 1.70	3.12	$N^{\epsilon}H$ 7.18
	8.40	4.10	1.31, 1.24	1.76, 1.70	3.12	$N^{\epsilon}H$ 7.18
Cys140	7.96	4.59	2.89, 2.68			
	7.96	4.59	2.89, 2.68			
Phe141	8.24	4.91	3.26, 3.02		7.33	$C^{\epsilon}H$ 6.81, $C^{\zeta}H$ 6.17
	8.24	4.91	3.26, 3.02		7.33	$C^{\epsilon}H$ 6.81, $C^{\zeta}H$ 6.13
Thr142	8.91	4.78	4.53	1.42		
	8.91	4.78	4.53	1.42		
Arg143	7.12	4.86	2.19, 1.62	1.78	3.32	$N^{\epsilon}H$ 8.46
	7.12	4.70	2.14, 1.59	1.74	3.30	$N^{\epsilon}H$ 8.52
Arg144	8.51		1.34	1.52, 1.22	3.05	$N^{\epsilon}H$ 7.29
	8.51		1.34	1.52, 1.22	3.05	$N^{\epsilon}H$ 7.34
Asp145	8.70	4.12	2.67, 2.56			
	8.66	4.12	2.67, 2.56			
Leu146	6.91	3.97	1.97, 1.52	1.65	1.14, 0.89	
	6.88	3.97	1.94, 1.52	1.65	1.14, 0.89	
Leu147	6.92	3.31	2.06, 1.48	1.57	1.14, 1.04	
	6.96	3.33	2.10, 1.48	1.57	1.14, 1.04	
Ile148	8.12	3.61	1.80	1.53, 1.23, 0.85	0.75	
	8.12	3.61	1.80	1.53, 1.23, 0.85	0.75	
Arg149	7.68	4.02	1.82	1.74, 1.66	3.17	$N^{\epsilon}H$ 7.22
	7.68	4.02	1.82	1.74, 1.66	3.17	$N^{\epsilon}H$ 7.28
His150	7.52	4.27	3.21, 2.78		7.19	$C^{\epsilon}H$ 7.94
	7.52	4.27	3.21, 2.78		7.16	$C^{\epsilon}H$ 7.97
Ala151	8.94	3.78	1.59			
	8.92	3.80	1.59			
Gln152	8.20	3.81	2.21, 2.03	2.56, 2.36		$N^{\epsilon}H$ 7.39, 6.75
	8.20	3.81	2.21, 2.03	2.56, 2.36		$N^{\epsilon}H$ 7.39, 6.75
Lys153	7.76	4.14	1.88, 1.82	1.48, 1.42	1.67	$C^{\epsilon}H$ 2.97
	7.76	4.14	1.88, 1.82	1.48, 1.42	1.67	$C^{\epsilon}H$ 2.97
Ile154	8.24	4.12	0.92	1.31, 0.98, 0.30	0.68	
	8.24	4.12	0.92	1.31, 0.98, 0.30	0.68	
His155	7.32	5.26	3.40, 3.22		6.47	$C^{\epsilon}H$ 7.89
	7.32	5.26	3.40, 3.22		6.50	$C^{\epsilon}H$ 7.91
Ser156	8.14	4.26	4.04, 3.97			
	8.14	4.26	4.04, 3.97			
Gly157	8.48	4.07, 3.54				
	8.48	4.07, 3.54				
Asn158	8.01	4.65	2.81, 2.72		7.47, 6.97	
	8.01	4.65	2.81, 2.72		7.47, 6.97	
Leu159	7.69	4.28	1.59	1.59	0.89, 0.87	
	7.90	4.20	1.59	1.59	0.90, 0.84	

[a] The upper line for each residue represents the chemical shifts of the major form; the second line represents the minor form. In the $C^{\gamma}H$ listing for isoleucines, the last (most upfield) entry is for the methyl group.

and the three-dimensional structure determined for ADR1b, the first cluster represents the Cys loop, the second cluster represents the α-helix, and the third cluster is the C-terminus.

What is the difference between the two species? As mentioned above, the backbone connectivity patterns followed for the two sets of resonances were almost identical, suggesting that the difference between the two forms is quite subtle. Indeed, chemical shift is such a sensitive indicator of proton environment that it is conceivable that one could observe two sets of resonances and, at least at a qualitative level of analysis, not discern any differences in the observed NOEs. However, when NOEs involving side chain protons were analyzed, several differences were observed, as summarized in Table III. In particular, there are a number of cross-peaks that were observed for one form of the peptide where the corresponding cross-peak was not observed for the other form. Of particular interest are the NOEs involving residues in the Zn^{2+} cluster. In the major species, the $C\epsilon H$ of the imidazole of His155 gives an NOE to one of the $C\beta$ protons of Cys137, while in the minor form, the imidazole proton gives an NOE to the other $C\beta$ proton of Cys137. This leads us to suggest that there is a small but detectable difference in the geometry of the ligands

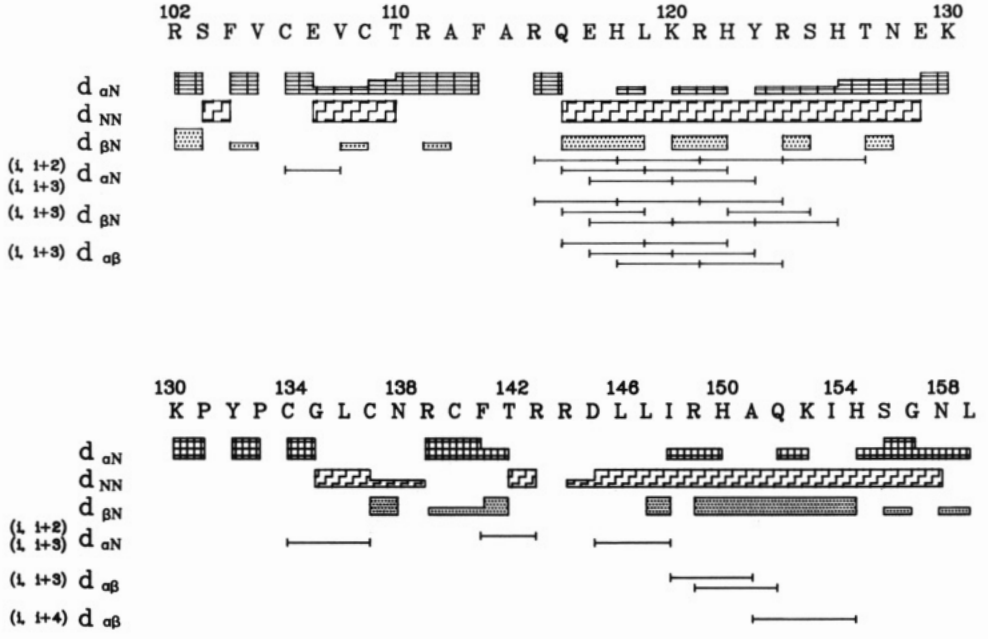

FIGURE 2: NOE connectivities observed for ADR1b (102–130) and ADR1a (130–159). For the residues preceding the prolines, $d_{\alpha\delta}$ instead of $d_{\alpha N}$ is given. The height of the bars reflects the intensities in the NOESY spectrum.

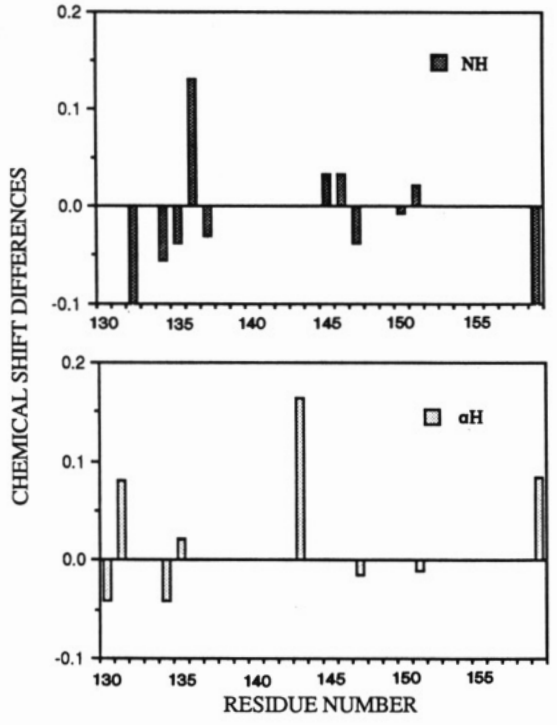

FIGURE 3: Histogram of the chemical shift differences ($\delta_{major} - \delta_{minor}$) observed for the two species in ADR1a.

in the Zn^{2+} clusters of the two forms. A change in the disposition of the cluster could result in small changes in the position of the α-helix relative to the rest of the molecule, thus giving rise to the chemical shift differences observed in the helix. As shown in Table III, several of the different NOEs between the two forms involve Asp145. As well, the CαH of Arg143 has a large chemical shift difference (0.16 ppm) in the two forms. These residues are in the "fingertip" of the motif that connects the Cys loop to the α-helix. Thus, it would appear that a small difference in ligand geometry in the Zn^{2+} cluster causes a detectable perturbation in the environment of the fingertip or vice verce.

The observation of two species in ADR1a leads us to reexamine the spectra of ADR1b. Under all conditions used, the spectra of ADR1b contain only a single set of lines. Such an observation could imply either that ADR1b exists only in a single conformation or that multiple conformations interconvert on a faster time scale than in ADR1a. The mutually exclusive NOEs involving the CϵH of His155 and the two Cβ protons of Cys137 are the clearest difference between the two forms. The analogous residues in ADR1b are His126 and Cys109. As indicated in Table III, NOEs to *both* Cβ protons are observed in ADR1b, but one is stronger than the other. If ADR1b were interconverting rapidly between the two forms, one would expect to see an averaging of the NOEs. Thus, it would appear that ADR1b may exist in a single form, but with a Zn^{2+} cluster geometry that is slightly different from either of the two ADR1a forms.

There are now several other reports of 2DNMR studies on zinc finger peptides in the literature (Lee et al., 1989; Omichinski et al., 1990; Neuhaus et al., 1990). As with ADR1b, the spectra for these peptides also consist of a single set of lines. What could be causing the different behavior in ADR1a? There are two obvious differences in the ADR1a sequence: the difference in the number of residues between the His residues that serve as Zn^{2+} ligands and the presence of two proline residues near the N-terminus of ADR1a (see sequence in Figure 2). Since prolines are known to undergo cis-trans isomerization at a slow but measurable rate in folded proteins and peptides, this could be the cause of the slow exchange behavior observed. However, *cis*-proline gives a distinctive sequential CαH–CαH NOE (Wuthrich, 1986), and such an NOE is not observed in either of the two forms of ADR1a, apparently ruling out this explanation. This leaves the extra residue between the histidines as the most likely cause of the different behavior of this peptide. We are currently preparing "mutant" ADR1a peptides with one residue removed from the His-X_4-His region to test this hypothesis.

It has been noted that in the putative transcription factor ZFY the 13 zinc fingers form a 2-finger repeating pattern in which 2 slightly different classes of zinc finger sequences occur (Weiss et al., 1990). The two consensus sequences for the different classes, with the two differences in these motifs highlighted in boldface, are Ar-X-Cys-X_{2-4}-Cys-$\mathbf{X_3}$-**Ar**-$\mathbf{X_5}$-Leu-X_2-His-$\mathbf{X_3}$-His ("odd") and Ar-X-Cys-X_{2-4}-Cys-$\mathbf{X_1}$-**Ar**-

Table II: Spectral Assignments for the Two Forms of A142I[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	other
Lys130		4.27	1.82, 1.36	1.29	1.62	$C^{\epsilon}H$ 2.93
		4.34	1.90	1.46	1.71	$C^{\epsilon}H$ 3.00
Pro131		4.51	2.09, 1.64	1.90	3.63, 3.48	
		4.46	2.11	1.92	3.68, 3.56	
Tyr132	7.98	5.22	3.00, 2.95		6.95	$C^{\epsilon}H$ 6.88
	8.81	4.51	2.96, 2.82		7.07	$C^{\epsilon}H$ 6.88
Pro133		4.63		2.11, 1.98	3.97, 3.41	
		4.63		2.11, 1.98	3.97, 3.41	
Cys134	8.67	4.49	3.37, 2.70			
	8.75	4.49	3.39, 2.77			
Gly135	9.14	4.26, 3.89				
	9.17	4.21, 3.91				
Leu136	9.32	4.47	1.02, 0.39	1.33	0.72, 0.54	
	9.16	4.46	1.02, 0.46	1.34	0.72, 0.55	
Cys137	7.92	4.94	2.88, 2.68			
	7.95	4.97	3.33, 3.28			
Asn138	8.14	4.83	3.24, 2.98		7.55, 6.84	
	8.16	4.88	2.98, 2.89		7.55, 6.84	
arg139	8.37	4.10	1.29, 1.25	1.75, 1.69	3.12	$N^{\epsilon}H$ 7.18
	8.37	4.13	1.76, 1.40	1.69, 1.23	3.10	$N^{\epsilon}H$ 7.18
Cys140	7.93	4.59	2.88, 2.67			
	7.93	4.59	2.88, 2.67			
Phe141	8.16	4.89	3.24, 2.90		7.32	$C^{\epsilon}H$ 6.82, $C^{\zeta}H$ 6.14
	8.16	4.89	3.24, 2.99		7.32	$C^{\epsilon}H$ 6.82, $C^{\zeta}H$ 6.10
Ile142	8.91	4.35	2.12	1.43, 1.05	0.94	
	8.87	4.50	2.12	1.43, 1.03	0.94	
Arg143	7.14	4.86	2.14, 1.64	1.75, 1.65	3.37, 3.31	$N^{\epsilon}H$ 8.47
	7.17	4.75	2.14, 1.64	1.75, 1.65	3.37, 3.30	$N^{\epsilon}H$ 8.47
Arg144	8.58	4.48	1.33, 0.99	1.50, 1.23	3.06	$N^{\epsilon}H$ 7.30
	8.58	4.33	1.33, 0.99	1.50, 1.24	3.06	$N^{\epsilon}H$ 7.31
Arg145	8.68	4.12	2.66, 2.58			
	8.68	4.12	2.66, 2.58			
Leu146	6.90	4.01	1.98, 1.55	1.67	1.15, 0.90	
	6.90	4.01	1.98, 1.55	1.67	1.15, 0.90	
Leu147	6.85	3.30	2.07, 1.44	1.56	1.11, 1.03	
	6.92	3.26	2.09, 1.44	1.56	1.11, 1.03	
Ile148	8.11	3.60	1.79	1.52, 1.23, 0.85	0.75	
	8.08	3.60	1.79	1.52, 1.23, 0.85	0.75	
Arg149	7.67	4.02	1.80	1.67	3.17	$N^{\epsilon}H$ 7.24
	7.65	4.02	1.80	1.67	3.17	$N^{\epsilon}H$ 7.29
His150	7.50	4.28	3.20, 2.78		7.19	$C^{\epsilon}H$ 7.94
	7.52	4.26	3.19, 2.78		7.14	$C^{\epsilon}H$ 7.98
Ala151	8.96	3.78	1.60			
	8.91	3.78	1.60			
Gln152	8.17	3.80	2.20, 2.03	2.57, 2.36		$N^{\epsilon}H$ 7.39, 6.75
	8.14	3.80	2.20, 2.03	2.57, 2.36		$N^{\epsilon}H$ 7.39, 6.75
Lys153	7.75	4.14	1.89, 1.81	1.48, 1.41	1.66	$C^{\epsilon}H$ 2.96
	7.77	4.14	1.89, 1.81	1.48, 1.41	1.66	$C^{\epsilon}H$ 2.96
Ile154	8.23	4.11	0.92	1.30, 0.98, 0.30	0.68	
	8.25	4.14	0.92	1.30, 0.98, 0.30	0.68	
His155	7.33	5.25	3.41, 3.23		6.49	$C^{\epsilon}H$ 7.88
	7.31	5.25	3.41, 3.23		6.50	$C^{\epsilon}H$ 7.92
Ser156	8.12	4.26	4.04, 3.96			
	8.12	4.26	4.04, 3.96			
Gly157	8.48	4.07, 3.54				
	8.48	4.07, 3.54				
Asn158	8.01	4.66	2.80, 2.72		7.48, 6.97	
	8.01	4.66	2.80, 2.72		7.48, 6.97	
Leu159	7.70	4.27	1.59	1.60	0.89, 0.85	
	7.89	4.20	1.59	1.60	0.89, 0.85	

[a] The upper line for each residue represents the chemical shifts of the major form; the second line represents the minor form. In the $C^{\gamma}H$ listing for isoleucines, the last (most upfield) entry is for the methyl group.

X_7-Leu-X_2-His-X_4-His ("even") (where "Ar" stands for an aromatic residue, and the names "odd" and "even" are derived from the fact that ZFY fingers 1, 3, 5, 7, 9, 11, and 13 belong to one class while fingers 2, 4, 6, 8, 10, and 12 belong to the other). Thus, there are two clear differences between the two subclasses: the placement of the aromatic residue that follows the second Cys and the number of residues between the two histidine residues. Weiss and Keutmann have reported studies in which the aromatic residue in an "even" ZFY finger peptide was placed in the position of the "odd" sequence. The two peptides had indistinguishable properties, as judged by CD and UV spectroscopies (Weiss & Keutmann, 1990). It would therefore appear on the basis of these studies that if the two subclasses of zinc fingers are indeed different, the difference must arise from the number of residues between the two histidines.

If the relevant difference in subclasses is whether there are three or four residues between the two histidines, one can identify other zinc finger proteins with alternating patterns. Thus, if we rename the "odd" pattern as "X_3" and the "even" pattern as "X_4", we note that two known DNA-binding proteins have the following patterns: ADR1 is X_3-X_4 and TFIIIA

Table III: Different NOEs Found in the Two Conformations of ADR1a and Corresponding NOEs Found in ADR1b

major form	minor form	ADR1b
no	NH Asp145-β' Arg144	NH Glu117-β' Gln116
no	NH Asp145-γ' Ile148	no
no	NH Asp145-δ' Leu147	no
no	ϵ His155-β' Cys137	ϵ His126-β' Cys109
ϵ His155-β Cys137	no	ϵ His126-β Cys109
β' His155-α Cys134	no	no
β' Leu147-α Cys134	no	β' Leu119-α Cys109
NH Ile148-β' Leu136	no	no

is X_3-X_3-X_4-X_3-X_3-X_4-X_3-X_4-X_3. If the conformational equilibrium observed in ADR1a is to be attributed to the X_4 sequence, we would expect other X_4 sequences to show similar behavior. Unfortunately, ADR1a is the only X_4 zinc finger peptide to be characterized by 2DNMR to date.[2] Thus, whether this will be a general phenomenon must await further characterization of other X_4 peptides. Furthermore, it is of interest to know whether the conformational flexibility observed for the ADR1a sequence when it exists as an isolated finger is maintained when the sequence is in the context of a larger polypeptide. Preliminary analysis of 2DNMR spectra of a 58-residue peptide that contains both zinc finger sequences of ADR1 indicates that the peak doubling is not observed for most of the ADR1a resonances. The merged ADR1a peaks, in the context of a two-finger fragment, have chemical shifts that are closer to those of the minor species reported here (Xu and Klevit, unpublished results).

As noted earlier, significant chemical shift differences are observed for Leu159, at the C-terminus of the peptide. However, the relative intensities of the Leu159 peaks are different from other paired peaks (closer to 50:50). Therefore, the doubling of peaks from Leu159 may be due to some other, as yet unidentified process.

Effects of a Fingertip Mutation on the Two Conformations of ADR1a. To aid in the process of assigning the spectrum of ADR1a, spectra were also obtained for a point mutant of ADR1a, in which residue 142 was changed from Thr to Ile (A142I). This mutation, in the context of the ADR1 protein, results in a 10-fold decrease in DNA-binding affinity (Blumberg et al., 1987). Complete assignments have been determined for this molecule as well and are given in Table II. The overall spectrum and the sequential connectivities observed for A142I are quite similar to the wild-type peptide. The most striking difference between the two peptides is the observation that the populations of the two forms have changed to 50%/50% for the mutant (Figure 4). Thus, a change at the fingertip of the domain appears either to stabilize the minor conformer or to destabilize the major conformer.

CONCLUSIONS

NMR studies on several isolated single zinc finger domains have begun to give a detailed picture of the structure of these independently folded domains. Since each zinc finger containing protein recognizes and binds to its own cognate DNA, for which there is no consensus sequence, it is likely that small differences in the structures of each zinc finger, brought about by differences in the sequence, are crucial to DNA binding. This was our main impetus for reexamining the C-terminal finger from ADR1. In the process, we discovered that the ADR1a sequence exists as two conformers that are in slow exchange with each other. Such dynamic equilibria do not appear to be a property of the other zinc finger peptides studied to date. The functional significance of these dynamics remains to be identified. We are currently comparing the results for the individual peptides with their behavior in a two-finger peptide, to begin to understand the properties and structures of zinc finger domains in the context of a larger molecule.

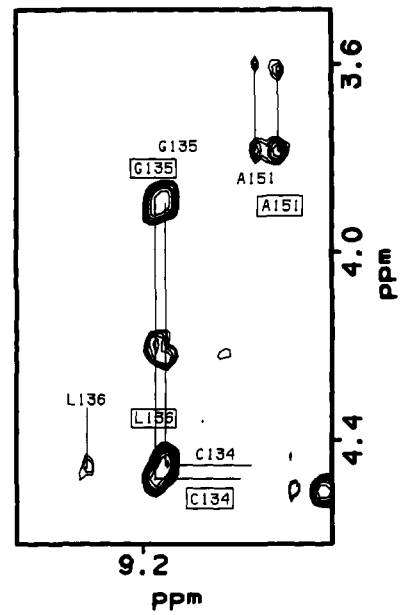

FIGURE 4: Portion of the $d_{\alpha N}$ region of a NOESY spectrum of A142I in H_2O (25 °C, pH 5.45), illustrating the doubling of peaks. This is the same region as that shown in Figure 1. The minor form is shown in boxes.

ACKNOWLEDGMENTS

We thank Dr. Jon Herriott for reading the manuscript.

Registry No. His, 71-00-1; Thr, 72-19-5.

REFERENCES

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Berg, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 99.
Blumberg, H., Eisen, A., Sledziewski, A., Bader, D., & Young, E. T. (1987) *Nature 328*, 443–445.
Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.
Brown, S. C., Weber, P. L., & Muller, L. (1988) *J. Magn. Reson. 77*, 166–169.
Bruist, M. F., Horvath, S. J., Hood, L. E., Steitz, T. A., & Simon, M. I. (1987) *Science 235*, 777–780.
Klevit, R. E., Herriott, J. R., & Horvath, S. J. (1990) *Proteins: Struct., Funct., Genet. 7*, 215–226.
Lee, M. S., Cavanagh, J., & Wright, P. E. (1989) *FEBS Lett. 254*, 159–164.
Marion, D., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Neuhaus, D., Nakaseko, Y., Nagai, K., & Klug, A. (1990) *FEBS Lett. 262*, 179–184.

[2] Weiss and Keutmann state in their recent report on "odd" and "even" ZFY peptides that the 2DNMR spectra of ZFY-6 (an X_4 sequence) "contain a single set of spin systems" (Weiss & Keutmann, 1990). No two-dimensional spectra are shown in this report, but we note in the published one-dimensional spectrum of ZFY-6 the presence of a number of small resonances that could in fact indicate the presence of another species in their sample.

Omichinski, J. G., Clore, G. M., Appella, E., Sakaguchi, K., & Gronenborn, A. M. (1990) *Biochemistry 29*, 9324–9334.
Parraga, G., Horvath, S., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., & Klevit, R. E. (1988) *Science 241*, 1489–1492.
Parraga, G., Horvath, S., Young, E. T., & Klevit, R. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 137–141.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Weiss, M. A., & Keutmann, H. (1990) *Biochemistry 29*, 9808–9813.
Weiss, M. A., Mason, K. A., Dahl, C. E., & Keutmann, H. (1990) *Biochemistry 29*, 5660–5664.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Alternating Zinc Fingers in the Human Male Associated Protein ZFY: 2D NMR Structure of an Even Finger and Implications for "Jumping-Linker" DNA Recognition†

Michel Kochoyan,‡,§ Timothy F. Havel,‡ Dzung T. Nguyen,‖ Charles E. Dahl,‡ Henry T. Keutmann,‖ and Michael A. Weiss*,‡,‖

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114

Received October 17, 1990; Revised Manuscript Received December 26, 1990

ABSTRACT: ZFY, a sex-related Zn-finger protein encoded by the human Y chromosome, is distinguished from the general class of Zn-finger proteins by the presence of a two-finger repeat. Whereas odd-numbered domains and linkers fit a general consensus, even-numbered domains and linkers exhibit systematic differences. Because this alternation may have fundamental implications for the mechanism of protein–DNA recognition, we have undertaken biochemical and structural studies of fragments of ZFY. We describe here the solution structure of a representative nonconsensus (even-numbered) Zn finger based on 2D NMR studies of a 30-residue peptide. Structural modeling by distance geometry and simulated annealing (DG/SA) demonstrates that this peptide folds as a miniglobular domain containing a C-terminal β-hairpin and N-terminal α-helix ($\beta\beta\alpha$ motif). These features are similar to (but not identical with) those previously described in consensus-type Zn fingers (derived from ADR1 and Xfin); the similarities suggest that even and odd ZFY domains bind DNA by a common mechanism. A model of the protein–DNA complex (designated the "jumping-linker" model) is presented and discussed in terms of the ZFY two-finger repeat. In this model *every other* linker is proposed to cross the minor groove by means of a putative finger/linker submotif HX_4HX_3–hydrophobic residue–X_3. Analogous use of a hydrophobic residue in a linker that spans the minor groove has recently been described in crystallographic and 3D NMR studies of homeodomain–DNA complexes. The proposed model of ZFY is supported in part by the hydroxyl radical footprint of the TFIIIA–DNA complex [Churchill, M. E. A., Tullius, T. D., & Klug, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5528–5532].

The Zn-finger motif defines a highly conserved class of eukaryotic nucleic acid binding proteins (Klug & Rhodes, 1987; Evans & Hollenberg, 1988). In this paper we describe 2D NMR[1] studies and structural modeling of an isolated Zn finger from the human male associated protein ZFY. This gene, originally identified from studies of sex reversal in man (de la Chapelle, 1972; Page et al., 1987), encodes a putative transcription factor proposed to participate in spermatogenesis (Palmer et al., 1989; Koopman et al., 1989). Genetic studies of intersex abnormalities of the newborn suggest that ZFY also plays an accessory role in the pathway of male sexual development in embryogenesis (Page et al., 1990).

Two-dimensional NMR studies have previously been conducted of single-finger peptides from *Saccharomyces* protein ADR1 and *Xenopus* protein Xfin (Parraga et al., 1988; Lee et al., 1989a,b). These fingers fold as compact globular minidomains in which the divalent metal is encaged. The N-terminal portion of the finger, containing the conserved cysteines, forms a β-sheet and β-turn (β-hairpin); the C-terminal

†Supported by the Lucille Markey Charitable Trust and by grants from the National Institutes of Health (HD 26465), the American Cancer Society, and the Whitaker Foundation to M.A.W. M.A.W. is supported in part by the Pfizer Scholars Program for New Faculty and a Junior Faculty Research Award from the American Cancer Society. T.H. is supported in part by NIH Grant GM 38221. M.K. is supported in part by a NATO postdoctoral fellowship.

*Address correspondence to this author at the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.
‡Harvard Medical School.
§Permanent address: CNRS URA 1254, BIOP Polytechnique, 91128 Palaiseau, France.
‖Massachusetts General Hospital.

[1] Abbreviations: CD, circular dichroism; DQF-COSY, double-quantum-filtered correlated spectroscopy; DG, distance geometry; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; RMS, root mean square; SA, simulated annealing; TOCSY, total correlation spectroscopy.

0006-2960/91/0430-3371$02.50/0