

(23). We suggest that a second cellular contribution to the unwinding reaction is a factor involved in modulating the functional state of T antigen—changing it from a static DNA binding protein to a mobile helicase. This factor may be the S phase activator.

We realize that we have presented the above model without reference to the cellular equivalent of T antigen—the cellular initiator protein. Cyclic control of initiator protein abundance may not be required to regulate effectively the onset of DNA replication. In support of this, we have demonstrated previously, in a specific model system, that in vivo the initiator protein SV40 T antigen remains stably associated with its template DNA for at least two cell cycles (24). Despite the continuous association of the initiator protein with its template, the model replicon replicated just once per cell cycle. Thus, the onset of replication was not regulated by the physical presence of the initiator protein (that is, T antigen). Similarly, it is possible that the cellular initiator is constitutively present and stably associated with replication origins. When the decision to replicate DNA is made, a factor that is necessary for the initiator to unwind DNA at the origin is synthesized. Only then does DNA replication begin.

ing the indicated amount of extract with activated calf thymus DNA for 10 minutes. Reactions were at least 80% inhibited by Aphidicolin at 10  $\mu\text{g}/\text{ml}$ . Most of the cellular DNA polymerase activity was present in our cytoplasmic extracts. DNA polymerase activity was measured under standard replication conditions except that T antigen was omitted, and 5  $\mu\text{g}$  of activated calf thymus DNA was substituted for p-SVori DNA as template. Activated DNA was prepared by incubating 500  $\mu\text{g}$  of calf thymus DNA with 100 ng of deoxyribonuclease I in 10 mM tris (pH 7.4), 10 mM  $\text{MgCl}_2$  for 10 minutes at 37°C [U. Hubscher, P. Gerschwiler, G. McMaster, *EMBO J.* 1, 1513 (1982)]. Topoisomerase was measured by relaxation of supercoiled DNA. Reaction products were visualized by ultraviolet illumination. One unit of activity relaxes 0.5  $\mu\text{g}$  of supercoiled DNA in 1 hour at 37°C. Visualization of reaction products on gels in saturating amounts of ethidium bromide indicated that the relaxation of supercoiled DNA in our reactions yielded closed circular, not nicked circular, products demonstrating that we were assaying a nicking-closing activity and not a nuclease. Also, relaxation was not inhibited by the addition of EDTA to the reaction. Both S and G<sub>1</sub> extracts contain approximately 0.25 unit of topoisomerase I per microgram of cell protein (as anticipated, most of the cellular topoisomerase I was not extracted in the preparation of our hypotonic cell lysates). Addition of purified PCNA (11) to G<sub>1</sub> reactions did not stimulate replication activity, demonstrating that the abundance of PCNA was not limiting the extent of G<sub>1</sub> replication. Visualization of replication products on agarose gels demonstrated that G<sub>1</sub> reactions did not preferentially accumulate replication intermedi-

ates such as catenated dimers or nicked circles, an indication that neither topoisomerase II nor DNA ligase was limiting.

17. J. Roberts and H. Weintraub, *Cell* 46, 741 (1986).
18. F. Dean *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84, 16 (1987); M. Dodson, F. Dean, P. Bullock, H. Echols, J. Hurwitz, *Science* 238, 964 (1987); A. Kornberg, *J. Biol. Chem.* 263, 1 (1988); H. Stahl, P. Droge, R. Knippers, *EMBO J.* 5, 1939 (1986); D. Bramhill and A. Kornberg, *Cell* 52, 743 (1988); T. Baker, S. Sekimizu, B. Funnell, A. Kornberg, *ibid.* 45, 53 (1986); M. Dodson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 83, 7638 (1986).
19. R. Umek and D. Kowalski, *Cell* 52, 559 (1988).
20. M. Wold, J. Li, T. Kelly, *Proc. Natl. Acad. Sci. U.S.A.* 84, 3643 (1987).
21. M. Fairman and B. Stillman, *EMBO J.* 7, 1211 (1988).
22. H. Stahl and R. Knippers, *J. Virol.* 47, 65 (1983).
23. H. Stahl, P. Droge, H. Zentgraf, R. Knippers, *ibid.* 54, 473 (1985); L. Tack and G. Proctor, *J. Biol. Chem.* 262, 6399 (1987).
24. J. Roberts and H. Weintraub, *Cell* 52, 397 (1988).
25. Supported in part by a grant from the Lucille P. Markey Charitable Trust (J.R.) and by NIH training grant AG0005 (G.D.). We thank B. Stillman for helping us begin these experiments and for supplying 293 cells and purified PCNA, T. Kelly and M. Wold for the purified human SSB (RF-A), B. Forrester for introducing us to centrifugal elutriation, and numerous colleagues for advice and suggestions during the course of this work.

29 April 1988; accepted 28 July 1988

## Zinc-Dependent Structure of a Single-Finger Domain of Yeast ADR1

GRACE PÁRRAGA, SUZANNA J. HORVATH, ARRI EISEN, WAYNE E. TAYLOR, LEROY HOOD, ELTON T. YOUNG, RACHEL E. KLEVIT

**In the proposed "zinc finger" DNA-binding motif, each repeat unit binds a zinc metal ion through invariant Cys and His residues and this drives the folding of each 30-residue unit into an independent nucleic acid-binding domain. To obtain structural information, we synthesized single and double zinc finger peptides from the yeast transcription activator ADR1, and assessed the metal-binding and DNA-binding properties of these peptides, as well as the solution structure of the metal-stabilized domains, with the use of a variety of spectroscopic techniques. A single zinc finger can exist as an independent structure sufficient for zinc-dependent DNA binding. An experimentally determined model of the single finger is proposed that is consistent with circular dichroism, one- and two-dimensional nuclear magnetic resonance, and visual spectroscopy of the single-finger peptide reconstituted in the presence of zinc.**

**T**HE ZINC FINGER MOTIF IS PRESENT in several transcription regulatory proteins (1, 2). The model for this protein domain was first proposed based on sequence analysis (3, 4), partial proteolysis (3), and zinc content (3) of *Xenopus* transcription factor TFIIIA. ADR1 is a positive transcription activator of the glucose-repressible alcohol dehydrogenase gene (*ADH2*) in the yeast *Saccharomyces cerevisiae* (5). The *ADR1* gene codes for a 1323-residue protein, and sequence analysis has revealed two zinc finger domains between residues 100 and 160 in the protein (6). In a genetic analysis of *ADR1* (7), point muta-

tions producing null alleles were clustered in the zinc finger sequences and also indicated that specific conserved residues within its two adjacent zinc finger domains were essential for protein function. Recently, an ADR1- $\beta$ -galactosidase (ADR1- $\beta$ -gal) fusion protein containing the first 229 residues of ADR1 was shown to footprint both upstream activator sequence 1 (UAS1) and upstream activator sequence 2 (UAS2) of

G. Párraga, A. Eisen, W. E. Taylor, E. T. Young, R. E. Klevit, Department of Biochemistry SJ-70, University of Washington, Seattle, WA 98195.  
S. J. Horvath and L. Hood, Division of Biology, California Institute of Technology, Pasadena, CA 91125.

### REFERENCES AND NOTES

1. L. Hartwell, R. Mortimer, J. Culotti, M. Culotti, *Genetics* 74, 267 (1973); L. Hartwell, J. Culotti, J. Pringle, B. Reid, *Science* 183, 46 (1974); L. Herford and L. Hartwell, *J. Mol. Biol.* 84, 445 (1974).
2. A. Pardee, *Proc. Natl. Acad. Sci. U.S.A.* 71, 1286 (1974); A. Yen and A. Pardee, *Exp. Cell Res.* 116, 103 (1978); M. Mendenhall, C. Jones, S. Reed, *Cell* 50, 927 (1987); V. Simanis and P. Nurse, *ibid.* 45, 261 (1986).
3. C. Thompson, P. Challoner, P. Neiman, M. Groudine, *Nature* 314, 363 (1985); J. Sherley and T. Kelly, *J. Biol. Chem.* 263, 8350 (1988).
4. R. Baserga, *Exp. Cell Res.* 151, 1 (1984); *The Biology of Cell Reproduction* (Harvard Univ. Press, Cambridge, MA, 1985); C. Thompson, P. Challoner, P. Neiman, M. Groudine, *Nature* 319, 374 (1986); D. Prescott, *Intern. Rev. Cytol.* 100, 93 (1987).
5. P. Farnham and R. Schimke, *J. Biol. Chem.* 260, 7675 (1985).
6. N. Heintz and R. Roeder, *Proc. Natl. Acad. Sci. U.S.A.* 81, 2713 (1984).
7. L. Johnston, J. White, A. Johnson, G. Lucchini, P. Plevani, *Nucleic Acids Res.* 15, 5017 (1987).
8. B. Beyers, unpublished observations.
9. J. Li and T. Kelly, *Proc. Natl. Acad. Sci. U.S.A.* 81, 6973 (1984); B. Stillman and Y. Gluzman, *Mol. Cell. Biol.* 5, 2051 (1985).
10. Y. Murakami *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 83, 2869 (1986).
11. G. Prelich, M. Kostura, D. Marshak, M. Mathews, B. Stillman, *Nature* 326, 517 (1987).
12. L. Yang, M. Wold, J. Li, T. Kelly, L. Liu, *Proc. Natl. Acad. Sci. U.S.A.* 84, 950 (1987).
13. R. Wobbe *et al.*, *ibid.*, p. 1834.
14. M. Meistrich, in *Cell Separation: Methods and Selected Applications*, T. Pretlow II and T. Pretlow, Eds. (Academic Press, New York, 1983), vol. 2, p. 33.
15. N. DeTerra, *Proc. Natl. Acad. Sci. U.S.A.* 57, 607 (1967).
16. Activity of  $\alpha$  and  $\delta$  DNA polymerases in G<sub>1</sub> and S phase Manca cell extracts was measured by incubat-

*ADH2* (8). However, a fusion protein of the first 150 residues (lacking the last His  $Zn^{2+}$  ligand) did not activate transcription of *ADH2* nor did it bind to DNA (9). An 89-residue fragment containing the minimal sequences for the three zinc fingers in another yeast transcription regulatory protein (*SWI5*) can footprint DNA (10). In the absence of detailed structural information about the zinc finger domains themselves, this type of deletion and genetic information has been some of the strongest evidence in support of the zinc finger model.

Our approach to this structural problem involved first chemically synthesizing a single zinc finger (30-residue) peptide by the stepwise solid-phase method (11, 12). The peptide ADR1a (Fig. 1) contains residues 130 to 159 of the protein. ADR1a was purified by standard reversed-phase high-performance liquid chromatography (HPLC) techniques on a C4 column; pep-

tide composition was confirmed by amino acid composition and sequence analysis. The lyophilized, reduced, and purified peptide was then reconstituted in a tris-buffered zinc solution.

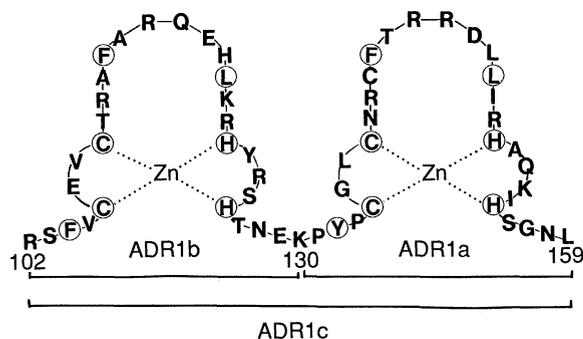
We needed to confirm that (i) metal binding occurred and that (ii) metal binding drove the folding of the peptide into a single major, stable conformation. A visual spectrum of the  $Co^{2+}$  form of ADR1a contained absorbance maxima indicative of tetrahedral coordination of the metal ion (635 nm) and of thiol-containing ligands (325 nm) and was similar to previous results with another zinc finger peptide (13). The  $Co^{2+}$  absorption band disappeared when a slight molar excess of  $ZnCl_2$  was added, suggesting that  $Zn^{2+}$  occupies the same metal pocket and ligand geometry as the  $Co^{2+}$  metal ion. These results also agree with an extended x-ray absorption fine structure (EXAFS) study of TFIIIA (14) that indicated that the  $Zn^{2+}$

metal ion was tetrahedrally coordinated by His and Cys residues.

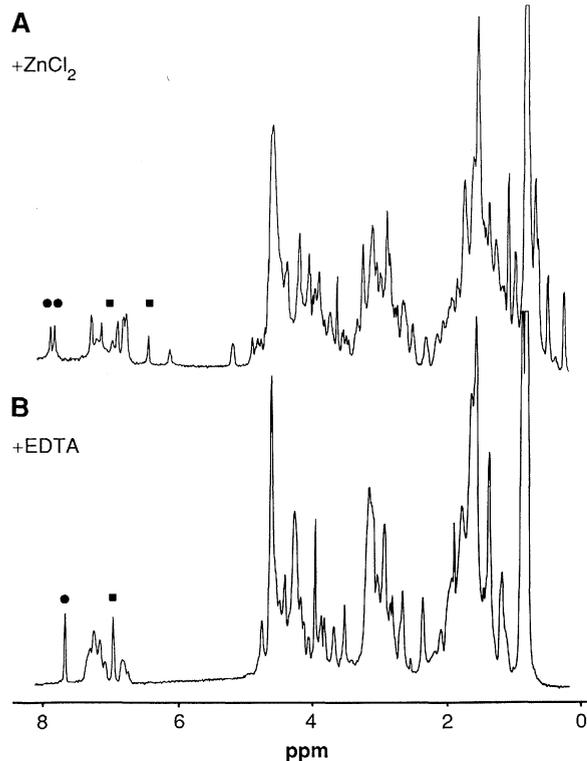
Given that ADR1a binds zinc in a tetrahedral geometry, we asked whether zinc binding was coincident with a single major conformational form of the peptide. This question was assessed by one-dimensional (1D)  $^1H$  nuclear magnetic resonance (NMR); NMR spectra of ADR1a in the presence and absence of  $Zn^{2+}$  are compared in Fig. 2. In Fig. 2A the aromatic region of the spectrum shows considerable rearrangement of aromatic proton resonances. For example, His C2H and C4H have shifted and are inequivalent in the zinc form, suggesting that His residues are involved in zinc binding or that a major conformational change in the peptide results in ring-current shifts of the His residues. The aliphatic region of the spectrum exhibits major rearrangement as well; in particular, several methyl proton resonances are shifted upfield from the main methyl proton cluster at 0.9 ppm. These chemical shifts are compatible with a major conformational change in the peptide as a consequence of  $Zn^{2+}$  binding. The two spectra have similar line widths, indicating that zinc is not promoting any intermolecular association. Thus, even at the millimolar concentrations used for the NMR experiments, the peptide probably exists as a monomer. Furthermore, 2D and 1D NMR experiments were performed at various peptide concentrations (0.5 to 6 mM), and no changes dependent on peptide concentration were found in either the 1D or 2D spectrum. NMR spectra obtained for the peptide in buffers lacking zinc or in 0.5% acetic acid and 5 mM  $ZnCl_2$  (in which His residues are protonated and unable to bind zinc), were identical to spectra obtained in the presence of EDTA. Given that the folded domain exhibits slow exchange behavior (15), the absence of minor peaks in Fig. 2 indicates that only one major conformational form of the peptide exists under conditions where it is folded in the presence of zinc. This form is stable from pH 4.5 to 8, as judged by 1D NMR.

Circular dichroism (CD) spectroscopy was used to monitor secondary structural changes in ADR1a upon metal binding (Fig. 3). The CD spectrum of ADR1a reconstituted in the presence of EDTA (curve a) differs from that for ADR1a refolded in  $ZnCl_2$  solutions (curve b). The large increase in negative ellipticity at 222 and 208 nm, shown for ADR1a refolded in  $ZnCl_2$ , indicates that the peptide contains  $\alpha$ -helical structure. The fraction of helix induced under these conditions was calculated (16) to be 32%, or about 10 of 30 residues, which would correspond to approximately three turns of an  $\alpha$  helix. This large increase in

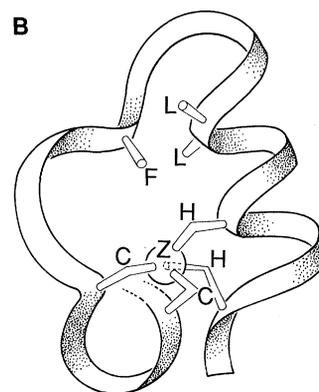
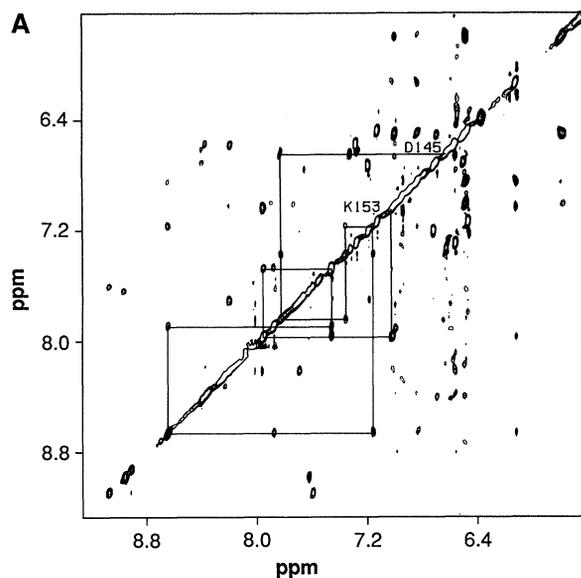
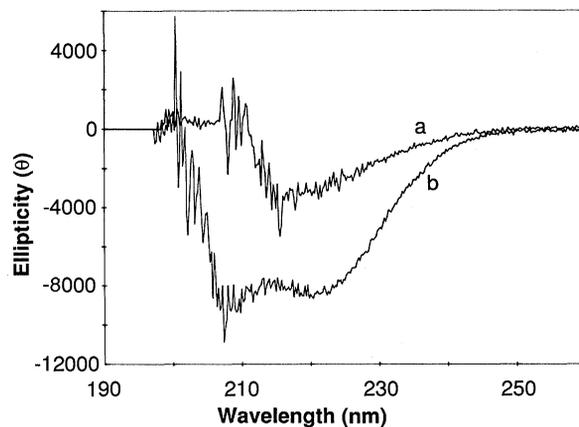
**Fig. 1.** Amino acid sequence of ADR1 zinc fingers and the sequence of ADR1a (22). The entire sequence is located between residues 102 and 159 in the ADR1 protein and is drawn according to the model proposed by Miller *et al.* (3). The sequence of the peptide ADR1a is underlined. Invariant residues are circled.



**Fig. 2.** 500-MHz  $^1H$  NMR spectra of ADR1a in the presence and absence of zinc. Spectra are shown of 1 mM ADR1a reconstituted in the presence of (A) 1.1 mM  $ZnCl_2$  and 50 mM tris and (B) 1 mM EDTA and 50 mM tris: (●), His C2H; and (■), His C4H. The peptide was synthesized by the stepwise solid-phase method and was reconstituted at pH 7.8 in the presence of  $ZnCl_2$  or EDTA in 50 mM tris. Proton NMR spectra of  $D_2O$  solutions were obtained on a Bruker WM500 at 30°C with a spectral width of 6410 Hz, acquisition time of 0.693 s, and a relaxation delay of 1 s.



**Fig. 3.** Circular dichroism spectra of ADR1a. CD experiments were conducted at 22°C on an ON-LINE SYSTEMS (Jefferson, Georgia)-modified Cary 61 spectropolarimeter. Peptide concentrations were 0.2 mM, and spectra were obtained in a 0.1-cm path length cuvette in which the peptide was reconstituted in the presence of: (a) 1 mM EDTA and 50 mM tris and (b) 5 mM ZnCl<sub>2</sub> and 50 mM tris (23).



**Fig. 4.** (A) NOESY spectrum of ADR1a peptide indicating NH-to-NH connectivities. The downfield region of a phase-sensitive NOESY (24) spectrum was obtained for 4 mM ADR1a refolded in 5 mM ZnCl<sub>2</sub>, 50 mM deuterated tris in 90% H<sub>2</sub>O–10% D<sub>2</sub>O (v/v), pH

4.5. The longest segment of sequential NH-NH connectivities was followed and assigned as: LLIRHAQK (22). (B) Ribbon model of single zinc finger domain (ADR1a) incorporating tetrahedral coordination of zinc by conserved Cys residues 134 and 137 and His residues 150 and 155 and  $\alpha$ -helical conformation between Leu<sup>146</sup> and Lys<sup>153</sup>, as suggested by the 2D NOESY spectrum. The conserved residues Leu<sup>147</sup>, His<sup>150</sup>, and Phe<sup>141</sup>, as well as Leu<sup>146</sup>, interact in the central core of the domain.

helical content is consistent with the proposal that metal ion binding drives the overall folding of the zinc finger domain. This large structural change is in contrast with the monomer structure (17); in this case, metal ions do not alter the global folding for the entire protein, but promote dimerization. Our observation that zinc binding promotes  $\alpha$ -helix formation in ADR1a is consistent with CD data on a reconstituted single-finger fragment from TFIIIA (13).

Given that we have a single, stable structure that binds zinc, a 2D NMR data set sufficient to complete the sequential assignments of the domain was obtained. To confirm the presence of  $\alpha$ -helical structure in the zinc form and to identify its location in the primary structure, 2D NMR nuclear Overhauser exchange spectroscopy (NOESY) spectra were examined. Sequential, nearest-neighbor NH-to-NH nuclear Overhauser effects (NOEs) are indicative of

$\alpha$ -helical conformation (18). In addition, NOEs between C $\beta$  protons and NHs between nearest-neighbor residues ( $i$  to  $i + 1$ ), and those three residues apart in the sequence ( $i$  to  $i + 3$ ), strengthen the interpretation that an  $\alpha$  helix is present. The sequential NH-NH connectivities for the zinc form of the ADR1a peptide from a phase-sensitive experiment are shown in Fig. 4A. The segments involved have been assigned with spin system assignments from  $J$ -correlated spectroscopy (COSY), relayed coherence transfer spectroscopy (RELAY), and total coherence (TOCSY) spectra. The longest NH-NH segment is for a contiguous sequence of the peptide (Fig. 4A) from Leu<sup>146</sup> through Lys<sup>153</sup>, which would fit about three turns of an  $\alpha$  helix. We also observed NOEs between C $\beta$  protons and NHs among nearest-neighbor residues ( $i$  to  $i + 1$  and  $i$  to  $i + 3$ ) for this sequence; NH-NH NOEs also were observed for the sequence Gly<sup>135</sup>

to Cys<sup>137</sup>, suggesting that these residues are involved in some form of a turn. Cys<sup>134</sup> and Cys<sup>137</sup> are possible zinc ligands, and a turn may be necessary in order to accommodate tetrahedral coordination of the zinc ion.

A ribbon model for the ADR1a structure that incorporates the tetrahedral organization of conserved His and Cys ligands to one zinc metal ion is shown in Fig. 4B. The  $\alpha$  helix is shown between residues Leu<sup>146</sup> and Lys<sup>153</sup>, as is a turn between Cys zinc ligands. Another specific feature of the ribbon model, which is supported by 1D and 2D NOEs, is a hydrophobic core that contains the three conserved residues Phe<sup>141</sup>, Leu<sup>147</sup>, and His<sup>150</sup>, as well as Leu<sup>146</sup>. Moreover, an examination of the helix shows that certain residues that are charged (Arg<sup>149</sup> and Lys<sup>153</sup>) or capable of hydrogen bonding (Gln<sup>152</sup>) would be located on one side of the helix facing away from the hydrophobic core. This organization may be important in facilitating DNA binding. The placement of an  $\alpha$  helix in this region agrees with a previous prediction (19) based upon Chou-Fasman analysis (20) of amino acid sequences of several zinc finger proteins, including ADR1. A model similar to this ribbon model has been proposed (21), based upon the comparison of crystal structures of metalloproteins with similar primary sequences proximal to the ligand site as found in the consensus zinc finger primary sequence. However, the low-resolution structure proposed here is based on the 2D NMR NOESY and 1D NMR NOEs of the folded single-finger peptide from ADR1. Our model differs from a previous one (21) in that we have no evidence for the antiparallel  $\beta$  turn depicted in the Cys-Cys region of the structure. The helix interpretation is consistent with our NOEs, however.

We have shown that an ADR1- $\beta$ -gal fusion protein (containing residues 17 to 229, with both zinc fingers of ADR1 intact) binds *ADH2* promoter elements sequence specifically in a zinc-dependent manner and yields deoxyribonuclease I (DNase I) and Fe(II)-MPE (methidium propyl-EDTA) footprints (8). DNA filter-binding and DNase I footprinting experiments were performed with ADR1a under conditions such that the fusion protein showed sequence-specific DNA binding. Unlike the ADR1- $\beta$ -gal fusion protein, ADR1a showed no sequence-specific DNA binding. Although the 30-residue peptide is stably folded, it did not participate in UAS-specific DNA binding, but its affinity for DNA was increased when reconstituted with zinc. DNase footprint analysis of a TFIIIA single-finger fragment also indicated that this domain was not sufficient for sequence-specific binding (13). Nagai *et al.* (10) suggest that this result

indicates incorrect folding, although we know of no evidence to indicate that a correctly folded single zinc finger is capable of sequence-specific binding. Indeed, point mutations in ADR1 (7) indicated that both fingers were essential for protein activity. In addition, a peptide fragment containing three zinc fingers (10) has been shown to footprint DNA. Taken together, these peptide-fragment DNA-binding results might suggest that multiple fingers are required to contribute the correct number and type of base contacts for sequence-specific binding to occur, or that a single-finger domain requires a larger protein context in order to maintain the proper geometry for specific DNA binding. Further analysis of the 2D NMR spectra of the single- and double-finger peptides should allow us to determine the complete structure of this nucleic acid-binding motif, and should be an important step toward an understanding of how multiple fingers bind DNA.

#### REFERENCES AND NOTES

1. A. Klug and D. Rhodes, *Trends Biochem. Sci.* **12**, 464 (1987).
2. R. M. Evans and S. M. Hollenberg, *Cell* **52**, 1 (1988).
3. J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **4**, 1609 (1985).
4. R. S. Brown, C. Sander, P. Argos, *FEBS Lett.* **186**, 271 (1985).
5. C. L. Denis and E. T. Young, *Mol. Cell. Biol.* **3**, 360 (1983).
6. T. A. Hartshorne, H. Blumberg, E. T. Young, *Nature* **320**, 283 (1986).
7. H. Blumberg, A. Eisen, A. Sledziewski, D. Bader, E. T. Young, *ibid.* **328**, 443 (1987).
8. A. Eisen, W. E. Taylor, H. Blumberg, E. T. Young, unpublished results.
9. S. Thrukal, W. E. Taylor, A. Eisen, E. T. Young, unpublished results.
10. K. Nagai, Y. Nakaseko, K. Nasmyth, D. Rhodes, *Nature* **332**, 284 (1988).
11. D. Roise, S. J. Horvath, J. M. Tomich, J. H. Richards, G. Schatz, *EMBO J.* **5**, 1327 (1986).
12. M. F. Bruist, S. J. Horvath, L. E. Hood, T. A. Steitz, M. I. Simon, *Science* **235**, 777 (1987).
13. A. D. Frankel, J. M. Berg, C. O. Pabo, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4841 (1987).
14. G. P. Diakun, L. Fairall, A. Klug, *Nature* **324**, 698 (1986).
15. Spectra of the folded form of ADR1a have been obtained from 15° to 35°C. No significant temperature-dependent chemical shifts have been observed, indicating that the system is in slow exchange. Furthermore, in samples containing a mixture of unfolded and folded species (oxidized and reduced species), separate resonances were observed for each form of the peptide, again indicating slow exchange behavior.
16. C. T. Chang, C. H. C. Wu, J. T. Yang, *Anal. Biochem.* **91**, 13 (1978).
17. A. D. Frankel, D. S. Brecht, C. O. Pabo, *Science* **240**, 70 (1988).
18. M. Billeter, W. Braun, K. Wüthrich, *J. Mol. Biol.* **155**, 321 (1982).
19. R. S. Brown and P. Argos, *Nature* **324**, 215 (1986).
20. P. Y. Chou and G. D. Fasman, *Biochemistry* **13**, 222 (1974).
21. J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 99 (1988).
22. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro;

- Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
23. Raw data were obtained and are shown without any curve smoothing with a resolution of 0.2 nm per data point; each data point represents the average of 40 samples. Ellipticity values at 222 nm were converted to mean molar residue ellipticity values with the equation  $[\theta]_{222} = \theta M_r / (100Lc)$ , where  $\theta$  = degrees,  $L$  = cell path length (dm),  $c$  = concentration (g/ml), and  $M_r$  = mean residue molecular weight calculated from the amino acid sequence. Fraction of helix  $f_H$  was calculated as  $f_H = [\theta]_{222} / ([\theta]_{222}^H (1 - k/n))$ , where  $[\theta]_{222}^H = -37,400$ ,  $k = 2.5$ , and  $n = 9$  [ $k$  is the chain length dependence factor, and  $n$  is the number of residues in a typical  $\alpha$ -helical segment

(16)].

24. NOESY spectra have been obtained in the phase-sensitive mode with mixing times of 200, 300, 400, and 600 ms to assess the degree of spin diffusion. Spectra obtained for samples in 90% H<sub>2</sub>O solutions were acquired with a presaturation pulse of 2.0 s.
25. We thank B. R. Reid and G. Drobny and their groups for use of NMR facilities; H. Charbonneau, S. Kumar, M. Harrylock, R. Wade, and K. Walsh for assistance with amino acid compositional analysis and for sequencing the peptide; and our colleagues D. Allison, J. Herriott, and B. M. Shapiro for their helpful comments on the manuscript.

15 June 1988; accepted 12 August 1988

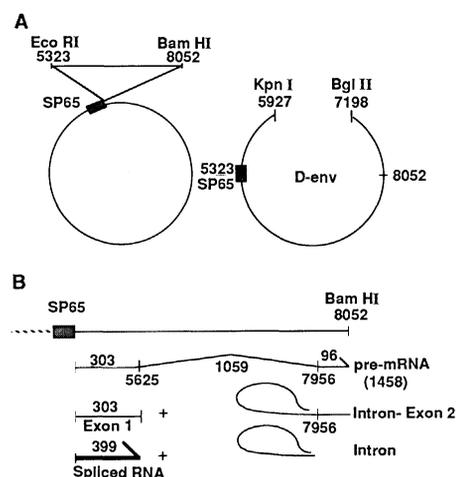
## Virus-Specific Splicing Inhibitor in Extracts from Cells Infected with HIV-1

DELIA GUTMAN AND CARLOS J. GOLDENBERG\*

Human immunodeficiency virus type 1 (HIV-1), in contrast with most other retroviruses, encodes trans-regulatory proteins for virus gene expression. It is shown in this study, by means of an *in vitro* splicing system, that nuclear extracts obtained from cells infected with HIV-1 contain a factor (or factors) that specifically inhibits splicing of a synthetic SP6/HIV pre-messenger RNA (pre-mRNA)-containing donor and acceptor splice sites in the coding region for the envelope protein. It is also shown that the SP6/HIV pre-mRNA is not capable of assembly in a ribonucleoprotein complex, spliceosome, in extracts from infected cells. These findings raise the possibility that specific inhibition of pre-mRNA splicing in the envelope protein coding region by HIV-1 trans-regulatory factors might be one control mechanism for efficient production of structural viral proteins and virion assembly.

THE HUMAN IMMUNODEFICIENCY virus type 1 (HIV-1) contains, in addition to genes *gag*, *pol*, and *env*, at least five other genes termed *vif*, *vpr*, *nef*, *tat-3*, and *rev* (1). The *tat-3* and *rev* genes are essential for virus expression (2-10), but their precise functions remain to be defined. The *tat-3* gene codes for a protein of 86 amino acids (9, 10) that interacts directly or indirectly with sequences located downstream from the HIV-1 initiation site for transcription (TAR element) (11). Several studies with viral deletion mutants suggest that *tat-3* is required for translation of viral mRNAs (8, 12); it appears to function by increasing the steady-state level of mRNA by direct transcriptional activation (3, 13), antitermination (14), or mRNA stabilization (3-5, 15). The product of the *rev* gene is a protein of 116 amino acids (2, 7, 8, 10). It may also operate at several levels, it may relieve a translational block of *gag* and *env* proteins (7), affecting the relative abundance of viral mRNAs (8), and it may also have a negative trans-regulatory role in transcription (16). Feinberg *et al.* (8) reported an abnormal pattern of HIV-1 viral mRNAs in

monkey (COS) cells transfected with *rev*-defective mutants. The levels of the 9-kb (*gag/pol*) and 4.3-kb (*env*) mRNAs were very low and that of the 2-kb (*tat-3/rev*)



**Fig. 1.** Cloning of HIV-*env* DNA into SP65 vectors. (A) The HIV-1 DNA Eco RI-Bam HI fragment (nucleotides 5323 to 8052) was cloned into the SP65 vector. This DNA was isolated and sequences between Kpn I-Bgl II (5927 to 7198) were deleted. (B) The expected RNAs from *in vitro* splicing of SP65/HIV pre-mRNA transcribed from the D-*env* DNA linearized at the Bam HI site (line with stippled box) are schematically represented. The numbers above the lines represent the nucleotide length of the RNAs.

Department of Microbiology and Immunology, University of Miami School of Medicine, Miami, FL 33101.

\*To whom correspondence should be addressed.