posed trenches of Latona. In formulating their model, Sandwell and Schubert (*8, 9*) analyzed the topography of the eastern part of the southernmost trench. The full-resolution, mosaicked image data record for this region (FMIDR 25S, 174 at 75 m per pixel) shows that concentric fractures extend from the proposed inner topographic high across the trench to the outer high and seem related to the formation of all three proposed trenches. Radial fractures are poorly developed here, and their temporal relation to the concentric fractures is unclear. However, a few radial fractures can be traced across the proposed plate boundary, and overall there is no change in fracture density across this boundary.

20. R. J. Phillips *et al.*, *J. Geophys. Res.* **97**, 15923 (1992).

21. Supported by the National Aeronautics and Space Administration Planetary Geology and Geophysics Program under grant NAGW-2915 to Southern Methodist University and grant NAGW-3024 to Washington University. We thank J. W. Goodge, G. E. McGill, and an anonymous reviewer for their comments and suggestions.

1 December 1992; accepted 18 February 1993

Molecular Basis for Specific Recognition of Both RNA and DNA by a Zinc Finger Protein

Karen R. Clemens, Veronica Wolf, Steven J. McBryant, Penghua Zhang, Xiubei Liao, Peter E. Wright,* Joel M. Gottesfeld*

Transcription factor IIIA (TFIIIA) from *Xenopus* oocytes binds both the internal control region of the 5*S* ribosomal RNA genes and the 5*S* RNA transcript itself. The nucleic acid binding domain of TFIIIA contains nine tandemly repeated zinc finger motifs. A series of precisely truncated forms of this protein have been constructed and assayed for 5*S* RNA and DNA binding. Different sets of zinc fingers were found to be responsible for high affinity interactions with RNA and with DNA. These results explain how a single protein can exhibit equal affinities for these two very different nucleic acids.

Developing *Xenopus* oocytes accumulate massive quantities of ribosomal RNAs (rRNAs). Transcription factor IIIA (TFIIIA) serves both as a positive transcription factor for the synthesis of 5S RNA (*1*) and as a storage protein for 5S RNA prior to the assembly of ribosomes during the late stages of oogenesis (*2*). Since the discovery of the zinc finger nucleic acid–binding motif in TFIIIA (*3*), the mode of interaction of this protein with both DNA and RNA has been the subject of intensive investigation. TFIIIA contains a tandem arrangement of nine zinc finger motifs; each motif contains approximately 30 amino acids (*3, 4*) (Fig. 1) with invariant cysteine and histidine residues that coordinate a single zinc atom (*5*). A COOH-terminal 10-kD domain is required for the transcriptional activity of the protein (*6*); however, this domain does not interact directly with either 5S RNA (*7*) or DNA (*6, 8*) but rather is responsible for protein-protein interactions with other components of the transcription complex (*8*).

The DNA binding site for TFIIIA is contained within the 5S RNA coding sequence and consists of three short promoter elements: A 5′ A-block, an intermediate element, and a 3′ C-block (*9*). With the use of a series of recombinant fragments of TFIIIA, we have shown that the first three NH_2-terminal zinc fingers bind with full specificity to the 3′ C-block promoter element of the gene and contribute 95% of the binding energy for DNA (*10*). This interaction occurs through base-specific major groove and phosphate contacts (*10, 11*). Finger 5 binds in the major groove of the intermediate promoter element and fingers 7 to 9 bind in the major groove of the promoter A-block. Fingers 4 and 6 each bind across the minor groove, spanning these promoter elements (*10, 12*). In contrast to our knowledge of the mode of interaction of TFIIIA with DNA, relatively little is known about the mechanisms of interaction of TFIIIA with 5S RNA. RNA footprinting and chemical-crosslinking experiments (*13*) have been performed to map the binding site for TFIIIA on 5S rRNA and the results of these experiments show that TFIIIA interacts with a large portion of the 5S RNA molecule. In addition, site-specific mutagenesis studies have shown that base changes that disrupt the secondary structure of 5S RNA impair TFIIIA binding (*14*). However, the location of specific zinc fingers on 5S RNA is unknown. Because TFIIIA exhibits comparable affinities for both the 5S gene promoter and 5S rRNA (dissociation constant $K_d \cong 1$ nM) (*7, 14*), this raises the question of whether the same or different zinc fingers are utilized for high affinity binding to DNA and 5S rRNA.

To identify the zinc fingers of TFIIIA that are responsible for binding 5S rRNA we have generated two series of polypeptides that contain precise deletions of zinc fingers from either the COOH-terminus or both termini of the protein (Fig. 1). Zinc finger deletions were generated from the full-length cDNA for TFIIIA (*4*) by polymerase chain reaction (PCR) cassette mutagenesis with the use of a series of specific oligonucleotide primers (*10, 15*). The PCR products were cloned into the expression vector pRK172 (*15*) and the truncated polypeptides were purified from *Escherichia coli* cell lysates by heparin-Sepharose ion-exchange chromatography (*10*). The amino acid sequence of these polypeptides was deduced from the coding sequence in the expression plasmids. A composite sequence, indicating the start and stop amino acids of each of the polypeptides, is shown (Fig. 1). Each of these polypeptides was tested for 5S RNA binding activity with radiolabeled RNA synthesized in vitro with T7 RNA polymerase and a synthetic 5S RNA gene template (*16*). Binding was assayed by gel electrophoresis on nondenaturing polyacrylamide gels (*10*) with increasing amounts of each of the polypeptides in separate reactions. Figure 2 provides examples of binding experiments which show the contrast between high affinity and low affinity binding. Figures 3A and 3B provide a graphical representation of the data for the complete set of zinc finger polypeptides. Dissociation constants determined from the protein titrations are given (Table 1).

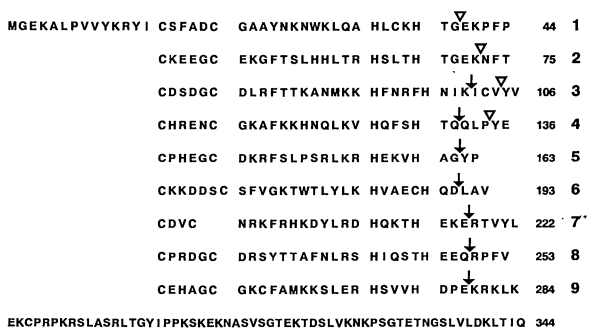

Fig. 1. Amino acid sequence of TFIIIA (*4*) with start (denoted with open triangles) and stop (denoted with arrows) positions of each of the truncated polypeptides. The start of each polypeptide was either the natural NH_2-terminus for the COOH-terminal deletions (zf1-N, where N = 3 through 9) or a Met before the indicated start amino acid for the NH_2-terminal deletions (zfN-7, where N = 2 through 5). The COOH-terminal amino acid for the latter series was Glu^{217}. The gaps within the sequence indicate regions of the zinc fingers, the loop between zinc coordinating cysteines, the finger region, the histidine loop, and the linkers between fingers.

Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037.

*To whom correspondence should be addressed.

Beginning with the COOH-terminal deletion series, zinc finger polypeptides, zfs 1-9, 1-8, and 1-7 each exhibited binding affinities similar to that of TFIIIA purified from *Xenopus* oocytes ($K_d \cong 1$ to 2 nM) (Table 1). This result confirms the finding that the 23-kD trypsin fragment of *Xenopus* TFIIIA (6), which contains zinc fingers 1 to 7, binds 5S RNA with full affinity (7). Deletion of finger 7, however, abolished high affinity 5S RNA binding. Polypeptides containing six or fewer NH_2-terminal zinc fingers all exhibited binding affinities 50- to 180-fold lower than that of zf1-7 (Figs. 2 and 3A and Table 1). Graphical representation of the data (Fig. 3A) clearly shows two groups of titration curves; one for polypeptides which exhibited high affinity binding (zfs 1-9 to 1-7) and a second group for those polypeptides which exhibited much lower binding affinities (zfs 1-6 to 1-3). In contrast, all of these polypeptides exhibit high affinity DNA binding (10). This suggests that each of these polypeptides was correctly folded and that the lack of high affinity RNA binding was not due to contaminants in the preparations of these polypeptides (17). A further experiment was performed to rule out the possibility that impurities in the zf1-6 preparation might have inhibited high affinity RNA binding. TFIIIA and zf1-6 were co-incubated with 5S RNA and assayed for binding as above. In these mixed reactions, each protein exhibited a dissociation constant similar to that observed in separate reactions. Moreover, even a 250-fold molar excess of zf1-6 did not inhibit binding of TFIIIA to 5S RNA. We conclude that the weak binding of polypeptides containing six or fewer NH_2-terminal zinc fingers is an intrinsic property of these proteins.

Because the two COOH-terminal fingers (8 and 9) are not required for high affinity RNA binding, we generated a series of four NH_2-terminal finger deletions of zf1-7 (15) (Fig. 1). These polypeptides, termed zfN-7 (where N = 2, 3, 4, 5), were assayed for RNA binding to determine the minimal set of zinc fingers that would retain high-affinity binding. Zfs 2-7, 3-7, and 4-7 each exhibited similar affinities for 5S rRNA (~2 nM) (Fig. 3B and Table 1). Zf4-7 was the shortest polypeptide to exhibit high affinity binding. Deletion of zinc finger 4 results in a 30-fold decrease in binding affinity of zf5-7 compared to zf4-7 (Figs. 2 and 3B and Table 1). Thus, zinc fingers 4-7 represent the minimal domain of TFIIIA which retained specific and high affinity RNA binding. This four-fingered polypeptide did not bind the 5S DNA promoter with high affinity or sequence specificity.

In one clone of zf1-7 (10), single base changes were found that affected the amino acid sequence in fingers 4 and 6 [Gln^{121} to Arg (Q121R) in finger 4 and Thr^{176} to Ile (T176I) in finger 6]. This polypeptide exhibited a markedly reduced affinity for 5S RNA when compared to either TFIIIA or zf1-7 with the wild-type sequence (Table 2). Both mutations involve surface residues at base contact positions (19). Substitutions at these sites are unlikely to disrupt the tertiary structure of these fingers. To assess the role of each amino acid in RNA binding, we separately converted each to the wild-type sequence by site-directed mutagenesis. The results of 5S RNA binding experiments for these mutants are summarized (Table 2). The finger 4 mutant showed a fourfold reduction in affinity when compared to wild-type zf1-7. The T176I mutation in finger 6 is clearly more deleterious than the finger 4 mutation, resulting in a 30-fold reduction in affinity. The magnitude and nature of this mutation suggest there is a specific role for this threonine in RNA binding. It is noteworthy that the sequence Thr-Trp-Thr (TWT) in finger 6 (Fig. 1) appears at the same position in finger 6 of the 5S rRNA binding protein p43 (20). p43 is similar to TFIIIA; it

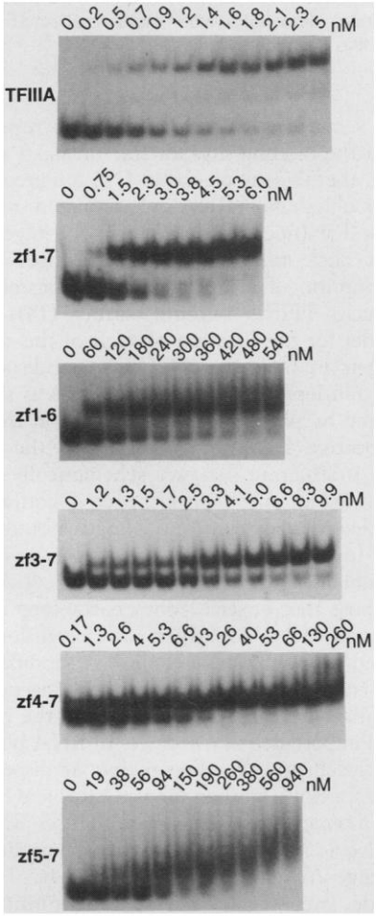

Fig. 2. Gel mobility shift analysis of the binding of TFIIIA and zinc finger polypeptides to synthetic radiolabeled 5S RNA (16). Measurements were made by protein titrations at constant 5S RNA concentrations (approximately 0.05 to 0.1 nM) in 20 mM Hepes (pH 7.8), 100 mM KCl, 1 mM $MgCl_2$, 10% glycerol, 25 μM $ZnCl_2$, 5 mM dithiothreitol, and bovine serum albumin (100 μg/ml) with 250 ng of poly dI-dC and 20 units of RNasin (Promega) per 20-μl reaction. Polypeptide concentrations are shown above the corresponding lanes. After incubation at ambient temperature for 20 min, samples were subjected to electrophoresis on a 6 or 8% nondenaturing polyacrylamide gel. The autoradiograms are shown.

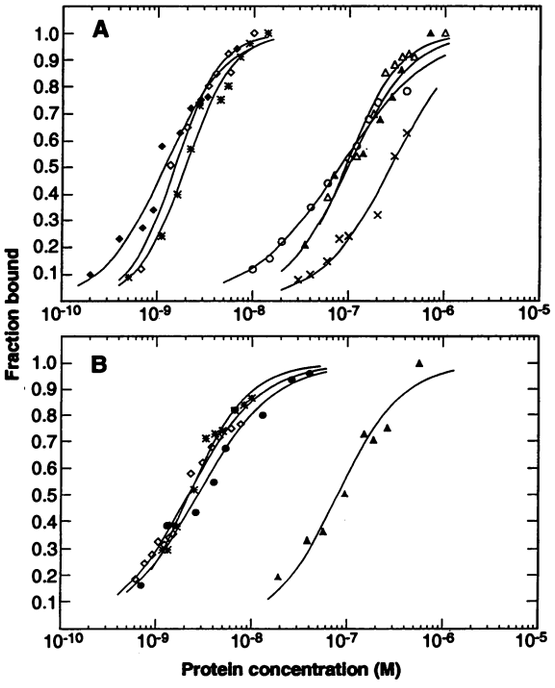

Fig. 3. Binding titrations for zinc finger polypeptides with 5S RNA. Autoradiograms of gel mobility shifts (taken within the linear range of the x-ray film) were scanned with a laser densitometer and the fraction of bound RNA was plotted against the concentration of polypeptide in each reaction. The curves shown were obtained by nonlinear fitting of the binding data to the Hill equation: fraction bound = $[Protein]^x/K_d^x + [Protein]^x$, with the use of Templegraph. In each instance the Hill coefficient (x) indicated that only one polypeptide molecule bound per RNA molecule. (**A**) COOH-terminal deletions: ◆, zf1-9; ✳, zf1-8; ◇, zf1-7; △, zf1-6; ×, zf1-5; ○, zf1-4; and ▲, zf1-3. (**B**) NH_2-terminal deletions of zf1-7: ◇, zf2-7; and ✳, zf3-7; ●, zf4-7; and ▲, zf5-7.

contains nine zinc fingers and also serves as a 5S rRNA storage particle protein in *Xenopus* oocytes. Mutations at any position within the TWT triplet significantly reduce the affinity of zf1-7 for 5S RNA (Table 2), consistent with specific RNA contacts. We would not expect the Gln (Q) to Arg (R) mutation in finger 4 to decrease affinity if Gln^{121} only participates in nonspecific backbone contacts. Thus, Gln^{121} may make a base-specific RNA contact. For finger 6, the threonine residues could be involved in direct hydrogen bonding or, as previously suggested (*21*), the tryptophan may participate in base stacking interactions.

Our present results and the results of DNA-binding studies with these zinc finger polypeptides (*10*) reveal a modular structure within TFIIIA. Zinc fingers 1 to 3 are required for specific and high affinity DNA binding, but are dispensable for RNA binding. Zinc fingers 8 and 9 are dispensable for both high affinity RNA and DNA binding; fingers 8 and 9 anchor the COOH-terminal three zinc fingers onto the 5′ A-block promoter element although they contribute little to the overall binding affinity (*10*). This interaction positions the transcriptional activation domain of the protein such that productive contacts with the other components of the transcription complex can form (*6*, *8*). In this report, we show directly that the central zinc fingers 4 to 7 are sufficient for high affinity RNA binding. This contrasts with a recent report that the COOH-terminal fingers of TFIIIA have the highest affinity for 5S RNA (*22*); however, that study was qualitative in nature whereas the current results provide a quantitative picture of the relative contributions made by different zinc fingers in 5S RNA binding.

After submission of this paper, a report appeared that claimed no single zinc finger of TFIIIA is essential for binding 5S RNA and implicated fingers 4 to 6 as the minimal RNA-binding domain (*21*). That study employed a qualitative indirect immunoprecipitation assay with large β-galactosidase fusion constructs and scored positive or negative binding under conditions of protein excess. In the present study, quantitative binding data reveal large differences in the relative affinities of different zinc fingers for 5S RNA. The finding that different sets of zinc fingers impart high affinity binding for DNA or RNA explains how TFIIIA can bind the two different nucleic acids with comparable affinities. We estimate that fingers 4 to 7 contribute 95% of the free energy of TFIIIA binding to 5S rRNA whereas fingers 1 to 3 make a similar contribution in binding the promoter of the 5S gene (*10*).

One question that remains unanswered is what distinguishes a DNA binding zinc finger from an RNA binding finger. Comparison of the sequences of the two 5S RNA binding proteins, TFIIIA and p43 (*20*), reveals few homologies other than the consensus zinc ligands (C and H), hydrophobic amino acids (*3*), and the TWT motif in finger 6 noted above. The conservation of amino acid residues involved in formation of the hydrophobic core (*19*) suggests that it is not the overall topology of individual finger domains which imparts specificity for DNA or RNA. As noted previously (*20*–*22*), the NH_2-terminal three zinc fingers of TFIIIA, which impart DNA specificity (*10*), are linked by the sequence Thr-Gly-Glu-Lys (TGEK); this sequence is commonly found in the linkers of many DNA binding zinc finger proteins (*23*). This linker sequence is not present in p43 or in TFIIIA beyond the NH_2-terminal three fingers. Indeed, mutation of the TGE linker amino acids between fingers 1 and 2 or between fingers 2 and 3 in zf1-3 to the corresponding amino acid sequences of p43 abolishes specific DNA binding (*24*).

The secondary structure of 5S RNA has been implicated as the major determinant of TFIIIA binding, although a degree of base specificity has been noted for the loop A (nucleotides 10 to 13) region of 5S RNA (*14*). Our current results indicate additional specific contacts. RNA mutagenesis studies (*14*, *21*) have shown that the major protein-RNA interactions occur between nucleotides 10 to 21, 57 to 77, and 99 to 109; these regions comprise loop A and helices II and V of 5S RNA. Inspection of the secondary structure of 5S RNA reveals that the binding sites for zinc fingers 5 and 7 in the 5S gene promoter (*10*) are present as double-helical RNA in base-paired stems II and V (nucleotides 69 to 71 for finger 5 and nucleotides 57 to 59 for finger 7). Disruption of either stem by mutagenesis reduces TFIIIA binding affinity (*14*). A model for the location of each of the zinc fingers in both the zf4-7 polypeptide and the full-length protein on 5S RNA is suggested by placing fingers 5 and 7 on their respective DNA binding sites in the 5S RNA structure (shown schematically in Fig. 4). In this model, finger 4 contacts nucleotides adjacent to the finger 5 binding site (nucleotides 73 to 76); mutations in this region also impair TFIIIA binding (*14*). We propose that finger 6 contacts the loop nucleotides 10 to 13, which have been implicated by mutagenesis studies as candidates for base-specific contacts (*14*). Mutagenesis of finger 6 (Table 2) demonstrates the pivotal importance of this finger in RNA binding specificity. Zinc fingers that are dispensable for RNA binding (fingers 1 to 3 and 8 to 9) interact with regions of the 5S RNA molecule which when mutated show little change in affinity for TFIIIA (*14*). This model is consistent with RNA footprinting experiments recently reported for TFIIIA deletion mutants (*21*). Future high-resolution nuclear magnetic resonance or crystallographic studies with zf4-7 and truncated versions of 5S RNA will be needed to verify and refine the details of this model.

Table 1. Binding constants for the interaction of zinc finger polypeptides with *Xenopus* 5S RNA.

COOH-terminal deletions		NH_2-terminal deletions	
Polypeptide	K_d (nM)*	Polypeptide	K_d (nM)*
zf1-9	1.2	zf2-7	2.2
zf1-8	1.9	zf3-7	2.2
zf1-7	1.5	zf4-7	2.7
zf1-6	92	zf5-7	77
zf1-5	270		
zf1-4	80		
zf1-3	98		

*Errors are estimated to be 30% of the presented values.

Table 2. Binding constants for finger 4 and 6 mutants of zf1-7 (*25*).

Polypeptide	K_d (nM)*
zf1-7 wild type	1.5
Q121R (finger 4)	6.7
T176I (finger 6)	46.6
T176I, Q121R	>>50
W177L, Q121R	16.7
T178R, Q121R	16.1

*Binding titrations were performed as described in Figs. 2 and 3.

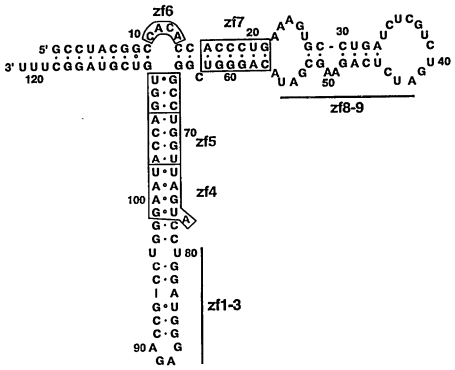

Fig. 4. Model for the interaction of individual zinc fingers with 5S RNA. The secondary structure of *Xenopus* oocyte-type 5S RNA was taken from Westhof *et al.* (*26*). Boxed nucleotides indicate regions that when mutated affect TFIIIA binding affinity in two studies (*14*).

REFERENCES AND NOTES

1. D. R. Engelke, S. Y. Ng, B. S. Shastry, R. G. Roeder, *Cell* **19**, 717 (1980); H. R. Pelham, W. M. Wormington, D. D. Brown, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 176 (1981).
2. B. M. Honda and R. G. Roeder, *Cell* **22**, 119 (1980); H. R. Pelham and D. D. Brown, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4170 (1980).
3. J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **4**,

1609 (1985); R. S. Brown, C. Sander, P. Argos, *FEBS Lett.* **186**, 271 (1985).
4. A. M. Ginsberg, B. O. King, R. G. Roeder, *Cell* **39**, 479 (1984); J. Y. T'so, K. J. van den Berg, L. J. Korn, *Nucleic Acids Res.* **14**, 2187 (1986).
5. G. P. Diakun, L. Fairall, A. Klug, *Nature* **324**, 698 (1986).
6. D. R. Smith, I. J. Jackson, D. D. Brown, *Cell* **37**, 645 (1984).
7. M. S. Sands and D. F. Bogenhagen, *Nucleic Acids Res.* **19**, 1791 (1992).
8. J. Hayes, T. D. Tullius, A. P. Wolffe, *J. Biol. Chem.* **264**, 6009 (1989).
9. T. Pieler, J. Hamm, R. G. Roeder, *Cell* **48**, 91 (1987).
10. X. Liao, K. R. Clemens, L. Tennant, P. E. Wright, J. M. Gottesfeld, *J. Mol. Biol.* **223**, 857 (1992); K. R. Clemens, X. Liao, V. Wolf, P. E. Wright, J. M. Gottesfeld, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10822 (1992); J. Hayes and K. R. Clemens, *Biochemistry* **31**, 11600 (1992).
11. S. Sakonju and D. D. Brown, *Cell* **31**, 395 (1982).
12. J. J. Hayes and T. D. Tullius, *J. Mol. Biol.* **227**, 407 (1992).
13. J. Anderson and N. Delihas, *J. Biol. Chem.* **261**, 2912 (1986); F. Baudin, P. Romby, P. J. Romaniuk, B. Ehresmann, C. Ehresmann, *Nucleic Acids Res.* **17**, 10035 (1989); J. Christiansen, R. S. Brown, B. S. Sproat, R. A. Garrett, *EMBO J.* **6**, 453 (1987); T. Pieler, V. A. Erdmann, B. Appel, *Nucleic Acids Res.* **12**, 8393 (1984); P. J. Romaniuk, *Biochemistry* **28**, 1388 (1989).
14. M. S. Sands and D. F. Bogenhagen, *Mol. Cell. Biol.* **7**, 3985 (1987); Q. You, N. Veldhoen, F. Baudin, P. J. Romaniuk, *Biochemistry* **30**, 2495 (1991).
15. K. D. MacFerrin, M. P. Terranova, S. L. Schreiber, G. L. Verdine, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2304 (1990). The sequences of primers for the native NH_2-terminus and terminating primers for the COOH-terminal deletions are given in Liao *et al.* and in Clemens *et al.* [in (*10*)]. Initiating oligonucleotides for the NH_2-terminal deletions contained an Nde I restriction site of which the last three bases form the start codon (underlined) and were: (zf2-7) 5′-CGGCCAAGCTTAGGAGGAATATCATATGGAGAAACCATTTCCATGTAAGGAAG; (zf3-7) 5′-CGGCCAAGCTTAGGAGGAATATCATATGGAGAAACCATTTCCATGTAAGGAAG; (zf4-7) 5′-GTTGCTGAAGGACATATGTATGTGTGCCATTTTG; (zf5-7) 5′-GTT GCTGAAGGACATATGTACGAATGTCCTC. After amplification, polymerase chain reaction products were cleaved with Nde I and Eco RI and were ligated into similarly digested plasmid pRK172 [M. McLeod, M. Stein, D. Beach, *EMBO J.* **6**, 729 (1987)] and amplified in *Escherichia coli* strain DH5α and transferred to strain BL21(DE3) for protein expression [F. W. Studier and B. Moffat, *J. Mol. Biol.* **189**, 113 (1986)]. The sequence of each of the truncated polypeptides was determined from the expression clones. The zf1-9 had a single base change, G → T, which causes the mutation of Ala^{257} to Ser. In the original cloning of zf1-7 [Clemens *et al.*, in (*10*)] single base mutations that affect the amino acid sequence of the protein were found in fingers 4 and 6. These mutations were converted to the wild-type sequence by oligonucleotide-directed mutagenesis [T. A. Kunkel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488 (1985)], and the sequence of zf1-7 was confirmed by DNA sequence analysis. Similar methods were used to introduce additional mutations into finger 6. Mutagenesis was performed in the vector pTZ19U (Bio-Rad) and the gene products were expressed from this vector.
16. A synthetic somatic-type 5*S* RNA gene cloned into pUC18 was the generous gift of P. J. Romaniuk [P. J. Romaniuk, I. Leal de Stevenson, H.-H. A. Wong, *Nucleic Acids. Res.* **15**, 2737 (1987)]. Plasmid DNA was cleaved with Dra I (Stratagene) and used as a template for transcription with T7 RNA polymerase (Stratagene) with [α-^{32}P]GTP (guanosine triphosphate) and unlabeled nucleoside triphosphates. The run-off transcript was 121 nucleotides long. RNA was purified by phenol extraction and concentrated by ethanol precipitation before use in binding reactions.
17. Note that any impurities in the recombinant polypeptide preparations would likely be similar for each protein because they were all isolated from the same bacterial strain by the same protocol.
18. Protein concentrations were determined by Coomassie blue staining of SDS-polyacrylamide gels and densitometry with known amounts of bovine serum albumin as a standard. After purification on a single heparin-Sepharose column (*10*), the purity of the recombinant polypeptides ranged from 25% (zfs 2-7 and 3-7) to near homogeneity. Concentrations of some of the polypeptides were determined by Western (immuno) blotting with a polyclonal serum raised against *Xenopus* TFIIIA with known amounts of the purest polypeptides (zfs 1-5 and 1-6) as standards. Protein concentrations used in K_d determinations were adjusted appropriately.
19. M. S. Lee, G. P. Gippert, K. V. Soman, D. A. Case, P. E. Wright, *Science* **245**, 635 (1989); N. P. Pavletich and C. O. Pabo, *ibid.* **252**, 809 (1991).
20. K. E. Joho, M. K. Darby, E. T. Crawford, D. D. Brown, *Cell* **61**, 293 (1990).
21. O. Theunissen, F. Rudt, U. Guddat, H. Mentzel, T. Pieler, *ibid.* **71**, 679 (1992).
22. M. K. Darby and K. E. Joho, *Mol. Cell. Biol.* **12**, 3155 (1992).
23. T. J. Gibson, J. P. Postma, R. S. Brown, P. Argos, *Protein Eng.* **2**, 209 (1988).
24. K. R. Clemens, V. Wolf, P. Zhang, S. J. McBryant, J. M. Gottesfeld, P. E. Wright, in preparation.
25. 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.
26. E. Westhof *et al.*, *J. Mol. Biol.* **207**, 417 (1989).
27. Supported by grants from the National Institute of General Medical Sciences to J.M.G. (GM26453 and GM47530) and to P.E.W. (GM36643). X.L. was supported by the Cancer Research Institute/ Miriam and Benedict Wolf Fellowship and K.R.C. received support from a National Institutes of Health postdoctoral fellowship award (F32 CA09023). We thank B. Blair for her assistance in the preparation of this manuscript.

9 November 1992; accepted 8 February 1993

Knotting of a DNA Chain During Ring Closure

Stanley Y. Shaw and James C. Wang

The formation of knotted species on random ring closure of two DNAs that are 5.6 kilobase pairs (kbp) and 8.6 kbp in length was measured, and these data were used to calculate the effective DNA helix diameter as a function of sodium ion and magnesium ion concentration. In the presence of more than 50 mM magnesium ion, interactions between DNA segments appear to be attractive rather than repulsive. The free energy of formation of relaxed trefoil and figure-eight DNA knots and of supercoiled trefoil DNA knots was also evaluated.

Descriptions of knots in 19th century atomic physics (*1*) presaged the formal development of knot theory, the study of the topology of deformable but nonbreakable closed curves in three-dimensional space. Concepts from knot theory have permeated such fields as algebraic topology, physics, and synthetic and polymer chemistry, sometimes revealing relations among seemingly disparate disciplines (*2*). In biology, the abundance of large ring-shaped molecules in living organisms has provided fertile ground for the study of knots, particularly since the discovery of knotted DNA rings (*3*). A number of DNA knots have been characterized in detail, and their knot-types have provided insights into the mechanisms of the enzymatic reactions that produced them (*4*).

When a particular DNA transaction can yield either a knotted or unknotted product, the energetics of DNA knotting might influence the course of the reaction (*5*). In contrast to the elegant and extensive work on knot typing and its mechanistic implications (*4*), however, no experimental datum exists for the probability (and hence the Gibbs free energy) of formation of even the simplest knotted DNA ring, the trefoil. It has been shown that for a polygon of n sides, the probability that it is unknotted approaches zero as n approaches infinity (*6*); thus, the probability of knotting in very large DNA rings is likely to be high. In solution, DNA behaves as a stiff molecule with a Kuhn statistical segment length (*7*) of about 1000 Å or 300 base pairs (bp). Theoretical models of knot formation in polymer chains suggest that for DNA molecules thousands of base pairs in length, the probability of knotting during ring closure is small but significant (*8–11*). Thus, in addition to yielding information on the energetics of knotted DNA rings, measurements of knot formation in randomly cyclized DNA should also provide experimental tests of the various theoretical treatments of knotting probabilities.

To determine experimentally the probability of knotting, we used two DNA molecules 5.6 kbp and 8.6 kbp in length. The DNAs were designed to have single-stranded ends identical to those of phage λ (*12*); because the chemical step of pairing such ends is the rate-limiting step in cyclization (*13*), the fraction of knotted rings in the product represents that in the equilibrium population.

Harvard University, Department of Biochemistry and Molecular Biology, 7 Divinity Avenue, Cambridge, MA 02138.