cDNA into plasmid pBATneo (*10*) in place of the BgI II–Hpa I fragment of the neomycin-resistance gene, and the cells expressing mouse N-cadherin were selected and cloned (*10*). One of these clones, mNLβ1, expressed mouse N-cadherin, showing N-cadherin–mediated aggregating activity. Using similar vectors, we also isolated an L cell clone expressing chicken N-cadherin, cNLβ1 (M. Hashimoto and M. Takeichi, unpublished).
19. S. A. Roth and J. A. Weston, *Proc. Natl. Acad. Sci. U.S.A.* **58**, 974 (1967).
20. M. S. Steinberg, *Science* **141**, 401 (1963).
21. ———, *ibid.* **137**, 762 (1962).
22. A. A. Moscona, *Proc. Natl. Acad. Sci. U.S.A.* **43**, 184 (1957); *Exp. Cell Res.* **22**, 455 (1961); B. Garber and A. A. Moscona, *J. Exp. Zool.* **155**, 179 (1964); *ibid.* **168**, 455 (1968).
23. S. G. Grady and E. J. McGuire, *J. Cell Biol.* **71**, 96 (1986).
24. M. L. Burdick and M. S. Steinberg, *Proc. Natl. Acad. Sci. U.S.A.* **63**, 1169 (1969); M. L. Burdick, *J. Exp. Zool.* **175**, 357 (1970).
25. B. Obrink, M. S. Kuhlenschmidt, S. Roseman, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1077 (1977).
26. S. B. Needleman and C. D. Wunsch, *J. Mol. Biol.* **48**, 443 (1970).
27. A. H. Rosenberg *et al.*, *Gene* **56**, 125 (1987).
28. F. W. Studier and B. A. Moffatt, *J. Mol. Biol.* **189**, 113 (1986).
29. R. A. Watt, A. R. Shatzman, M. Rosenberg, *Mol. Cell. Biol.* **5**, 448 (1985).
30. Supported by the Ministry of Education, Science, and Culture of Japan and the Science and Technology Agency of the Japanese government.

30 December 1988; accepted 19 May 1989

Three-Dimensional Solution Structure of a Single Zinc Finger DNA-Binding Domain

MIN S. LEE, GARRY P. GIPPERT, KIZHAKE V. SOMAN, DAVID A. CASE, PETER E. WRIGHT*

The three-dimensional solution structure of a zinc finger nucleic acid binding motif has been determined by nuclear magnetic resonance (NMR) spectroscopy. Spectra of a synthetic peptide corresponding to a single zinc finger from the *Xenopus* protein *Xfin* yielded distance and dihedral angle constraints that were used to generate structures from distance geometry and restrained molecular dynamics calculations. The zinc finger is an independently folded domain with a compact globular structure in which the zinc atom is bound by two cysteine and two histidine ligands. The polypeptide backbone fold consists of a well-defined helix, starting as α and ending as 3_{10} helix, packed against two β strands that are arranged in a hairpin structure. A high density of basic and polar amino acid side chains on the exposed face of the helix are probably involved in DNA binding.

THE ZINC FINGER MOTIF, WHICH recognizes specific DNA sequences (*1–3*), was first identified as a repeated motif in transcription factor IIIA from *Xenopus* oocytes and was proposed to be an independently folded DNA-binding domain (*4*). The zinc is either coordinated by two Cys and two His ligands (TFIIIA-type) or by four Cys ligands (steroid receptors and yeast transcriptional factors) (*2, 3, 5, 6*). Zinc is required for correct folding and is essential for specific DNA recognition (*6–10*). Circular dichroism and NMR spectroscopy have been used to demonstrate zinc-dependent folding of synthetic single fingers with sequences from TFIIIA and ADR1 (*11, 12*). These isolated single fingers bind in a nonspecific but zinc-dependent manner to DNA. A previous NMR study of a synthetic zinc finger from ADR1 (*12*) indicated the presence of a helix in the zinc complex and led to a qualitative model for the finger structure. In this report, we describe complete three-dimensional (3-D) solution structure determination for a synthetic peptide corresponding to the 31st zinc finger from the *Xenopus* protein *Xfin* (*13*) [denoted *Xfin-31*] based on an extensive set of experimental NMR constraints and the use of distance geometry and molecular dynamics (MD) calculations for structural analysis.

We selected *Xfin-31* for structural analysis since it corresponds closely to a consensus sequence derived from 148 zinc finger domains (*14*). A synthetic 25-residue peptide (Fig. 1) spanning the putative zinc-binding residues was shown to be more than 90% bound to zinc at *p*H 5.5 (*15*). Substantial changes occurred in the one-dimensional ^{1}H NMR spectrum upon zinc binding; two dimensional (2-D) NMR spectra indicate the formation of a single folded conformation in aqueous solution (*p*H 5.5, 5° to 25°C). Gel mobility assays show that the *Xfin-31* peptide binds nonspecifically to DNA in the presence of zinc. No detectable binding occurs if zinc is absent.

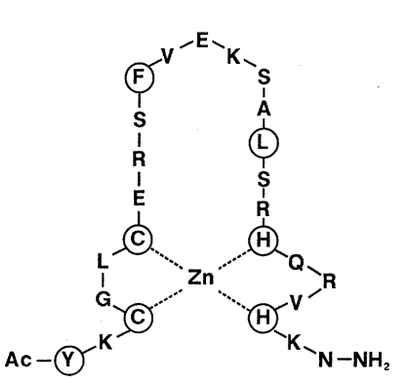

Fig. 1. Amino acid sequence of finger 31 from *Xfin* (*13*) drawn to illustrate the finger motif (*4*). Invariant and highly conserved residues are circled. The amino and carboxyl termini of the peptide used for the present studies are blocked by acetyl (Ac) and amide groups, respectively.

Complete sequence-specific assignments for all backbone and side chain protons were obtained with the use of 2-D double-quantum filtered correlated spectroscopy (COSY), double-quantum spectroscopy, total correlated spectroscopy (TOCSY), and 2-D nuclear Overhauser effect spectroscopy (NOESY) (*16*). A total of 120 interresidue and 24 intraresidue distance constraints derived from NOEs (*17*) was used for structure determination. Backbone dihedral angle restraints were obtained from $^3J_{HN\alpha}$ coupling constants measured from a high-resolution double-quantum filtered COSY spectrum recorded at 25°C. The backbone torsion angle ϕ was constrained to $-160° \leq \phi \leq -80°$ for five residues with $^3J_{HN\alpha} > 8$ Hz and to $-90° \leq \phi \leq -20°$ for seven residues with $^3J_{HN\alpha} < 6$ Hz. No explicit hydrogen-bond constraints were used in the structure calculations.

The zinc was constrained to be tetrahedrally coordinated by Cys^3, Cys^6, His^{19}, and His^{23} (*6, 11*), and Zn–S and Zn–N bond distances were constrained to 2.30 ± 0.05 and 2.0 ± 0.05 Å, respectively (*5*). These distance and angle (109 ± 10°) constraints agree with those observed in x-ray structures of zinc amino acid complexes (*18*). Because it is not known a priori whether zinc coordinates to the imidazole Nδ or Nε of the His ligands, distance geometry calculations were performed with a combination of distance and chirality constraints that allow coordination to either Nε or Nδ and keep the zinc within the plane of the imidazole ring. These calculations showed unambiguously that both His residues coordinate through their Nε atoms.

A family of structures was calculated with the distance geometry program DISGEO (*19*). Pseudoatoms (*20*) were used wherever necessary and interproton distances were corrected accordingly. A total of 50 struc-

Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037.

*To whom correspondence should be addressed.

Fig. 2. Superposition of 37 structures of the *Xfin-31* zinc finger derived from distance geometry calculations and molecular dynamics refinement. A stereo view of the Cα chain tracing is shown. The amino terminus is at the bottom right and the carboxyl terminus at the upper left. The Cys^3 and Cys^6 ligands are shown in yellow, the His^{19} and His^{23} side chains in blue, and the zinc in white. The structures are superimposed for minimum rms deviation between backbone heavy atoms.

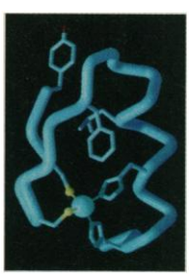

Fig. 3. Structure of the *Xfin-31* zinc finger with fewest residual constraint violations. The overall shape is rather like a right hand, with the thumb representing the amino-terminal β strand. The molecule is rotated 180° relative to the view in Fig. 2 to emphasize its shape. The Cα chain tracing is represented as a tube of diameter 1 Å. The side chains of the Cys and His ligands, Tyr^1, Phe^{10}, and Leu^{16} are shown. [Figure generated with the program MCS written by M. Connolly]

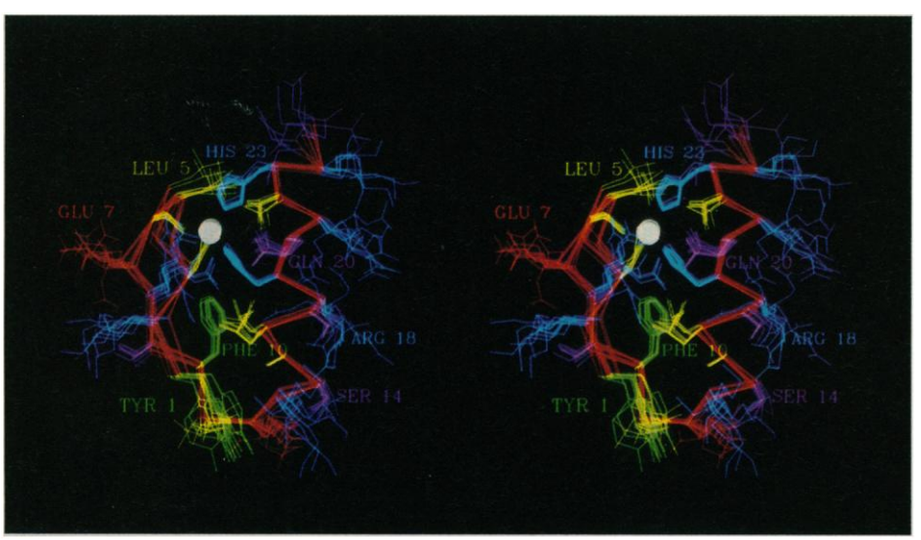

Fig. 4. Stereo view of the zinc finger structure showing Cα backbone tracing and side chains. For clarity only the ten best refined structures (see text) are displayed, superimposed for minimum rms deviation between backbone heavy atoms. The side chains are color-coded as follows: Phe and Tyr, green; Ala, Val, and Leu, yellow; Ser, Asn, and Gln, purple; Glu, red orange; Arg and Lys, blue. The Cys ligands are shown in yellow, the His ligands in blue, and the zinc in white.

tures was embedded at the substructure level, 37 of which satisfactorily completed minimization and were used as starting structures for MD refinement.

Our refinement procedure uses the AMBER all-atom force field (*21*) to compute the intrinsic strain energy and adds a half-parabola penalty for proton-proton distances that violate the NMR constraints. Dihedral angles falling outside the limits listed above are penalized by a function of the form $K[1 - \cos(\phi - \phi_0)]$, where ϕ_0 is the end point of the "allowed" range and $K = 32$ kcal/mol. Based on our experience and that of others (*22*), we increased the torsional constant K_ω for peptide bonds from 10 to 50 kcal/mol, which keeps nearly all of the peptide bonds within 7° of the planar trans conformation. We also reduced the net charges on Glu, Lys, and Arg side chains from unity to ±0.2 to partially mimic dielectric effects of the solvent and to reduce the propensity for the formation of intramolecular salt bridges. Parameters for the metal coordination (*23*) were designed to maintain a roughly tetrahedral geometry consistent with the extended x-ray absorption fine structure (EXAFS) data (*5*) and model crystal structures (*18*) and yet still allow sampling of a variety of metal-binding geometries. All 37 distance geometry structures were energy minimized in this combined potential, then heated to 1200 K during 1 ps with a temperature-regulated MD algorithm (*24*) with a temperature relaxation time τ of 0.3 ps. This procedure was followed by 4 ps of equilibration at 1200 K and a 6-ps cooling run with $\tau = 2$ ps (*25*). A final step of energy minimization produced the refined structures.

The conformation of the polypeptide backbone (Fig. 2) is well defined by the NMR constraints (*26*). The average root-mean-square (rms) deviation from the mean structure is 0.81 Å for the backbone and 1.81 Å for the side chain heavy atoms when all 37 refined structures are considered. Ten of the refined structures had final AMBER energies less than −305 kcal/mol and NMR constraint violation energies less than 7 kcal/mol; only one structure had any distance violation greater than 0.2 Å (0.23 Å). These structures, for which the average backbone rms deviation is only 0.58 Å, were used for detailed conformational analysis. Theoretical NOESY spectra calculated from the refined structures with a full relaxation matrix approach (*27*) are in good qualitative agreement with the experimental spectra.

The overall fold of the zinc finger is evident from Fig. 2. Residues 1 to 10 form a hairpin structure that encompasses Cys^3 and Cys^6. A well-defined helix extends from Glu^{12} to Lys^{24} and includes both His li-

gands. The zinc atom is buried in the interior of the molecule. The zinc finger folds into a compact structure shaped rather like a right hand, with the zinc and the carboxyl terminus at the base of the hand and the amino terminus extending like the thumb toward the fingertips (Fig. 3). The amino- and carboxyl-terminal Cα atoms are at opposite ends of the molecule and are separated by 17.3 ± 1.7 Å [37 structures]. This has important implications for modeling the interaction of multifinger domains with DNA.

The conformations of many of the interior side chains are well defined by the NMR constraints (Fig. 4). The side chains of the highly conserved (*14*) Phe^{10} and Leu^{16} are packed in a hydrophobic cluster with Ala^{15}. The phenyl ring of Phe^{10} is in van der Waals contact with the imidazole group of His^{19}, as are the side chains of Gln^{20} and Val^{22}. The zinc finger can be described as a "miniglobular" protein, with a close-packed, predominantly hydrophobic core and with polar side chains on the surface.

The backbone dihedral angles and hydrogen-bonding patterns in the ten best structures were examined to better define the common elements of secondary structure. Considerable commonality occurs in terms of backbone amide–carbonyl hydrogen bonding (*28*). Note that no explicit hydrogen-bond constraints were included in generating the structures. All ten structures have hydrogen bonds from Tyr^1 NH to Phe^{10} CO and from Phe^{10} NH to Tyr^1 CO, a pattern found in antiparallel β sheets. The backbone conformation between the two Cys ligands has not been determined uniquely, although all ten structures have an internal hydrogen bond characteristic of a tight turn, either from Leu^5 NH to Cys^3 CO (γ turn) or Cys^6 NH to Cys^3 CO (β turn). In the helical region, backbone NH–CO hydrogen bonds from residues 16 → 12, 17 → 13, 18 → 14, and 19 → 15 are found in all ten structures, as are backbone ϕ and φ dihedral angles that are characteristic of α helix. From this point, five of the structures continue the α-helical motif with a 20 → 16 hydrogen bond, while the other five tighten into a 3_{10} helix containing a 20 → 17 hydrogen bond. The 3_{10} helical motif continues with hydrogen bonds from residues 21 → 18 (nine structures), 22 → 19 (ten structures), and 23 → 20 (ten structures). In eight of the ten structures, the helical region ends with a bifurcated hydrogen bond from the Gln^{20} CO to the NH groups of His^{23} and Lys^{24}. No other internal hydrogen bonds (backbone or side chain) are present in more than five of the structures. In summary, the hydrogen-bonding patterns indicate an antiparallel β interaction between residues 1 and 10, a tight turn beginning at residue 3, and a helix from residues 12 to 24, which is α helical at the beginning and 3_{10} helical toward the end. These hydrogen bonds are also present in the unrefined distance geometry structures. The change from α to 3_{10} helix may be a consequence of coordination of the two His residues to the zinc.

Previous zinc finger models bear a striking resemblance to the actual 3-D structure determined by NMR. The Berg model (*29*) is most similar to the NMR structure and differs primarily in the length of the helix. In the model proposed by Gibson *et al.* (*14*), the alignment of the β strands is offset by two residues, but the helix is of the correct length. Neither model predicted the presence of 3_{10} helix. The 3-D structure reported here also differs from the qualitative 2-D NMR model proposed by Klevit and coworkers (*12*) for an ADR1 finger, in that no evidence was found for the β hairpin between residues 1 and 10 and the presumed α helix was shorter.

How do zinc fingers bind to DNA? There is evidence that binding occurs in the major groove of DNA (*30*), and it has been proposed that it is primarily residues on the helix of the finger motif that determine binding specificity (*14, 29, 31*). The *Xfin-31* finger structure has a high density of basic amino acids (Lys^{13}, Arg^{18}, Arg^{21}, and Lys^{24}) on the exposed face of the helix (Fig. 4) where they might make contact with either the phosphate backbone or specific bases. In addition, there is a distribution of polar side chains (Glu^{12}, Ser^{14}, Ser^{17}, and Gln^{20}) on the surface of the helix, some or all of which might be involved in specific base recognition.

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. F. Payre and A. Vincent, *FEBS Lett.* **234**, 245 (1988).
4. J. Miller *et al.*, *EMBO J.* **4**, 1609 (1985).
5. G. P. Diakun *et al.*, *Nature* **324**, 698 (1986).
6. L. P. Freedman *et al.*, *ibid.* **334**, 543 (1988).
7. K. Nagai *et al.*, *ibid.* **332**, 284 (1988).
8. H. Blumberg *et al.*, *ibid.* **328**, 443 (1987).
9. N. Redemann *et al.*, *ibid.* **332**, 90 (1988).
10. Y. Severne, S. Wieland, W. Schaffner, S. Rusconi, *EMBO J.* 7, 2503 (1988).
11. A. D. Frankel, J. M. Berg, C. O. Pabo, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4841 (1987).
12. G. Párraga *et al.*, *Science* **241**, 1489 (1988).
13. A. R. Altaba *et al.*, *EMBO J.* **6**, 3065 (1987).
14. T. J. Gibson *et al.*, *Protein Eng.* **2**, 209 (1988).
15. The peptide was synthesized on an Applied Biosystems Synthesizer and purified by means of a Phenomenex C4 semipreparative reversed-phase column on a Hitachi high-performance liquid chromatography (HPLC) system. The peptide was eluted with an acetonitrile-water solvent gradient containing 0.05% trifluoroacetic acid and was >95% pure by analytical HPLC. The sequence was confirmed by the sequential assignment procedure. Samples for NMR experiments were prepared by dissolving the free peptide in H_2O and adjusting the *p*H to 7.2 with 1*M* NaOH. A 1.5-fold molar excess of $ZnCl_2$ in ^2H_2O was added to the peptide and the *p*H readjusted to *p*H 5.5. The concentration of the $ZnCl_2$ solution was such that the final solvent composition was 90% H_2O/10% ^2H_2O. All NMR experiments were performed on a Bruker AM500 spectrometer with a single 5.8 m*M* sample.
16. M. S. Lee, J. Cavanagh, P. E. Wright, *FEBS Lett.*, in press.
17. NOESY spectra [J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **71**, 4546 (1979)] were recorded at 5°C with mixing times of 100, 200, and 300 ms. Distance constraints were obtained from analysis of cross-peak intensities in the 200-ms NOESY spectrum. Upper bounds were determined by empirical calibration versus known sequential distances [M. P. Williamson, T. F. Havel, K. Wüthrich, *J. Mol. Biol.* **182**, 295 (1985)]. The NOE cross peaks were classified as strong, medium, weak, and very weak, and upper bounds were assigned as 2.5, 3.0, 3.7, and 4.2 Å for backbone-backbone NOEs and 3.0, 3.5, 4.2, and 5.0 Å for backbone–side chain and side chain–side chain NOEs.
18. H. C. Freeman, in *Inorganic Biochemistry*, G. L. Eichorn, Ed. (Elsevier, Amsterdam, 1973), vol. 1, pp. 121–166.
19. T. F. Havel and K. Wüthrich, *Bull. Math. Biol.* **46**, 673 (1984).
20. K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* **169**, 949 (1983).
21. S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comput. Chem.* 7, 230 (1986).
22. A. T. Brünger, M. Karplus, G. A. Petsko, *Acta Crystallogr.* **A45**, 50 (1989).
23. Zinc–ligand bonds were assigned equilibrium distances of 2.30 Å [Zn–S] and 2.00 Å [Zn–N] and a force constant of 200 kcal/(mol Å^2). Angles centered on the metal were modeled with tetrahedral equilibrium angles and force constants of 40 kcal/(mol rad^2); these are rather weak constraints that allow distortions of 10° about the tetrahedral angle with a thermally allowed amount (0.6 kcal/mol) of strain energy. The Nε atoms of the His ligands were constrained to lie in the plane of the Cδ2, Cε1, and zinc atoms with a function of the form $K_d\delta^2$, where δ is the perpendicular distance of the nitrogen from the plane defined by the other three atoms and K_d = 50 kcal/(mol Å^2).
24. H. J. C. Berendsen *et al.*, *J. Chem. Phys.* **81**, 3684 (1984).
25. During the initial minimization and the first 3 ps of the dynamics, the force constant for distance violations was slowly increased from 1 to 32 kcal/(mol Å^2); this final value was kept constant for the cooling and final minimization steps. Refinement of each structure required about 30 min of cpu time on a Convex C240 supercomputer.
26. Coordinates have been deposited with the Brookhaven Protein Data Bank; a list of distance and dihedral angle constraints used in structure determination is available from the authors.
27. Theoretical NOESY spectra were calculated with the use of the program CORMA [J. W. Keepers and T. L. James, *J. Magn. Reson.* **57**, 404 (1984)]; We assumed a fixed correlation time. Details of these calculations will be published elsewhere.
28. A hydrogen bond is deemed to be formed if the O–H distance is less than 2.5 Å and the O–H–N angle is greater than 135°.
29. J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 99 (1988).
30. L. Fairall, D. Rhodes, A. Klug, *J. Mol. Biol.* **192**, 577 (1986).
31. R. S. Brown and P. Argos, *Nature* **324**, 215 (1986).

32. We thank M. Pique for assistance with computer graphics; P. Yip for relaxation matrix calculations; J. Gottesfeld for helpful discussions and for assays of DNA binding; M. Buchmeier for peptide synthesis; J. Cavanagh, J. Moore, and J. Waltho for valuable discussions; and L. Harvey for preparation of the manuscript. Supported by the National Institutes of Health (grants GM 36643 and GM38794). M.S.L. is the recipient of an American Cancer Society Postdoctoral Fellowship (Grant PF-3265).

1 May 1989; accepted 27 June 1989