

9. G. H. Denton and T. J. Hughes, in *The Last Great Ice Sheets*, G. H. Denton and T. J. Hughes, Eds. (Wiley, New York, 1981), pp. 440–467.
10. T. J. Hughes *et al.*, *Nature* **266**, 596 (1977).
11. The mass continuity equation is solved with the finite-element method. Boundary conditions are a mixture of mass flux conditions. At seaward boundaries, horizontal ice flux is passed freely to the model exterior as if balanced by iceberg calving. Land boundaries in Greenland and across North America are set along natural ice divides of the Greenland ice sheet and the past Laurentide ice sheet. Ice flux along these land boundaries is thus set to 0. Basal melting is assumed to be 0 throughout the model domain.
12. The stress-equilibrium equations are solved with the finite-element method. Our treatment of grounded ice is nearly identical to that employed by Oerlemans in a simulation of the Antarctic ice sheet (7). Our treatment of the more spatially complex stress balance in floating ice is identical to that employed previously in studies of Antarctic ice streams and ice shelves [D. R. MacAyeal, *J. Geophys. Res.* **94**, 4071 (1989)].
13. The effective viscosity of ice used in the simulations is based on laboratory ice-deformation tests [P. Barnes, D. Tabor, J. C. F. Walker, *Proc. R. Soc. London Ser. A* **324**, 1557 (1971)].
14. The time constant for isostatic rebound depends on properties of the solid earth and on the horizontal length scale of the applied surface load. The time constant we used was determined with a simple, viscous half-space earth model and an applied load that varied sinusoidally in one horizontal direction with a wavelength of 2000 km.
15. Present-day mean annual and summer temperatures were obtained from H. Walter and H. Lieth [*Klimadiagramm Weltatlas* (Fischer Verlag, Jena, German Democratic Republic, 1967)]. Present-day net accumulation rates were obtained from contour plots of present-day precipitation and evaporation rates [V. I. Korzoun, *Atlas of World Water Balance* (Gidrometeoizdat, Leningrad, 1977)].
16. Net ice accumulation rates on the surfaces of present-day ice sheets is comparable to, or less than, the changes in high-latitude precipitation produced by some GCM simulations when CO₂ concentration was reduced from its present value to the value at the LGM. A CO₂-induced change of approximately 0.15 m per year per square meter, for example, was reported by Manabe and Bryan (17) for the latitude band encompassing the Eurasian ice sheet.
17. S. Manabe and K. Bryan, *J. Geophys. Res.* **90**, 11689 (1985).
18. J. Oerlemans and C. J. van der Veen, *Ice Sheets and Climate* (Reidel, Hingham, MA, 1984).
19. We derived $\partial T/\partial c$ and $\partial T_s/\partial c$ from ice-free land gridpoints of the VC (variable cloudiness) GCM experiment conducted by Manabe and Broccoli in which CO₂ concentration was 200 ppm and ice-sheet surface topography was specified [S. Manabe and A. J. Broccoli, *J. Atmos. Sci.* **42**, 2643 (1985)]. We derived α from Manabe and Bryan's (17) $c = \sqrt{2X}$ coupled atmosphere-ocean model experiment in which CO₂ concentration was 212 ppm. We would prefer to have derived $\partial T/\partial c$, $\partial T_s/\partial c$, and α from a single GCM climate study, but we used two studies because Manabe and Broccoli's study incorporated the effect of ice-sheet topography on atmospheric circulation (which we consider important in determining temperature sensitivity) and Manabe and Bryan's study incorporated oceanic circulation and oceanic temperature and salinity budgets (which we consider important in determining atmospheric water transport). The values of $\Delta c(\partial T/\partial c)$ (under LGM conditions) shown in Table 1 reflect the observation that GCM predictions of cooling reduced CO₂ concentration are amplified at high latitudes. For comparison, the globally averaged difference in surface temperature between LGM and present given by Manabe and Broccoli's VC simulation was -4.7°C . Uncertainty in this global difference evident by other GCM simulations (such as Manabe and Broccoli's FC simulation, which specified fixed cloud distribution) is approximately 1°C .
20. As an additional experiment, we ran the model with $T(\phi, \lambda)$ and $A(\phi, \lambda)$ specified directly from output of Kutzbach and Guetter's (3) GCM simulation that incorporated specified LGM sea-surface temperature, reduced atmospheric CO₂ concentration, insolation effects due to orbital geometry, and specified ice-sheet topography. The ice-sheet configuration that resulted agreed well with geological reconstructions. The ice shelf thickness over the Arctic Ocean, however, became unrealistically large.
21. J. Jouzel *et al.*, *Nature* **329**, 403 (1987).
22. The lapse rate over Antarctica, for example, is between 7° and 10°C km^{-1} [J. Oerlemans, in *Dynamics of the West Antarctic Ice Sheet*, C. J. Van der Veen and J. Oerlemans, Eds. (Reidel, Hingham, MA, 1987), pp. 287–292].
23. D. H. Bromwich, *Rev. Geophys.* **26**, 149 (1988).
24. Evidence includes geomorphology of the sea-floor sediments and continental slope debris fans [M. G. Grosswald, *Quat. Res.* **13**, 1 (1980); T. O. Vorren, M. Hald, E. Lebesbye, *Paleoceanography* **3**, 601 (1988)], polar-wander paths and relative sea-level rise [W. R. Peltier, *Science* **240**, 895 (1988)], and deep-sea oxygen-isotope records [G. A. Jones and L. D. Keigwin, *Nature* **336**, 56 (1988)].
25. Evidence includes push-tectonic overthrusts and the distribution of permafrost [M. G. Grosswald, in *Data of Glaciological Studies*, G. A. Avsyok and V. M. Kotlyakov, Eds. (Publ. 63, Academy of Science of the U.S.S.R. Soviet Geography Committee, Moscow, 1988), pp. 3–25].
26. W. F. Ruddiman and A. McIntyre, *Paleogeogr. Paleoclimatol. Paleocol.* **35**, 145 (1981).
27. The effect of the simulated ice sheet on eustatic sea level may be 30 to 50% higher than previously estimated [G. S. Boulton, G. D. Smith, A. S. Jones, J. Newsome, *J. Geol. Soc. London* **142**, 447 (1985)]. Trial simulations in which we accounted for ice streams that flow across deformable sub-glacial sediment yielded smaller ice-sheet volumes. The total ice volume and its effect on eustatic sea level thus depends strongly on basal boundary conditions associated with the stress equilibrium.
28. We thank A. Anderson, V. Barclon, T. J. Hughes, C. S. Lingle, M. Monaghan, G. W. Platzman, and A. Ziegler for critical review of the Ph.D. dissertation on which this work is based. J. Kutzbach and R. Selin kindly provided detailed background information concerning their GCM experiments. Model computations were performed on the computers of the National Center for Atmospheric Research. Financial support was provided by National Science Foundation grant DPP 85-09451.

17 April 1989; accepted 21 June 1989

Neural Cadherin: Role in Selective Cell-Cell Adhesion

SEIJI MIYATANI, KENJI SHIMAMURA, MASAYUKI HATTA, AKIRA NAGAFUCHI, AKINAO NOSE, MAYUMI MATSUNAGA, KOHEI HATTA, MASATOSHI TAKEICHI*

Cadherins are a family of Ca²⁺-dependent intercellular adhesion molecules. Complementary DNAs encoding mouse neural cadherin (N-cadherin) were cloned, and the cell binding specificity of this molecule was examined. Mouse N-cadherin shows 92 percent similarity in amino acid sequence to the chicken homolog, while it shows 49 percent and 43 percent similarity to epithelial cadherin and to placental cadherin of the same species, respectively. In cell binding assays, mouse N-cadherin did not cross-react with other mouse cadherins, but it did cross-react with chicken N-cadherin. The results indicate that each cadherin type confers distinct adhesive specificities on different cells, and also that the specificity of N-cadherin is conserved between mammalian and avian cells.

THE ADHESIVE SELECTIVITY OF ANIMAL cells is thought to be crucial for controlling the association or movement of cells involved in embryonic morphogenesis (1). Cadherins are a family of intercellular adhesion receptors that may play a role in selective cell adhesion (2). Two members of this family, epithelial cadherin (E-cadherin), which is also called uvomorulin or cell CAM120/80 (3, 4), and placental cadherin (P-cadherin) (5), were characterized molecularly in the mouse. Two other members, liver cell adhesion molecule (L-CAM) (6) and neural cadherin (N-cadherin) (7), were identified in the chicken. The different cadherin types exhibit distinct tissue distribution patterns (8). Both E-cadherin and L-CAM are thought to be the interspecies homologs since they show a similar pattern of tissue distribution in the respective species (8).

Transferring cadherin cDNAs into heterologous cells allowed us to examine the role

of cadherins in cell adhesion (3, 7, 9). We found that L cells transfected with E-cadherin cDNA sorted out from those transfected with P-cadherin cDNA when mixed (10), indicating that these cadherins have binding specificities. Moreover, when the E-cadherin-transfected L cells were added to a suspension of embryonic lung cells, they preferentially attached to the epithelial cells of this tissue, which also express E-cadherin (10). These results suggest that cadherins participate in cell sorting in embryonic tissues.

N-cadherin is expressed in various neural tissues (11, 12) and has been implicated in the attachment of axons to other cells (13), raising the possibility that this molecule is involved in neuronal recognition mechanism. To determine whether this cadherin

Department of Biophysics, Faculty of Science, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606, Japan.

*To whom correspondence should be addressed.

Table 1. Similarities in the amino acid sequences of different cadherins. Sequences from these molecules are divided into three regions, the extracellular (EC), the transmembrane (TM), and the cytoplasmic (CP) regions. Percent similarity was calculated for each region as well as for the entire region of the mature proteins (MP). N, N-cadherin; E, E-cadherin; P, P-cadherin; L, L-CAM.

Cadherin types	Similarity (%)			
	EC	TM	CP	MP
Mouse-mouse				
N to E	46	44	62	49
N to P	43	22	56	43
E to P	53	41	80	58
Chicken-chicken				
N to L	46	50	63	50
Mouse-chicken				
N to N	89	100	99	92
E to L	58	79	89	65

weakly with liver but not with placenta. The molecular size and the tissue distribution pattern of these bands are consistent with the results of RNA blot analysis and are the same as those of chicken N-cadherin (16). N-cadherin in liver is probably derived from fibroblasts of this tissue, as deduced from the chicken data (17).

Sequence similarity of the cloned mouse N-cadherin to E- and P-cadherin of the same species is 49 and 43%, respectively, and that of E- to P-cadherin is 58%; thus, the similarity among different cadherin types in a single species resides in a range of 43 to 58% (Table 1). Mouse E-cadherin shows 65% similarity to chicken L-CAM; this finding contrasts with the high similarity (92%) between the N-cadherin homologs of these species.

In order to verify the proposed interspecies relations among these cadherins, we performed DNA blot analyses. A cDNA probe of mouse N-cadherin gave several bands in mouse DNA and also in chicken DNA digested with Eco RI or Bam HI, under washing conditions with $2\times$ saline sodium citrate (SSC), revealing DNA fragments with 65% sequence similarity (Fig. 3B, the third panel). This probe also gave positive bands in chicken DNA when washed at higher stringency with $0.1\times$ SSC to detect 90% similarity (Fig. 3B, the fourth panel). A cDNA probe of mouse E-cadherin gave positive bands in chicken DNA under washing conditions for detecting 60% sequence similarity in which $5\times$ SSC was used (Fig. 3B, the first panel), but the chicken bands became faint under washing conditions with $2\times$ SSC (Fig. 3B, the second panel). These results suggest that E-cadherin is not as highly conserved as N-cadherin in the mouse and as in the chicken, supporting the notion that L-CAM is the closest relative

to mouse E-cadherin in the chicken cadherin family.

For determining the binding specificity of mouse N-cadherin, we transfected L cells with cDNA encoding the entire length of this molecule, which was inserted into an expression vector containing the β -actin promoter (18). As with the other cadherin cDNAs, the transfected L cells expressed functional N-cadherin molecules, as judged by the cell aggregation assay and immunoblot analysis. We then mixed a single-cell suspension of the transfectants of mouse N-cadherin with that of the L cell lines transfected with mouse E-cadherin or P-cadherin cDNA and examined these cells for adherence to each other. The N-cadherin-transfected L cells tended to aggregate separately from the P-cadherin- or E-cadherin-transfected L cells (Fig. 4), as found previously in the combination of E- and P-cadherin-positive cells (10). Therefore, all three cadherins, N-, E-, and P-type, have independent specificities and thus preferentially interact with their own types through the homophilic mechanism in cell-cell binding reactions.

We tested mouse N-cadherin for its ability to cross-react with chicken N-cadherin by mixing the L cells transfected with cDNA-encoding N-cadherin from each species. In

this combination, the two cell lines randomly intermixed, forming chimeric aggregates (Fig. 4) and indicating that they share the same adhesive specificity.

The above findings suggest that the members of the cadherin family participate in selective cell adhesion and explain how cells of one tissue segregate from those of others when mixed in vitro. For example, the segregation of cells observed in such cell mixtures as liver and neural retina (19), liver and limb bud (20), or liver and heart ventricle (20) can be ascribed to the expression of different cadherins in the different cell types that were combined. Nevertheless, some observations appear incompatible. For example, heart ventricle was sorted out from neural retina (21), although they both expressed N-cadherin. This kind of phenomenon could be explained by assuming the presence of other unidentified cadherins; actually, neural retina has some of the activity of unidentified cadherins (17).

In embryos, the expression pattern of cadherins always correlates with the pattern of segregation of different cell populations (8). Together with these observations, the present findings suggest that the cadherin-mediated adhesive specificities play a significant role in cell sorting processes not only in vitro but also in vivo. These findings also

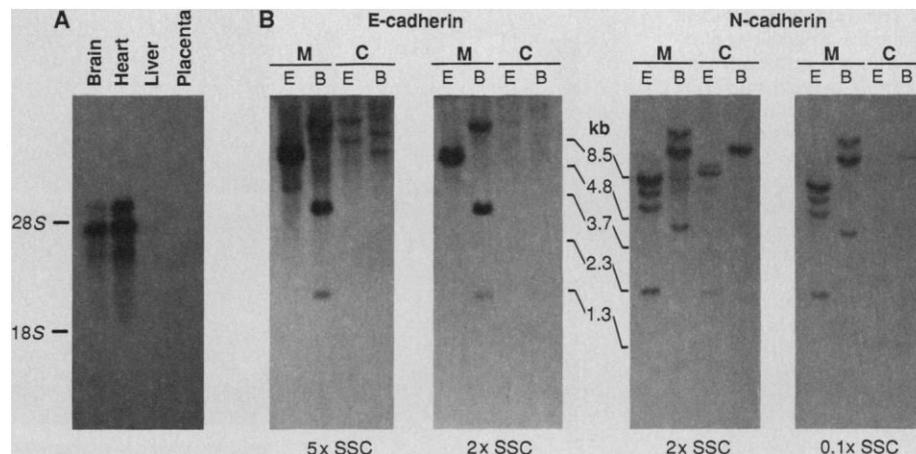


Fig. 3. RNA and DNA blot analysis. (A) RNA blot analysis. Total RNA extracted from brain, heart, and liver of 14-day-old mouse embryos and from placental cones of pregnant mice was subjected to electrophoresis on a 1% agarose-6% formaldehyde gel and transferred directly in $20\times$ SSC to a nitrocellulose filter. The transferred nitrocellulose filter was kept at 80°C at reduced pressure for 3 hours and hybridized with a nick-translated probe of mouse N-cadherin cDNA (mn-1 in Fig. 1). Each lane contains $10\ \mu\text{g}$ of total RNA. The positions of 28S and 18S ribosomal RNA are shown. (B) DNA blot analysis. DNA isolated from mouse (M) and chicken (C) embryos was digested completely with Eco RI (E) or Bam HI (B), separated by electrophoresis on a 0.7% agarose gel, and blotted. To adjust the genome copy number of each species, the amount of DNA on the gels was $10\ \mu\text{g}$ of DNA per lane for the mouse and $3\ \mu\text{g}$ of DNA per lane for the chicken. For hybridization, nick-translated probes of mouse 980-bp N-cadherin cDNA (mn-2 in Fig. 1) and mouse 619-bp E-cadherin cDNA (me-1 in Fig. 1), each corresponding to the NH_2 -terminal region of the mature protein, were used. Hybridization was done for 24 hours at 42°C in a solution consisting of $6\times$ SSC, $50\ \text{mM}\ \text{NaH}_2\text{PO}_4$ (pH 6.5), $5\times$ Denhardt's solution, 1% SDS, 10% dextran sulfate, denatured salmon sperm DNA at $10\ \mu\text{g}/\text{ml}$, and 30% formamide. After hybridization, the filters were washed at 60°C in the indicated SSC solution, which contained 0.1% SDS. When washed with $5\times$ SSC, $2\times$ SSC, and $0.1\times$ SSC, it is assumed that hybridized DNA fragments detected as bands have sequence similarities of about 60, 65, and 90%, respectively (26).

support the idea that N-cadherin might participate in the selective attachment of neurites to particular cell types and in guiding the growth cone to specific targets (13).

What differences among different cadherin molecules generate their binding specificities? Our data showed that the cross-reactivity among different cadherins appears to depend on the extent that their amino acid sequences are similar; that is, only the molecules with a high degree of similarity can cross-react with each other. The amino-terminal region of the cadherins must be particularly important in their specific binding interactions, since some monoclonal antibodies that can interfere with the cadherin-mediated adhesion recognize this region. Monoclonal antibody NCD-2 to chicken N-cadherin shows such activity and recognizes a portion of this molecule, which is located between the amino terminus and the amino acid residue at position 272 (7) (Fig. 2), suggesting that this region may contain the cell binding sites. The amino acid sequence in this region of N-cadherin is highly conserved between the chicken and the mouse, whereby only four amino acid residues out of 108 total residues are different. Therefore, this degree of similarity in the cell binding sites might be necessary for

positive binding interactions between two cadherin molecules. Whether mouse E-cadherin and chicken L-CAM with only 58% similarity in their extracellular domain (Table 1) can bind to each other remains to be determined.

The cross-binding capability between mouse and chicken N-cadherin is consistent with the early observations that some embryonic cells derived from different animal species can intermix with each other (22). We can now ascribe at least in part the nonselective adhesion between cells of different animal species to the presence of the conserved N-cadherin molecules. In fact, the cells used to confirm the occurrence of nonselective adhesion of mouse and chicken cells (23) were brain cells that express N-cadherin.

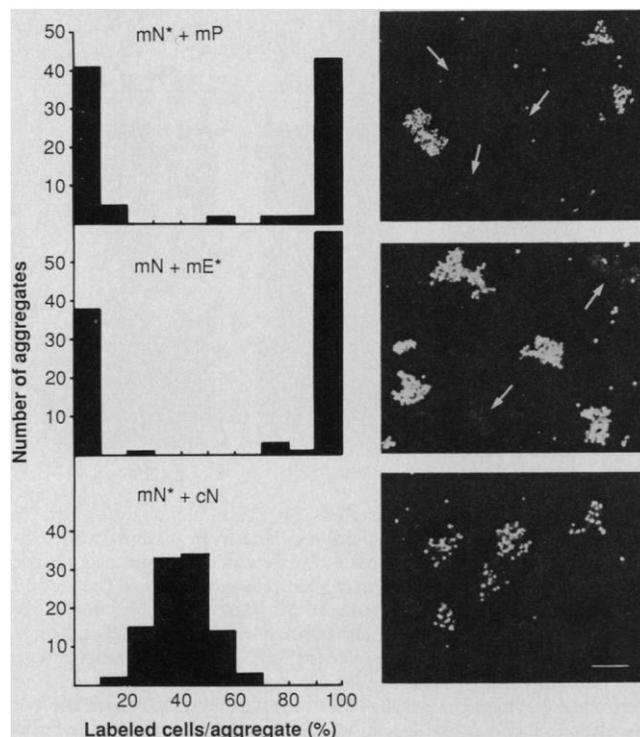
It was, however, reported that cells of some tissues exhibit species-dependent adhesive specificities (24). For example, mammalian liver cells do not adhere to avian liver cells in aggregation experiments (23, 25). The major cadherin in liver parenchymal cells is E-cadherin in the mouse and L-CAM in the chicken. The lower degree of similarity between mouse E-cadherin and chicken L-CAM than between mouse and chicken N-cadherin suggests the possibility that

those heterologous cadherins might have a weaker binding affinity for each other than the others. DNA blot analysis suggests the absence of E-cadherin homologs with a high degree of similarity, which is comparable to the case for the N-cadherin homologs in the chicken. It thus seems that the present findings explain, at least in part, the molecular basis for the mechanism of the species specificity as well as of the tissue specificity of intercellular adhesion. Subsequently, it should be interesting to know the biological significance of the differential diversification of different cadherin types.

REFERENCES AND NOTES

1. P. L. Townes and J. Holtfreter, *J. Exp. Zool.* **128**, 53 (1955); J. P. Trinkaus and P. W. Groves, *Proc. Natl. Acad. Sci. U.S.A.* **41**, 787 (1955); A. A. Moscona, *Proc. Soc. Exp. Biol. Med.* **92**, 410 (1956); T. S. Okada, *J. Embryol. Exp. Morphol.* **13**, 299 (1965).
2. M. Takeichi, T. Atsumi, C. Yoshida, K. Uno, T. S. Okada, *Dev. Biol.* **87**, 340 (1981); M. Takeichi, K. Hatta, A. Nagafuchi, in *Molecular Determinants of Animal Form*, G. M. Edelman, Ed. (Liss, New York, 1985), pp. 223-233.
3. A. Nagafuchi, Y. Shirayoshi, K. Okazaki, K. Yasuda, M. Takeichi, *Nature* **329**, 341 (1987).
4. M. Ringwald *et al.*, *EMBO J.* **6**, 3647 (1987); C. Damsky, J. Richa, D. Solter, K. Knudsen, C. A. Buck, *Cell* **34**, 455 (1983).
5. A. Nose, A. Nagafuchi, M. Takeichi, *EMBO J.* **6**, 3655 (1987).
6. W. J. Gallin, B. C. Sorkin, G. M. Edelman, B. A. Cunningham, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2808 (1987).
7. K. Hatta, A. Nose, A. Nagafuchi, M. Takeichi, *J. Cell Biol.* **106**, 873 (1988).
8. M. Takeichi, *Development* **102**, 639 (1988).
9. G. M. Edelman, B. A. Murray, R. Mege, B. A. Cunningham, W. J. Gallin, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8502 (1987).
10. A. Nose, A. Nagafuchi, M. Takeichi, *Cell* **54**, 993 (1988).
11. K. Hatta, S. Takagi, H. Fujisawa, M. Takeichi, *Dev. Biol.* **120**, 215 (1987).
12. M. Matsunaga, K. Hatta, M. Takeichi, *Neuron* **1**, 289 (1988).
13. M. Matsunaga, K. Hatta, A. Nagafuchi, M. Takeichi, *Nature* **334**, 62 (1988); K. J. Tomaselli, K. M. Neugebauer, J. L. Bixby, J. Lilien, L. Reichardt, *Neuron* **1**, 33 (1988).
14. Filed with GenBank **M22556.
15. A plasmid that encodes a 16-kD mouse N-cadherin fusion protein was constructed by inserting the 359-bp Pvu II fragment (mn-3 in Fig. 1) of mouse N-cadherin cDNA into the Bam HI site of pET3c vector in-frame (27). After BL21 (DE3) host cells were transfected with this plasmid for protein expression (28), the synthesized fusion protein was then fractionated (29), separated by SDS-polyacrylamide gel electrophoresis (PAGE), and electroeluted from gel slices. An antiserum to the mouse N-cadherin fusion protein was obtained as follows: the fusion protein (1 mg) was emulsified with Freund's complete adjuvant and injected subcutaneously into a rabbit. The rabbit was then given this fusion protein with Freund's incomplete adjuvant at three 2-week intervals. The rabbit antiserum yielded N-cadherin-specific antibodies, which were affinity-purified.
16. K. Hatta and M. Takeichi, *Nature* **320**, 447 (1986).
17. S. Hirano, A. Nose, K. Hatta, A. Kawakami, M. Takeichi, *J. Cell Biol.* **105**, 2501 (1987).
18. L cells were transfected with plasmid pBATMNC for the expression of mouse N-cadherin; this plasmid was constructed by inserting a 4.3-kb Eco RI fragment of the full-length mouse N-cadherin

Fig. 4. Selective aggregation of cells expressing different cadherins. Adhesive selectivity among L cells expressing mouse N-cadherin (mNLβ1), chicken N-cadherin (cNLβ1), mouse E-cadherin (ELs8) (3), and mouse P-cadherin (PLs5) (10) was assayed as below. Monolayers of each cell line were completely dissociated into single cells by treatment with 0.01% trypsin in the presence of 1 mM CaCl₂ and the subsequent washing with Ca²⁺-free saline, preserving cadherins (8). When necessary, the monolayers of these cells were labeled with a fluorescent dye, 5- (and 6-) carboxyfluorescein diacetate succinimyl ester (CFSE) before the trypsin treatment (10). The dispersed cells, labeled with CFSE, were then mixed with the unlabeled cells expressing different cadherins in a ratio of approximately 1:1, and allowed to aggregate for 45 min as described previously



(10). Aggregates that were composed of 20 to 60 cells were randomly selected, and the percentage of labeled cells in each aggregate was determined for 100 aggregates in each set of experiments. mN, mNLβ1 cells; cN, cNLβ1 cells; mP, PLs5 cells; mE, ELs8 cells. Asterisks show the cells labeled with CFSE. Fluorescence photomicrographs of cell aggregates are shown for each set of experiments. Arrows indicate cell aggregates, which are composed mainly of unlabeled cells. Note that the labeled and unlabeled cells were segregated from each other in the upper two combinations but not in the bottom combination. Bar, 100 μm.

- cDNA into plasmid pBATneo (10) in place of the Bgl I 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, KIZHAK 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 pH 5.5 (15). Substantial changes occurred in the one-dimensional ^1H NMR spectrum upon zinc binding; two dimensional (2-D) NMR spectra indicate the formation of a single folded conformation in aqueous solution (pH 5.5, 5° to 25°C). Gel mobility assays show that the *Xfin-31* peptide binds nonspecifically to

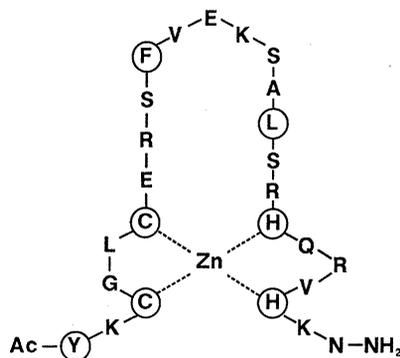


Fig. 1. Amino acid sequence of finger 31 from *Xfin* (13) drawn to illustrate the finger motif (4). Invariant and highly conserved residues are circled. The amino and carboxyl termini of the peptide used for the present studies are blocked by acetyl (Ac) and amide groups, respectively.

DNA in the presence of zinc. No detectable binding occurs if zinc is absent.

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 $^3\text{J}_{\text{HN}\alpha}$ coupling constants measured from a high-resolution double-quantum filtered COSY spectrum recorded at 25°C. The backbone torsion angle ϕ was constrained to $-160^\circ \leq \phi \leq -80^\circ$ for five residues with $^3\text{J}_{\text{HN}\alpha} > 8$ Hz and to $-90^\circ \leq \phi \leq -20^\circ$ for seven residues with $^3\text{J}_{\text{HN}\alpha} < 6$ Hz. No explicit hydrogen-bond constraints were used in the structure calculations.

The zinc was constrained to be tetrahedrally coordinated by Cys³, Cys⁶, His¹⁹, and His²³ (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 ($109 \pm 10^\circ$) 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 N δ or N ϵ of the His ligands, distance geometry calculations were performed with a combination of distance and chirality constraints that allow coordination to either N ϵ or N δ and keep the zinc within the plane of the imidazole ring. These calculations showed unambiguously that both His residues coordinate through their N ϵ 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.