mineralization can also be used to reconstruct variations in seasonal diet quality and growth rate over the course of an animal's life. This technique has many applications. Cementum bands appear in a wide variety of mammals with both grazing and nongrazing diets (*7*, *8*), and aspects of cementum microstructure are visible in many fossils because Sharpey's fiber orientation is sometimes preserved through either collagen preservation or mineral replacement. In addition, because cementum is rarely remodeled or resorbed, it can be used for longitudinal studies of the effects of growth rate and mechanical stress on the structure of mineralized tissues. These results support the hypotheses that collagen orientation in mineralized tissues can be affected by the orientation and degree of tensile forces (*24*, *25*) and that the rates of collagen matrix production and collagen mineralization may be relatively independent (*26*).

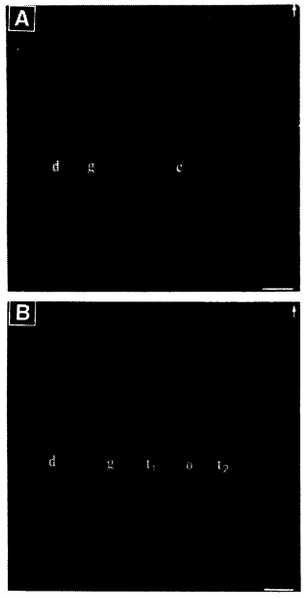


Fig. 4. X-ray microradiographs of variation in acellular cementum mineralization on ground thin sections (50 μm) through goat M_1's just below the enamel-dentine-cementum junction on the lingual surface. Scale bar, 10 μm; the arrow points toward the tooth crown. (**A**) Goat fed the control diet for the entire experiment. The density of the cementum (c) is constant. Cementum-dentine border (granular layer of Tomes) (g) and dentine (d). (**B**) Goat fed a hard food, maintenance-level diet for the first and last 4 months and a hard, low protein and mineral diet for the middle 4 months. First band corresponding to control diet for first 4 months (t_1); more opaque band corresponding to reduced nutrition diet for middle 4 months (o); and band corresponding to control diet for last 4 months (t_2). The more opaque band has a higher density and is relatively more narrow. The teeth were prepared for the micrographs as described (*28*).

REFERENCES AND NOTES

1. A. Boyde and S. Jones, in *Developmental Aspects of Oral Biology*, H. Slavkin and L. Bavetta, Eds. (Academic Press, New York, 1972), pp. 243–274.
2. K. Selvig, *Acta Odontol. Scand.* **23**, 423 (1965).
3. S. Fancy, *Wildl. Soc. Bull.* **8**, 242 (1980).
4. R. M. Laws, *Nature* **169**, 972 (1952).
5. A. Saxon and C. Higham, *Am. Antiq.* **34**, 303 (1969).
6. B. Bourque, K. Morris, A. Speiss, *Science* **199**, 530 (1978).
7. G. Klevezal and S. Kleinenberg, *Age Determination of Mammals by Layered Structure in Teeth and Bone* (Canada Foreign Languages Translations, Quebec, 1967).
8. H. Grue and B. Jensen, *Dan. Rev. Game Biol.* **11**, 1 (1979).
9. D. Lieberman and R. Meadow, *Mammal Rev.* **22**, 57 (1992).
10. F. Peabody, *J. Morphol.* **108**, 11 (1961).
11. W. Hylander, *Arch. Oral Biol.* **31**, 149 (1986).
12. S. Jones, *Scanning Microsc.* **1**, 2003 (1987).
13. H. Lowenstam and S. Weiner, *On Biomineralization* (Oxford Univ. Press, Oxford, 1989).
14. R. Martin and D. Burr, *Structure, Function, and Adaptation of Compact Bone* (Raven, New York, 1989).
15. M. Portigliatti-Barbos, P. Bianco, A. Ascenzi, A. Boyde, *Metab. Bone Dis. Relat. Res.* **5**, 309 (1984).
16. S. Weiner and W. Traub, *FASEB J.* **6**, 879 (1992).
17. For consistency and comparability, nine measurements were taken on each individual at 10-μm intervals on the lingual surface of M_1 near the enamel-dentine-cementum junction. Differences are statistically significant (t test = −34.06, $P < 0.0001$, $n = 18$). Correlations between diet phases and cementum bands were made with fluorochrome dyes.
18. D. Lieberman, thesis, Harvard University (1993).
19. W. Schmidt and A. Kiel, *Polarizing Microscopy of Dental Tissues* (Pergamon, Oxford, 1972).
20. S. Jones, in *Dental Anatomy and Embryology*, J. Osborne, Ed. (Blackwell, Oxford, 1981), pp. 66–209.
21. A. de Ricqlès, F. Meunier, J. Castanet, H. Francillon-Veillot, in *Bone*, B. Hall, Ed. (CRC Press, Boca Raton, FL, 1991), vol. 3, chap. 1.
22. J. Castanet, H. Francillon-Veillot, F. Meunier, A. de Ricqlès, in (*21*), vol. 7, chap. 9.
23. M. Weinstock and C. Leblond, *J. Cell Biol.* **56**, 838 (1973).
24. L. Lanyon, in (*21*), vol. 3, chap. 2.
25. E. Burger and J. Veldhuijzen, in (*21*), vol. 7, chap. 2.
26. M. Glimcher, in *The Nature of the Mineral Component of Bone and the Mechanism of Calcification*, P. Griffin, Ed. (American Academy of Orthopedic Surgeons, Park Ridge, IL, 1987), pp. 49–69.
27. A fresh tooth was cut out of the mandible with a rotary saw and the periodontal ligament was stripped. The tooth was then immersed in 1 M acetone for 10 min, fractured in the mesio-distal plane through the tooth root, rinsed with distilled H_2O, air-dried, coated with 200 nm of gold, and photographed at 20 kV.
28. Fresh teeth were cut out of the mandible with a rotary saw, and the periodontal ligaments were stripped. Teeth were fixed in 10% formaldehyde for 24 hours, rinsed in distilled H_2O for 1 hour, dehydrated in 1 M ethanol, embedded in epoxy, sectioned in the mesio-distal plane, ground to a thickness of 50 μm, and polished. Specimens were photographed at 110 KV for 20 min.

29. I thank O. Bar-Yosef, A. W. Crompton, F. A. Jenkins, W. Landis, R. Meadow, D. Pilbeam, K. Reeve, and S. Weiner for their suggestions and help. This work was supported by grants from the L. S. B. Leakey Foundation and National Science Foundation.

14 April 1993; accepted 9 July 1993

Repair of DNA Methylphosphotriesters Through a Metalloactivated Cysteine Nucleophile

Lawrence C. Myers, Michael P. Terranova, Ann E. Ferentz, Gerhard Wagner,* Gregory L. Verdine*

The *Escherichia coli* Ada protein repairs methylphosphotriesters in DNA by direct, irreversible methyl transfer to one of its own cysteines. Upon methyl transfer, Ada acquires the ability to bind specific DNA sequences and thereby to induce genes that confer resistance to methylating agents. The amino-terminal domain of Ada, which comprises the methylphosphotriester repair and sequence-specific DNA binding elements, contains a tightly bound zinc ion. Analysis of the zinc binding site by cadmium-113 nuclear magnetic resonance and site-directed mutagenesis revealed that zinc participates in the autocatalytic activation of the active site cysteine and may also function as a conformational switch.

Methylation of DNA can occur enzymatically and nonenzymatically. Whereas enzymatic methylation fulfills an essential role in many organisms, nonenzymatic methylation insults the genome with a variety of toxic and mutagenic adducts (*1*, *2*). To counter this threat, cells express a variety of proteins that recognize and repair aberrantly methylated DNA. A noteworthy example is *Escherichia coli* Ada, which repairs the mutagenic adduct 6-O-methylguanine by direct, irreversible methyl transfer to a cysteine residue in the COOH-terminal domain (*3*). Ada also repairs the S_p diastereomer of DNA methylphosphotriesters (MePs) by direct methyl transfer to a second cysteine, Cys^{69}, located in the NH_2-

L. C. Myers, Program for Higher Degrees in Biophysics, Harvard University, Cambridge, MA 02138.
M. P. Terranova and A. E. Ferentz, Department of Chemistry, Harvard University, Cambridge, MA 02138.
G. Wagner, Program for Higher Degrees in Biophysics, Harvard University, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.
G. L. Verdine, Program for Higher Degrees in Biophysics and Department of Chemistry, Harvard University, Cambridge, MA 02138.

*To whom correspondence should be addressed.

terminal domain (Fig. 1A). Methyl transfer to Cys^{69} unmasks a sequence-specific DNA binding activity in the NH_2-terminal domain, conferring on Ada the ability to activate several methylation-resistance genes. Thus, Ada functions as a chemosensor for methylation damage in the cell, inducing resistance genes in response to challenging dosages.

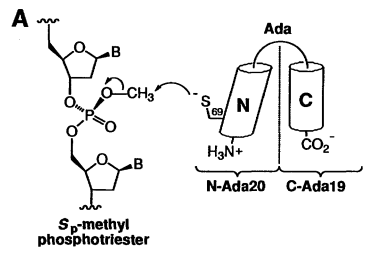
The NH_2-terminal domain (N-Ada20), which contains the MeP and sequence-specific DNA binding domains, possesses a tightly bound zinc ion that is essential for folding in vivo and in vitro (*4*). To identify the metal ligand residues of Ada, we used ^{113}Cd nuclear magnetic resonance (NMR) spectroscopy, which has been widely used to study the coordination environments of zinc-binding proteins (*5, 6*). Although Cd(II) is slightly larger than Zn(II), both ions possess the same valence-shell electronic configuration (d_{10}) and thus exhibit similar preferences with respect to ligand atom types and arrangement (*7*). Moreover, zinc-binding proteins are generally functional in the Cd(II) form (*6*). The ^{113}Cd-substituted N-Ada20 (Cd-N-Ada20) (Fig. 1), prepared by biosynthetic labeling (*8*), retained the ability to repair MeP and bind DNA in a sequence-specific manner. However, its rate of MeP repair was one-fourth that of Zn-N-Ada20, suggesting that a change in the metal affects the structure of the protein or that the metal is directly involved in the rate-determining step for methyl transfer (*9*).

The ^{113}Cd NMR spectrum of Cd-N-Ada20 (*8*) exhibited a single peak at 684 parts per million (ppm) (Fig. 2A), which is characteristic of a metal tetrahedrally coordinated to four Cys residues (*5, 10*). A virtually identical ^{113}Cd peak was obtained with an Ada fragment comprising residues 1 to 92 (Cd-N-Ada10), which indicates that the metal ligand residues are contained entirely within the first 92 residues of Ada (Fig. 2B). Of the six putative ligand Cys residues present in N-Ada10 (Fig. 1B), four are conserved (Cys^{38}, Cys^{42}, Cys^{69}, and Cys^{72}) and two are not (Cys^{6} and Cys^{91}) (*4*).

To test whether the nonconserved Cys residues function as metal ligands, we examined the effect of mutations at these positions on the ^{113}Cd NMR signal (*11*). The N-Ada20 domains containing Cys^{6} to Ala (C6A N-Ada-20) or Cys^{91} to Ala (C91A N-Ada20) mutations exhibited ^{113}Cd NMR spectra similar to that of the wild-type protein (Fig. 2A). Because the replacement of a cysteine ligand with virtually any other residue should have a dramatic effect on the ^{113}Cd chemical shift, these results rule out Cys^{6} and Cys^{91} as ligands and thus implicate the remaining four candidates.

Scalar coupling between ^{113}Cd and ligand protons can permit the sequence-specific identification of ligand residues, provided that the ^{1}H NMR spectrum of the protein has been assigned. Sequence-specific ^{1}H resonance assignments have been obtained for Zn-N-Ada10 (*12*). The two-dimensional ^{1}H total correlation spectroscopy (TOCSY) and nuclear Overhauser and exchange spectroscopy (NOESY) spectra of Zn-N-Ada10 and Cd-N-Ada10 (*12*) were almost identical, except that the β protons of Cys^{38}, Cys^{42}, Cys^{69}, and Cys^{72} were shifted by as much as 0.2 ppm. This similarity provided direct evidence that the Cd protein is closely related in structure to the Zn protein. A one-dimensional ^{1}H-^{113}Cd heteronuclear multiple quantum coherence (HMQC) NMR spectrum of Cd-N-Ada10 (Fig. 3A) revealed three-bond coupling between the metal and the β protons of Cys^{38}, Cys^{42}, and Cys^{72} (*13*). These assignments were confirmed in a two-dimensional ^{113}Cd-edited NOESY-HMQC experiment (*14*) (Fig. 3B). This method detects nuclear Overhauser effect (NOE) cross peaks of protons close in space to protons that are coupled to ^{113}Cd. Although the nature of the metal and residue at position 69 affect each other's chemical shifts, a signal for Cys^{69} was not readily observed in either ^{113}Cd-edited experiment. Failure to observe coupling between ^{113}Cd and the β protons of a ligand Cys is not uncommon and can result from peak overlap (in this case with Cys^{72}), a small scalar ^{113}Cd-^{1}H coupling constant, or kinetic lability of the metal-ligand bond (*6*).

To demonstrate direct interaction of Cys^{69} with the metal ion, we generated mutant proteins in which Cys^{69} was



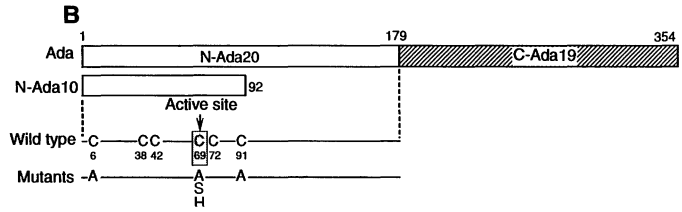


Fig. 1. Schematics of the Ada protein, its constituent domains, and the autocatalytic DNA repair reaction of the NH_2-terminal domain. (**A**) Ada consists of two domains. The 20-kD NH_2-terminal domain (N-Ada20) repairs the S_p diastereomer of methylphosphotriesters and, in its methylated form, binds DNA in a sequence-specific manner; the 19-kD COOH-terminal domain (C-Ada19) repairs the mutagenic lesion 6-*O*-methylguanine by direct methyl transfer to Cys^{321}. Each domain retains its activity when separated from the other. (**B**) The N-Ada20 domain comprises residues 1 to 179. A 10-kD fragment generated by gene truncation, N-Ada10, comprises residues 1 to 92. All of the ligand residues of Ada are contained within N-Ada10; the six potential ligand cysteines are indicated by a C, with their position in the sequence indicated below. The positions of mutations made in N-Ada20 are as indicated below the corresponding cysteines. Ada, N-Ada20, and N-Ada10 bind 1 mol of zinc per mole of protein.

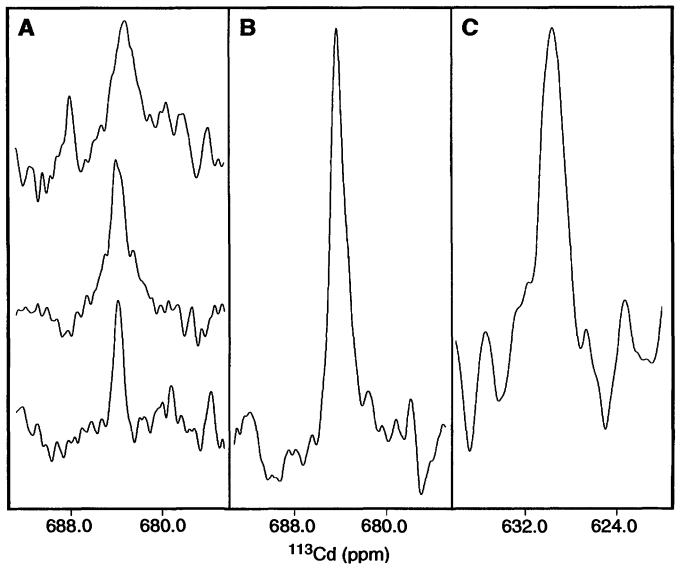


Fig. 2. Direct-detection ^{113}Cd NMR spectra of (**A**) wild-type N-Ada20 (bottom trace; peak at 684 ppm) and the C91A (middle trace; peak at 684 ppm) and C6A (top trace; peak at 683.1 ppm) mutants; (**B**) wild-type N-Ada10 (peak at 684 ppm); and (**C**) the C69H mutant of N-Ada20 (peak at 629.5 ppm). For each spectrum, 1,024 to 12,288 transients were collected and a sweep width from 0 to 750 ppm was scanned.

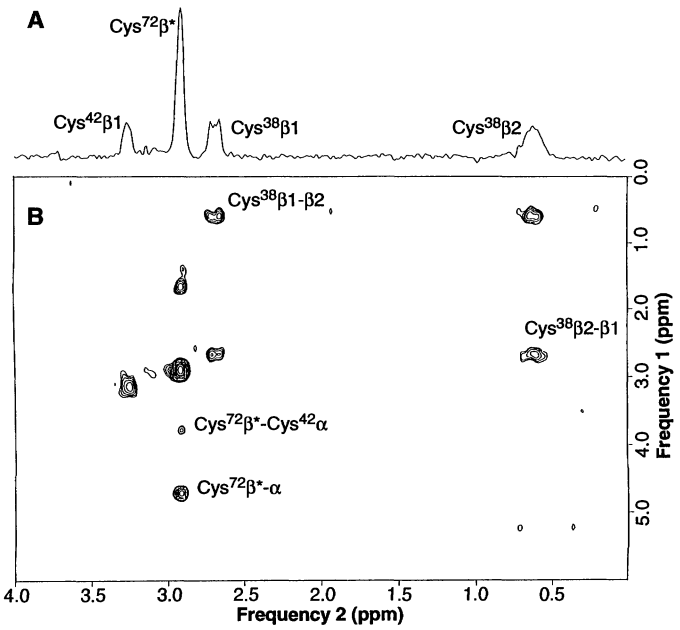


Fig. 3. NMR spectra resulting from (**A**) a ^{1}H-^{113}Cd HMQC experiment and (**B**) a two-dimensional ^{113}Cd-edited NOESY-HMQC experiment on N-Ada10. The defocusing and refocusing delay in (A) was set to 10 ms and 4096 transients were collected. The defocusing and refocusing delay in (B) was 11 ms and the NOESY mixing time was 125 ms. Denoted in (A) are the identities of β-methylene protons that are strongly coupled to ^{113}Cd; denoted in (B) are relevant NOE cross peaks, indicating the spatial proximity between protons. The horizontal scales in (A) and (B) are identical.

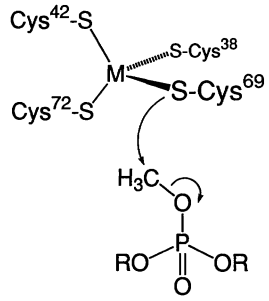


Fig. 4. A model for metalloactivated repair of a DNA methylphosphotriester by Ada. Residue Cys^{69}, present either as a metal-bound thiolate or a transiently free thiolate, attacks the methyl carbon of a DNA MeP.

changed to Ala, Ser, or His (C69A, C69S, and C69H, respectively) (Fig. 1B) (*11*). Whereas wild-type N-Ada20, C6A N-Ada20, and C91A N-Ada20 were expressed as soluble, active proteins in zinc- or cadmium-supplemented media, the C69A and C69S mutant N-Ada20 proteins were expressed as insoluble, inactive aggregates. Furthermore, partially purified inclusion bodies containing the mutant proteins did not contain substantial amounts of metal (*15*). On the other hand, when Cys^{69} was mutated to a potential metal ligand residue, His, the resulting protein was expressed in a soluble but inactive form. Because the folding of N-Ada20 depends on Zn(II) or Cd(II), the insolubility of the C69A and C69S mutants can be ascribed to their failure to bind the metal, whereas the solubility of the C69H mutant suggests that this protein does bind the metal. Consistent with this hypothesis, the C69H mutant of N-Ada20 exhibited a well-defined ^{113}Cd NMR spectrum. Furthermore, the ^{113}Cd signal at 629.5 ppm (Fig. 2C) is similar to that observed for proteins possessing $(Cys)_3$His-coordinated cadmium (*16*). These results lend strong support to the notion that Cys^{69} is coordinated to the metal ion of Ada.

Our finding that the active site nucleophile Cys^{69} is ligated to the metal in Ada may offer insight into the protein's switch from a DNA MeP repair agent to a transcriptional regulator. The conversion of a thiolate-*S* ligand to the more weakly bound thioether-*S* ligand may drive ligand reorganization at the metal center (*17*). Such reorganization may in turn propel structural changes that reveal the sequence-specific DNA binding conformation of the protein. The chemosensory mechanism used by Ada may resemble that of matrix metalloproteases, in which chemical modification of a zinc-bound Cys results in dissociation of that ligand and the unveiling of the masked zinc-dependent protease activity of the enzyme (*18*).

Many proteins are known to coordinate zinc through four Cys residues. However, except for Ada the metal has invariably been shown merely to stabilize the folded structure of the protein (*19*). Ada differs even from the numerous transcription factors that contain structural $(Cys)_4$-ligated zinc (*20–22*) in that Ada's sequence-specific DNA binding activity is potentiated only after methylation of one of its Cys ligands. The catalysis of biochemical reactions by protein-bound zinc ions has also received a great deal of attention (*19*). Ada is unusual in that the zinc serves to activate the nucleophilicity of one of its amino acid ligands (Fig. 4). The activation of Cys^{69} by Ada may be viewed as analogous to the activation of H_2O by carbonic anhydrase because the metal assists deprotonation of the protonucleophile, presenting a metal-bound anion for reaction with an electrophile (*23*). The kind of reaction observed with Ada—the selective alkylation of a metal-bound thiolate—is also widely documented in transition metal complexes (*24–26*). The irreversibility of the methyl transfer reaction catalyzed by Ada dictates that such metalloactivation by the protein be termed autocatalytic (*27*) rather than catalytic. However, there is no a priori reason why other proteins could not use ligand metalloactivation to accelerate the formation of reversible bonds.

REFERENCES AND NOTES

1. E. Li, T. H. Bestor, R. Jaenisch, *Cell* **69**, 915 (1992); H. Cedar, *ibid.* **53**, 3 (1989).
2. B. Singer and D. Grunberger, *Molecular Biology of Mutagens and Carcinogens* (Plenum, New York, 1983).
3. T. Lindahl, B. Sedgewick, M. Sekiguchi, Y. Nakabeppu, *Annu. Rev. Biochem.* **57**, 133 (1988); B. Demple, in *Protein Methylation*, W. K. Paik and S. Kim, Eds. (CRC Press, Boca Raton, FL, 1990), pp. 285–304.
4. L. C. Myers, M. P. Terranova, H. M. Nash, M. A. Markus, G. L. Verdine, *Biochemistry* **31**, 4541 (1992).
5. M. F. Summers, *Coord. Chem. Rev.* **86**, 43 (1988).
6. J. E. Coleman and P. Gettins, *Zinc Enzymes* (Birkhauser, Boston, 1986), pp. 77–99.
7. F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry* (Wiley, New York, ed. 5, 1988), pp. 597–609.
8. The ^{113}Cd-N-Ada20 was generated in *E. coli* grown on minimal media supplemented with $^{113}CdCl_2$ (10 μM; MSD Isotopes, Merck, Rahway, NJ) and purified as described for Zn-N-Ada20 (*4*). The protein was exchanged into a buffer containing 25 mM sodium phosphate (pH 7.8), 50 mM NaCl, and 10 mM 2-mercaptoethanol and adjusted to a concentration of 1 to 2 mM. The concentrated protein was lyophilized, resuspended in an equal volume of degassed D_2O (99.99%), and measured at 25°C under argon. The ^{113}Cd-N-Ada10 was prepared by the same procedure. Measurements were made in a Bruker (Rhein Stetten, Germany) AMX600 spectrometer equipped with a 5-mm inverse geometry broadband probe. All ^{113}Cd chemical shifts were referenced to a ^{113}Cd-EDTA complex with a signal at 86 ppm.
9. F. Jackow, L. C. Myers, G. L. Verdine, unpublished data.
10. The ^{113}Cd chemical shifts for proteins with a tetrahedral $(Cys)_4$ arrangement range from 750 ppm [horse liver alcohol dehydrogenase: B. R. Bobsein and R. J. Myers, *J. Biol. Chem.* **256**, 5313 (1981)] to 669 ppm [GAL4 (a cluster composed of six cysteines, two of which are bridging, bound tetrahedrally to two zinc ions; each zinc is surrounded by four Cys residues): T. Pan and J. E. Coleman, *Biochemistry* **29**, 3023 (1990)]. Complexes formed between ^{113}Cd and simple alkanethiol ligands typically give ^{113}Cd chemical shifts between 650 and 700 ppm (*5*).
11. Mutant proteins were generated by site-directed mutagenesis (Muta-Gene Kit, Bio-Rad) of the single-stranded form of a phagemid (*4*) that overproduced N-Ada20; mutations were confirmed by sequencing. Labeling with ^{113}Cd and purification was carried out and NMR spectra were determined as in (*4*), except that the C6A and C91A N-Ada20 mutants were adjusted to a concentra-

tion of 1 to 2 mM in 25 mM sodium phosphate (pH 8.0), 50 mM NaCl, 10 mM dithiothreitol, and 20 mM ammonium sulfate added to 5% D_2O and measured directly at 5°C.
12. L. C. Myers, G. L. Verdine, G. Wagner, unpublished data.
13. D. H. Live, I. M. Armitage, D. C. Dalgarno, D. Cowburn, *J. Am. Chem. Soc.* **107**, 1775 (1985); M. H. Frey *et al.*, *ibid.*, p. 6847.
14. E. Wörgötter, G. Wagner, K. Wüthrich, *ibid.* **108**, 6162 (1986).
15. M. P. Terranova, L. C. Myers, G. L. Verdine, unpublished data.
16. The nucleocapsid protein of human immunodeficiency virus, which has two $(Cys)_3His$ metal binding sites, gives ^{113}Cd chemical shifts of 659 and 640 ppm [D. W. Fitzgerald and J. E. Coleman, *Biochemistry* **30**, 5195 (1991)]; a small peptide corresponding to the NH_2-terminal finger of this same viral protein gives a ^{113}Cd chemical shift of 653 ppm [T. L. South, B. Kim, M. F. Summers, *J. Am. Chem. Soc.* **111**, 395 (1989)]. In the T4 gene 32 protein, three known Cys residues and an unidentified His ligate the metal, resulting in a ^{113}Cd chemical shift of 639 ppm [D. P. Giedroc, B. A. Johnson, I. M. Armitage, J. E. Coleman, *Biochemistry* **28**, 2410 (1989); D. P. Giedroc *et al.*, *ibid.* **31**, 765 (1992)].
17. C. G. Kuehn and S. S. Iseid, *Prog. Inorg. Chem.* **27**, 154 (1980).
18. E. B. Springman, E. L. Angleton, H. Birkedal-Hansen, H. E. Van Wart, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 364 (1990).
19. B. L. Vallee and D. S. Auld, *Biochemistry* **29**, 5647 (1990); *Proc. Natl. Acad. Sci. U.S.A.* **87**, 220 (1990).
20. J. M. Berg, *Curr. Opin. Struct. Biol.* **3**, 11 (1993).
21. R. Kaptein, *ibid.* **1**, 63 (1991).
22. B. L. Vallee, J. E. Coleman, D. S. Auld, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 999 (1991).
23. J. E. Coleman, *Prog. Bioorg. Chem.* **1**, 159 (1971).
24. L. F. Lindoy and D. H. Busch, *Inorg. Chem.* **13**, 2494 (1974).
25. R. C. Elder, G. J. Kennard, M. D. Payne, E. Deutsch, *ibid.* **17**, 1296 (1978).
26. B. Rajanikanth and B. Ravindranath, *Indian J. Chem. B* **23**, 1043 (1984).
27. T. R. Cech, *Science* **236**, 1532 (1987).
28. We thank B. Vallee, B. Shapiro, and L. Kunze for carrying out the atomic absorption analyses and for related discussions; F. Jackow for communication of unpublished results and valuable discussions; and J. Baleja for assistance with the ^{113}Cd NMR. Supported by the Chicago Community Trust (Searle Scholars Program) and the National Science Foundation (Presidential Young Investigator Program) (G.L.V.) and the National Institutes of Health (G.W.). L.C.M. is a Howard Hughes Medical Institute predoctoral fellow.

27 April 1993; accepted 11 June 1993

Deletion of the Paired α5(IV) and α6(IV) Collagen Genes in Inherited Smooth Muscle Tumors

Jing Zhou,* Toshio Mochizuki, Hubert Smeets, Corinne Antignac, Pekka Laurila, Anne de Paepe, Karl Tryggvason, Stephen T. Reeders

The gene encoding α6(IV) collagen, *COL4A6*, was identified on the human X chromosome in a head-to-head arrangement and within 452 base pairs of the α5(IV) collagen gene, *COL4A5*. In earlier studies, intragenic deletions of *COL4A5* were detected in a subset of patients with Alport syndrome (AS), a hereditary defect of basement membranes. In some families, AS cosegregates with diffuse leiomyomatosis (DL), a benign smooth muscle tumor diathesis. Here it is shown that patients with AS-DL harbor deletions that disrupt both *COL4A5* and *COL4A6*. Thus, type IV collagen may regulate smooth muscle differentiation and morphogenesis.

Basement membranes (BMs) compartmentalize tissues and provide important signals for the differentiation of the cells they support. Type IV collagen, the major structural component of BM, is a triple-helical molecule composed of three α chains (*1*). To date, five genetically distinct type IV isoforms have been described in mammals (*1–4*). The α1(IV) and α2(IV) chains encoded by *COL4A1* and *COL4A2*, respectively, are ubiquitous, whereas α3(IV), α4(IV), and α5(IV) have restricted tissue distributions (*3*, *5*). On the basis of sequence similarities, the chains fall into two classes: α1(IV), α3(IV), and α5(IV) compose the α1-like class, and α2(IV) and α4(IV) compose the α2-like class. The human genes for the α1(IV) and α2(IV) chains are located in a head-to-head configuration on chromosome 13 (*6*), and the α3(IV) and α4(IV) genes are similarly arranged on chromosome 2 (*7*). Thus, it appears that the type IV collagens evolved by duplication of an ancestral α chain gene, giving rise to a pair of α chain genes with closely apposed 5′ ends. The pair presumably underwent additional duplication, giving rise to the ancestors of the α1(IV)-α2(IV) and α3(IV)-α4(IV) gene pairs. We predicted that the α5(IV) gene, a member of the α1(IV)-like class, might also be paired with an α2(IV)-like gene that had not yet been identified.

Mutations in *COL4A5* are estimated to be present in ~50% of X-linked AS cases (*8*). In males AS is characterized by progressive renal failure, sensorineural deafness, and ocular lesions; female carriers are mildly affected. In some families AS cosegregates with DL (*9*), a benign proliferation of smooth muscle in the esophagus, female genitalia, and trachea. Both sporadic and hereditary cases have been reported (*10*). We have shown that AS-DL patients have deletions that include the 5′ end of *COL4A5* (*11*), whereas AS patients without DL have internal deletions or point mutations of *COL4A5* (*8*). These results suggested that DL is caused by the deletion of an unidentified gene located upstream of *COL4A5*.

To isolate the putative type IV collagen gene upstream of *COL4A5*, we screened an X-chromosome library with JZ-4, an α5(IV) cDNA clone (*4*), and isolated a 14.1-kb clone, λLA226 (Fig. 1). It contained exon 1 of *COL4A5* and an upstream 2.8-kb Hind III fragment, LA226-H6, that displayed cross-species hybridization. We therefore used LA226-H6 to probe an adult kidney cDNA library. Three identical clones, JZK-1, JZK-2, and JZ-3, contained an open reading frame (1643 bp) encoding a 21–amino acid signal peptide, a 25–amino acid noncollagenous segment, and a 502–amino acid collagenous domain with nine interruptions (Fig. 2) that are believed to confer flexibility in type IV collagens. The deduced translation product, which we have termed α6(IV), is a type IV collagen that has not previously been detected genetically or biochemically (*12*). Sequence analysis clearly places α6(IV) in the α2(IV)-like class (Fig. 2). The head-to-head arrangement of *COL4A5* and *COL4A6* resembles that of *COL4A1* and *COL4A2*.

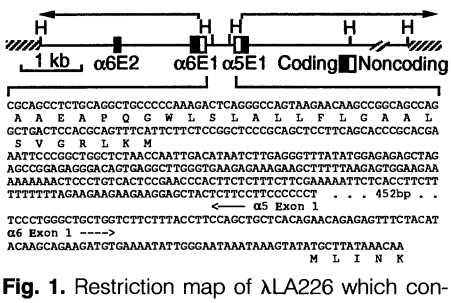


Fig. 1. Restriction map of λLA226 which contains the 5′ ends of both *COL4A6* and *COL4A5*. The striped bars represent phage arms; H, Hind III. The DNA sequence of the region containing the first exon of each gene is shown as well as the deduced amino acid sequences of the first exons (*24*).

J. Zhou, T. Mochizuki, S. T. Reeders, Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, CT 06536-0812.
H. Smeets, Clinical Genetics Center, University Hospital, Nijmegen, Netherlands.
C. Antignac, INSERM Unité 192, Hôpital Necker–Enfants Malades, Paris, France.
P. Laurila, Department of Pathology, University of Helsinki, Helsinki, Finland.
A. de Paepe, Department of Human Genetics, University Hospital, Ghent, Belgium.
K. Tryggvason, Biocenter and Department of Biochemistry, University of Oulu, Oulu, Finland.

*To whom correspondence should be addressed.