J. M. Keller, *Biochemistry* **28**, 8100 (1989).
27. U. K. Laemmli, *Nature* **227**, 680 (1970).
28. G. Carpenter and S. Cohen, *J. Cell. Physiol.* **88**, 227 (1976).
29. We stained with the antibody to BrdU following manufacturer's instructions (Becton-Dickinson).
30. We thank S. Gandy and H. Monis for expert assistance in these experiments, C. Hart of Zymogen for mouse IgG monoclonal antibodies to bFGF, and C. Basilico, New York University School of Medicine for k-FGF. Supported by NIH grants HD21881 to A.C.R. and AR39467 to B.B.O. B.B.O. is the recipient of PEW and SHAW scholarships. A.K. is supported by NIH training grant 5T32H007118.

4 January 1991; accepted 18 April 1991

Identification of p53 as a Sequence-Specific DNA-Binding Protein

SCOTT E. KERN, KENNETH W. KINZLER, ARTHUR BRUSKIN, DAVID JAROSZ, PAULA FRIEDMAN, CAROL PRIVES, BERT VOGELSTEIN*

The tumor-suppressor gene p53 is altered by missense mutation in numerous human malignancies. However, the biochemical properties of p53 and the effect of mutation on these properties are unclear. A human DNA sequence was identified that binds specifically to wild-type human p53 protein in vitro. As few as 33 base pairs were sufficient to confer specific binding. Certain guanines within this 33–base pair region were critical, as methylation of these guanines or their substitution with thymine abrogated binding. Human p53 proteins containing either of two missense mutations commonly found in human tumors were unable to bind significantly to this sequence. These data suggest that a function of p53 may be mediated by its ability to bind to specific DNA sequences in the human genome, and that this activity is altered by mutations that occur in human tumors.

THE GENE FOR THE NUCLEAR PHOSphoprotein p53 is the most commonly mutated gene yet identified in human cancers (*1*); missense mutations occur in tumors of the colon, lung, breast, ovary, bladder, and several other organs (*2, 3*). One of the important challenges of current cancer research is the elucidation of the biochemical properties of the p53 gene product and the way in which mutations of the p53 gene affect these properties. Although some biological characteristics of p53 have been defined, such as its ability to suppress the growth of in vitro–transformed murine cells (*4*) or human cancer cells (*5*), the biochemical basis of this suppression remains unknown.

As a step toward understanding such properties, we have attempted to determine whether p53 binds to specific DNA sequences within the human genome. Two previous lines of evidence stimulated these studies. First, p53 can provide a transcriptional activation function when fused to a DNA-binding polypeptide such as GAL4 (*6*). Second, p53 can bind nonspecifically to DNA (*7*), and such nonspecific binding was altered in each of 15 human tumor-derived or murine-transforming forms of mutant p53 tested (*8, 9*). Because many proteins with a specific DNA-binding function also bind nonspecifically to DNA (*10*), we have suggested that a sequence-specific binding ability of p53 might exist and be a functional target of p53 mutations (*9*).

To identify a sequence-specific binding site, cloned DNA sequences were screened by means of an immunoprecipitation technique (*11, 12*). Two classes of clones were tested. The first consisted of 400 clones containing inserts of 300 to 1000 bp obtained randomly from the human genome. The second class consisted of cosmid and plasmid clones chosen because they contained sequences that might be important in normal growth control (*12*). Each clone was digested with an appropriate restriction endonuclease, end-labeled with ^{32}P, and incubated with p53 from a lysate of cells infected with a recombinant vaccinia virus expressing p53 (*11*). Labeled DNA fragments that bound to p53 were then recovered by immunoprecipitation with monoclonal antibodies against p53 (anti-p53). Of the more than 1400 restriction fragments tested, only two, both from the second class of clones, bound reproducibly to p53 under the experimental conditions used: a 259-bp Hinf I fragment (fragment A) of clone 772 C_{BE} (Fig. 1A, panel 2), and a 190-bp Hinf I fragment (fragment B) of clone λ 5R (Fig. 1A, panel 3); these fragments bound to a far greater extent (at least tenfold more) than any of the other labeled fragments of larger or smaller size present in the same assay mixes.

Subsequent efforts were concentrated on fragment A. First, it was demonstrated that detection of the binding of fragment A was dependent on both p53 protein and anti-p53. Lysates from cells infected with wild-type vaccinia virus (devoid of p53) were not able to specifically immunoprecipitate fragment A (Fig. 1B). Similarly, the precipitation of fragment A was dependent on the presence of anti-p53 (Fig. 1B). The binding was evident in lysates prepared from either human HeLa cells or monkey BSC40 cells

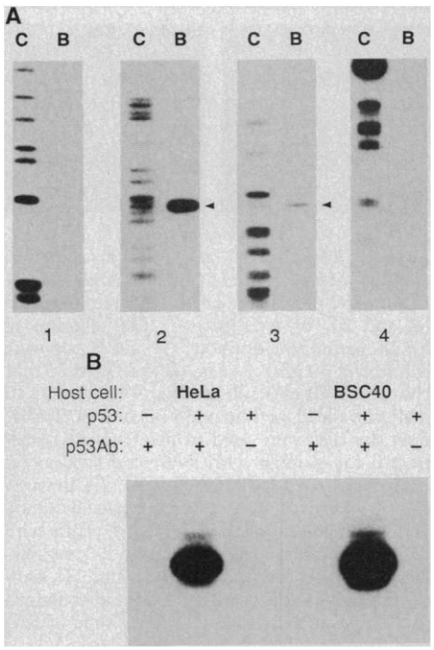

Fig. 1. (**A**) Screening for fragments bound by p53 by an immunoprecipitation assay (*11*). Clones (*12*) were cleaved by a restriction endonuclease, end-labeled, and incubated with lysates of vaccinia virus–infected cells synthesizing wild-type p53. After immunoprecipitation with anti-p53 (*17*), bound DNA fragments were recovered and separated on a nondenaturing polyacrylamide gel. Panel 1 contains the hFosAva2 clone; panel 2, 772 C_{BE}; panel 3, λ 5R; panel 4, a pool of clones with inserts of randomly cloned human genomic sequences. The 772 C_{BE} and λ 5R contain Hinf I fragments (259 and 190 bp, respectively) that bound p53 relatively strongly (arrowheads). C, control lane, containing 2% of the labeled DNA used in the binding reactions. B, bound DNA recovered from the immunoprecipitate. (**B**) Tests for dependence on p53 and specific antibody. Cell lysates were produced by infection with vaccinia virus that did (+) or did not (−) contain an insert of wild-type p53 cDNA. Immunoprecipitation was performed with anti-p53 (+) or normal mouse immunoglobulin G (−).

S. E. Kern, The Oncology Center and Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21231.
K. W. Kinzler and B. Vogelstein, The Oncology Center, The Johns Hopkins University School of Medicine, Baltimore, MD 21231.
A. Bruskin and D. Jarosz, Applied Biotechnology, Cambridge, MA 02142.
P. Friedman and C. Prives, Department of Biological Sciences, Columbia University, New York, NY 10027.

*To whom correspondence should be addressed.

infected with vaccinia virus and expressing wild-type p53 (Fig. 1B). Second, we wished to determine whether p53 binding was dependent on factors other than p53 that might be present in the vaccinia virus-infected cell lysates. Affinity-purified, baculovirus-produced, wild-type p53 (*11*) was therefore tested and found to bind fragment A strongly (Fig. 2). This suggested that the binding to fragment A was an intrinsic property of the p53 polypeptide.

Next, we asked whether mutations of p53 altered the ability of the protein to bind to fragment A. The proportion of fragment A bound to wild-type p53 increased in tandem with the amount of p53 added to the assay mixture (Fig. 2A). In contrast, fragment A did not specifically bind to a mutant form of p53 (273^{his}) even at the highest protein concentration used. The 273^{his} mutation is the most common p53 mutant identified in human tumors (*1–3*). Another p53 mutant (175^{his}) commonly found in human tumors also was unable to bind to fragment A (Fig. 2B).

We then attempted to define the sequences within fragment A that allowed it to bind to wild-type p53. Fragment A was subcloned, and the 259-bp insert from subclone 10d (Fig. 3A) bound to p53 as expected. A strategy based on the polymerase chain reaction (PCR) and restriction endonuclease digestion was used to generate subfragments of this clone (*13*). Subfragment 3, including bp 106 to 294 (Fig. 4B, panel 2) bound well to p53 as did subfragment 4, containing bp 1 to 141 (Fig. 4B, panel 3). This localized the critical sequences to bp 106 to 141. This segment contained three repeats of the sequence TGCCT (Fig. 3A). Digestion of subfragment 3 with Hae III (cleaving between bp 125 and 126 and removing two of the repeats) greatly reduced this binding (Fig. 4B, subfragment 3A) suggesting that a critical sequence lay at or near this restriction site and that a single TGCCT repeat was not sufficient for binding. Additional subfragments were tested (#5, Figs. 4A and 5B; #6, Figs. 4A and 4B, panel 4), and established that a 33-bp subfragment (bp 106 to 138) provided binding capability.

We then studied the requirements for binding at the single nucleotide level by means of a methylation interference assay (*14*). Subfragment 5, demonstrating efficient binding (Figs. 4A and 5B), was methylated in vitro and immunoprecipitated after binding to p53. The bound DNA was then cleaved with piperidine at methylated residues and separated by electrophoresis on a sequencing gel. Assay of one strand (Fig. 5A, right) demonstrated that methylation at the G at bp 120 significantly interfered with binding. On the opposite strand (Fig. 5A, left), the most effective interference was produced by methylation at G residues at bp 117, 121, and 122. Partial interference was also produced by methylation at nearby G sites (bp 110 to 112, 114, and 115). Thus, the methylation interference assay pinpointed one of the repeats (centered at bp 121) and adjacent residues as critical for binding.

To obtain independent evidence of the specificity for the G residues identified by methylation interference, in vitro mutagenesis was used. A DNA fragment was generated that was identical to subfragment 5 except for the substitution of G at bp 120, 121, and 122 with T residues. This mutant subfragment (5mut2) bound poorly to p53 (Fig. 5B). A fragment identical to subfragment 5 except for a single base pair (T substituted for G at bp 120) was then tested. This fragment (5mut1) also did not bind appreciably (Fig. 5B).

Finally, we turned our attention to the second fragment that bound to p53 in the initial assay (Fig. 1A). The binding fragment was subcloned (Fig. 3B). This fragment had two repeats of the TGCCT motif (centered at bp 135 and 152). A PCR strategy similar to that used for fragment A was used to demonstrate that a 95-bp subfragment (bp 104 to 198), which contained both of these repeats, was sufficient for binding (Fig. 4B, panel 5).

The two clones found to contain p53 binding sequences have both been associated with replication origins. They were chosen for examination because previous studies had demonstrated that wild-type p53 expression can inhibit the entry into or continuation of DNA synthesis in transfected cells (*5*). Fragment A contains sequences near a putative replication origin of the ribosomal gene cluster (*15*), while fragment B contains sequences that may allow adjacent sequences to replicate as extrachromosomal circles in HeLa cells (*12*). The TGCCT repeat present in the DNA binding region of both of these clones has been observed in other potential replication origins (*16*). At present, we do not know whether the p53 binding ability of frag-

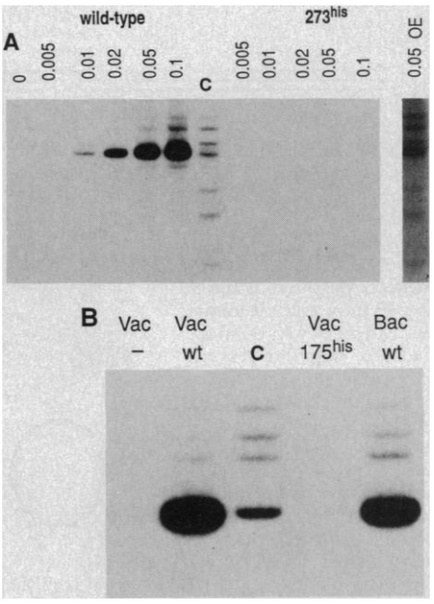

Fig. 2. Relative abilities of wild-type and mutant p53 to precipitate fragment A. C, control lane, containing 2% of the labeled DNA used in the binding reaction, other lanes contained bound DNA recovered from the immunoprecipitate. (**A**) Increasing quantities (amount indicated in micrograms) of wild-type and mutant 273^{his} p53, affinity-purified from a baculovirus expression system (*11*), were used to precipitate labeled C_{BE} fragments. An overexposure (OE) of a representative lane containing mutant-bound fragment shows that, of the small amount of DNA that is bound by 273^{his}, the 259-bp fragment (fragment A) was not significantly favored. (**B**) Lysates from a vaccinia virus system (Vac) producing the wild-type (wt), mutant (175^{his}), or no p53 protein (–), were used to immunoprecipitate labeled C_{BE} fragments. Equivalent quantities of p53 were present in the wild-type and mutant p53 lysates, as assessed by immunoblot. In the Bac lane, affinity-purified wild-type p53 produced in baculovirus-infected insect cells (*11*) was used in place of the vaccinia-infected lysates.

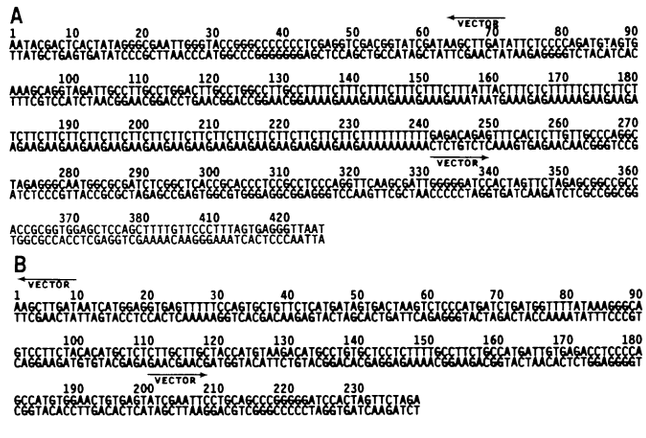

Fig. 3. (**A**) Fragment A and adjacent vector sequences in the 10d subclone of 772 C_{BE}. This differs from the published 772 C_{BE} sequence (GenBank M25718) in the number of CTT repeats (bp 173 to 229) and in the presence of A instead of C at bp 116. (**B**) Fragment B and adjacent vector sequences in a subclone of Lambda 5R. A related sequence (GenBank X05913) varied somewhat from the Lambda 5R subclone studied (*12*).

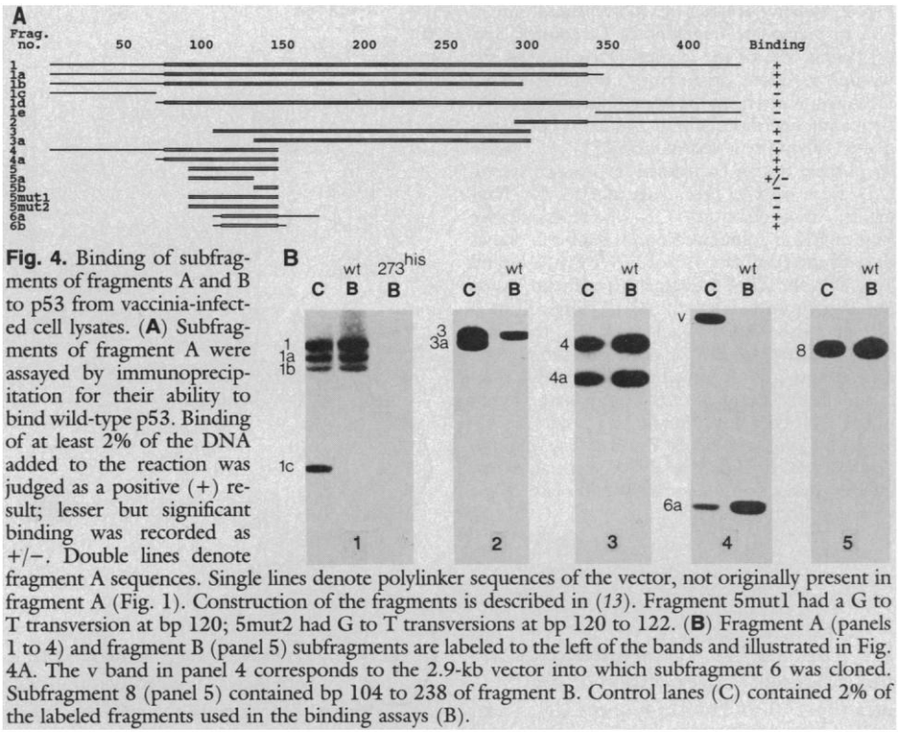

Fig. 4. Binding of subfragments of fragments A and B to p53 from vaccinia-infected cell lysates. (**A**) Subfragments of fragment A were assayed by immunoprecipitation for their ability to bind wild-type p53. Binding of at least 2% of the DNA added to the reaction was judged as a positive (+) result; lesser but significant binding was recorded as +/−. Double lines denote fragment A sequences. Single lines denote polylinker sequences of the vector, not originally present in fragment A (Fig. 1). Construction of the fragments is described in (*13*). Fragment 5mut1 had a G to T transversion at bp 120; 5mut2 had G to T transversions at bp 120 to 122. (**B**) Fragment A (panels 1 to 4) and fragment B (panel 5) subfragments are labeled to the left of the bands and illustrated in Fig. 4A. The v band in panel 4 corresponds to the 2.9-kb vector into which subfragment 6 was cloned. Subfragment 8 (panel 5) contained bp 104 to 238 of fragment B. Control lanes (C) contained 2% of the labeled fragments used in the binding assays (B).

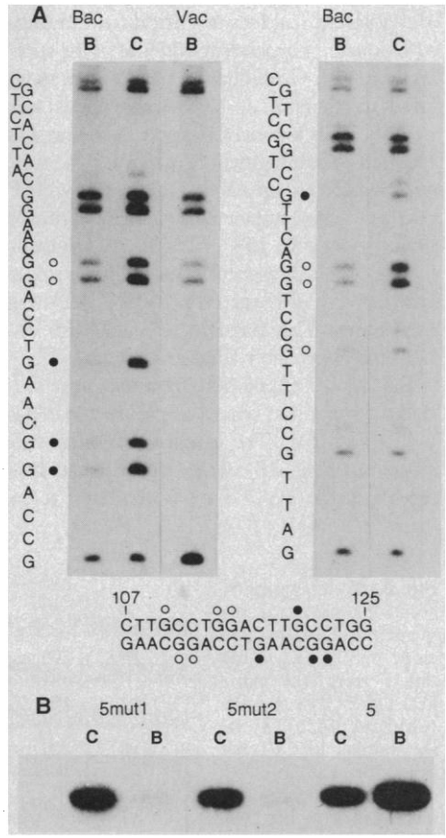

Fig. 5. Effects of methylation and point mutations in fragment A on DNA-binding. (**A**) Methylation interference assay. PCR was used to generate a 55-bp subfragment of fragment A (fragment #5, Fig. 4A, bp 87 to 141), which was end-labeled on either of the two strands. The subfragment was partially methylated at guanine residues by dimethyl-sulfate. Fragments retaining p53-binding ability were selected by immunoprecipitation with purified baculovirus-produced wild-type p53 (Bac) or vaccinia virus-infected cell lysates of BSC40 cells expressing wild-type p53 (Vac). B; bound DNA recovered from the immunoprecipitate. C; equivalent amount of control DNA fragments, not subject to binding reaction. Bound and control DNA samples were cleaved at methylated Gs and equal amounts separated by electrophoresis on a 6% denaturing gel. Dots represent methylation-sensitive sites (open for partial, solid for strong interference); some variation in band intensities occurred between assays, and only the reproducible changes are marked. (**B**) Binding of mutant subfragments of fragment A (5mut1 and 5mut2) to purified baculovirus-produced p53 is compared to that of the normal subfragment 5 sequence. The 5mut1 contains a T instead of G at bp 120, and 5mut2 contains Ts in place of Gs at bp 120, 121, and 122.

ments A and B is related to the possible participation of these sequences in the initiation of DNA synthesis, or simply coincidental. However, the data suggest some interesting avenues for further research.

In summary, p53 binds specifically to DNA sequences, and G residues of a 5-bp repeat within these sequences appear critical for this interaction. The proteins encoded by p53 mutants found in human tumor cells lose this specific DNA-binding ability. The fragments identified in this work will almost certainly not be the only ones in the human genome with the capacity to bind p53. However, the demonstration that p53 has the capacity to specifically bind these fragments will aid in analyzing the putative target sequences that mediate the biologic actions of p53 in tumor suppression.

REFERENCES AND NOTES

1. B. Vogelstein, *Nature* **348**, 681 (1990).
2. S. J. Baker *et al.*, *Science* **244**, 217 (1989).
3. J. M. Nigro *et al.*, *Nature* **342**, 705 (1989); T. Takahashi *et al.*, *Science* **246**, 491 (1989); J. W. Romano *et al.*, *Oncogene* **4**, 1483 (1989); A. G. Menon, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5435 (1990); R. Iggo *et al.*, *Lancet* **335**, 675 (1990); T. Takahashi *et al.*, *J. Clin. Invest.* **86**, 363 (1990); L. M. Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5863 (1990); Bartek *et al.*, *Oncogene* **5**, 893 (1990); M. R. Stratton *et al.*, *ibid.* p. 1297.
4. D. Eliyahu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8763 (1989); C. A. Finlay, P. W. Hinds, A. J. Levine, *Cell* **57**, 1083 (1989).
5. S. J. Baker *et al. Science* **249**, 912 (1990); W. E. Mercer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6166 (1990); L. Diller *et al.*, *Mol. Cell Biol.* **10**, 5772 (1990).
6. S. Fields and S. K. Jang, *Science* **249**, 1046 (1990); L. Raycroft, H. Wu, G. Lozano, *ibid.*, p. 1049.
7. D. P. Lane and J. Gannon, *Cell Biol. Int. Rep.* **7**, 513 (1983).
8. K. Steinmeyer and W. Deppert, *Oncogene* **3**, 501 (1988).
9. S. E. Kern *et al.*, *ibid.* **6**, 131 (1991).
10. S.-Y. Lin and A. D. Riggs, *J. Mol. Biol.* **72**, 671 (1972); *Cell* **4**, 107 (1975).
11. Immunoprecipitation assay for DNA-binding, modified from R. D. G. McKay, *J. Mol. Biol.* **145**, 471 (1981). Binding reactions included 5 μl of vaccinia-infected cell lysate or purified baculovirus p53 preparations, 95 μl of binding buffer (20 mM tris, pH 7.2, 100 mM NaCl, 10% glycerol, 1% NP40, 5 mM EDTA), 6×10^4 to 20×10^4 dpm ^{32}P-labeled DNA, 4 μl (0.4 μg) of pAb421 and 4 μl (0.4 μg) of pAb1801 (nine purified monoclonal antibodies to p53; Oncogene Science), at 4°C for 30 min. Protein A–Sepharose (1.5 mg) (Sigma) and 12.5 μg of poly(dIdC)poly(dIdC) (Pharmacia), in 25 μl of binding buffer, were added and the reactions rotated end-over-end at 4°C for 30 min. The pellet was washed twice with binding buffer and the proteins digested with SDS–proteinase K, then extracted with phenol and chloroform. The DNA was ethanol-precipitated and dissolved in electrophoresis sample buffer. The fragments were separated on a tris-borate nondenaturing polyacrylamide (7, 10, or 12%) gel. The gel was fixed and dried for autoradiography. Recombinant vaccinia viruses expressing p53 from the vaccinia 40K promoter were isolated and purified as described by J. Lyons *et al.*, *Infect. Immun.* **58**, 4089 (1990). To prepare vaccinia-infected cell lysates the indicated cells (about 2.5×10^7) were infected (at a multiplicity of infection of 2), harvested at 24 hours, and lysed in 2 ml of lysis buffer [PBS containing 5 mM EDTA, 0.5% NP-40, 0.5 mM PMSF, TPCK (10 μg/ml), aprotinin (1 μg/ml), trypsin inhibitor (10 μg/ml), leupeptin (1 μg/ml)]. Cell debris was pelleted at 16,000 *g* and the supernatant was ultraviolet-irradiated (0.5 joule). The supernatant was frozen in aliquots at −80°C and found to be stable for at least 6 months. Baculovirus-produced p53 was purified as described by P. N. Friedman, S. E. Kern, B. Vogelstein, C. Prives, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9275 (1990); while the presence of an accessory cofactor in the purified preparation cannot be excluded, SDS–polyacrylamide gel electrophoresis of the preparation revealed only a single polypeptide migrating at 53 kD upon silver staining.
12. Random human genomic clones were prepared from a partial MboI digest of human DNA; fragments (300 to 1000 bp) were purified and ligated to Eco RI linkers, then cloned into the Eco RI site of pBlue-script II (Stratagene). Other clones used included SP65hFosAva2, containing regulatory sequences from *fos*, obtained from T. Curran; 772 C_{BE} containing nucleotide −5088 to −509 relative to the transcription start site of the ribosomal gene cluster, from J. Sylvester [J. E. Sylvester, R. Petersen, R. D. Schmeckel, *Gene* **84**, 193 (1989)], a 4.4-kb

Bgl II subclone of Lambda 5R containing THE-1 sequences found in the discrete extrachromosomal circles of HeLa cells, from C. Schmid [R. Misra, A. G. Matera, C. W. Schmid, M. G. Rush, *Nucleic Acids Res.* **17**, 8327 (1989)]; the *c-myc* clone HSR-1 [K. Alitano, M. Schwab, C. C. Lin, H. E. Varmus, J. M. Bishop, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1707 (1983)] from J. Bishop. Cosmids containing the entire p53 genomic region were from Y. Nakamura; cosmids including the entire dihydrofolate reductase amplicon from hamster cells [B. Anachkova and J. L. Hamlin, *Mol. Cell. Biol.* **9**, 532 (1989)] were from J. Hamlin.

13. Binding site definition by PCR: One primer for each reaction was labeled with ^{32}P at the 5′ end with T4 polynucleotide kinase in a 5-μl reaction, and the kinase inactivated at 70°C for 5 min. PCR was done with 350 ng of each of the appropriate primers and approximately 50 ng plasmid template in a 50-μl volume; 25 cycles were used with the PCR conditions specified in S. J. Baker *et al.* [*Cancer Res.* **50**, 7717 (1990)]. The products were extracted with phenol and chloroform, ethanol-precipitated, and dissolved in 3 mM tris, 0.2 mM EDTA prior to binding. Subfragment 1 contained bp 1 to 425 of subclone 10d of fragment A (Fig. 3A); subfragments 1a, 1b, 1c, 1d, and 1e were generated by digestion of subfragment 1 with Bam HI, Mbo I, Hind III, Hind III, and Bam HI, respectively. Subfragment 2 contained bp 283 to 425. Subfragment 3a was generated by digestion of subfragment 3 (bp 106 to 294) with Hae III. Subfragment 4a was produced from subfragment 4 (bp 1 to 141) by Hind III digestion. Subfragments 5a and 5b were products of the Hae III digestion of subfragment 5 (bp 87 to 141). Mutant subfragments 5mut1 and 5mut2 were produced with primers P3m1 (5′-GAAAGAAAAGGCAAGGCCAGGAAAGT-3′) and P3mut2 (5′-GAAAGAAAAGGCAAGGCCA*TTA*-AAGT-3′) and were identical to subfragment 5 except for the positions italicized in the primers. Subfragment 6 contained bp 106 to 138, and the insert was excized by digestion with Hind III and Bam HI to generate 6a or with Hind III and Eco RI to generate 6b.
14. Methylation interference assay: PCR products labeled at one end were generated with primers labeled with T4 polynucleotide kinase (U.S. Biochemicals). The product of each PCR reaction was purified by polyacrylamide gel electrophoresis, eluted from a crushed gel slice in 500 mM ammonium acetate, extracted with phenol and chloroform, and precipitated with ethanol. DNA (2×10^6 dpm) was methylated at G residues by means of dimethylsulfate as described [T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982), p. 477], ethanol-precipitated and dissolved in 10 μl of 3 mM tris, 0.2 mM EDTA. A volume of 0.5 μl was removed as the DNA control. A total of 4.5 μl was added to a binding reaction containing baculovirus-produced p53 or vaccinia-infected cell lysates. The immunoprecipitated DNA was purified by SDS–proteinase K digestion, extracted with phenol and chloroform, and ethanol-precipitated. The control DNA and precipitates of bound DNA were cleaved with piperidine at the methylated sites. Equivalent amounts of labeled DNA were loaded and separated on a denaturing polyacrylamide (14.5%) sequencing gel, which was fixed and dried for autoradiography.
15. J. Huberman, personal communication.
16. W. R. Jelinek *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1398 (1980).
17. Fragments A and B were precipitated more efficiently when a combination of the two monoclonal antibodies (*11*) were used rather than one antibody alone. While this is attributable in part to an increased effectiveness of immunoprecipitation of p53, a potential role for antibodies in stabilization of the DNA-protein interaction cannot be excluded. [V. Zimarino, S. Wilson, C. Wu, *Science* **249**, 546 (1990)].
18. Supported by The Clayton Fund, the McAshan Fund, and NIH grants CA06973, CA43460, and CA33620. We thank K. Johnson for construction of recombinant plasmids and confirmation of p53 expression in the vaccinia system and E. Fearon for reading the manuscript.

4 April 1991; accepted 24 May 1991

Mapping of DNA Instability at the Fragile X to a Trinucleotide Repeat Sequence p(CCG)*n*

E. J. KREMER, M. PRITCHARD, M. LYNCH, S. YU, K. HOLMAN, E. BAKER, S. T. WARREN, D. SCHLESSINGER, G. R. SUTHERLAND,* R. I. RICHARDS

The sequence of a Pst I restriction fragment was determined that demonstrates instability in fragile X syndrome pedigrees. The region of instability was localized to a trinucleotide repeat p(CCG)*n*. The sequences flanking this repeat were identical in normal and affected individuals. The breakpoints in two somatic cell hybrids constructed to break at the fragile site also mapped to this repeat sequence. The repeat exhibits instability both when cloned in a nonhomologous host and after amplification by the polymerase chain reaction. These results suggest variation in the trinucleotide repeat copy number as the molecular basis for the instability and possibly the fragile site. This would account for the observed properties of this region in vivo and in vitro.

FRAGILE X SYNDROME IS THE MOST common form of familial mental retardation (*1*). In addition to characteristic clinical symptoms the disorder has an associated chromosomal abnormality at Xq27.3 in the form of a rare fragile site. We and others (*2, 3*) have isolated large fragments of human DNA in the form of yeast artificial chromosomes (YACs) that span the fragile X region. Previously we localized the fragile X region within this cloned DNA by in situ hybridization and identified the location of breakpoints of two somatic cell hybrids, constructed to break at the fragile X (*4*), to a common 5-kb Eco RI restriction fragment (pfxa1). This restriction fragment was found to be unstable in fragile X pedigrees, the instability segregating with genotype (*5*). The unstable sequences were further localized to a 1-kb Pst I restriction fragment (pfxa2) (Fig. 1A). A search through the GenBank DNA sequence collection revealed limited identity with a functionally heterogenous group of proteins encoded by genes which all contain $p(CCG)_n$ sequences. This identity ranged from 60% for a 433-nucleotide sequence of the chicken protamine gene to 85% for a 105-nucleotide sequence of the human androgen receptor. Other notable proteins within this range were several fish antifreeze proteins and mammalian keratins. The functional significance, if any, of these sequence identities is not clear.

The pfxa2 sequence contains two regions of particular note. One is a CpG-rich region that contains seven recognition sites for CpG-dependent, infrequently cutting restriction endonucleases (positions 1 to 357). Three of these sites have been demonstrated (*6, 7*) to be targets for methylation in fragile X–affected individuals. CpG-rich regions are also noteworthy in that they frequently identify the promoter regions of eukaryotic genes (*8*). The second outstanding feature of pfxa2 is the $p(CCG)_n$ repeat (position 358 to 476). This sequence also contains CpG and may itself be the subject of methylation.

To ascertain which region of pfxa2 gave rise to the observed instability, oligodeoxyribonucleotide primers suitable for polymerase chain reaction (PCR) analysis of the 1-kb Pst I restriction fragment were designed (Fig. 2B). Lymphocyte DNA from normal and affected individuals was used as template for the reaction. While the reactions primed by the oligos 201 and 204 and 209 and 214 gave constant products from both normal and fragile X–genotype individuals, the reaction primed by oligos 203 and 213 did not produce product in any reaction except for one that contained high molar concentrations of the subcloned XTY26 YAC from a fragile X genotype individual (*5*) as template (*9*). The composition of the sequence between the 203 and 213 primers in XTY26 is 92% GC, which presents a formidable problem to the PCR. A variety of PCR conditions and additives (including trimethylammonium chloride, dimethyl sulfoxide, formamide, and glycerol) have been utilized unsuccessfully in attempts to facilitate amplification across this region. The products generated from XTY26 were also noted to be heterogeneous (*9*).

In attempts to obtain sequence data for this region from normal individuals and additional fragile X genotype individuals, two approaches were undertaken. The first used two-stage PCR. Starting material was either total chromosomal DNA or, in one

E. J. Kremer, M. Pritchard, M. Lynch, S. Yu, K. Holman, E. Baker, G. R. Sutherland, R. I. Richards, Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, North Adelaide, South Australia, Australia.
S. T. Warren, Departments of Biochemistry and Pediatrics, Emory University School of Medicine, Atlanta, GA.
D. Schlessinger, Center for Genetics in Medicine, Washington University School of Medicine, St. Louis, MO.

*To whom correspondence should be addressed.