to 120 (TH primer: 5′-GCTCTGCCTGCGCCCAATGAACCGCGGGGA-3′). For both reactions, cDNA was processed for 35 cycles at 94°C for 1.5 min, at 68°C for 0.5 min, and at 72°C for 1.5 min. To further ensure specificity, the products were digested with Hind III [producing a 160–base pair (bp) fragment containing the TH sequences] and subjected to Southern (DNA) analysis with a radiolabeled nucleotide (TH nucleotides 36 to 65: 5′-GGGCTTCCGCAGGGCCGTGTCTGAGCTGGA-3′). All analyses were conducted without knowledge of the injection conditions of the animal.

25. K. Watson, J. G. Stevens, M. L. Cook, J. H. Subak-Sharpe, *J. Gen. Virol.* **49**, 149 (1980).
26. In a defective HSV-1 vector, three IE promoters (IE1, IE3, and IE 4/5) support expression of the *LacZ* gene for up to 3 months in cultured rat sensory neurons (R. Smith, A. I. Geller, C. Wilcox, in preparation). Also, in stable transformants of fibroblast cell lines, the IE3 promoter is stably expressed at a low level [J. D. Mosca, G. R. Reyes, P. M. Pitha, G. S. Hayward, *J. Virol.* **56**, 867 (1985)]. In contrast, in a latent infection of wild-type HSV-1, IE promoters are not transcribed [J. G. Stevens, E. K. Wagner, G. B. Devi-Rao, M. L. Cook, L. T. Feldman, *Science* **235**, 1056 (1987)]; IE promoters may be regulated by other sequences in the HSV-1 genome or by transcripts from wild-type HSV-1. Similarly, although the murine leukemia virus long terminal repeat promoter can be stably expressed in retrovirus vectors or in transgenic mice, this promoter is not stably expressed in the HSV-1 genome [A. T. Dobson, T. P. Margolis, F. Sedaerati, J. G. Stevens, L. T. Feldman, *Neuron* **5**, 353 (1990)]. Understanding the mechanism or mechanisms underlying these paradoxes may provide information about HSV-1 latency but is not required in order to use IE promoters in this defective HSV-1 vector.
27. These rats were not analyzed; however, HSV-1 vectors prepared with this packaging system (*17*) were injected into the midbrain, and the rats that died within 2 weeks after gene transfer contained HSV-1 particle immunoreactivity (*28*) in multiple brain areas (S. Song, Y. Wang, A. I. Geller, unpublished results).
28. R. L. Adam, D. R. Springall, M. N. Levene, T. E. Bushell, *J. Pathol.* **143**, 241 (1984). Cells containing HSV-1 particle immunoreactivity were detected 4 days after gene transfer.
29. Cortical or pallidal cells expressing TH may contribute to striatal TH concentrations by axonal transport of TH. Injection of pHSVth into other brain areas was not evaluated; however, data from rat pHSVth no. 30 suggest that expression of TH in cortical cells is not effective. Cell transplantation approaches to treatment of PD use striatal delivery (*4–8*), and the location of the site in the striatum can be critical for obtaining behavioral recovery (*6*).
30. J. Wang, K. S. Bankiewicz, R. J. Plunkett, E. H. Oldfield, *J. Neurosurg. 80*, 484 (1994).
31. C. van Horne, B. J. Hoffer, I. Stromberg, G. A. Gerhardt, *J. Pharmacol. Exp. Ther.* **263**, 1285 (1992); W. A. Cass, N. R. Zahniser, K. A. Flach, G. A. Gerhardt, *J. Neurochem.* **61**, 2269 (1993); J. S. Schneider, D. S. Rothblat, L. DiStefano, *Brain Res.* **643**, 86 (1994).
32. We thank K. Burns, K. Davis, S. Harmon, M. Kake, P. Leone, E. Lukacsi, G. Mirchandani, and D. Ullrey for technical assistance. We thank J. Haycock and R. Roth for critical readings of the manuscript. Supported by NIH grants NS28227 and NS06208 and the Parkinson's Disease Foundation (M.J.D.); by NIH grants EY09749 and MH49351 and the Tourette Association (J.R.N.); by the American Parkinson's Disease Association and NIH grant AG10827 (A.I.G. and K.L.O'M.); by NIH grant 50081 (K.L.O'M.) and by Alkermes Inc., the American Health Assistance Foundation, the Burroughs Wellcome Fund, and the National Parkinson Foundation (A.I.G.).

22 July 1994; accepted 20 September 1994

Binding of Mismatched Microsatellite DNA Sequences by the Human MSH2 Protein

Richard Fishel,* Amy Ewel, Suman Lee, Mary Kay Lescoe, Jack Griffith

Alteration of the human mismatch repair gene *hMSH2* has been linked to the microsatellite DNA instability found in hereditary nonpolyposis colon cancer and several sporadic cancers. This microsatellite DNA instability is thought to arise from defective repair of DNA replication errors that create insertion-deletion loop-type (IDL) mismatched nucleotides. Here, it is shown that purified hMSH2 protein efficiently and specifically binds DNA containing IDL mismatches of up to 14 nucleotides. These results support a direct role for hMSH2 in mutation avoidance and microsatellite stability in human cells.

Hereditary nonpolyposis colon cancer (HNPCC) affects about 1 in 200 people in industrialized nations (*1*). Four genes cosegregate with and are the likely cause of HNPCC in at least half of the identified kindreds (*2*, *3*): Current estimates suggest that *hMSH2* accounts for 50%, *hMLH1* for 30%, *hPMS1* for 5%, and *hPMS2* for 5% of the cancers in these kindreds. These genes are members of two families of postreplication mismatch repair genes, *mutS* and *mutL*, that are conserved from bacteria to humans.

Postreplication mismatch repair corrects DNA polymerase misincorporation errors, which are a major source of spontaneous mutations in dividing cells (*4*). Misincorporation errors produce nucleotide mismatches that are recognized by the MutS protein as a first step in targeting a strand-specific excision repair reaction directed at the newly synthesized DNA strand (*5*). Such targeted mismatch repair results in increased replication fidelity and a reduced spontaneous mutation rate. The human MutS homolog, hMSH2, has been proposed to perform this initial mismatch recognition function (*2*).

Instability of simple repetitive (microsatellite) DNA sequences has been observed in sporadic tumors of the colon, stomach, pancreas, bladder, hematopoietic system, lung, ovary, breast, endometrium, prostate, brain, and skin (*6*, *7*), in several tumor cell lines (*8*), and in most tumors from HNPCC kindreds (*9*). Consistent with a role in tumor initiation, microsatellite instability appears to occur early in the development of colorectal tumors (*7*, *10*). One postulated model for microsatellite instability invokes DNA polymerase slippage during replication, resulting in insertion-deletion loop-type (IDL) mismatched nucleotides that contain integral numbers of the repeated sequence (*11*). In the absence of mismatch repair, these IDL mismatched nucleotides may be fixed into the genome in a subsequent round of replication, producing the phenotype of microsatellite instability. A direct role for hMSH2 in maintaining microsatellite stability, which would require that the protein recognize IDL mismatched nucleotides, has been questioned because such mismatches are poorly recognized and repaired by the bacterial homolog MutS (*12*).

We previously demonstrated that purified hMSH2 binds efficiently and with high specificity to single base pair mismatched nucleotides (*13*). These results suggested that hMSH2 was similar to the bacterial and yeast MutS homologs (*14*) and were consistent with a role for hMSH2 in postreplication mismatch repair. To investigate the role of hMSH2 in microsatellite stability, we examined the ability of the purified protein (*15*) to bind IDL mismatched nucleotides.

A gel shift assay was used to assess the binding of hMSH2 to IDL mismatched oligonucleotides (Fig. 1A) (*16*). Binding of hMSH2 produced a shift of the radioactively labeled oligonucleotide to a species of lower mobility. The results of these experiments suggest that purified hMSH2 binds efficiently to all of the IDL mismatched nucleotides (Fig. 1B). Many of the binding reactions produced multiple species, some of which correspond in mobility to monomers or homopolymeric multimers of hMSH2 bound to the mismatched DNA (*13*, *15*). To compare the binding of hMSH2 to each IDL mismatched substrate, we prepared labeled probes for each that were of identical specific activities. Quantitation of hMSH2 specific binding activity (phosphorimager pixel counts per square millimeter in the bound mismatch complex per microgram of hMSH2) did not vary more than fivefold for all of the IDL mismatched substrates tested (Fig. 1C). The hMSH2 does not bind efficiently to single-stranded DNA; nevertheless, we undertook several steps to eliminate unannealed oligonucleotides from the IDL

R. Fishel, A. Ewel, M. K. Lescoe, Department of Microbiology and Molecular Genetics, Markey Center for Molecular Genetics, University of Vermont School of Medicine, Burlington, VT 05405, USA.
S. Lee and J. Griffith, Department of Microbiology, Lineberger Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA.

*To whom correspondence should be addressed.

substrates (*13*). These results suggest that the types of mismatches proposed to be intermediates in microsatellite instability are effective substrates for hMSH2 protein recognition and binding.

To test the specificity of hMSH2 binding to IDL mismatches, we performed competition studies with an unlabeled oligonucleotide containing a G/T mismatch that binds with high affinity to hMSH2 (*13*). The hMSH2 protein was allowed to bind to the G/C homoduplex (gel shift not shown), G/T, $(CA)_4$, and SCR_{14} radioactively labeled oligonucleotide substrates, and then unlabeled G/T competitor was added (Fig. 2). About 40% of the initially bound mismatched DNA substrates remained bound to hMSH2 in the presence of the competitor (100-fold molar excess), whereas homoduplex binding was almost completely disrupted at half that amount of competitor. The competition curve was biphasic, which suggests that there were at least two forms of hMSH2 bound to the mismatched substrates (Fig. 2B). One form appears similar to that characterizing low-specificity homoduplex binding, whereas the other appears to reflect high-affinity, specific binding. The relative order of binding specificity was G/C $\ll$ G/T $\cong$ $(CA)_4$ < SCR_{14}.

We observed that hMSH2 bound homoduplex as well as heteroduplex DNA substrates. Although competition with unlabeled homoduplex and heteroduplex DNAs suggested that the binding was at least 100-fold more specific for mismatched oligonucleotides (*13*), the apparent binding to homoduplex substrates called into question the exact location of hMSH2 on these DNAs. We therefore used electron microscopy (EM) to directly visualize the hMSH2-DNA complexes. Three IDL mismatches, each containing three cytosines and each separated by three base pairs (bp), were inserted into a 1.1-kb DNA fragment, and the resulting construct was used as a substrate for mismatch binding (*17*).

EM analysis revealed abundant DNA-protein complexes with the protein bound at a site corresponding to the position of the mismatch. In some cases (Fig. 3, A and C), the DNA was bent around the protein; however, a single three-nucleotide IDL mismatch will bend DNA by nearly 90° (*18*), so this bending may be a result of the intrinsic DNA structure and not of the bound hMSH2. The size of the smaller protein-DNA complexes (Fig. 3A) was consistent with the binding of an hMSH2 monomer, whereas the size of the larger complexes was of a size consistent with the binding of hMSH2 dimers (Fig. 3B) or higher order oligomers (Fig. 3C). When the incubations were carried out in the absence of adenosine triphosphate (ATP), 11% of the DNA molecules contained hMSH2 at the site of the mismatch (n = 100), whereas the inclusion of 5 mM ATP increased this fraction to 29% (n = 100). About 5 to 10% of the DNA fragments appeared to have hMSH2 bound to the ends of the DNA, a common property of many DNA binding

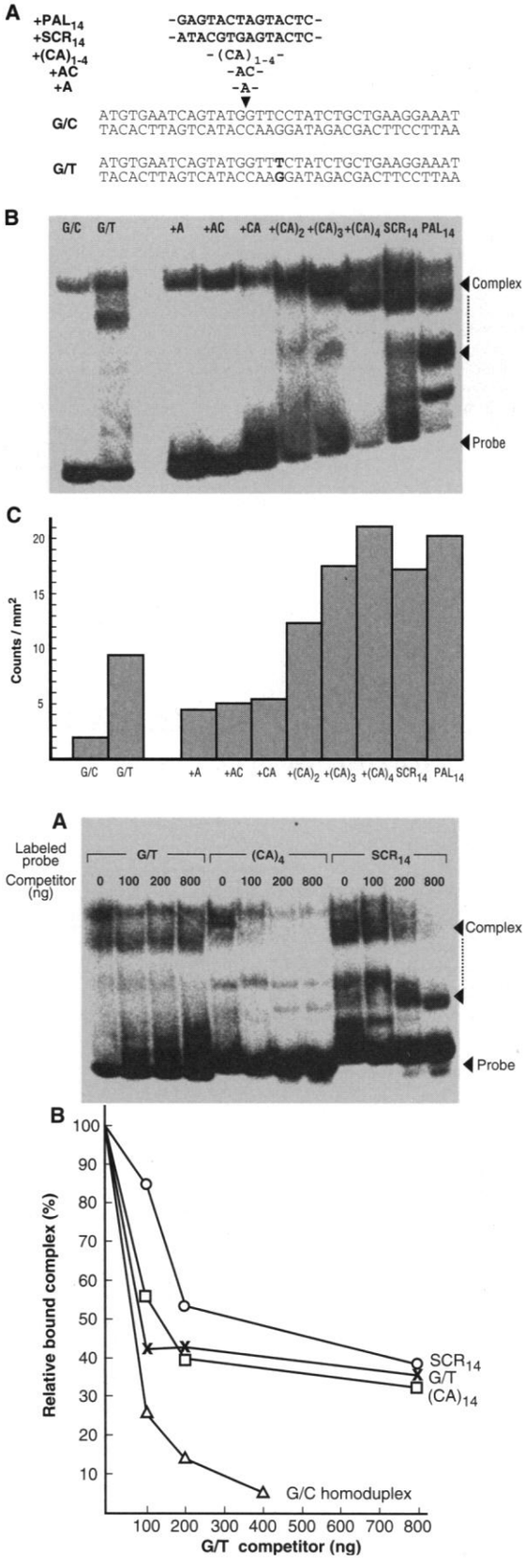

Fig. 1. Binding of hMSH2 to IDL mismatched nucleotides. (**A**) Structure of IDL mismatched DNA substrates (*16*). (**B**) Gel shift analysis of hMSH2 binding to IDL mismatched substrates. Binding was performed with the indicated IDL substrates (10 ng, containing 100,000 cpm) and purified hMSH2 (120 ng) in 20 mM potassium phosphate (pH 7.5), 50 mM KCl, 1 mM dithiothreitol, 1 mM EDTA, and 5% glycerol (*13*). (**C**) Quantitation of bound IDL substrates with Bio-Rad Molecular Analyst 2.0 software. Data are expressed in pixel counts per square millimeter (Bio-Rad, Hercules, California). Bound complex was identified as shifted material not present in untreated controls or not specifically competed by unlabeled mismatch competitor. MSH2 binding to single-stranded regions of the IDL substrates is unlikely because the protein does not bind effectively to single-stranded DNA. All of the mismatched DNA substrates were titrated with multiple concentrations of hMSH2 to develop a saturation curve before the design of the single protein concentration experiment shown.

Fig. 2. Specificity of mismatch binding by hMSH2 protein. Competition of hMSH2 bound to the G/C homoduplex and G/T, $(CA)_4$, and SCR_{14} mismatched DNA substrates was performed with increasing concentrations of unlabeled G/T mismatch. (**A**) Gel shift analysis of the relative specificity of mismatch binding. Binding was performed as in Fig. 1 for 10 min at 25°C; the G/T mismatch–containing oligonucleotide was then added and the mixture incubated for an additional 10 min at 25°C. (**B**) Amount of bound DNA probe as a function of the amount of G/T unlabeled competitor introduced into the binding reaction (*13*). The results shown are the average of four independent experiments.

proteins (*19*), and perhaps this may be the reason that we detected hMSH2 binding to homoduplex DNA substrates. All of the internally bound hMSH2 molecules were within 55 bp of the centrally located IDL mismatch site (Fig. 4).

Our results support a role for hMSH2 in the maintenance of microsatellite stability in normal human cells. Loss of hMSH2 function could result in the loss of IDL mismatch recognition, leading to the DNA sequence expansion and contraction that is characteristic of the microsatellite instability observed in many human tumors. The recent finding that repair of IDL mismatched nucleotides is deficient in extracts from a hMSH2 mutant cell line (*23*) further supports this idea. Loss of hMSH2 function could also produce a generalized increase in spontaneous mutation rates similar to that found in MutS-deficient bacteria, perhaps contributing to the multiple mutations that characterize multistage carcinogenesis (*20*).

The binding of hMSH2 to IDL mismatched nucleotides distinguishes it from its bacterial homolog MutS. Interestingly, bacterial DNA contains little of the rapidly renaturing DNA sequences that are indicative of the microsatellite sequences present in higher eukaryotes (*21*). Conceivably, MutS has evolved to recognize the relatively simple replication errors of the bacterial genome, whereas hMSH2 has evolved to recognize the more complex replication errors associated with the repetitive sequences found in the human genome. Our results are consistent with the hypothesis that the repair machinery of an organism matches the needs of its particular genome (*22*).

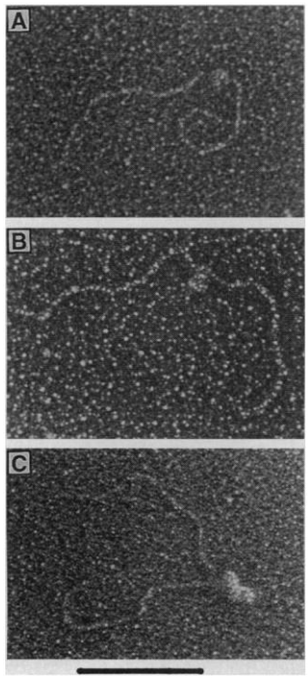

Fig. 3. Visualization of hMSH2 bound to DNA containing central IDL mismatches. The hMSH2 protein (*13*) was incubated with a 1.1-kb DNA fragment (*17*) containing three (3-cytosine) IDL mismatches at its center. DNA bound with a presumptive monomer, dimer, or homopolymeric oligomer of hMSH2 is shown in (**A**), (**B**), and (**C**), respectively. Samples were prepared for EM as described (*17*). Bar = 100 nm.

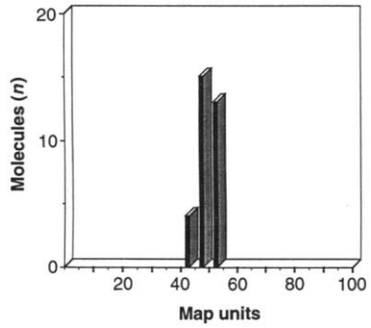

Fig. 4. Distribution of hMSH2 particles on DNA containing a central IDL mismatch. The distance of the hMSH2 protein complex from the left end of the DNA was measured in 32 molecules selected from micrographs that contained the representative molecules shown in Fig. 3. One map unit equals 1% of the DNA length. The hMSH2 bound to the ends of DNA (approximately 5 to 10% of the DNA molecules) is not shown here.

REFERENCES AND NOTES

1. T. D. Bishop and H. Thomas, *Cancer Surv.* **9**, 585 (1990); H. T. Lynch *et al.*, *Semin. Oncol.* **18**, 337 (1991); H. T. Lynch *et al.*, *Gastroenterology* **104**, 1535 (1993); P. Peltomäki *et al.*, *Science* **260**, 810 (1993).
2. R. Fishel *et al.*, *Cell* **75**, 1027 (1993).
3. F. Leach *et al.*, *ibid.*, p. 1215; C. E. Bronner *et al.*, *Nature* **368**, 258 (1994); N. Papadopoulos *et al.*, *Science* **263**, 1625 (1994); N. C. Nicolaides *et al.*, *Nature* **371**, 75 (1994).
4. P. M. Leong, H. C. Hsia, J. H. Miller, *J. Bacteriol.* **168**, 412 (1986).
5. P. Modrich, *J. Biol. Chem.* **264**, 6597 (1989); *Annu. Rev. Genet.* **25**, 229 (1991).
6. Y. Ionov, M. A. Peinado, S. Malkbosyan, D. Shibata, M. Perucho, *Nature* **363**, 558 (1993); S. N. Thibodeau, G. Bren, D. Schaid, *Science* **260**, 816 (1993); J. Young *et al.*, *Hum. Mutat.* **2**, 351 (1993); H.-J. Han, A. Yanagisawa, Y. Kato, J.-G. Park. Y. Nakamura, *Cancer Res.* **53**, 5087 (1993); M. G. Rhyu, W. S. Park, S. J. Meltzer, *Oncogene* **9**, 29 (1994); M. Gonzalex-Zulueta *et al.*, *Cancer Res.* **53**, 5620 (1993); C. Wada *et al.*, *Blood* **83**, 3449 (1994); V. Shridhar, J. Siegfried, J. Hunt, M. del Mar Alonso, D. I. Smith, *Cancer Res.* **54**, 2084 (1994); A. Merlo *et al.*, *ibid.*, p. 2098; R. Wooster *et al.*, *Nature Genet.* **6**, 152 (1994); C. J. Yee, N. Roodi, C. S. Verrier, F. F. Parl, *Cancer Res.* **54**, 1641 (1994); J. I. Risinger *et al.*, *ibid.* **53**, 5100 (1993); R. T. Burks, T. D. Kessis, K. R. Cho, L. Hedrick, *Oncogene* **9**, 1163 (1994); M. P. Schoenberg *et al.*, *Biochem. Biophys. Res. Commun.* **198**, 74 (1994); R. Honchel, K. C. Halling, D. J. Schaid, M. Pittelkow, S. N. Thibodeau, *Cancer Res.* **54**, 1159 (1994).
7. D. Shibata, M. A. Peinado, Y. Ionov, S. Malkhosyan, M. Perucho, *Nature Genet.* **6**, 273 (1994).
8. N. P. Bhattacharyya, A. Skandalis, A. Ganesh, J. Groden, M. Meuth, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6319 (1994).
9. L. A. Aaltonen *et al.*, *Science* **260**, 812 (1993); P. Peltomäki *et al.*, *Cancer Res.* **53**, 5853 (1993).
10. L. A. Aaltonen *et al.*, *Cancer Res.* **54**, 1645 (1994).
11. T. A. Kunkel, *Biochemistry* **29**, 8003 (1990); *Nature* **365**, 207 (1993).
12. S.-S. Su, R. S. Lahue, K. G. Au, P. Modrich, *J. Biol. Chem.* **263**, 6829 (1988); B. A. Learn and R. H. Grafstrom, *J. Bacteriol.* **171**, 6473 (1989); B. O. Parker and M. G. Marinus, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1730 (1992).
13. R. Fishel, A. Ewel, M. K. Lescoe, *Cancer Res.* **54**, 5539 (1994).
14. S.-S. Su and P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5057 (1986); N.-W. Chi and R. Kolodner, *J. Biol. Chem.*, in press; *ibid.*, in press.
15. R. Fishel, A. Ewel, M. K. Lescoe, unpublished results.
16. IDL mismatched substrates (E. Alani, N.-W. Chi, R. Kolodner, unpublished data) contain mismatched sites at position 15 of a model 36-mer oligonucleotide. The oligonucleotide was designed to contain an invariable bottom strand sequence; the mismatched nucleotides were always introduced into the top strand. As a reference point to previous studies, we also tested a single G/T mismatch (at position 19) that is bound with high affinity by purified hMSH2 and by *Saccharomyces cerevesiae* MSH1 and MSH2 (*14*). In the present studies, we examined hMSH2 binding to a +1 adenine (+A) nucleotide insertion, a +2 adenine-cytosine (+AC) nucleotide insertion, a series of cytosine-adenine (CA) dinucleotide insertions containing one, two, three, or four dinucleotide repeats [$(CA)_{1-4}$] within the insertion loop, a scrambled 14-mer nucleotide sequence (SCR_{14}) containing no obvious internal homologies, and a perfect 14-mer palindromic sequence (PAL_{14}) inserted into the model 36-mer oligonucleotide. The results on selective binding to multiple base insertion mispairs are qualitatively similar to those obtained with *S. cerevisiae* MSH2 (E. Alani, N.-W. Chi, R. Kolodner, unpublished data).
17. Construction of the DNA templates will be described in detail elsewhere. In brief, 50-bp oligonucleotides were synthesized and annealed to produce 46-bp duplexes containing three IDL mismatches, each containing three cytosines and each spaced by three duplex nucleotides. DNA fragments (542 bp, with one overhanging end that was complementary to the ends in the 46-mer and one blunt end) were ligated to each end of the 46-mer, and the desired 1130-bp product was purified after separation on agarose gels. The DNA samples were prepared in binding buffer (*13*), hMSH2 was added at a molar ratio of five protein monomers per DNA fragment, and the mixtures were incubated at ambient temperature for 10 min. The reaction mixtures were then treated with 0.6% glutaraldehyde for 5 min before filtration through Biogel A5m to remove unbound protein and fixative. Samples were mixed with buffer containing 2 mM spermidine, applied to carbon foils treated with high-voltage discharge in a mild vacuum, rotary shadow–casted with tungsten, and then processed for EM [J. D. Griffith and G. Christiansen, *Annu. Rev. Biophys. Bioeng.* **7**, 10 (1978)].
18. Y.-H. Wang, P. Baker, J. Griffith, *J. Biol. Chem.* **267**, 4911 (1992).
19. J. Griffith, unpublished observations.
20. E. C. Cox, *Annu. Rev. Genet.* **10**, 135 (1976); A. G. Knudson, *ibid.* **20**, 231 (1986); E. R. Fearon and B. Vogelstein, *Cell* **61**, 759 (1990); L. A. Loeb, *Cancer Res.* **51**, 3075 (1991); M. J. Renan, *Mol. Carcinogen.* **7**, 139 (1993).
21. J. G. Wetnur and N. Davidson, *J. Mol. Biol.* **31**, 349 (1968).
22. J. W. Drake and R. H. Baltz, *Annu. Rev. Biochem.* **45**, 11 (1976); J. W. Drake, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7160 (1991).
23. A. Umar, J. C. Boyer, T. A. Kunkel, *Science*, in press.

24. We thank R. Kolodner for sharing unpublished results, for mismatched oligonucleotide substrates, and for advice; and L. Kallal and G. Gilmartin for advice and comments. Supported by NIH grant CA56542 and a grant from the Lake Champlain Cancer Research Organization (R.F.) and NIH grants GM42342 and GM31819 (J.G.).

27 September 1994; accepted 26 October 1994