specific for CD25 and CD117 (Pharmingen) and were analyzed with a FACScan (Becton Dickinson, Mountain View, CA) flow cytometer and Lysis II software. The specificity of antibody staining was confirmed with isotype-matched control antibodies. CD117 expression was similar to that previously reported (*3*).

16. HSA$^-$CD25$^-$ day 14 to 15 fetal and neonatal SCID thymocytes were obtained by antibody- and complement-mediated lysis. Single-cell suspensions of thymocytes ($25 \times 10^6$ to $30 \times 10^6$ cells) were incubated on ice with 300 μl of culture supernatant of J11d.2 (anti-HSA) and 7D4 (anti-CD25) for 15 min, Low-Tox rabbit complement (Cedar Lane, Hornby, Ontario) was added, and cells were incubated at 37°C for 30 min. After complement-mediated lysis, viable cells were recovered by density centrifugation with Lympholyte-M (Cedar Lane). CD25$^-$HSA$^-$ thymocytes represented 4% of total day 15 fetal or neonatal SCID thymocytes. CD25$^+$HSA$^+$ thymocytes represented freshly isolated cells that were not treated with anti-CD25 or anti-HSA.
17. HSA$^-$CD25$^-$ day 14 to 15 fetal and neonatal SCID thymocytes ($1 \times 10^5$ to $3 \times 10^5$) (*16*) were cultured for 24 hours in round-bottom 96-well microtiter plates in 200 μl of Click's medium (Biofluids) with 10% fetal calf serum. All cytokines were obtained from R&D Systems (Minneapolis, MN). TNF-α, IL-1α, and IL-7 were used at 50 ng/ml, IL-6 and lymphotoxin (LT) were used at 100 ng/ml, and SCF was used at 20 ng/ml. Cell recovery was not affected by the addition of TNF-α or IL-1α; however, addition of IL-7 increased total cell recovery by 30 to 40% in both the control and treated cultures. CD25 expression on control (SCF only) cultures was 5 to 10%. The increase of CD25 expression compared to its expression in control cultures was determined by flow cytometry. CD25 induction was not a result of outgrowth of a CD25$^+$ subpopulation because (i) cell yields in cultures with or without TNF-α treatment were similar, (ii) no cellular turnover was observed, and (iii) CD25 expression could be detected as early as 8 hours after treatment (Fig. 2A) (*8*). In Fig. 2B, SCID thymocytes gave better cell yield in culture, resulting in greater recovery of CD25$^+$ cells. The data shown correspond to the average of triplicate cultures and are representative of at least five independent trials.
18. Day 15 fetal thymuses (BALB/c) were placed in organ culture for 6 days in the presence of 1.0 mM deoxyguanosine (*14*) and then washed for 1 day before reconstitution. Day 13 fetal liver cells (C57BL/6) ($3 \times 10^4$) were used to reconstitute the thymic shells in a hanging drop setup for 24 hours (*14*). Reconstituted dGuoFTOC were then cultured, under standard FTOC conditions (*14*), for 5 or 9 days before flow cytometry analysis. Antibodies to mouse TNF-α, mouse IL-1α, and human interferon γ (control) were obtained from Genzyme (Cambridge, MA). Antibodies were used at a final dilution in stock of 1:80. Antibodies were added every other day, starting with day 1, to the FTOC. No changes in cell recovery were observed by day 5 of culture; however, by day 9, 30% fewer cells were usually recovered per thymic lobe from the antibody-treated FTOC.
19. Day 14 to 15 fetal thymocytes, which had been gated for lack of CD3 expression to exclude mature γδ T cells, were sorted by flow cytometry into CD117$^+$ thymocytes with or without coexpression of CD25. Day 13 fetal liver cells ($1 \times 10^5$) or sorted day 14 to 15 fetal thymocytes ($2 \times 10^5$ to $3 \times 10^5$) (>98% purity; FACStar-plus, Becton Dickinson) were intravenously injected into Ly5 congenic host mice irradiated with 7.5 Gy (750 rads). After 22 days, peripheral lymphocytes (pooled spleen and lymph nodes) were collected and analyzed for donor-derived lymphocytes.
20. We thank C. Eigst and D. Stephany for fluorescent cell sorting and R. N. Germain, B. J. Fowlkes, P. Schwartzberg, and M. H. Julius for reading the manuscript. Supported by the Jane Coffin Childs Memorial Fund for Medical Research (J.C.Z.-P.) and Cancer Research Institute (M.J.L.).

6 February 1995; accepted 24 April 1995

# Isolation of an hMSH2-p160 Heterodimer That Restores DNA Mismatch Repair to Tumor Cells

James T. Drummond,* Guo-Min Li,* Matthew J. Longley, Paul Modrich†

A mismatch-binding heterodimer of hMSH2 and a 160-kilodalton polypeptide has been isolated from HeLa cells by virtue of its ability to restore mismatch repair to nuclear extracts of hMSH2-deficient LoVo colorectal tumor cells. This heterodimer, designated hMutSα, also restores mismatch repair to extracts of alkylation-tolerant MT1 lymphoblastoid cells and HCT-15 colorectal tumor cells, which are selectively defective in the repair of base-base and single-nucleotide insertion-deletion mismatches. Because HCT-15 cells appear to be free of hMSH2 mutations, this selective repair defect is likely a result of a deficiency of the hMutSα 160-kilodalton subunit, and mutations in the corresponding gene may confer hypermutability and cancer predisposition.

Certain sporadic cancers and virtually all tumors that occur in patients with hereditary nonpolyposis colorectal cancer (HNPCC) are characterized by a high incidence of mutation in microsatellite repeat sequences [reviewed in (*1*)], and cell lines derived from such tumors are genetically unstable (*2–4*). Cancer predisposition in most HNPCC kindreds is attributable to defects in any one of four genes, all of which encode homologs of the microbial mismatch repair proteins MutS and MutL. The *hMSH2* gene specifies a MutS homolog (*5*), whereas *hMLH1*, *hPMS1*, and *hPMS2* encode homologs of MutL (*6*, *7*).

J. T. Drummond and P. Modrich, Howard Hughes Medical Institute and Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA.
G.-M. Li and M. J. Longley, Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed.

As judged by biochemical assay, tumor cells that display microsatellite instability are typically defective in mismatch correction (*2*, *8*), thus providing a direct link between the HNPCC genes and genetic stability afforded by this DNA repair system. For example, the H6 and LoVo colorectal tumor cell lines, which are defective in both alleles of *hMLH1* and *hMSH2*, respectively (*6*, *8*), are both defective in mismatch repair. An activity that restores mismatch repair to extracts of *hMLH1*-deficient H6 cells has been isolated from HeLa cells and shown to be a heterodimer of hMLH1 and hPMS2 (*9*).

Although hMSH2 has been shown to bind mismatched base pairs (*10*, *11*), the form of the protein active in mismatch repair has not been defined. To clarify this issue, we isolated from HeLa cells a component that restores mismatch repair to nuclear extracts of LoVo cells (*12*). This complementing activity, designated hMutSα, is associated with two polypeptides of 105 kD and 160 kD (Fig. 1A). The 105-kD protein was shown by immunoblot to be hMSH2 (Fig. 1B). The molar equivalence of the two polypeptides (*13*) suggested that they might interact. To test this possibility, we monitored the protein composition and activity of purified hMutSα (*12*) during gel permeation chromatography and band sedimentation through sucrose density gradients (*14*). Association of the two polypeptides and their relative stoichiometry ($0.92 \pm 0.05$ mol per 1 mol, $n = 5$) were preserved during these procedures.

We estimated the native relative molecular mass of hMutSα on the basis of its Stokes' radius ($R_S$; 67 Å), its sedimentation coefficient (9.0S), and the partial specific volume of the subunits (*14*, *15*). Assuming a typical protein partial specific volume of 0.725 cm$^3$/g for p160 (*15*) and a calculated partial specific volume (*16*) of 0.741 cm$^3$/g for hMSH2 based on the predicted amino acid sequence (*5*), these hydrodynamic parameters indicate a native molecular mass of 255 kD. Because this value is close to that expected for a 1:1 complex of the two polypeptides, we conclude that hMSH2 functions in mismatch repair as one subunit of a heterodimer, the other component of which is p160.

To identify p160, we sequenced five internal tryptic peptides derived from the protein (*17*). None of the peptides showed significant homology to proteins in the National Center for Biotechnology Information (NCBI) nonredundant sequence database. This finding, and the highly specific response of hMSH2 antibodies to the 105-kD polypeptide of hMutSα, demonstrate that p160 and hMSH2 subunits are distinct at the sequence level and are thus

encoded by distinct genes.

Bacterial MutS (*18*), a human mismatch binding factor designated GTBP (*10*), recombinant forms of yeast MSH1 and MSH2 (*19*, *20*), and recombinant human MSH2 (*11*) bind heteroduplex DNAs with variable degrees of specificity relative to homoduplex DNA. Mobility-shift assays revealed that hMutSα binds efficiently to a G-T mismatch and to heteroduplexes containing a $(dT)_1$ or $(dT)_3$ insertion in one strand, but binds to an A-T homoduplex with less than one-tenth that efficiency (Fig. 2). The binding of hMutSα to heteroduplex DNA was reduced in the presence of adenosine 5′-triphosphate (ATP), but only slightly so in the presence of AMPPNP, the nonhydrolyzable β,γ-imido analog of ATP. This ATP effect is similar to that observed with *Escherichia coli* MutS (*21*) and GTBP (*10*) and may distinguish hMutSα from recombinant hMSH2, which has been reported to bind more efficiently to mismatches in the presence of ATP (*11*).

Extracts of the hypermutable, *hMSH2*-defective LoVo cell line, which we used to assay hMutSα during isolation (*12*), have been shown to be defective in the repair of a G-G heteroduplex and several insertion-deletion mismatches (*8*, *22*). We found LoVo cells to be deficient in the repair of several base-base mismatches (*23*), as well as one-, two-, three-, or four-nucleotide insertion-deletion heteroduplexes (Table 1), which confirms these earlier studies. For example, repair activity of LoVo extracts was less than 3% of that observed with extracts prepared from the genetically stable S0 colorectal tumor cell line (*2*), and this defect was evident whether the strand break directing the reaction was located 3′ or 5′ to the mispair (*24*). The substrate specificity of the LoVo cell repair defect is thus similar to that of the hMLH1-deficient H6 cell line (Table 1) (*2*, *9*).

Our analysis of two additional hypermutable cell lines, MT1 and HCT-15, revealed that they are selectively defective in the repair of base-base and single-nucleotide insertion-deletion mismatches. The MT1 cell line, a derivative of TK6 lymphoblastoid cells that is resistant to the cytotoxic effects of *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine, is deficient in the repair of the eight base-base mismatches, and the hypermutable and alkylation-tolerant phenotypes have been attributed to this defect (*25*). MT1 cells are also defective in the repair of single-nucleotide insertion-deletion mismatches, but repair two-, three-, or four-nucleotide mismatches at 20 to 50% the rate observed with the TK6 parental line (Table 1). A similar partial defect in the repair of this set of insertion-deletion mismatches was observed with HCT-15 colorectal tumor cells (Table 1), which are as defective in the repair of the eight base-base mismatches as are H6 and LoVo cells (*23*).

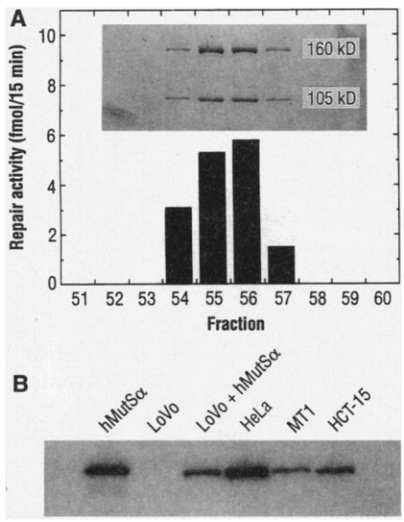


**Fig. 1.** Isolation and protein composition of an hMutSα heterodimer. (**A**) Repair activity of hMutSα. We determined the activity profile of the MonoQ eluate (*12*) by assaying 2.5 μl of each fraction for the ability to restore mismatch repair to nuclear extracts of MT1 cells (shown) or *hMSH2*-defective LoVo cells (*30*, *31*). Samples (20 μl) were also analyzed by electrophoresis through a 6% SDS-polyacrylamide gel (inset), and molecular mass was estimated with protein standards. As judged by several gel staining methods, the purity of the MonoQ eluate exceeded 95%. (**B**) Identification of hMSH2 by immunoblot analysis. Lanes contained 75 ng of purified hMutSα (fraction III) or 50 μg of a nuclear extract from LoVo, HeLa, MT1, or HCT-15 cell lines. After electrophoresis through a 6% SDS-polyacrylamide gel, proteins were transferred to a nylon membrane (Biotrans, ICN) and probed with monoclonal antibody EH12, which recognizes the COOH-terminus of hMSH2. Immune complexes were visualized with an enhanced chemiluminescence reagent (Amersham).

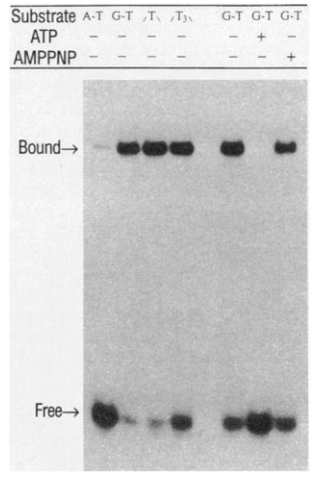


**Fig. 2.** Binding of hMutSα to mismatched base pairs. Band shift assays (*31*) were performed in 20-μl reactions containing 10 mM Hepes-KOH (pH 7.5), 1 mM EDTA, 1 mM DTT, 5 mM $MgCl_2$, bovine serum albumin (50 μg/ml), 4% glycerol, 0.13 pmol of the indicated 5′-$^{32}$P-oligonucleotide duplex, 2.5 pmol of unlabeled oligonucleotide homoduplex competitor, and 106 ng (0.4 pmol) of hMutSα. One millimolar ATP or AMPPNP (Calbiochem) was used. After a 20-min incubation on ice, 5 μl of 50% sucrose was added, and the samples were subjected to electrophoresis at room temperature at 9 V/cm through 6% polyacrylamide in 6.7 mM tris-acetate (pH 7.5) and 1 mM EDTA with buffer recirculation.

Because HCT-15 and MT1 cells repair two-, three-, and four-nucleotide insertion-deletion mismatches, whereas H6 and LoVo cells do not, we conclude that hMLH1 and hMSH2 are required for the repair of these insertion-deletion mismatches. Two observations support the idea that correction of small insertion-deletion mispairs in HCT-15 and MT1 extracts is mediated by the same

**Table 1.** Differential repair defects in extracts of HCT-15 and MT1 cells (*30*, *31*). Heteroduplexes contained a strand-specific incision 125 nucleotides 5′ to the mismatch (5′-substrate) or 181 nucleotides 3′ to the mispair (3′-substrate). When substantial repair occurred, repair bias to the incised strand was typically better than 10-fold. As observed previously (*2*), the 3′-⟋CTG⟍ heteroduplex was a poor substrate for correction by human cell nuclear extracts. ND, not determined.

| Cell extract | Repair of heteroduplex (fmol/50 μg) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | G-T | ⟋A⟍ | ⟋T⟍ | ⟋CA⟍ | ⟋TG⟍ | ⟋CTG⟍ | ⟋CAG⟍ | ⟋CACA⟍ |
| | *Repair of 5′-substrate* | | | | | | | |
| S0 | 13 | 11 | 10 | 10 | 7.8 | 11 | 13 | 11 |
| TK6 | 8.9 | 11 | 7.0 | 7.0 | 9.2 | 7.0 | 8.5 | 7.3 |
| H6 | <0.3 | <0.3 | <0.3 | <0.3 | 0.3 | 0.3 | <0.3 | 0.8 |
| LoVo | <0.3 | <0.3 | <0.3 | <0.3 | <0.3 | <0.3 | <0.3 | <0.3 |
| HCT-15 | <0.3 | <0.3 | <0.3 | 4.9 | 3.6 | 4.9 | 3.2 | 2.4 |
| MT1 | <0.3 | <0.3 | <0.3 | 4.6 | 5.6 | 5.5 | 2.9 | 2.7 |
| | *Repair of 3′-substrate* | | | | | | | |
| S0 | 10 | ND | ND | 12 | 8.4 | 2.4 | 10 | 8.9 |
| TK6 | 9.0 | ND | ND | 8.7 | 4.7 | 0.9 | 7.2 | 6.9 |
| H6 | <0.3 | ND | ND | <0.3 | <0.3 | <0.3 | <0.3 | <0.3 |
| LoVo | <0.3 | ND | ND | <0.3 | <0.3 | <0.3 | <0.3 | <0.3 |
| HCT-15 | <0.3 | ND | ND | 6.2 | 1.2 | 2.2 | 2.4 | 3.5 |
| MT1 | <0.3 | ND | ND | 5.6 | 1.7 | 2.3 | 2.5 | 2.7 |

**Table 2.** In vitro complementation of HCT-15 and MT1 nuclear extracts. Repair assays contained 50 μg of each extract and 24 fmol of a G-T or ⟨TG⟩ heteroduplex that contained a single-strand break 125 nucleotides 5′, or 181 nucleotides 3′, to the mispair. Assays with purified proteins contained 30 ng of hMutLα (9) or 100 ng of hMutSα (fraction III). Blank spaces are entries that would be redundant with data in Table 1.

| Cell extract | Repair (fmol) with addition of | | | | | |
|---|---|---|---|---|---|---|
| | LoVo extract | MT1 extract | HCT-15 extract | hMutLα | hMutSα | None |
| | | | *Repair of 5′ G-T* | | | |
| H6 | 3.9 | 5.1 | 4.8 | 7.7 | <0.3 | <0.3 |
| LoVo | | <0.3 | <0.3 | <0.3 | 2.4 | <0.3 |
| MT1 | | | 0.3 | <0.3 | 7.2 | <0.3 |
| HCT-15 | | | | <0.3 | 4.0 | <0.3 |
| | | | *Repair of 3′ ⟨TG⟩* | | | |
| H6 | 1.1 | 2.6 | 4.2 | 5.3 | <0.3 | <0.3 |
| LoVo | | <0.3 | <0.3 | <0.3 | 2.5 | <0.3 |
| MT1 | | | 2.2 | <0.3 | 7.0 | 1.7 |
| HCT-15 | | | | <0.3 | 6.7 | 1.2 |

pathway as that operative in wild-type cells. First, like the repair reaction in HeLa cell extracts (26), correction of two-, three-, and four-nucleotide insertion-deletion mispairs by HCT-15 and MT1 extracts was inhibited by aphidicolin and depended on the presence of a strand break (23). Second, the excision intermediates (24) produced in HCT-15 extracts were similar to those generated in extracts of HeLa or S0 cells (2, 23).

The in vivo mutational spectra of H6 and HCT-15 cells correlate well with the differential nature of their mismatch repair defects. The rate of $(CA)_n$ microsatellite mutation is elevated several hundredfold in H6 cells (2, 4), and a comparable increase in mutability occurs at the hypoxanthine-guanine phosphoribosyltransferase (HPRT) locus (4). Although the *HPRT* mutability of HCT-15 cells is comparable to that of H6 cells, HCT-15 cells display only a modest increase in the rate of $(CA)_n$ microsatellite mutation (4). *HPRT* mutability in the MT1 cell line is also elevated about 60-fold, and the exon mutations that occur are base substitution and single-nucleotide frameshifts (25). The rate of microsatellite mutation in MT1 cells has not been measured in a rigorous fashion. However, the incidence of $(CA)_n$ microsatellite mutations, although somewhat higher than that found in parental TK6 cells (8), is much lower than that found in H6 cells (4).

The four cell lines described above fall into two classes on the basis of in vitro complementation analysis (Table 2). HCT-15, MT1, and LoVo extracts do not complement one another in vitro, but mismatch repair is restored when any of these extracts is mixed with H6 extract. A heterodimer of hMLH1 and hPMS2 (hMutLα) restores mismatch repair to H6 extracts (9), consistent with the finding that H6 cells are defective in both alleles of *hMLH1* (6), whereas the hMutSα heterodimer of hMSH2 and p160 restores mismatch repair to HCT-15, MT1, and LoVo extracts (Table 2). In addition to restoring repair of base-base and single-nucleotide insertion-deletion mismatches to MT1 and HCT-15 extracts, hMutSα increases the proficiency with which these extracts process heteroduplexes with two-, three-, or four-nucleotide insertion-deletion mismatches (Table 2) (23). Because attempts to identify *hMSH2* mutations in HCT-15 cells have not been successful (27), we infer that HCT-15 and MT1 cells are biochemically deficient in the 160-kD subunit of hMutSα. The failure to observe in vitro complementation between these two lines and LoVo cells can be reconciled with this conclusion if p160 is unstable in extracts of LoVo cells because of the absence of the hMSH2 subunit (Fig. 1B). In fact, native hMutSα is unstable because of the degradation of p160 (12).

Palombo *et al.* (28) and Papadopoulos *et al.* (29) demonstrate that the HCT-15 and MT1 cell lines harbor defects in the gene encoding the 160-kD MutS homolog GTBP. Given the similar sizes of the p160 subunit of hMutSα and GTBP, we have searched the GTBP sequence for the presence of the peptides identified in p160 (17). Because nearly all of these peptides are specified by the partial GTBP complementary DNA (28), p160 and GTBP appear to be identical. Thus, hMutSα is a heterodimer of MutS homologs. The selective nature of repair defects of the HCT-15 and MT1 cell lines suggests that mismatch recognition may be at least partially differentiated between the two subunits of this complex.

## REFERENCES AND NOTES


1. L. A. Loeb, *Cancer Res.* **54**, 5059 (1994); P. Modrich, *Science* **266**, 1959 (1994).
2. R. Parsons *et al.*, *Cell* **75**, 1227 (1993).
3. D. Shibata, M. A. Peinado, Y. Ionov, S. Malkhosyan, M. Perucho, *Nature Genet.* **6**, 273 (1994).
4. N. P. Bhattacharyya, A. Skandalis, A. Ganesh, J. Groden, M. Meuth, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6319 (1994); J. R. Eshleman *et al.*, *Oncogene* **10**, 33 (1994).
5. R. Fishel *et al.*, *Cell* **75**, 1027 (1993); F. S. Leach *et al.*, *ibid.*, p. 1215.
6. N. Papadopoulos *et al.*, *Science* **263**, 1625 (1994).
7. C. E. Bronner *et al.*, *Nature* **368**, 258 (1994); N. C. Nicolaides *et al.*, *ibid.* **371**, 75 (1994).
8. A. Umar *et al.*, *J. Biol. Chem.* **269**, 14367 (1994).
9. G.-M. Li and P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1950 (1995).
10. M. J. Hughes and J. Jiricny, *J. Biol. Chem.* **267**, 23876 (1992); F. Palombo, M. Hughes, J. Jiricny, O. Truong, J. Hsuan, *Nature* **367**, 417 (1994).
11. R. Fishel, A. Ewel, M. K. Lescoe, *Cancer Res.* **54**, 5539 (1994); R. Fishel, A. Ewel, S. Lee, M. K. Lescoe, J. Griffith, *Science* **266**, 1403 (1994).
12. Isolation of hMutSα was performed at 4°C. Buffers contained 0.025 M Hepes-KOH (pH 7.5), 0.1 mM EDTA, and 0.1 M KCl (buffer A), 0.2 M KCl (buffer B), or 0.65 M KCl (buffer C). Nuclear extract (4 mg/ml of protein) prepared from HeLa $S_3$ cells (*26*) was treated with solid $(NH_4)_2SO_4$ (0.215 g/ml), and the precipitate was removed by centrifugation (20 min at 15,000*g*). The supernatant was then treated with 0.192 g/ml of $(NH_4)_2SO_4$ in a similar manner, the precipitate collected by centrifugation, dissolved in 0.025 M Hepes-KOH (pH 7.6), 0.05 M KCl, 0.1 mM EDTA, 2 mM dithiothreitol (DTT), leupeptin (1 μg/ml), and 0.1% phenylmethylsulfonyl fluoride (PMSF; concentration was relative to a saturated solution in isopropanol at 23°C), and dialyzed against this buffer until the conductivity was approximately that of buffer containing 0.11 M KCl. Samples of 1 to 1.5 ml (25 to 40 mg/ml of protein) were frozen in liquid $N_2$ and stored at −80°C (fraction I). Fraction I (30 to 50 mg of protein) was thawed, diluted to 12 ml with buffer B, centrifuged at 10,000*g* for 10 min, and applied at a rate of 1 to 1.5 ml/min to a single-stranded DNA cellulose column (1.8 $cm^2$ by 1.6 cm; Sigma; 3.6 mg of DNA per gram of cellulose) equilibrated with buffer B. The material that passed through the column was reloaded to ensure complete binding, and the column was eluted stepwise at a rate of 1 to 1.5 ml/min with 50 ml of buffer B and 50 ml of buffer B containing 2.5 mM $MgCl_2$, followed by 50 ml of the latter buffer containing 1 mM ATP. Human MutSα, which eluted in the presence of ATP, was immediately loaded onto a 0.1-ml column of Q-Sepharose (Pharmacia) equilibrated in buffer B. After washing with 2 ml of buffer B, hMutSα (~80% pure) was step-eluted with 0.4 ml of buffer C (fraction II; 0.4 ml). Fraction II was immediately diluted with 1.6 ml of 0.025 M Hepes-KOH (pH 7.5) to a conductivity equivalent to that of buffer A and loaded at a rate of 0.5 ml/min onto a 1.0-ml Pharmacia HR 5/5 Mono Q column equilibrated with buffer A at a flow rate of 0.5 ml/min. After washing with 10 ml of buffer A, the column was eluted with a 20-ml linear gradient running from 100% buffer A to 100% buffer C. Fractions containing hMutSα (fraction III) were either used directly or concentrated as described above and dialyzed against buffer A containing 2 mM DTT and leupeptin (1 μg/ml). This procedure yielded 30 to 60 μg of purified hMutSα and could be completed in 8 hours. Because the p160 subunit was unstable because of degradation, preparations were typically used within 24 hours of isolation. Fractions were stabilized by addition of bovine serum albumin (1 mg/ml) and dialysis against 0.025 M Hepes-KOH (pH 7.5), 0.1 M KCl, 1 mM EDTA, 0.1% PMSF, pepstatin A (1 μg/ml), leupeptin (1 μg/ml), and 10% sucrose. Dialyzed material was frozen in small samples in liquid $N_2$ and stored at −80°C.
13. The protein content of individual electrophoretic species was estimated with a cooled charge-coupled device imager (Photometrics) after staining with Coomassie brilliant blue. This method indicated a stoichiometry of 1.05 (± 0.16, $n$ = 3) mol of the 160-kD protein (p160) per mole of hMSH2 in fraction III (*12*).
14. Gel filtration of fraction III (*12*) was performed at 4°C with a 26 cm by 0.38 $cm^2$ column of Sephacryl S300 (Pharmacia) equilibrated with buffer A (*12*) and calibrated with bovine thyroglobulin ($R_S$ = 85.8 Å), apoferritin ($R_S$ = 67.3 Å), catalase ($R_S$ = 52.2 Å), yeast alcohol dehydrogenase ($R_S$ = 46 Å), and ovalbumin

($R_S$ = 30.5 Å). Purified hMutSα eluted at a value for $R_S$ of 67 Å. Band sedimentation through 10 to 30% sucrose gradients in buffer A was performed at 4°C at 68,000*g* for 24 hours in a Beckman SW 50.1 rotor. Sedimentation markers run in parallel tubes were catalase [sedimentation coefficient ($s_{20,w}$) = 11.3*S*], γ-globulin ($s_{20,w}$ = 7.1*S*), and ovalbumin ($s_{20,w}$ = 3.55*S*). Sedimentation relative to these markers yielded a value of $s_{20,w}$ for hMutSα of 9.0*S*.


15. L. M. Siegel and K. J. Monty, *Biochim. Biophys. Acta* **112**, 346 (1966).
16. E. J. Cohn and J. T. Edsall, in *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, E. J. Cohn and J. T. Edsall, Eds. (Hafner, New York, 1943), pp. 370–381.
17. Peptide sequence analysis was performed as described (*9*), after resolution of subunits (about 100 pmol) by electrophoresis in the presence of SDS. Attempts to obtain the $NH_2$-terminal sequence of either subunit were unsuccessful; however, sequence information was obtained for seven internal tryptic peptides of p160. Three of these were unambiguous, and the mass of each peptide was consistent with the mass predicted from the sequence. These were PT48 (GGHFYSALPEILR), PT56 (VHVQFFDDSPTR), and PT64 (TLLEEEYFR). Two others gave secondary sequences in early cycles, and only the unambiguous sequence information is given here: PT36 (DAAWSEAGPGPR) and PT40 (SVLEGDPSENYR). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
18. S.-S. Su, R. S. Lahue, K. G. Au, P. Modrich, *J. Biol. Chem.* **263**, 6829 (1988).
19. N.-W. Chi and R. D. Kolodner, *ibid.* **269**, 29984 (1994).
20. E. Alani, N.-W. Chi, R. Kolodner, *Genes Dev.* **9**, 234 (1995).
21. M. Grilley, K. M. Welsh, S.-S. Su, P. Modrich, *J. Biol. Chem.* **264**, 1000 (1989).
22. A. Umar, J. C. Boyer, T. A. Kunkel, *Science* **266**, 814 (1994).
23. G.-m. Li, M. J. Longley, J. T. Drummond, unpublished results.
24. W.-h. Fang and P. Modrich, *J. Biol. Chem.* **268**, 11838 (1993).
25. V. S. Goldmacher, R. A. Cuzick, W. G. Thilly, *ibid.* **261**, 12462 (1986); A. Kat *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6424 (1993).
26. J. Holmes, S. Clark, P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5837 (1990); D. C. Thomas, J. D. Roberts, T. A. Kunkel, *J. Biol. Chem.* **266**, 3744 (1991).
27. L. T. da Costa *et al.*, *Nature Genet.* **9**, 10 (1995).
28. F. Palombo *et al.*, *Science*, **268**, 1912 (1995).
29. N. Papadopoulos *et al.*, *ibid.*, p. 1915.
30. Cell lines HeLa $S_3$, H6, S0, TK6, and MT1 were grown as described (*2, 25, 26*). The HCT-15 cell line obtained from the American Type Culture Collection was cultured in roller bottles (850 $cm^2$) in RPMI 1640 medium supplemented with 20% fetal bovine serum (Hyclone, Logan, UT). LoVo cells were cultured in a similar manner in McCoy's 5A medium containing glutamine (4 mM final concentration), 10% fetal bovine serum, penicillin (100 U/ml), and streptomycin (0.1 mg/ml).
31. Phage f1 heteroduplexes representing eight base-base mispairs and two-, three-, or four-nucleotide loopout, slipped-strand insertion-deletion mismatches were prepared as described (*2, 18*). Bacteriophage f1MR21 and f1MR22 were prepared by insertion of synthetic oligonucleotide duplexes (5′-AGCTGCAGCCAAAAAGGCTTGG-3′ annealed with 3′-CGTCGGTTTTTCCGAACCGATC-5′ or 5′-AGCTGCAGCCAAAAAAGGCTTGG-3′ annealed with 3′-CGTCGGTTTTTCCGAACCGATC-5′, respectively), into f1MR1 replicative form DNA (*18*) that had been cleaved with Hind III and Xba I. Resulting isolates were used to prepare heteroduplexes containing /A\ and /T\ insertion-deletion mismatches within overlapping restriction sites for Bgl I and Xcm I. Mismatch repair assays (*2, 26*) were performed with 50 μg of nuclear extract and 24 fmol (0.1 μg) of heteroduplex DNA in 10 μl at 37°C for 15 min. All results are averages of at least two determinations. For band shift assays, oligonucleotide duplexes containing a G-T mispair, a /T\ insertion mismatch, a /TTT\ insertion mismatch, or a control A-T base pair were formed by annealing radiolabeled 5′-$^{32}$P-GCTAGCAAGCTTTCGATTCTAGAAATTCGGC-3′ with unlabeled 5′-GCCGAATTTCTAGAATCGAGAGCTTGCTAGC-3′, 5′-GCCGAATTTCTAGAATCGAAGCTTGCTAGC- 3′, 5′-GCCGAATTTCTAGAATCGGCTTGCTAGC-3′, or with 5′-GCCGAATTTCTAGAATCGAAAGCTTGCTAGC-3′, respectively.
32. J.T.D. isolated and characterized hMutSα and demonstrated that HCT-15 and MT1 cells are deficient in this activity. G.-M.L. identified the selective repair defects in HCT-15 and MT1 cell lines and demonstrated that hMutSα binds to mismatched base pairs. We thank F. Leach, D. Hill, and B. Vogelstein for monoclonal antibody EH12; E. Korytynski and S. Larson for culturing the cell lines; B. Vogelstein and J. Jiricny for sharing unpublished results; and the Harvard Microchemistry Facility for peptide analysis. Supported by grant GM45190 from the National Institute of General Medical Sciences, translational funds from the Duke Comprehensive Cancer Center, and by an American Cancer Society postdoctoral fellowship PF-3940 (M.J.L.).


6 January 1995; accepted 4 May 1995

# GTBP, a 160-Kilodalton Protein Essential for Mismatch-Binding Activity in Human Cells

Fabio Palombo, Paola Gallinari, Ingram Iaccarino, Teresa Lettieri, Melya Hughes, Antonello D'Arrigo, Oanh Truong, J. Justin Hsuan, Josef Jiricny*

DNA mismatch recognition and binding in human cells has been thought to be mediated by the hMSH2 protein. Here it is shown that the mismatch-binding factor consists of two distinct proteins, the 100-kilodalton hMSH2 and a 160-kilodalton polypeptide, GTBP (for G/T binding protein). Sequence analysis identified GTBP as a new member of the MutS homolog family. Both proteins are required for mismatch-specific binding, a result consistent with the finding that tumor-derived cell lines devoid of either protein are also devoid of mismatch-binding activity.

Cells from human hereditary nonpolyposis colorectal cancers (HNPCC) exhibit a mutator phenotype with a marked instability of microsatellite sequences, a phenotype that has been traced to defects in DNA mismatch repair [reviewed in (*1*)]. The serendipitous discovery of open reading frames (ORFs) encoding murine and human polypeptide homologs of the *Escherichia coli* mismatch-binding protein MutS (*2*) paved the way for the identification of an ever-growing family of MutS homolog (MSH) genes (*3*). Proteins encoded by three members of this family (*Saccharomyces cerevisiae* MutS homologs MSH1 and MSH2, and the human homolog hMSH2) have been shown to bind to DNA mismatches in vitro (*4*). The link between the biological function of hMSH2 and the phenotype of the HNPCC tumors was forged when the *hMSH2* gene was shown to segregate with a known HNPCC locus on chromosome 2p (*5*), and the *hMSH2*-deficient colorectal tumor cell line LoVo was shown to be deficient in mismatch repair (*6*) as well as in mismatch-binding activity (*7*) and to exhibit a marked instability of microsatellite sequences (*8*).

We identified a mismatch-binding factor in HeLa cells (*9*) that binds preferentially to heteroduplexes containing G/T mispairs and one- and two-nucleotide loops (*7*). Purification of this DNA binding activity by G/T mismatch affinity chromatography (*10*) yielded a mixture consisting of a 100-kD protein that we later showed to be hMSH2 (*11*) and a second protein with an apparent molecular weight of 160 kD. As it was this latter polypeptide that became covalently bound to the G/T substrate in cross-linking experiments, we named it GTBP (*10*). To identify GTBP and elucidate its role in mismatch binding, we subjected a purified sample to proteolysis and microsequence analysis (*12*). Seven peptide sequences were obtained: VRVHVQFFDD, KLPDLERLXSK, LSRGIGVMLPQVL, TLDTLLEEEYFREK, SYGFNAARLANLPEEVIQ, NPEGRFPDLTVELN, and IIDFLSALEGFK (*13*). Following the strategy described in (*14*), we identified a unique DNA sequence encoding the central eight amino acids of the peptide SYGFNAARLANLPEEVIQ (*13*). We then prepared a double-stranded DNA probe based on this sequence and used it to screen a HeLa complementary DNA (cDNA) library in phage lambda (*15*). The longest identified clone, C1, contained an insert of 3992 nucleotides with an ORF encoding a polypeptide of 1292 amino acids and a predicted molecular mass of 142 kD (Fig. 1A). As all seven peptides obtained from the microsequence analysis of the 160-kD protein were located within this ORF, we concluded

F. Palombo, P. Gallinari, I. Iaccarino, T. Lettieri, M. Hughes, A. D'Arrigo, J. Jiricny, Istituto di Ricerche di Biologia Molecolare P. Angeletti, Via Pontina km 30.600, 00040-Pomezia, Italy.
O. Truong and J. J. Hsuan, Ludwig Institute for Cancer Research, University College London Medical School, 91 Riding House Street, London W1P 8BT, UK.
*To whom correspondence should be addressed.