DNA Template Requirements for Human Mismatch Repair in Vitro*

The human mismatch repair pathway is competent to correct DNA mismatches in a strand-specific manner. At present, only nicks are known to support strand discrimination, although the DNA end within the active site of replication is often proposed to serve this role. We therefore tested the competence of DNA ends or gaps to direct mismatch correction. Eight G-T templates were constructed which contained a nick or gap of 4, 28, or ~200 nucleotides situated ~330 bp away in either orientation. A competition was established in which the mismatch repair machinery had to compete with gap-filling replication and ligation activities for access to the strand discontinuity. Gaps of 4 or 28 nucleotides were the most effective strand discrimination signals for mismatch repair, whereas double strand breaks did not direct repair to either strand. To define the minimal spatial requirements for access to either the strand signal or mismatch site, the nicked templates were linearized closed to either site and assayed. As few as 14 bp beyond the nick supported mismatch excision, although repair synthesis failed using 5'-nicked templates. Finally, asymmetric G-T templates with a remote nick and a nearby DNA end were repaired efficiently.

The known contributions of the human DNA mismatch repair (MMR)1 pathway to genomic stability have expanded greatly over the past decade (1–3). A clear relationship between the loss of a functional DNA MMR pathway and tumor formation in both hereditary and sporadic settings has emerged (1, 2). Loss of one of the protein activities essential for mismatch correction results in a mutator phenotype, in which mononucleotide repeats within coding sequences of genes that contribute to growth suppression or apoptosis represent important mutational hot spots. In addition to its role in repairing biosynthetic errors, the MMR pathway plays an oversight role in DNA recombination events, limiting nonhomologous strand exchange (4, 5). Lesion recognition and processing have also been implicated in apoptotic events mediated by p53 and p73, although the specific sequence of events subsequent to lesion recognition is not well defined (for review, see Ref. 6). Two commonly cited models are the futile repair model, in which the protein components interact to effect mismatch correction results in strand excision and resynthesis remains unresolved (1, 3). This is especially true for the events subsequent to mismatch recognition, including the identification of the strand discrimination mechanism. One model posits that mismatches are bound by the MSH2/MSH6 heterodimer (MutSo), followed by an ATP-dependent, bidirectional translocation along the DNA to identify a strand discrimination signal and orchestrate downstream events (8). A more recent analysis suggests that MutSo and MutLα (MLH1/PMS2) together form a complex capable of translocating along the DNA helix (9). A second model suggests that MSH2/MSH6 is an activity analogous to a G protein, recognizing mismatches when an ADP molecule is bound (10, 11). Exchange of ADP for ATP in the ternary complex results in a hydrolysis-independent sliding clamp that allows diffusion along the DNA, again presumably to initiate later events in the repair pathway. In a very different formulation, it has also been argued that mismatch-binding proteins, in the presence of a heterodimer of MutL homologs and ATP, remain at the mismatch site (3). This model was formulated using prokaryotic proteins and proposes a role for ATP in mismatch verification (12), which appears to be inconsistent with the two models cited above. Proliferating cell nuclear antigen is also known to participate in the early stages of mismatch correction (13), most recently as a modulator of mismatch recognition (14) and by its ability to disrupt the complex of MutS and MutL homologs from the mismatch site in Saccharomyces cerevisiae (15).

In contrast to the diverse efforts directed toward understanding how the protein components interact to effect mismatch correction, the known DNA template requirements for human MMR are few. A successful template for in vitro assays can be summarized as a circular molecule containing a mismatch and nick positioned from 100 to 1,000 bp away in either DNA strand (16, 17), where repair efficiency is diminished as the distance separating the two sites is increased (18). The nick is essential for targeting repair to the strand where it resides (16, 17, 19), although the ultimate strand signal in human cells (or any organism outside a small number of related Gram-negative prokaryotes) remains unknown. Recent biochemical analyses suggest that mismatch processing in human cell extracts operates independently of the CpG methylation status on the template, whether it is fully, hemi-, or unmethylated (20; for an analysis using Xenopus laevis extracts, see Ref. 21).

The possibility that a nick was sufficient to act as a strand signal was drawn from the better characterized Escherichia coli model system (3, 22). In this organism, a site-specific DNA nick represents an intermediate in mismatch correction (23, 24). The action of MutH protein on a hemimethylated GATC sequence introduces a strand-specific nick, and this reaction is largely dependent upon the presence of a nearby mismatch (25). Furthermore, the presence of a nick eliminates the need for MutH activity in a reconstituted assay system (24). In

1 This work was supported by National Institutes of Health Genetics, Cellular, and Molecular Sciences Training Grant GM 07757-21 (to K. I. and E. D. L.) and National Institutes of Health Grant R01-GM61586. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

2 To whom correspondence should be addressed: Dept. of Biology, Indiana University, 1001 E. Third St., Bloomington, IN 47405. Tel.: 812-856-4184; Fax: 812-855-6705; E-mail: jdrummon@bio.indiana.edu.

3 The abbreviations used are: MMR, mismatch repair; dsDNA, double strand DNA; ssDNA, single strand DNA.

Keith Iams, Erik D. Larson, and James T. Drummond‡

From the Department of Biology, Indiana University, Bloomington, Indiana 47405

Received for publication, January 28, 2002, and in revised form, June 18, 2002
Published, JBC Papers in Press, June 19, 2002, DOI 10.1074/jbc.M200846200

* This work was supported by National Institutes of Health Genetics, Cellular, and Molecular Sciences Training Grant GM 07757-21 (to E. D. L.) and National Institutes of Health Grant R01-GM61586. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Biology, Indiana University, 1001 E. Third St., Bloomington, IN 47405. Tel.: 812-856-4184; Fax: 812-855-6705; E-mail: jdrummon@bio.indiana.edu.

‡ The abbreviations used are: MMR, mismatch repair; dsDNA, double strand DNA; ssDNA, single strand DNA.
parallel with the bidirectional capability of mismatch correction in *E. coli* (26), extracts from human cells can access a single strand nick placed in either strand and in either orientation with respect to the mismatch (18). Mismatch processing appears to function equally well when exonucleolytic removal of the mismatch occurs 5′→3′ or 3′→5′.

Although a nick may act as a strand discrimination device *in vitro*, its relationship with the authentic strand signal *in vivo* remains unclear. Strand discontinuities, including nicks, are present during replication, but they are much more frequent in lagging strand biosynthesis. Another potential means of identifying the daughter strand is via the DNA end where strand elongation occurs, even though it is sequestered by the replicative polymerase. This possibility is supported by the fact that proliferating cell nuclear antigen participates in an early step in mismatch correction, prior to resynthesis, and it has been shown to interact directly with mismatch recognition proteins (13, 27). Models for strand discrimination which invoke the site of DNA synthesis are superficially inconsistent with nick-directed repair because displacement of the replicative polymerase to allow access to the MMR machinery is likely to reveal a DNA end (3′-OH) within a gap rather than a nick. Restated, the ability to identify and access a nick does not represent the same capability as accessing the active site of replication.

In this work, we assessed the contribution of two major features of the DNA template upon which the human MMR pathway acts. First, we asked whether strand discontinuities beyond a single strand nick are competent to target mismatch correction. Heteroduplexes containing a single G·T mismatch and a site- and size-specific single strand gap, positioned in either orientation (5′→3′ or 3′→5′ from the discontinuity to the mismatch) were assayed using nuclear extracts prepared from HeLa cells. The gap sizes chosen were intended to represent those that might arise from specific DNA repair or synthesis events and would plausibly represent targets for the human DNA polymerases α, β, δ, or ε. Nicks and gaps of 4–210 nucleotides were found to be effective strand signals, and intermediate sized gaps were the most efficient in either orientation. Second, we asked whether DNA ends were competent to direct MMR, and we found that neither blunt ends nor small overhangs supported mismatch correction targeted to either strand. MMR was supported efficiently on linear templates, and we used this result to ask systematically whether DNA ends limited mismatch correction as they were moved close to either the strand signal or mismatch. In this way, we were able to define some of the spatial and structural relationships required for efficient mismatch correction *in vitro*.

**MATERIALS AND METHODS**

**Construction of M13 Derivatives Competent for Assaying MMR Using Restriction Analysis**—Manipulation of M13mp18 bacteriophage and its derivatives was performed using standard methodology for culturing filamentous phage (28) using the XL-1 Blue strain of *E. coli* (Stratagene). Similarly, all DNA manipulation steps such as precipitation, digestion, and ligation were carried out using standard methodology (28). Before introducing a site competent for mismatch construction, the unique HindIII site was eliminated from M13mp18. A sample of the DNA was digested with HindIII, the overhanging ends filled in with the Klenow fragment of DNA polymerase I, and the molecule was recircularized using a blunt end ligator. The original HindIII site was thus replaced with an Nhel site, giving a molecule we called M13AH.

An oligonucleotide containing the mismatch was designed to replace the sequence intervening between the M13AH Kpnl and *Aval* sites and generate two prophage phages for substrate construction. The sequence, described below, included a mismatch embedded within overlapping restriction sites to facilitate strand-specific identification of MMR in the final substrates (29; see Fig. 1). After digestion with *Aval* and Kpnl, removal of the putative intervening fragment, and ligation in the presence of the replacement heteroduplex, transformation yielded single plaques that were sequenced across the targeted region. It became apparent that the Kpnl digestion failed, and a cryptic but useful ligation reaction occurred. Plaques were isolated in which the *AvaI* site was not corrected, but the Kpnl overhang was trimmed from the oligonucleotide and ligated to the second *AvaI* overhang, which was appropriately modified, by a 5′ end by replication driven by pUC19 (C), whereas M13AH clone 7 was sensitive to HindIII (T). Both included an additional StyI site (CCTAGG) located adjacent to the mismatch site to facilitate template remodeling.

To construct G·T heteroduplexes with 5′→3′ strand discontinuities, large scale preparations of both double (ds) and single strand (ss) DNA were performed such that clone 4 was used as the double strand species and clone 7 as the single strand species. Because the heteroduplex region was introduced at the *AvaI* site of M13AH, such constructs place the mismatch 329 bp from the EcoRI site and a rich source of other nearby unique digestion sites useful for the construction of gapped regions. The DNA sequence reads 5′→3′ from nick or gap toward the mismatch. Repair directed by the nick or gaps is therefore scored with HindIII, whereas mismatch correction in the continuous strand is identified with XhoI (16, 29).

To construct comparable phages useful for preparing 3′-nicked or gapped substrates, a 394-bp segment containing the mismatch region and polycytosine site of M13 was excised and inverted. The fragment was first digested with Nhel and StyI, which recognize distinct sequences but yield complementary overhanging ends upon digestion. The fragment was reinserted *in situ* by adding a 1/30 volume of 10× ligation buffer (New England Biolabs) and 80 units of T4 DNA ligase, then incubating for 24 h at 22 °C in the presence of Nhel and StyI. The presence of these activities selects against regeneration of the parent phage. The reaction was transformed into XL1-Blue cells, and the resulting plaques were screened for loss of StyI and Nhel sensitivity. Two selected phages, designated 4F (XhoI site) and 7F (HindIII site), were sequenced through the targeted region to demonstrate sequence inversion between the StyI and Nhel sites.

For construction of 3′-substrates, clone 4F was used to prepare the single strand species, and DNA derived from clone 7F was used to amplify the double strand species. The substrate preparation protocol described below results in the production of G·T heteroduplexes with a fragment of DNA sequence between the mismatch and strand discontinuity is preserved compared with the 5′-substrates. Mismatch removal and resynthesis events that are nick- or gap-directed now yield products that are sensitive to XhoI, whereas mismatch correction in the continuous strand is identified with HindIII (16, 29).

**Construction of a Nicked Circular DNA Containing a G·T Mispair**—The preparation of heteroduplexes was performed using minor modifications to published protocols (29). To construct a nicked substrate, 1.5 μg of the appropriate dsDNA was linearized by digestion with EcoRI and precipitated by the addition of 1/10 volume of 3 M sodium acetate and 2 volumes of absolute ethanol. The DNA was recovered by centrifugation for 20 min at 16,000 g in a microcentrifuge at room temperature. This precipitation protocol was followed for each ethanol precipitation described subsequently, unless otherwise specified. The pelleted DNA was then resuspended in 2.6 ml of buffer containing 50 mM Tris-HCl, pH 8.0, 12.5 mM NaCl, 0.8 mM EDTA, and 750 μg of ssDNA. The strands were denatured for 5 min at room temperature by the addition of 60 μl of 10× NaOH, and the reaction was neutralized by the addition of a mixture of 200 μl of 2.9 M acetic acid, 248 μl of 1 M potassium phosphate buffer, pH 7.4, and 90 μl of 3 M KCl. The renaturation reaction was incubated at 65 °C for 1 h and then 37 °C for a further 3 h. The solution was diluted with 1.5 volumes of distilled water and the DNA precipitated with ethanol. The DNA was recovered by centrifugation in an SW41 swinging bucket rotor at 68,000 × *g* for 1 h at 20 °C. To remove the excess ssDNA, the pellet was resuspended in 2.5 ml of 1 M NaCl and passed three times through a column containing 50 μl of benzylated naphthyl-Sepharose-DEAE-cellulose (Sigma) that had been equilibrated in 1 M NaCl. The DNA was then eluted directly from this solution by adding 2 volumes of absolute ethanol and centrifugation as described above. To digest any residual linear DNA, the pellet was resuspended in a 400-μl reaction containing 66 mM glycine at pH 9.4, 5 mM MgCl₂, 0.5 mM ATP, and 35 units of exonuclease V enzyme (U. S. Biochemical Corp.) and incubated at 37 °C for 1 h. The enzyme was then inactivated by a 20-min incubation at 65 °C. The
circular substrate was fractionated from nucleotides and small oligonucleotides by ultrafiltration; each reaction was concentrated to less than 100 μl in volume using a Microcon 100 centrifugal filter as described by the manufacturer (Millipore). The filtrate containing small molecules and buffer was discarded, and the process was repeated twice by adding 1 ml to the solution to 400 μl with a buffer of 10 mM Tris-HCl, pH 8.0, containing 1 mM EDTA.

Construction of a Circular DNA Containing a G7 Mispair and a 4-Nucleotide Gap—To construct substrates containing 4-nucleotide gaps, two additional phage templates were constructed which were expanded by 4 bp. The double strand form of clones 7 (5’-template) and 4F (3’-template) were digested independently with EcoRI and the resulting overhangs were filled in with the Klenow fragment, as described above. After blunt end ligation and transformation, individual plaques were screened for lack of sensitivity to EcoRI digestion and sequenced to confirm the addition of 4 nucleotides. These two constructs, called 7E and 4F, were used to produce the ssDNA partner in the annealing reaction described above. The annealing reactions yield a gapped molecule containing the single strand sequence 5’-ATTG in place of the EcoRI recognition site. All further steps were carried out as per the nicked circular DNA substrate.

Construction of a Circular DNA Containing a G7 Mispair and a 28-Nucleotide Gap—To construct the templates containing gaps of 28 nucleotides, the protocol for preparing nicked substrates was followed with one modification. Before digestion with EcoRI, 100 μg of the appropriate dsDNA was digested with 1 unit/μg XbaI, and the product was precipitated twice with ethanol. The digestion was validated by gel electrophoresis, and DNA sequencing was used to verify the second digestion. One-fifth of the EcoRI digestion was passed through each of five High Pure PCR columns as described by the manufacturer to remove the 28-bp fragment resulting from the double digest, and the eluates were combined. All subsequent steps were performed as described for templates containing a nick.

Direct repair of Mismatch—Mismatch Repair-directing Structures

Mismatch Repair Assays—Mismatch repair repair assays employing LoVo nuclear extracts were carried out essentially as with HeLa repair assays with the exception that 50 μg of LoVo nuclear extract was used in place of the HeLa extract. In the assays that were supplemented with purified MutSα, 300 ng of purified protein was added to each reaction. This represents a saturating amount of MutSα; adding more than ~100 ng to the assay did not enhance repair efficiency.

Effect of Salt Concentration and Nuclear Extract Amount on MMR—Assays were set up so that the final salt concentration of components entering the assay was 100 mM with respect to added cations (primarily potassium). When complete assays were tested, the ionic strength was found to be equivalent to a solution of 125 mM KCl, based on a standard curve. In our hands, maximal repair was observed over a concentration range from 125 to 150 mM. Repair assays failed when the salt concentration exceeded 200 mM KCl. Although the assays reported in this work were done at the lower end of the range, no improvement in repair activity for any substrate was observed at higher salt concentrations. The amount of nuclear extract used in each assay was determined by titrating the repair efficiency using 100 ng of heteroduplex DNA. The maximal level of repair was reached at ~20–30 μg of extract, and 50 μg was chosen as a standard assay amount to ensure that extract components did not limit repair efficiency. Time courses were performed to verify that no significant amount of mismatch correction occurred beyond the 20-min assay time course.

The MMR assay employed here (16) has several limitations that must be recognized. In the absence of a template, it is a nonlinear assay (18) that reports the maximal amount of repair that may be obtained under a specific set of reaction conditions. For a given amount of template, no increase in the extent of MMR can be obtained by increasing the amount of nuclear extract that provides the requisite activities (16, 18). The substrates rely on a strand break to direct repair, and sealing the required break by a DNA ligase can preclude further mismatch processing (18). Similarly, in the reconstituted repair pathway from E. coli (34), high concentrations of ligase can suppress mismatch correction. HeLa extracts prepared as described are relatively unstable at 37 °C, which complicates the critical assessment of parameters beyond the magnitude of mismatch correction. Under the conditions used here, patterned on earlier characterizations of this assay method (16, 31), typical values reported for nick-directed repair of G7 mismatches ranged from 38 to 50% (9.4–12 fmol).

Mismatches and Purified MutSα—Mismatch repair repair assays employing LoVo nuclear extracts were carried out essentially as with HeLa repair assays with the exception that 50 μg of LoVo nuclear extract was used in place of the HeLa extract. In the assays that were supplemented with purified MutSα, 300 ng of purified protein was added to each reaction. This represents a saturating amount of MutSα; adding more than ~100 ng to the assay did not enhance repair efficiency.

Analysis of Mismatch Repair on Linear Templates—Repair assays using linear substrates were performed exactly as they were for circular substrates, except that the DNA was digested in a buffer compatible with both enzyme and the assay (New England Biolabs buffer 2 or 4) and introduced directly into the assay. For example, 450 ng of substrate was digested for 1 h at 37 °C with the indicated enzyme in a 25-μl volume as specified by the supplier. A sample containing 50 ng was removed and evaluated for completeness of the reaction by agarose gel analysis. Aliquots containing 100 ng (20 fmol) were added directly to each assay and the concentration of monovalent cations adjusted to 125 mM. Similar assay results were obtained when the substrates were precipitated from the restriction digest or concentrated by ultrafiltration and introduced into the assay (not shown). When the relative efficiency of the MMR pathway on linear and circular templates was compared (Fig. 4), the substrate containing a 28-nucleotide gap entering each assay was treated under restriction digest conditions. AluNI restriction enzyme (1 unit/assay) was added where indicated to generate the linear substrate. When nicked templates were mock digested the repair efficiency dropped modestly, which is reflected in Figs. 5 and 7.

In experiments in which the substrates were linearized near the mismatch, a modification of the scheme for scoring repair was required. The 5’- nicked template was digested with BspHI in addition to HindIII, and nucleotide fragments of 2,600 and 2,400 nt were isolated from the template linearized prior to the assay with DrdI. The 3’- nicked substrate was digested with PstI in addition to XhoI, generating repaired and unrepairped fragments of 2,165 and 2,299 bp, respectively, on tem-
plate linearized prior to the assay with PvuI. The repair value in each assay was calculated by comparing band intensities of the unique repaired and unrepaired fragments, with consideration given to the small repair-specific fragments not resolved in the assay.

Analysis of Denatured DNA by Alkaline Gel Electrophoresis—Mismatch-containing templates were assayed for mismatch correction as described above and digested with AluNI. Precipitated DNA was re-suspended in 15 μl of aqueous solution containing 50 mM NaOH and 1 mM EDTA, and 15-cm alkaline 0.9% agarose gels were poured and run at 4 °C (60 V for 12.5 h) as described (28). After gel neutralization, the DNA was transferred onto a nylon membrane (Osmonics, Inc.) and cross-linked using a 45-s exposure to short wave UV illumination using an Alpha Innnotech transilluminator. The membranes were pretreated at 65 °C for 1.5 h in 6× SSC buffer (0.15 M NaCl and 0.15 M sodium citrate at pH 7.0 in 1× SSC) containing 0.25% nonfat milk powder, whereupon the radiolabeled probe (5'-GGTTCTGTTGCGGCCTCTGGA-3') was introduced and the hybridization reaction allowed to proceed at 52 °C (28). The membranes were washed with three changes of 2× SSC buffer containing 0.5% SDS at room temperature and exposed to Kodak X-Omat AR film. The probe sequence was complementary to the discontinuous DNA strand near the AluNI site (see Fig. 6).

RESULTS

A set of eight dsDNA molecules that all contain a single G-T mismatch was constructed to address three mechanistic questions about the human MMR pathway. First, because DNA nicks represent the only characterized strand discrimination signal available for targeting MMR, we wanted to determine whether other types of DNA discontinuities or ends were also competent. This allows us to investigate whether specific DNA templates that might arise during DNA replication or recombination are capable of directing mismatch excision and resynthesis. Second, we wanted to ask whether the MMR pathway could compete effectively with different classes of gap-filling replication events. In models that seek to explain strand discrimination based on the DNA terminus where elongation occurs (see Refs. 27 and 32), mismatch correction must be able to compete effectively with DNA synthesis using polymerases that function during either leading or lagging strand biosynthesis. Third, we sought to characterize a minimal DNA template for the human MMR pathway, using DNA constructs that systematically limited the DNA template close to either the mismatch or the strand discrimination signal. This deletion-based mapping strategy allows an approximation of the physical template size required for specific stages in mismatch correction, such as the initiation of an excision tract.

Comparison of 3′- and 5′-Substrates—The eight G-T templates can be broken down into two groups based on strand polarity. Four templates represent nicks or gaps where the DNA sequence reading from the nick or gap to the mismatch reads 5′ → 3′ (5′-substrates; Fig. 1, top), whereas four represent the opposite polarity (3′-substrates; Fig. 1, bottom). Inversion of the 394-bp region within the replicative phage form of M13 does confer substantial differences between the 3′- and 5′-substrates which must be considered. The heteroduplexes are constructed from a single strand and double strand phage partner, and the discontinuous strand is always derived from the duplex. This means that one substrate is technically a G mismatch and the later stages of gap-filling replication on the 5′-substrates shown in the upper portion, and with the early stages of gap-filling synthesis for the 3′-substrates in the lower portion.

![Fig. 1. Relationship between DNA mismatches and strand discontinuities](http://www.jbc.org/)
and reconstituted the HindIII restriction enzyme sequence (Fig. 2, upper panel, even numbered lanes). No digestion of these substrates with XhoI restriction enzyme was detected (Fig. 2, upper panel, odd numbered lanes). Under the assay conditions used, the mean assay value for each gap size represents a significant difference in MMR efficiency compared with nick-directed repair using a Student's independent samples t-test ($p = 0.0005$).

Parallel assays were performed for the substrates containing a 3'-nick or gapped region, and the results are shown in the bottom panel of Fig. 2. In these substrates, the overlapping restriction sites used to score repair (HindIII and XhoI) were maintained, but the strand signals now reside in the T-containing strand (refer to Fig. 1). Note that a replication event that fills the gap must initiate at the DNA end that serves as a strand discrimination signal. Mismatch correction events therefore compete with replication initiation and elongation for the available 3'-terminus. As in the 5'-substrates above, repair of mismatches present in molecules with gaps of either 4 or 28 nucleotides ($\sim 80\%$) was significantly more efficiently than nick-directed repair. The template containing a gap of 210 nucleotides was repaired with an efficiency similar to the nick-containing molecules (44% compared with 41% of the molecules, respectively). In each instance, only repair directed to the nicked or gapped strand that reconstituted the XhoI restriction site, and no repair was detectable in the continuous strands. In all cases, the repair values are given ± 1 S.D. from the mean value determined, based on 3–11 independent assays.

FIG. 3. MMR assays using LoVo extracts in the absence and presence of added MutSα. Mismatch correction was not detectable (nd) in LoVo extracts alone (odd numbered lanes), which lack the MSH2 protein (35) and functional MutS heterodimers (30, 34). The addition of 300 ng of purified MutSα (even numbered lanes) restored mismatch correction to a detectable level that mirrored the pattern found using HeLa extracts.

and reconstituted the HindIII restriction enzyme sequence (Fig. 2, upper panel, even numbered lanes). DNA fragments resulting from digestion by HindIII at the corrected mismatch site are visible as a doublet below (see “Materials and Methods”). No mismatch correction targeted to the continuous (closed) strands was detectable with XhoI (odd numbered lanes), whereas mismatch removal and resynthesis in the open strands varied from 49 to 80% (even numbered lanes). nd indicates not detectable; na means not attempted. In the bottom panel, the repair efficiency of the analogous 3'-substrates is compared after passage through HeLa extract. Strand-specific repair directed to the discontinuous strand resulted in reconstitution of the XhoI restriction site, and no repair was detectable in the continuous strands. In all cases, the repair values are given ± 1 S.D. from the mean value determined, based on 3–11 independent assays.

and reconstituted the HindIII restriction enzyme sequence (Fig. 2, upper panel, even numbered lanes). DNA fragments resulting from digestion by HindIII at the corrected mismatch site are visible as a doublet below (see “Materials and Methods”). No mismatch correction targeted to the continuous (closed) strands was detectable with XhoI (odd numbered lanes), whereas mismatch removal and resynthesis in the open strands varied from 49 to 80% (even numbered lanes). nd indicates not detectable; na means not attempted. In the bottom panel, the repair efficiency of the analogous 3'-substrates is compared after passage through HeLa extract. Strand-specific repair directed to the discontinuous strand resulted in reconstitution of the XhoI restriction site, and no repair was detectable in the continuous strands. In all cases, the repair values are given ± 1 S.D. from the mean value determined, based on 3–11 independent assays.

and reconstituted the HindIII restriction enzyme sequence (Fig. 2, upper panel, even numbered lanes). DNA fragments resulting from digestion by HindIII at the corrected mismatch site are visible as a doublet below (see “Materials and Methods”). No mismatch correction targeted to the continuous (closed) strands was detectable with XhoI (odd numbered lanes), whereas mismatch removal and resynthesis in the open strands varied from 49 to 80% (even numbered lanes). nd indicates not detectable; na means not attempted. In the bottom panel, the repair efficiency of the analogous 3'-substrates is compared after passage through HeLa extract. Strand-specific repair directed to the discontinuous strand resulted in reconstitution of the XhoI restriction site, and no repair was detectable in the continuous strands. In all cases, the repair values are given ± 1 S.D. from the mean value determined, based on 3–11 independent assays.
nucleotides were assayed under standard conditions. Lanes DNA treated under mock restriction digest conditions and assayed was digested with HindIII to identify molecules repaired at the mismatch site shown in lane 3.

Our next objective was to truncate the template molecules for strand excision events using denaturing alkaline gels (see “Materials and Methods”), assuming that the early stages of mismatch correction might be successful but that strand resynthesis might fail. No evidence for any mismatch-dependent processing events was found (data not shown).

Dependence of Mismatch Correction Activity on Template Length Adjacent to the Nick—The 5’- and 3’-nicked substrate molecules were singly truncated with a series of restriction enzymes that create a double strand break close to the nick. These experiments were directed toward establishing the minimal spatial requirements of the MMR apparatus for initiating an excision event from the nick site and for the polymerization event that restores the excision tract. Samples of the 5’-nick substrate were digested independently with Smal, AccI, and Nhel (Fig. 5), whereas samples of the 3’-nick substrate were digested with only the former two enzymes. The Nhel site in this template was disrupted during construction of the 3’-substrates and was therefore unavailable. Digestion produces a linear molecule with ends located 14 (blunt), 33 (5’-overhang), and 53 bp (5’-overhang) away from the 5’-nick, respectively. The linear templates were assayed directly for effectiveness as a repair substrate under standard assay conditions (see “Materials and Methods”), using the same template digested at a remote site (AlwNI) as a control. Note that in this and subsequent experiments, where any of the DNA templates was digested or treated under mock digest conditions prior to assay, the overall magnitude of nick-directed repair efficiency was reduced from 50% to ~40%.

Shown in Fig. 5 (left panel) are assays using 5’-nick substrate linearized close to the nick site. A diagram drawn to scale showing the spatial relationship between these sites is shown below it. No mismatch correction was detectable on templates linearized 14 or 33 bp away from the nick (Smal or AccI), whereas 38% repair was measurable with the Nhel linearized substrate. For comparison, similar repair efficiencies were obtained in control assays using 5’-nicked substrate molecules digested with AlwNI so that the closest linear end was located 3,540 bp from the nick. This result indicates that, in substrates containing a 5’-nick, restoration of the correct base pair at the mismatch site is inhibited by a double strand break.
placed within 33 bp of the nick. With the slightly greater flanking site size of 53 bp, the mismatch site was repaired at the control level.

An analogous experiment was performed to assess the viability of templates truncated near the 3'-nick for competence in mismatch correction. Digestion with Smal or AccI was used to produce linear ends located 18 or 33 bases from the nick (refer to the diagram in Fig. 5, right panel). As a benchmark, 35% of the 3'-nicked substrate linearized at a remote site (AlwNI) was repaired in this assay. Equivalent repair efficiency was found when only 33 bp extended beyond the 3'-nick (39%), and nearly half-of the control level of repair was detectable when only 18 bases extend beyond the nick (16%). This suggests that the MMR machinery can access a nick placed within 18 bp of a blunt DNA end, and sufficient template is present to allow loading and function of the machinery required for mismatch-dependent, nick-directed strand excision.

The failure to observe mismatch correction using a restriction-based assay, as in the case of the molecules linearized proximal to the 5'-nick site described above, does not imply that either mismatch recognition or subsequent processing steps failed. This assay simply quantifies the extent to which the repair machinery can reach the discontinuity with a 32P-labeled oligonucleotide complementary to the G-containing strand (refer to the diagram below the figure). Probing the denatured substrate in the absence of HeLa treatment reveals a fragment that runs from the AlwNI-generated DNA end to the gap or nick region and includes the mismatch site (4,069 nucleotides; see lane 1). The mismatch site lies 330 nucleotides from the discontinuity in these 5'-heteroduplexes. Upon HeLa treatment, ligation of the nick (lane 2) or a gap-filling replication followed by ligation (lane 3) results in a full-length M13 template of 726 nucleotides which represents both repaired and unrepairable molecules. Substrates linearized 14 bp beyond the discontinuity with Smal before HeLa treatment are unable to support resynthesis and yield only a family of excision products (the smeared band in lanes 4 and 5) which extends beyond the mismatch site (indicated by a dashed line). The broad excision tract is centered ~300 nucleotides past the mismatch site, as depicted in the diagram on the right.

As shown in Fig. 6, full-length G-T substrates containing either a 5'-nick or a 5'-gap of 4 nucleotides were assayed and probed after resolution of MMR intermediates and products on a denaturing agarose gel. Prior to HeLa treatment, the probe identified a ssDNA fragment of 4,069 nucleotides (lane 1); this represents the sequence from the AlwNI site to the discontinuity, including the mismatch site. After HeLa treatment, a full-length product appeared where the nick or gap is sealed (lanes 2 and 3, respectively). Based on restriction analysis performed in parallel, this larger band comprises both repaired and unrepairable molecules where the strand break has been sealed (18). This experimental approach was used to verify that the substrate nicks or gaps were sealed after treatment with either HeLa or LoVo nuclear extracts.

In contrast, when the same substrates were linearized 14 bp away from the discontinuity with Smal prior to the assay, evidence for an excision tract large enough to remove the mismatch is apparent for both the nicked and gapped substrates (Fig. 6, lanes 4 and 5). Such intermediates have been documented previously when repair synthesis has been inhibited with aphidicolin (18) and can extend several hundred nucleotides beyond the mismatch. These products appear as a smeared family of fragments degraded from the nick site (lane 4) or gap site (lane 5) to yield smaller products. When the substrate contains a nick, a substantial fraction of the molecules is apparently unprocessed, consistent with the proportion of molecules that is not repaired when the molecules are assayed by restriction analysis (refer to Fig. 2). When the 4-nu-
show a parallel experiment performed on the 3'-/H11032 used for linearization.
molecules, the sites present in the modified M13 template were unique restriction sites were incorporated into the template restriction sites in the gel-based assay. Because no additional note that the unrepaired molecules are 150 bp larger. We found that linearization close to the mismatch site had little effect on 5'-overhang) and NgoMIV (5'-overhang) that linearized the substrates 165 and 311 bases from the mismatch, respectively (Fig. 7, left panel). Identification of repaired and un repaired DNA products after HeLa treatment was accomplished by digestion with enzymes BspHI and HindIII (see “Materials and Methods”). In the case of the NgoMIV-linearized substrate, note that the un repaired molecules are 150 bp larger. We found that linearization close to the mismatch site had little effect on the rate of repair. Substrate digested at the remote, control site (AlwNI) or the NgoMIV site was repaired with equal efficiency. Substrate digested at the closest site (DrdI) was repaired with an efficiency similar to the undigested (circular) substrate. These results indicate that for the 5'-substrates, a linear end positioned as close as 165 bases away from the G-T mismatch is sufficient to allow the MMR pathway to act upon these substrates, including resynthesis through the mismatch site. We next determined the efficiency with which the 3'-nicked substrate was repaired when a linear end was placed either 136 bp (PvuI, 3'-overhang) or 165 bp (BglII, 3'-overhang) away from the mismatch (see Fig. 7, right panel). Shown in Fig. 7 are typical results of assays using PvuI (lane 4) and BglII (lane 5) linearized substrates possessing a 3'-nick after scoring digests with PacI and XhoI (see “Materials and Methods”). Control experiments in which the substrate was digested with AlwNI more than 3,000 bp away were repaired efficiently (37%). Substrate digested 165 bp away from the mismatch yielded detectable repair (25%), whereas the closest truncation did not yield a repair band that could be reliably scored. These data suggest that mismatch correction can be supported on a 3'-nicked template truncated as close as 165 bp away from the mismatch. Based on the results described in Fig. 6, we suggest that the early stages of mismatch correction, up to strand resynthesis, are likely to be accommodated on this substrate but that resynthesis is limited because of the excision of a required priming template.

DNA Ends Close to a Mismatch Do Not Deter Repair Directed from a Remote Nick—We designed an experiment to determine whether the MMR pathway was competent to access a remote nick as a strand discrimination signal in the presence of a proximal DNA end. Fig. 8 shows the efficiency of mismatch correction when a 5'-nick was positioned 884 bp away from a 3'-nick.

![Mismatch Repair-directing Structures](image)
G-T mismatch and linearized 165, 311, or 3,537 bp beyond the mismatch. Because a relatively small region of DNA separates the mismatch and one end in two of the assays, all substrates were linearized with BspHI after HeLa treatment to facilitate discrimination of repaired and unrepaired molecules by HindIII sensitivity (see the diagram in Fig. 8). Fragments that correspond to repaired and unrepaired substrate are designated R and U, respectively, in Fig. 8. Fragments not designated R or U were common to digests of both repaired and unrepaired molecules. When assayed, no substantial difference in competence among these three templates was observed. These data indicate that a linear end did not deter communication between the mismatch and nick site, even when the mismatch was 5-fold closer to the nearest free end than it was to the nick. Quantitatively similar results were obtained in assays conducted with a 5'-nicked substrate where the nick and mismatch were separated by 1,037 bp, although the overall repair efficiency was lower (data not shown).

**DISCUSSION**

At the outset of this work, a DNA nick was the only structure known to be competent to allow strand discrimination for MMR assays using human nuclear extracts in vitro. We wanted to determine whether other types of DNA discontinuities or ends were also competent. We therefore compared the efficiency with which a nick can direct the repair of a G-T mismatch placed ∼300 bp away with the efficiency of single strand gaps of 4, 28, and 202 (or 210) bp. In each case, the efficiency of mismatch correction was higher for the gapped species than for the nick. In contrast, neither blunt ends nor single strand overhangs of 4 nucleotides were competent to direct MMR to either strand.

Although we did not discriminate among models that might explain these repair differences, we offer two simple working models to account for them. One possibility is that nicks can be sealed by the action of a single enzymatic activity, a DNA ligase, whereas the gaps all require polymerase activity prior to strand sealing by ligase. Alternatively, the DNA ends available within the gaps represent sites where activities involved in DNA repair or replication may be loaded, which might involve prior loading of proliferating cell nuclear antigen at the 3'-end (38). In this model, gaps might facilitate recruiting repair machinery to a site near the mismatch. Such a model offers the testable prediction that the rate at which the strand signal, either nick or gap, is accessed might be fundamentally different between the nicked and gapped molecules. A critical kinetic analysis of key steps in the MMR pathway is essential to make such a distinction.

Our second objective was to determine whether the MMR pathway competes effectively with different classes of gap-filling repair events. In models that seek to explain strand discrimination based on the DNA terminus where elongation occurs (27, 32), mismatch correction must be able to compete with DNA synthesis performed by polymerases that function during either leading or lagging strand biosynthesis. Although we are unable to generate DNA mismatches by replication and ask whether the polymerase active site may be used as a strand discrimination signal, our assays test this hypothesis indirectly. By providing DNA gaps of different sizes, which presumably serve as templates for DNA polymerases with different tract length specificities, the MMR pathway was allowed to compete with replication for an available DNA end. In the case of 5'-discontinuities, the polymerase must copy the template toward the mismatch, and mismatch processing presumably competes with the late stages of replication and ligation for the DNA terminus. In the case of 3'-discontinuities, polymerase loading and synthesis must compete with the MMR pathway for access to the site of elongation.

The choice of the gap sizes was intended as a device to recruit different eukaryotic polymerases to compete with mismatch correction. For example, we expected that human polymerase β would be most likely to fill in the smallest gap of 4 nucleotides, whereas replicative polymerases such as δ or ε might fill in the largest gaps (42). The latter two polymerases have been associated with specific gap-filling repair events in nucleotide excision repair (39), so we also tested an intermediate gap size of 28 nucleotides. Polymerase α has not been associated with any specific repair events (39), but it also represents a polymerase capable of synthesis within the intermediate size gaps. Although it is beyond the scope of this work to characterize the extent to which each DNA polymerase contributed to the gap-filling events tested, this represents the rationale for the experimental design.

It is clear that gap sizes of 4 or 28 nucleotides, presented in either strand polarity, were the most effective in strand discrimination using the HeLa nuclear extract assay system. The largest gaps (≥202 nucleotides) were the least efficient strand discrimination signals. These data suggest that not all DNA ends are equally effective signals or that the largest gaps are accessed more efficiently by a replication apparatus than they are by MMR. It should be noted, however, that the substrates all presented an arbitrary distance of ∼300 bp separating discontinuity and mismatch. This situation may not reflect the optimal physical layout for communication between the sites of mismatch recognition and ongoing replication.

One hypothesis for the mechanism by which strands may be discriminated posits that MMR is directly coupled with genomic replication (3, 40). The intimate involvement of proliferating cell nuclear antigen in both early and late stages of the MMR process supports this notion (13, 14). It deserves mention that if the polymerase active site represents the primary strand discrimination information for mismatch correction, tract excision to remove the mismatch is expected to proceed primarily from 3' to 5'. This raises the question of why the human pathway, as judged using assays from nuclear extracts in vitro, is competent to access strand discontinuities efficiently when the excision proceeds from 5' to 3'. This stands in contrast to the better characterized E. coli model system, in which strand discrimination may be separated from replication, accessing transiently hemimethylated sites in either orientation to define the daughter strand. Both pathways are nonetheless bidirectional (18, 26), able to operate with similar efficiencies given a strand discontinuity in either orientation.

One often cited mechanistic possibility that preserves the requirement for a bidirectional competence of the human pathway is based on the fact that lagging strand biosynthesis is discontinuous and that nicks might be used as strand markers in this strand (4, 40). This possibility suggests that mismatch correction is targeted in a fundamentally different manner between the leading and lagging strands. The bidirectional capability of the human MMR pathway as well as the efficiency of gaps presented in either strand polarity as strand signals might be unified based on the observation that, in the lagging strand, the succeeding Okazaki fragment is synthesized toward the mismatch-containing fragment. Both a DNA gap and an approaching replication complex are available close to the mismatch, with a 5'→3' polarity reading toward the mismatch. Such a model might require that the 5'→3' exonuclease recruited to remove the mismatch, e.g. human exonuclease I (41), be able to excise both an RNA primer and the intervening DNA. At a minimum, it represents a situation in which a DNA gap could be accessed by a MMR pathway with a bidirectional capability.
One striking finding resulting from the deletion-based mapping strategy was that heteroduplex molecules containing a nick or 4-nucleotide gap, digested to leave a blunt end only 14 bp beyond the break (5’-orientation), were processed efficiently. Communication between the DNA mismatch site and the discontinuity must result in the loading of activities with a mandate to excise one DNA strand in the direction of the mismatch. Our data suggest that little more than one helical turn of DNA beyond the strand break is sufficient to direct mismatch excision in a strand-specific manner. This requirement held true for either strand polarity with respect to the mismatch. However, resynthesis to restore the mismatch site in these assays depended upon the presence of a DNA primer to initiate resynthesis, which was presumably removed in the case of the 5’-nicked species.

Finally, we infer from our data that ssDNA ends within a gap are at least as effective as a DNA nick in strand discrimination. By contrast, linear DNA ends with short overhangs, such as those generated by EcoRI digestion, do not contain enough information to allow targeting of repair to either strand. These observations might be important in contexts such as recombination, where the MMR pathway is known to participate (4, 5). Based on our results, a DNA end generated by strand invasion into a homologous duplex is predicted to be an effective means of targeting mismatch correction to the invading strand. This assumes that the mismatch and DNA end are close enough for effective communication and that mismatch excision leaves a priming sequence for replication. The ability of small gaps to direct MMR might also suggest that mismatch removal can be coupled to other DNA repair events, such as base excision repair or nucleotide excision repair processes that remove from 1 up to ~35 nucleotides to form a transient single strand gap (42). It is beyond the scope of this work to consider all of the possible implications, but the fact that MMR can be activated by small gaps may also be important within the context of lesion-induced apoptosis mediated by the MMR pathway (43). Such a scenario might arise when two lesions that can be removed by multiple repair pathways occur close to each other because processing of the first lesion may generate a strand signal appropriate to trigger a MMR-dependent processing event of the second lesion.

Acknowledgment—We recognize the excellent cell culture service provided by the National Cell Culture Center (Minneapolis).

REFERENCES