DNA Mismatch Correction in a Defined System

R. S. LAHUE, K. G. AU, P. MODRICH

DNA mismatch correction is a strand-specific process involving recognition of noncomplementary Watson-Crick nucleotide pairs and participation of widely separated DNA sites. The *Escherichia coli* methyl-directed reaction has been reconstituted in a purified system consisting of MutH, MutL, and MutS proteins, DNA helicase II, single-strand DNA binding protein, DNA polymerase III holoenzyme, exonuclease I, DNA ligase, along with ATP (adenosine triphosphate), and the four deoxynucleoside triphosphates. This set of proteins can process seven of the eight base-base mismatches in a strand-specific reaction that is directed by the state of methylation of a single d(GATC) sequence located 1 kilobase from the mispair.

ENZYME SYSTEMS CAPABLE OF RECOGNITION AND CORRECTION of base pairing errors within the DNA helix have been demonstrated in bacteria, fungi, and mammalian cells, but the mechanisms and functions of mismatch correction are best understood in the *Escherichia coli* system (*1–5*). In *E. coli*, the processing of mismatches within recombination intermediates is the basis of several marker effects associated with crossing over, while correction of mispairs resulting from DNA biosynthetic errors or the deamination of 5-methylcytosine in a m^5C·G base pair results in the elimination of lesions that would otherwise be fixed as mutations.

Of the several mismatch repair systems that have been identified in *E. coli* (*1–4, 6, 7*) the most interesting with respect to mechanism is the methyl-directed pathway for repair of DNA biosynthetic errors. This system processes base pairing errors within the helix in a strand-specific manner by exploiting patterns of DNA methylation (*8*). Since DNA methylation is a postsynthetic modification, newly synthesized strands temporarily exist in an unmethylated state (*9*), with the transient absence of adenine methylation on d(GATC) sequences directing mismatch correction to new DNA strands (*10–12*). In addition to its ability to respond to patterns of DNA methylation, this system is also capable of recognition of a number of different base pair mismatches (*13–15*), with in vitro analysis suggesting that of the eight possible base-base mispairs, only C·C may be refractory to repair (*16*). Since G·T, A·C, G·A, A·A, T·C, and T·T can assume intrahelical conformations (*3, 17*) and since several arguments suggest that it is this conformation that is recognized (*3, 18*), the enzymatic system responsible for correction must be capable of detecting subtle perturbations in helix structure associated with the presence of the different mispairs.

In order to address the biochemistry of methyl-directed mismatch correction, we have assayed the reaction in vitro with the use of the type of substrate illustrated in Fig. 1 (*11, 16*). Application of this method to cell-free extracts of *E. coli* (*11*) confirmed in vivo findings (*1–4*) that methyl-directed repair requires the products of four mutator genes, *mutH*, *mutL*, *mutS*, and *uvrD* (also called *mutU*), and also demonstrated a requirement for the *E. coli* single-strand DNA binding protein (SSB) (*19*). The dependence of in vitro correction on *mutH*, *mutL*, and *mutS* gene products has permitted isolation of these proteins in near homogeneous, biologically active forms. The 97-kilodalton MutS protein binds to mismatched DNA base pairs (*16, 20*); the 70-kD MutL protein binds to the MutS-heteroduplex complex (*21*); and the 25-kD MutH protein has a latent endonuclease that incises the unmethylated strand of a hemimethylated d(GATC) sequence and either strand of an unmethylated site (*3, 22*), with activation of this activity depending on interaction of MutS and MutL with a heteroduplex in the presence of adenosine triphosphate (ATP) (*5, 23*). However, these three Mut proteins together with SSB and the DNA helicase II product of the *uvrD* (*mutU*) gene (*24*) are not sufficient to mediate methyl-directed repair. We describe below identification of the remaining required components and reconstitution of the reaction in a defined system.

Protein and cofactor requirements for mismatch correction. Methyl-directed mismatch correction occurs by an excision repair reaction in which as much as several kilobases of the unmethylated DNA strand is excised and resynthesized (*8, 19, 25*). DNA polymerase I, an enzyme that functions in a number of DNA repair pathways, does not contribute in a major way to methyl-directed correction since extracts from a *polA* deletion strain exhibit normal levels of activity (*26*). However extracts derived from a *dnaZ*^{ts} strain are temperature-sensitive for methyl-directed repair in vitro (Table 1). The *dnaZ* gene encodes the τ and γ subunits of DNA polymerase III holoenzyme (*27, 28*), and mismatch correction activity is largely

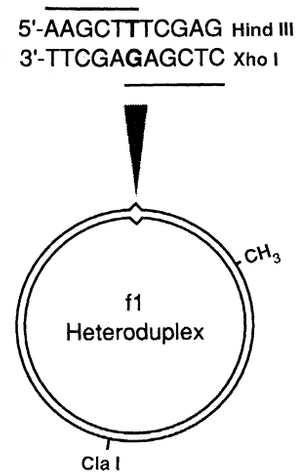


Fig. 1. Heteroduplex substrate for in vitro mismatch correction. Each substrate was a 6440-bp, covalently closed, circular heteroduplex derived from bacteriophage f1 and containing a single base-base mismatch located within overlapping recognition sites for two restriction endonucleases at position 5632 (*16*). In the example shown, a G·T mismatch resides within overlapping sequences recognized by Hind III and Xho I endonucleases. Although the presence of the mispair renders this site resistant to cleavage by either endonuclease, repair occurring on the complementary (C) DNA strand yields an A·T base pair and generates a Hind III–sensitive site, while correction on the viral (V) strand results in a G·C pair and Xho I sensitivity. The heteroduplexes also contain a single d(GATC) sequence 1024 bp from the mismatch (shorter path) at position 216. The state of strand methylation at this site can be controlled, thus permitting evaluation of the effect of DNA methylation on the strand specificity of correction.

The authors are in the Department of Biochemistry, Duke University Medical Center, Durham, NC 27710.

restored to heated extracts of the temperature-sensitive mutant strain by addition of purified polymerase III holoenzyme. A requirement for DNA polymerase III has also been indicated by recent experiments suggesting that the ε subunit of the holoenzyme is involved in *mutHLS*-dependent repair in vivo (*29*). Since DNA polymerase III holoenzyme is highly processive, incorporating thousands of nucleotides per DNA binding event (*30*), the involvement of this activity is consistent with the large repair tracts associated with the methyl-directed reaction.

We have obtained data indicating that purified MutH, MutL, and MutS proteins, DNA helicase II, SSB, and DNA polymerase III holoenzyme support methyl-directed mismatch correction, but this reaction is inhibited by DNA ligase (*31*), an enzyme that is shown below to be required to restore covalent continuity to the repaired strand. This observation led to isolation of a 55-kD stimulatory protein that obviates ligase inhibition. The molecular size and NH_2-terminal sequence of this protein indicated identity to exonuclease I (*32*), and homogeneous exonuclease I readily substitutes for the 55-kD stimulatory activity (Table 2). Thus, exonuclease I and the six activities mentioned above mediate efficient methyl-directed mismatch correction in the presence of ligase to yield product molecules in which both DNA strands are covalently continuous.

The requirements for repair of a covalently closed G·T heteroduplex (Fig. 1) are summarized in Table 3 (closed circular). No detectable repair was observed in the absence of MutH, MutL, or MutS proteins or in the absence of DNA polymerase III holoenzyme, and omission of SSB or exonuclease I reduced activity by 85 to 90 percent. These findings are in accord with previous conclusions concerning requirements of the methyl-directed reaction. However, in contrast to observations in vivo (*1–4*) and in crude extracts (*11*) indicating a requirement for the *uvrD* product, the reconstituted reaction proceeded readily in the absence of the added DNA helicase II (Table 2). Nevertheless, the reaction was abolished by antiserum to homogeneous helicase II, suggesting a requirement for this activity and that it might be present as a contaminant in one of the other proteins. Analysis of these preparations for their ability to restore mismatch repair to an extract derived from a *uvrD* (*mutU*) mutant and for the physical presence of helicase II by immunoblot assay revealed that the DNA polymerase III holoenzyme preparation contained sufficient helicase II (13 to 15 percent of total protein by weight) to account for the levels of mismatch correction observed in the defined system. Similar results were obtained with holoenzyme preparations obtained from two other laboratories. The purified system therefore requires all the proteins that have been previously implicated in methyl-directed repair.

The rate of correction of the closed circular heteroduplex was

Table 1. Requirement for τ and γ subunits of DNA polymerase III holoenzyme in mismatch repair. Extracts from strains AX727 (*lac thi* str^R *dnaZ*20-16) and AX729 (as AX727 except *purE dnaZ*$^+$) were prepared as described (*11*). Samples (110 μg of protein) were mixed with 0.8 μl of 1*M* KCl and water to yield a volume of 7.2 μl, and incubated at 42° or 34°C for 2.5 minutes. All heated (42°C) samples were then placed at 34°C and supplemented with 2.2 μl of a solution containing 0.1 μg (24 fmol) of hemimethylated G·T heteroduplex DNA, 16 ng of MutL protein, 50 ng of MutS protein, and buffer and nucleotide components of the mismatch correction assay (*11*). DNA polymerase III holoenzyme (57 ng in 0.6 μl) or enzyme buffer was then added, and incubation at 34°C was continued for 60 minutes. Heated extracts were supplemented with purified MutL and MutS proteins because these components are labile at 42°C. Activity measurements reflect the correction of heteroduplex sites. Heteroduplex correction activities shown are the result of two or more measurements. With the exception of the value shown as the range, variation between individual determinations was less than 15 percent.

Extract genotype	DNA pol III addition (ng)	Mismatch correction* (fmol^{-1} hour^{-1} mg^{-1})		Activity ratio (42°C/34°C)
		42°C	34°C	
dnaZ^{ts}		5 to 11	91	0.09
	57 ng	75	160	0.47
dnaZ$^+$		150	160	0.94
	57 ng	160	160	1.0

*Values obtained from extracts first incubated at either 42°C or 34°C, as indicated.

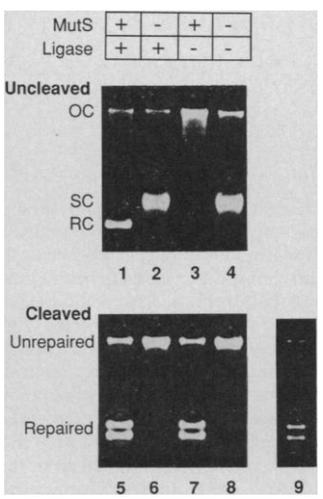


Fig. 2. Requirement for DNA ligase in mismatch correction. Hemimethylated G·T heteroduplex DNA [Fig. 1, 0.6 μg, d(GATC) methylation on the complementary DNA strand] was subjected to mismatch repair under reconstituted conditions in a 60-μl reaction (Table 3, closed circular heteroduplex), or in 20-μl reactions (0.2 μg of DNA) lacking MutS protein or ligase, or lacking both activities. A portion of each reaction (0.1 μg of DNA) was treated with EDTA (10 m*M* final concentration) and subjected to agarose gel electrophoresis in the presence of ethidium bromide (1.5 μg/ml; top panel, lanes 1 to 4). Positions are indicated for the unreacted, supercoiled substrate (SC), open circles containing a strand break (OC) and covalently closed, relaxed circular molecules (RC). A second sample of each reaction containing 0.1 μg of DNA was hydrolyzed with Xho I and Cla I endonucleases (Fig. 1) to score G·T → G·C mismatch correction and subjected to electrophoresis in parallel with the samples described above (bottom panel, lanes 5 to 8). The remainder of the complete reaction (0.4 μg of DNA, corresponding to the sample analyzed in lane 1) was made 10 m*M* in EDTA, and subjected to electrophoresis as described above. A gel slice containing closed circular, relaxed molecules was excised, and the DNA was eluted. This sample was cleaved with Xho I and Cla I, and the products were analyzed by electrophoresis (lane 9).

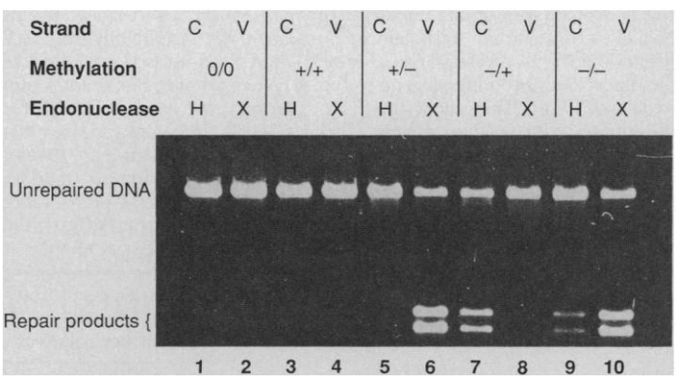


Fig. 3. Methyl-direction of mismatch correction in the purified system. Repair reactions with the G·T heteroduplex (Fig. 1) were performed as described in Table 3 (closed circular heteroduplex), except that reaction volumes were 20 μl (0.2 μg of DNA) and the incubation period was 60 minutes. The reactions were heated to 55°C for 10 minutes, and each was divided into two portions to test strand specificity of repair. The G·T → A·T mismatch correction, in which repair occurred on the complementary (C) DNA strand, was scored by cleavage with Hind III and Cla I endonucleases, while hydrolysis with Xho I and Cla I were used to detect G·T → G·C repair occurring on the viral (V) strand. Apart from the samples shown in the left two lanes, all heteroduplexes were identical except for the state of methylation of the single d(GATC) sequence at position 216 (Fig. 1). The state of modification of the two DNA strands at this site is indicated by + and – notation. The G·T heteroduplex used in the experiment shown in the left two lanes (designated 0/0) contains the sequence d(GATT) instead of d(GATC) at position 216 (*33*), but is otherwise identical in sequence to the other substrates.

unaffected by omission of DNA ligase (Table 3), but the presence of this activity results in production of a covalently closed product. Incubation of a hemimethylated, supercoiled G·T heteroduplex with all seven proteins required for correction in the presence of DNA ligase resulted in extensive formation of covalently closed, relaxed, circular molecules. Production of the relaxed DNA was dependent on MutS (Figs. 1 and 2) and MutL (*31*) proteins, and the generation of this species was associated with heteroduplex repair (Fig. 2). Correction also occurred in the absence of ligase, but in this case repair products were open circular molecules, the formation of which depended on the presence of MutS (Fig. 2). Since MutS has no known endonuclease activity but does recognize mispairs (*16, 20*), we infer that open circular molecules are the immediate product of a mismatch-provoked excision repair process. Ligase closure of the strand breaks present in this species would yield the covalently closed, relaxed circular product observed with the complete system.

The set of purified activities identified here as being important in methyl-directed repair support efficient correction. In the experiments summarized in Table 3, the individual proteins were used at the concentrations estimated to be present in the standard crude extract assay for correction (*11*) as calculated from known specific activity determinations. Under such conditions, the rate and extent of mismatch repair in the purified system are essentially identical to those observed in cell-free extracts (*31*).

DNA sites involved in repair by the purified system. The single d(GATC) sequence within the G·T heteroduplex shown in Fig. 1 is located 1024 base pairs from the mispair. Despite the distance separating these two sites, correction of the mismatch by the purified system responded to the state of modification of the d(GATC) sequence as well as its presence within the heteroduplex (Fig. 3). A substrate bearing d(GATC) methylation on both DNA strands did not support mismatch repair nor did a related heteroduplex in which the d(GATC) sequence was replaced by d(GATT). However, each of the two hemimethylated heteroduplexes were subject to strand-specific correction, with repair in each case being restricted to the unmodified DNA strand. With a heteroduplex in which neither strand was methylated, some molecules were correct-

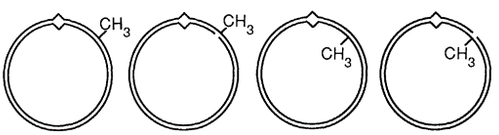
Table 2. Stimulation of in vitro methyl-directed correction by exonuclease I. Reactions (10 μl) contained 0.05*M* Hepes (potassium salt, *p*H 8.0), 0.02*M* KCl, 6 m*M* $MgCl_2$, bovine serum albumin (0.05 mg/ml), 1 m*M* dithiothreitol, 2 m*M* ATP, 100 μ*M* (each) dATP, dCTP, dGTP, and dTTP, 25 μ*M* β-NAD^+ (nicotinamide adenine dinucleotide), 0.1 μg of hemimethylated, covalently closed G·T heteroduplex DNA [Fig. 1, methylation on the **c** (complementary) strand, 24 fmol], 0.26 ng of MutH (*22*), 17 ng of MutL (*21*), 35 ng of MutS (*20*), 200 ng of SSB (*38*), 10 ng of DNA helicase II (*39*), 20 ng of *E. coli* DNA ligase (U.S. Biochemical), 95 ng of DNA polymerase III holoenzyme (*40*), and 1 ng of 55-kD protein (*31*) or exonuclease I as indicated. Reactions were incubated at 37°C for 20 minutes, quenched at 55°C for 10 minutes, chilled on ice, and then digested with Xho I or Hind III endonuclease to monitor correction. Repair of the G·T mismatch yielded only the G·C containing, Xho I–sensitive product. Except for the value shown as a range, results are the average of two or three measurements, with variation between individual determinants being less than 10 percent.

Protein added	Mismatch correction (fmol/20 min)
None	< 1 to 2
55-kD protein	18
Exonuclease I	18

Table 3. Protein and cofactor requirements for mismatch correction in a defined system. Reactions with covalently closed G·T heteroduplex (modification on **c** strand) were performed as described in the legend to Table 2 except that 1.8 ng of exonuclease I was used. Repair of open circular DNA was performed in a similar manner except that MutH, DNA ligase, and β-NAD^+ were omitted from all reactions, and the hemimethylated G·T heteroduplex (modification on **c** strand) had been incised with MutH protein as described in the legend to Fig. 4. When present, rabbit antiserum to helicase II or preimmune serum (5 μg of protein) was incubated at 0°C for 20 minutes with reaction mixtures lacking $MgCl_2$; the cofactor was then added and the assay was performed as above. Although not shown, antiserum inhibition was reversed by the subsequent addition of more helicase II. With the exception of the DNA polymerase III preparation, which contained about 15 percent (by weight) DNA helicase II (text), the purity of individual protein fractions was ≥95 percent. NT, not tested.

Reaction conditions	Mismatch correction (fmol/20 min) for circular heteroduplex	
	Closed	Open
Complete	15	17*
Minus MutH	<1	—
Minus MutL	<1	<1
Minus MutS	<1	<1
Minus DNA polymerase III holoenzyme	<1	<1
Minus SSB	2	1.4
Minus exonuclease I	2	<1
Minus DNA helicase II	16	15
Minus helicase II, plus immune serum	<1	<1
Minus helicase II, plus preimmune serum	14	NT
Minus ligase/NAD^+	14	
Minus $MgCl_2$	<1	NT
Minus ATP	<1	NT
Minus dNTP's	<1	NT

*No MutH, no ligase.



Reaction conditions	Repair (fmol/20 min)			
Complete	15 (<1)	17 (<1)	8 (<1)	10 (<1)
- Mut H	<1	18	1	9
- Mut L	<1	<1	<1	<1
- Mut S	<1	<1	<1	1
- SSB	2	<1	<1	<1
- pol III holoenzyme	<1	<1	<1	<1

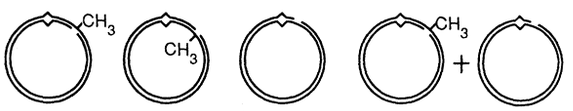
Fig. 4. Strand-specific repair of heteroduplexes containing a single-strand scission in the absence of MutH protein. Hemimethylated G·T heteroduplex DNA's (Fig. 1, 5 μg) bearing d(GATC) modification on the viral or complementary strand were subjected to site-specific cleavage with near homogeneous MutH protein (*22*). Because the MutH-associated endonuclease is extremely weak in the absence of other mismatch repair proteins (*23*), cleavage at d(GATC) sites by the purified protein requires a MutH concentration 80 times that used in reconstitution reactions. After removal of MutH by phenol extraction, DNA was ethanol-precipitated, collected by centrifugation, dried under vacuum, and resuspended in 10 m*M* tris-HCl (*p*H 7.6), 1 m*M* EDTA. Mismatch correction of MutH-incised and covalently closed, control heteroduplexes was performed as described in the legend to Table 2 except that ligase and NAD^+ were omitted. Outside and inside strands of the heteroduplexes depicted here correspond to complementary and viral strands respectively. Values in parentheses indicate repair occurring on the methylated, continuous DNA strand. The absence of MutH protein in preparations of incised heteroduplexes was confirmed in two ways. Preparations of incised molecules were subject to closure by DNA ligase (>80 percent) demonstrating that MutH protein does not remain tightly bound to incised d(GATC) sites. Further, control experiments in which each MutH-incised heteroduplex was mixed with a closed circular substrate showed that only the open circular form was repaired if MutH protein was omitted from the reaction, whereas both substrates were corrected if MutH protein was present.

ed on one strand, and some were corrected on the other. As can be seen, the hemimethylated heteroduplex bearing methylation on the complementary DNA strand was a better substrate than the alternative configuration in which modification was on the viral strand, with a similar preference for repair of the viral strand being evident with the substrate that was unmethylated on either strand. This set of responses of the purified system to the presence and state of modification of d(GATC) sites reproduce effects documented in vivo and in crude extract experiments (*10, 11, 13, 14, 16, 33*).

The efficiency of repair by the methyl-directed pathway depends not only on the nature of the mispair, but also on the sequence environment in which the mismatch is embedded (*1–3, 14*). To assess the mismatch specificity of the purified system under conditions where sequence effects are minimized, we have used a set of heteroduplexes in which the location and immediate sequence environment of each mispair are essentially identical (*16*). This analysis (Table 4) showed that the purified system is able to recognize and repair in a methyl-directed manner seven of the eight possible base-base mismatches, with C·C being the only mispair that was not subject to significant correction. Table 3 also shows that the seven corrected mismatches were not repaired with equal efficiency and that in the case of each heteroduplex, the hemimethylated configuration modified on the complementary DNA strand was a better substrate than the other configuration in which the methyl group was on the viral strand. These findings are in good agreement with patterns of repair observed with this set of heteroduplexes in *E. coli* extracts (*6, 16, 34*).

Strand-specific repair directed by a DNA strand break. Early experiments on methyl-directed repair in *E. coli* extracts led to the proposal that the strand specificity of the reaction resulted from endonucleolytic incision of an unmethylated DNA strand at a d(GATC) sequence (*3, 19*). This idea was supported by the finding that purified MutH protein has an associated, but extremely weak, d(GATC) endonuclease (*3, 22*) that is activated in a mismatch-dependent manner in a reaction requiring MutL, MutS, and ATP (*23*). Incision at d(GATC) sequences as the basis of strand specificity has also been proposed by Längle-Rouault *et al.* (*35*) to account for the finding that a strand break can direct strand-specific rectification of a mismatch upon transfection of a *mutH* mutant provided that the strain also carries a ligase mutation. However, the "nick"-directed repair observed under conditions of MutH and ligase deficiency was not shown to depend on other proteins required for the methyl-directed pathway. We have utilized the purified system to explore this effect more completely.

The two hemimethylated forms of the G·T heteroduplex shown in Fig. 1 were incised using high concentrations of purified MutH protein to cleave the unmethylated DNA strand at the d(GATC) sequence (↓ pGpApTpC) (*22*). After removal of the protein, these open circular heteroduplexes were tested as substrates for the purified system in the absence of DNA ligase. Both open circular species were corrected in a strand-specific manner and at rates similar to those for the corresponding covalently closed heteroduplexes (Fig. 4). As observed with closed circular heteroduplexes, repair of the MutH-cleaved molecules required MutL, MutS, SSB, DNA polymerase III holoenzyme, and DNA helicase II (Fig. 4 and open circle entries of Table 2), but in contrast to the behavior of the closed circular substrates, repair of the mismatch within the open circular molecules occurred readily in the absence of MutH protein. Thus prior incision of the unmethylated strand of a d(GATC) site can bypass the requirement for MutH protein in strand-specific mismatch correction.



Ligase	MutH	Repair (fmol/20 min)				
−	−	19 (<1)	9 (<1)	11 (<1)	19 (<1)	9 (<1)
+	−	2	<1	1	2	1
+	+	20	7	2	15	1

Fig. 5. Requirements for MutH protein and a d(GATC) sequence for correction in the presence of DNA ligase. Hemimethylated G·T heteroduplexes incised on the unmethylated strand at the d(GATC) sequence were prepared as described in the legend to Fig. 4. A G·T heteroduplex devoid of d(GATC) sites (Fig. 4) and containing a single-strand break within the complementary DNA strand at the Hinc II site (position 1) was constructed as described previously (*33*). Mismatch correction assays were performed as described in Table 3, with ligase (20 ng in the presence of 25 μM NAD^+) and MutH protein (0.26 ng) present as indicated. Table entries correspond to correction occurring on the incised DNA strand, with parenthetic values indicating the extent of repair on the continuous strand. Although not shown, repair of the nicked molecule lacking a d(GATC) sequence (first entry of column 3) was reduced more than an order of magnitude upon omission of MutL, MutS, SSB, or DNA polymerase III holoenzyme.

Table 4. Correction efficiencies for different mismatches. Correction of the eight possible base-base mispairs was tested with the set of covalently closed heteroduplexes described previously (*16*) including the G·T substrate shown in Fig. 1. With the exception of the mispair and the variations shown at the fifth position on either side, all heteroduplexes were identical in sequence. Each DNA was tested in both hemimethylated configurations under complete reaction conditions (Table 3, closed circular heteroduplex) except that samples were removed at 5-minute intervals over a 20-minute period in order to obtain initial rates (fmol/min). C and V, complementary and viral DNA strands; bias, relative efficiency of mismatch repair occurring on the two DNA strands (ratio of unmethylated to methylated) as determined 60 minutes after the reaction was started; NS, not significant. With the exception of the C·C heteroduplexes, repair in the absence of MutS protein was less than 20 percent (in most cases <10 percent) of that observed in its presence.

Heteroduplex	Markers	Methylation state			
		C^+V^-		C^-V^+	
		Rate	Bias	Rate	Bias
C 5′-CTCGA **G** AGCTT V 3′-GAGCT **T** TCGAA	Xho I Hind III	1.2	>18	0.38	>5
C 5′-CTCGA **G** AGCTG V 3′-GAGCT **G** TCGAC	Xho I Pvu II	1.1	>17	0.38	>6
C 5′-ATCGA **T** AGCTT V 3′-TAGCT **T** TCGAA	Cla I Hind III	1.0	>16	0.24	3
C 5′-ATCGA **A** AGCTT V 3′-TAGCT **A** TCGAA	Hind III Cla I	0.88	>20	0.20	>7
C 5′-CTCGA **A** AGCTT V 3′-GAGCT **C** TCGAA	Hind III Xho I	0.61	17	0.28	>5
C 5′-GTCGA **C** AGCTT V 3′-CAGCT **T** TCGAA	Sal I Hind III	0.60	12	0.23	>4
C 5′-GTCGA **A** AGCTT V 3′-CAGCT **G** TCGAA	Hind III Sal I	0.44	>13	0.21	5
C 5′-CTCGA **C** AGCTG V 3′-GAGCT **C** TCGAC	Pvu II Xho I	0.04	NS	<0.04	NS

The nature of the MutH-independent repair was examined further to assess the effect of ligase on the reaction and to determine whether a strand break at a sequence other than d(GATC) can direct correction in the absence of MutH protein (Fig. 5). As mentioned above, a covalently closed G·T heteroduplex that lacks a d(GATC) sequence is not subject to repair by the purified system in the presence (Fig. 3) or absence of DNA ligase (*31*). However, the presence of one strand-specific, site-specific break is sufficient to render this heteroduplex a substrate for the purified system in the absence of ligase and MutH protein (Fig. 5). Repair of this open circular heteroduplex was limited to the incised, complementary DNA strand and required presence of MutL and MutS proteins, DNA polymerase III, and SSB (*31*); correction of the molecule was as efficient as that observed with the hemimethylated heteroduplex that had been cleaved by MutH at the d(GATC) sequence within the complementary strand. Although the presence of a strand break is sufficient to permit strand-specific correction of a heteroduplex in the absence of MutH and ligase, the presence of the latter activity inhibited repair not only on the heteroduplex lacking a d(GATC) sequence but also on both hemimethylated molecules that had been previously incised with MutH protein (Fig. 5). This inhibition by ligase was circumvented by the presence of MutH protein, but only if the substrate contained a d(GATC) sequence, with this effect being demonstrable when both types of heteroduplex were present in the same reaction (Fig. 5, last column). This finding proves that MutH protein recognizes d(GATC) sites and is consistent with the view that the function of this protein in mismatch correction is the incision of the unmethylated strand at this sequence.

Implications for the mechanism of the methyl-directed pathway. The purified system described above supports the entire methyl-directed mismatch correction reaction and yields a covalently closed repair product. Although further analysis of this complex reaction is necessary, several features of methyl-directed repair can be inferred from the results summarized above and previous analyses of *mutH*, *mutL*, and *mutS* gene products. Provocation of a repair event is presumably initiated by binding of MutS protein to a mispair (*16, 20*). Interaction of MutL and MutH with this complex in the presence of ATP serves to activate the latent MutH-associated endonuclease which incises the unmethylated strand at a hemimethylated d(GATC) site (*3, 5, 21–23*). The incision induced by this activity serves to determine the strand specificity of correction as proposed earlier (*3, 19, 35*). The demonstration that a single-strand break can direct MutL- and MutS-dependent mismatch correction, but only under conditions of ligase deficiency, indicates an additional role for MutH in mismatch repair. Since a MutH incision is ligatable (*22*) (see also Fig. 5), this protein must (i) function in stabilizing the strand break it generates to ligation or (ii) facilitate entry of excision components, or both. Although the mechanism of excision and resynthesis steps have not been defined, it seems likely that function of the other required proteins (DNA helicase II, exonuclease I, SSB, DNA polymerase III holoenzyme, and ligase) are restricted to this phase of the reaction.

The demonstration that single-strand breaks can direct MutL- and MutS-dependent mismatch correction under conditions of ligase deficiency raises the possibility that this reaction may be of biological significance. For example, can nonligatable termini, like those present in gaps or at the ends of newly synthesized chains, promote strand-specific mismatch repair? In the case of DNA biosynthetic errors, Claverys and Lacks (*1*) have suggested that undermethylation of d(GATC) sequences within newly replicated DNA may determine the strand specificity of only a small fraction of repair events, with the majority being determined by DNA termini such as those present at the ends of newly synthesized strands. The nature of end-directed repair observed in the purified system would be consistent with this proposal, but biological data suggests that this reaction cannot be of major significance in the processing of biosynthetic errors. Since termini-directed correction requires MutL and MutS proteins, but not MutH (Fig. 4), this hypothesis predicts that the increased spontaneous mutability associated with *mutH* mutations would be much lower than that associated with mutations resulting in loss of *mutL* or *mutS* function. In fact, reported mutabilities of *mutH*, *mutL*, and *mutS* strains (*15, 36*) are of the same order and thus do not support this prediction. Another possibility is that termini-directed mismatch correction may be of significance in the processing of heteroduplex regions within recombination intermediates that contain exposed DNA ends. Although components of the methyl-directed pathway have been implicated in recombination-associated phenomena (*37*), the involvement of DNA termini in these effects has not been addressed.

REFERENCES AND NOTES

1. J.-P. Claverys and S. A. Lacks, *Microbiol. Rev.* **50**, 133 (1986).
2. M. Radman and R. Wagner, *Annu. Rev. Genet.* **20**, 523 (1986).
3. P. Modrich, *Annu. Rev. Biochem.* **56**, 435 (1987).
4. M. Meselson, in *Recombination of the Genetic Material*, K. B. Low, Ed. (Academic Press, San Diego, CA, 1988), pp. 91–113.
5. P. Modrich, *J. Biol. Chem.* **264**, 6597 (1989).
6. K. G. Au *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9163 (1988).
7. J. P. Radicella, E. A. Clark, M. S. Fox, *ibid.*, p. 9674.
8. R. Wagner and M. Meselson, *ibid.* **73**, 4135 (1976).
9. S. M. Lyons and P. F. Schendel, *J. Bacteriol.* **159**, 421 (1984).
10. P. J. Pukkila *et al.*, *Genetics* **104**, 571 (1983).
11. A.-L. Lu, S. Clark, P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4639 (1983).
12. R. Wagner *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 611 (1984).
13. C. Dohet, R. Wagner, M. Radman, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 503 (1985).
14. M. Jones, R. Wagner, M. Radman, *Genetics* **115**, 605 (1987).
15. R. M. Schaaper and R. L. Dunn, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6220 (1987).
16. S.-S. Su, R. S. Lahue, K. G. Au, P. Modrich, *J. Biol. Chem.* **263**, 6829 (1988).
17. F. H. Arnold, S. Wolk, P. Cruz, I. Tinoco, Jr., *Biochemistry* **26**, 4068 (1987).
18. G. V. Fazakerley *et al.*, *EMBO J.* **5**, 3697 (1986).
19. A.-L. Lu, K. Welsh, S. Clark, S.-S. Su, P. Modrich, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 589 (1984).
20. S.-S. Su and P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5057 (1986).
21. M. Grilley, K. M. Welsh, S.-S. Su, P. Modrich, *J. Biol. Chem.* **264**, 1000 (1989).
22. K. M. Welsh, A.-L. Lu, S. Clark, P. Modrich, *ibid.* **262**, 15624 (1987).
23. K. G. Au and P. Modrich, unpublished.
24. I. D. Hickson, H. M. Arthur, D. Bramhill, P. T. Emmerson, *Mol. Gen. Genet.* **190**, 265 (1983).
25. S.-S. Su, M. Grilley, R. Thresher, J. Griffith, P. Modrich, *Genome*, in press.
26. R. S. Lahue, S.-S. Su, K. Welsh, P. Modrich, *UCLA Symp. Mol. Cell. Biol. New Ser.* **47**, 125 (1987).
27. M. Kodaira, S. B. Biswas, A. Kornberg, *Mol. Gen. Genet.* **192**, 80 (1983).
28. D. A. Mullin, C. L. Woldringh, J. M. Henson, J. R. Walker, *ibid.*, p. 73.
29. R. M. Schaaper, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8126 (1988).
30. P. J. Fay, K. O. Johanson, C. S. McHenry, R. A. Bambara, *J. Biol. Chem.* **256**, 976 (1981).
31. R. S. Lahue and P. Modrich, unpublished.
32. G. J. Phillips and S. R. Kushner, *J. Biol. Chem.* **262**, 455 (1987).
33. R. S. Lahue, S.-S. Su, P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1482 (1987).
34. Although the patterns of substrate activity observed in extracts and in the purified system are qualitatively identical, the magnitude of variation observed differs for the two systems. Hemimethylated heteroduplexes modified on the complementary DNA strand are better substrates in both systems, but in extracts such molecules are repaired at about twice the rate of molecules methylated on the viral strand (*16*). In the purified system these relative rates differ by factors of 2 to 4. A similar effect may also exist with respect to mismatch preference within a given hemimethylated family. Although neither system repairs C·C, the rates of repair of other mismatches vary by factors of 1.5 to 2 in extracts (*16*) but by factors of 2 to 3 in the defined system.
35. F. Langle-Rouault, G. Maenhaut-Michel, M. Radman, *EMBO J.* **6**, 1121 (1987).
36. B. W. Glickman and M. Radman, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1063 (1980).
37. S. I. Feinstein and K. B. Low, *Genetics* **113**, 13 (1986).
38. T. M. Lohman, J. M. Green, R. S. Beyer, *Biochemistry* **25**, 21 (1986).
39. K. Kumura and M. Sekiguchi, *J. Biol. Chem.* **259**, 1560 (1984).
40. C. McHenry and A. Kornberg, *ibid.* **252**, 6478 (1977).

41. We thank A. Kornberg and C. McHenry for gifts of DNA polymerase III holoenzyme, J. Chase for a gift of SSB and an overproducing plasmid for preparation of the protein, and S. Kushner for near homogeneous exonuclease I; and C. McHenry for helpful advice in the purification of DNA polymerase III. Supported by NIH grant GM23719 and postdoctoral fellowships from the Damon Runyon–Walter Winchell Cancer Research Fund (DRG-877 to R.S.L.), NIH (F32 GM12684 to R.S.L.), and a Syntex Fellowship from the Life Sciences Research Foundation (K.G.A.).

15 February 1989; accepted 9 June 1989