The methyl-directed DNA repair efficiency of a set of M13mp18 heteroduplexes containing 1–8 or 22 unpaired bases was determined by using an in vitro DNA mismatch repair assay. The unpaired bases of each heteroduplex residing at overlapping recognition sites of two restriction endonucleases allow independent assay of repair on either DNA strand. Our results showed that the repair of small nucleotide heterologies in *Escherichia coli* extracts was very similar to base-base mismatch repair, being strand-specific and highly biased to the unmethylated strand. The *in vitro* activity was also dependent on products of *mutH*, *mutL*, *mutS*, and *uvrD* loci and was equally efficient on nucleotide insertions and deletions. The repair levels of small heterologies were affected by base composition of the heterologies. However, the extent of repair of heteroduplexes containing small heterologous sequences was found to decrease with an increase in the number of unpaired bases. Heteroduplexes containing an extra nucleotide of 22 bases provoked very low level of methyl-directed repair.

The *Escherichia coli* methyl-directed mismatch repair system monitors the fidelity of DNA replication and recombination in this organism (1). The methyl-directed reaction has been reconstituted in a purified system (2). The reaction is composed of hemimethylated DNA substrate containing a mispair, *E. coli* MutH, MutL, MutS, and DNA helicase II, along with single-stranded DNA-binding protein, DNA polymerase III holoenzyme, DNA ligase, and any one of the DNA exonucleases Exo I, Exo VII, or RecJ (2, 3). The major steps in the excision repair pathway for methyl-directed removal of mismatched DNAs have been well defined. The requisite strand specificity for processing of replication errors is provided by patterns of adenine methylation at d(GATC) sequences (2). Repair is initiated by binding of MutS to the mismatch (4), followed by the addition of MutL (5). Assembly of this complex leads to activation of a latent d(GATC) endonuclease of MutH protein, which incises the unmodified strand at a hemimethylated d(GATC) sequence (6). The resulting strand break, which can occur either 3' or 5' to the mismatch on the unmethylated strand, suffices to target correction to this strand (3). The ensuing excision reaction, which depends on MutS, MutL, and the cooperative action of DNA helicase II with an appropriate exonuclease, removes that portion of the unmodified strand spanning the d(GATC) site and the mismatch (7). Resynthesis of the excised strand by DNA polymerase III holoenzyme subsequently replaces the misincorporated nucleotide, with ligase restoring covalent integrity to the helix.

In *E. coli* mismatch correction, adenine methylation of d(GATC) sequences determines the strand on which repair occurs (2, 8). With hemimethylated heteroduplex, which is methylated at d(GATC) sequences on only one DNA strand, repair is highly biased to the unmethylated strand, with the methylated strand serving as template for correction. Mismatch repair also occurs on heteroduplex in which neither strand is methylated, but in this case correction shows little strand preference. Regions of DNA in which d(GATC) sequences are fully adenine-methylated are refractory to mismatch repair (2, 9). It appears to be the transient undermethylation of newly synthesized d(GATC) sequences in the region immediately following the replication fork that allows mismatch repair to operate only on newly synthesized strands and thereby remove replication errors (9, 10).

The *E. coli* mismatch repair system does not recognize and repair all base-base mismatches with equal efficiency. Generally, the transition G-T mispair is corrected most efficiently, with A-C, C-T, A-A, T-T, G-G, and A-G are repaired at different efficiency. C-C is refractory to the repair (2, 11).

The mutator effects observed in *E. coli* mutH, mutL, mutS, and mutU mismatch repair-deficient mutants, are primarily transition and frameshift mutations (12). The fact that mutants deficient in mismatch repair show increased frequencies of frameshift mutations suggests that the *E. coli* mismatch repair system can recognize and repair heteroduplexes with one or more unpaired bases. Transfections of *E. coli* with artificially constructed heteroduplexes and *in vitro* assays have demonstrated that the different heterologous are subject to correction with different efficiencies. The methyl-directed repair of heteroduplex with one-, two-, and three-base deletions is as efficient as the repair of G-T mismatches (13–15). Heteroduplexes with a four-base deletion are marginally repaired, and DNA with a five-base deletion is not detectably repaired by the MutHLS system (15, 16), but an alternative pathway such as recF-dependent activity may repair large deletions (17). The elements of heterologous structure that are recognized by mismatch binding proteins and features of the repair system that determine repair efficiency are not understood.

Mismatch repair genes are conserved in bacteria and higher organisms. Yeast MSH2 and human hMSH2 are analogous to the *E. coli* MutS protein, whereas yeast MLH1 and PMS1 and human hMLH1, hPMS1, and hPMS2 correspond to the bacterial MutL protein (reviewed in Ref. 18). A number of studies suggest a functional similarity between the postreplication repair pathway in prokaryotes and eukaryotes (reviewed in Ref. 1). Recently, Umar et al. (19) reported the correction of loops of five or more unpaired bases by human cell extracts. They
showed that the repair of 1–4-base loops was strongly dependent on hMLH1 and hMSH2 (19). However, in 5–16-base loops, repair was independent of hMLH1 (19). Human MSH2 had shown specific binding to loops as large as 14 nucleotides (20), although repair of loops larger than five bases has not been tested in the hMSH2-deficient extracts.

The E. coli methyl-directed repair of heteroduplexes containing more than a three-base insertion/deletion has not yet been examined in vitro. Thus, we used the in vitro repair assay and a set of M13mp18 derivatives containing 1–8 or 22 unpaired bases as substrates to determine the size constraints of the repair pathway. We demonstrate here that the methyl-directed mismatch repair pathway efficiently corrects heteroduplexes containing up to five unpaired bases. We also show that loop repair efficiency of heteroduplex is affected by base composition of the heterology.

EXPERIMENTAL PROCEDURES

Materials—E. coli strain NM522 (supE thy ΔlacIΔ215M15 Δlac-proAB) F′proAB lacF′ lacZΔM15) was from Kung King (Sinica Academia, Taipei). E. coli strains AB1157 (thr-1 leu-6 thi-1 lacY1 galK2 ara-14 xyl-5) and M15 (mutS201::kan) were subject to some cleavage by HpaI (Fig. 1) and the indicator restriction endonuclease, HpaII (Fig. 2) (21). They were prepared as described (21). E. coli DNA ligase, ATP-dependent DNase, dam methylase, thi methylase, T4 polynucleotide kinase, calf intestinal alkaline phosphatase, and restriction endonucleases were obtained from New England Biolabs.

Construction of M13 Mutants—A Synthetic 22-base pair oligonucleotide linker was inserted into the HindIII cleavage site of M13mp18 (Fig. 1). The product of this construction is dubbed M13LR1, which was further mutagenized with oligonucleotides partially complementary to the inserted linker (22) to create other derivatives. Each of the oligonucleotides used carried 1–8 additional bases, disrupting the recognition sequence of a unique restriction endonuclease site in vector and creating a new, unique restriction endonuclease recognition sequence (see Table I). Mutant M13LR phages were identified by restriction analysis of replication form minipreparation, and mutant sequences were confirmed by digestion sequencing (23).

Heteroduplex DNA—Replicating form DNAs fully methylated at dGATC sites were isolated from strain NM522 or JJ119. DNAs devoid of dam methylation at dGATC sites were prepared as described (21). E. coli DNA ligase, ATP-dependent DNase, dam methylase, thi methylase, T4 polynucleotide kinase, calf intestinal alkaline phosphatase, and restriction endonucleases were obtained from New England Biolabs or Amersham Corp.

Construction of M13 Mutants—A Synthetic 22-base pair oligonucleotide linker was inserted into the HindIII cleavage site of M13mp18 (Fig. 1). The product of this construction is dubbed M13LR1, which was further mutagenized with oligonucleotides partially complementary to the inserted linker (22) to create other derivatives. Each of the oligonucleotides used carried 1–8 additional bases, disrupting the recognition sequence of a unique restriction endonuclease site in vector and creating a new, unique restriction endonuclease recognition sequence (see Table I). Mutant M13LR phages were identified by restriction analysis of replication form minipreparation, and mutant sequences were confirmed by digestion sequencing (23).

Methyl-directed Loop Repair in vitro

The map of the M13mp18 shows restriction sites relevant to this study and the synthetic linker sequence used to construct the insertion derivative M13LR1. The DNA of this phage contains six dGATC sites recognized by dam methylase (positions 1382, 1714, 6253, 6406, 6502, and 6836). The methylation state of these sites can be tested by DpnII digestion.

DpnII. Only 25% of the heteroduplexes were resistant to DpnII; 25% were cleaved once, and 50% were cleaved more than once. Fully unmethylated heteroduplexes (V−C−) were formed by annealing unmethylated complementary strands to unmethylated viral strands. Fully methylated heteroduplexes (V+C+) were prepared by in vitro methylation of hemimethylated heteroduplexes with dam methylase as described (9, 15). In V−C− heteroduplexes, the complementary strands were fully methylated at more than 95% of the dGATC sequences, and the viral strands were methylated at about 80%.

Repair Assays—Growth of cells and preparation of cell extracts were as described by Lu et al. (9). The repair reaction in concentrated E. coli lysate was carried out in 10 μl containing 0.8 μM Tris-HCl (pH 7.6); 5 mM MgCl₂; 50 μg/ml bovine serum albumin; 1 mM ATP; 0.1 mM concentration each of dATP, dGTP, dTTP, and dCTP; and 0.1 μg (20 fmol) of methylated heteroduplex DNA. Cell-free E. coli extracts were included at optimal amounts at 7.5–10 μg of protein/ml. Incubation was at 37 °C for 1 h. Then 30 μl of 25 mM EDTA (pH 8.0) was added, and the DNA was purified by phenol extraction and ethanol precipitation. The DNA was then analyzed by restriction endonuclease digestion and agarose gel electrophoresis. The ethidium complexes of DNA products were quantitated using a gel documentation CCD camera (UVI Ltd.) (25).

RESULTS

Construction of Heteroduplexes in Which the Small DNA Heterologies Reside in a Similar Sequence Environment—Starting from phage M13mp18, we have prepared a set of M13LR derivatives that contain extra nucleotides within the polylinker region located between the single EcoRI and HindIII sites of M13mp18 (Fig. 1 and Table I). This set of M13LR derivatives permits construction of heteroduplexes representing base pair mismatches and 1–8- and 22-base insertion/deletion heterologies. In each of these heteroduplexes the heterology is located in a similar environment. Moreover, as shown in Table II, each insertion/deletion mispair is located within overlapping restriction endonuclease recognition sites.

This approach allows an independent assay of correction on either DNA strand. Digestion of the heteroduplex DNA with AlwNI (Fig. 1) and the indicator restriction endonuclease, whose recognition site is inactive because of the heterology, will
yield a 7.2-kilobase pair fragment only. Similar digestion of
DNA in which the recognition sequence has been restored by
repair reaction will yield 4.1- and 3.1-kilobase pair fragments.
In the case of the heteroduplexes constructed, almost all sub-
strates were refractory to the digestion by the indicator restric-
tion endonucleases (Fig. 2, B and D), except C2cc and C3
were susceptible to the star activity by one of the indicator restric-
tion endonuclease HindIII. In these cases, 20–40% of
untreated C2cc and C3 heteroduplexes were cleaved by HindIII.
Thus, the repair on the complementary strand of these two
substrates was not determined (Table III).

Methyl-directed Repair of Small DNA Heterologies by
MutHLS System in Vitro—To determine whether base inser-
tions/deletions were corrected via the methyl-directed mis-
mismatch repair pathway in vitro, hemimethylated heterodu-
plexes containing 1–8, or 22-base insertion/deletion
heterologies were tested for repair in a cell extract from mis-
mismatch repair-proficient E. coli strain. A G-T mismatch and an
A-C mismatch at position 6303 of M13LR1 were used as posi-
tive controls. We choose to compare the efficiency of correction
of DNA containing nucleotide heterologies with that of G-T and
A-C mismatches, since a methyl-directed repair system has
been shown to correct them with high efficiency (11). Table III
compares the efficiency of correction of all heteroduplexes as
scored by restriction endonuclease digestion, and Fig. 2 illus-
trates the behavior of the several repair classes in the restric-

| M13 mutant | Parent phage | Mutagenic oligonucleotide (5′ → 3′) | Marker |
|------------|--------------|-------------------------------------|--------|
| M13LR1     | M13mp18      | AGCTTCGAGGAGGCTGCTGCTGCT           | XhoI   |
| M13LR3     | M13LR1       | AGCTTCGAGGAGGCTGCTGCT              | XhoI   |
| M13LR5     | M13LR1       | AGCTTCGAGGAGGCTGCTGCT              | HindIII|
| M13LR6     | M13LR3       | AGCTTCGAGGAGGCTGCTGCT              | XhoI   |
| M13LR7     | M13LR3       | AGCTTCGAGGAGGCTGCTGCT              | XhoI   |
| M13LR8     | M13LR3       | AGCTTCGAGGAGGCTGCTGCT              | XhoI   |
| M13LR9     | M13LR5       | AGCTTCGAGGAGGCTGCTGCT              | EcoRV  |
| M13LR10    | M13LR5       | AGCTTCGAGGAGGCTGCTGCT              | NcoI   |
| M13LR12    | M13LR5       | AGCTTCGAGGAGGCTGCTGCT              | XhoI   |

TABLE II

M13LR heteroduplex substrates

Covalently closed, circular heteroduplexes containing a set of base pair mismatches or small nucleotide insertions/deletions were prepared using the phage DNAs shown. Heteroduplexes with unpaired bases are depicted by C or V, the complementary or viral strand where extra bases are located, followed by the number of unpaired bases. The aa, gg, tt, cc, ac, or gt of C2 or V2 substrates indicates the base composition of the unpaired bases. The n of 4- or 5-base insertion/deletion substrates indicates the marker restriction enzyme NcoI, c, complementary strand; v, viral strand. Relevant restriction endonuclease sequences are underlined (see Table I). The base-base mismatches and unmatched bases are shown in boldface type and hyphens, respectively.

| DNA Heteroduplex site (loop on complementary strand) | M13 DNA Heteroduplex site (loop on viral strand) |
|-----------------------------------------------------|-------------------------------------------------|

yield a 7.2-kilobase pair fragment only. Similar digestion of
DNA in which the recognition sequence has been restored by
repair reaction will yield 4.1- and 3.1-kilobase pair fragments.
In the case of the heteroduplexes constructed, almost all sub-
strates were refractory to the digestion by the indicator restric-
tion endonucleases (Fig. 2, B and D), except C2cc and C3
were susceptible to the star activity by one of the indicator restric-
tion endonuclease HindIII. In these cases, 20–40% of
untreated C2cc and C3 heteroduplexes were cleaved by HindIII.
Thus, the repair on the complementary strand of these two
substrates was not determined (Table III).

Methyl-directed Repair of Small DNA Heterologies by
MutHLS System in Vitro—To determine whether base inser-
tions/deletions were corrected via the methyl-directed mis-
mismatch repair pathway in vitro, hemimethylated heterodu-
plexes containing 1–8, or 22-base insertion/deletion
heterologies were tested for repair in a cell extract from mis-
mismatch repair-proficient E. coli strain. A G-T mismatch and an
A-C mismatch at position 6303 of M13LR1 were used as posi-
tive controls. We choose to compare the efficiency of correction
of DNA containing nucleotide heterologies with that of G-T and
A-C mismatches, since a methyl-directed repair system has
been shown to correct them with high efficiency (11). Table III
compares the efficiency of correction of all heteroduplexes as
scored by restriction endonuclease digestion, and Fig. 2 illus-
trates the behavior of the several repair classes in the restric-

TABLE I

Construction of M13 derivatives for heteroduplex preparation

Bacteriophage M13LR1, a derivative of M13mp18 was constructed by insertion into the HindIII site at position 6283 of the 22-base pair synthetic
duplex.

V 5′-AGCTAGCAGCAGCAGCCTCGAG

TGGTCTGCTGCGAGACCTGCTGAG-5′ C

With M13LR1 as a precursor, a set of additional M13LR phages has been constructed for this study. The underlined nucleotide in the column of
mutagenic oligonucleotides indicates the base change for each step. An asterisk indicates phage M13LR6 was mutagenized by the same
oligonucleotide as in M13LR7; however, a single C to A* base change was found.
Repair was determined in extracts of NM522 and RK1517 (7.5 mg/ml protein) as described under "Experimental Procedures." C+, repair occurring on the methylated complementary strand; V−, repair occurring on the unmethylated viral strand. Repair results are averages of several experiments. Complete repair would correspond to 20 fmol. ND, not determined.

| Substrates | Wild type extracts | mutS:Tn5 extracts |
|------------|--------------------|------------------|
|            | V−     | C+     | V−     | C+     |
| G-T        | 13.1   | <0.5   | 0.8    | <0.5   |
| A-C        | 12.7   | 1.3    | 2.3    | 0.6    |
| C1t        | 11.0   | 1.0    | 1.1    | 2.1    |
| Vlg        | 10.7   | 1.9    | 1.9    | 0.9    |
| C2cc       | 12.6   | ND     | 0.7    | ND     |
| V2gg       | 13.5   | <0.5   | 1.7    | <0.5   |
| C2cc       | 9.4    | 1.6    | 1.2    | 2.0    |
| V2gt       | 9.0    | 1.1    | 0.5    | 1.7    |
| C2tt       | 11.3   | 0.7    | 0.6    | 0.8    |
| V2aa       | 8.8    | 0.5    | 1.2    | 0.8    |
| C3         | 7.5    | ND     | <0.5   | ND     |
| V3         | 8.5    | 1.2    | 0.5    | <0.5   |
| C4         | 7.8    | <0.5   | 0.5    | 0.5    |
| V4         | 7.2    | 0.6    | 0.7    | 0.8    |
| C4n        | 9.4    | 1.8    | 0.5    | 0.8    |
| V4n        | 4.7    | <0.5   | 0.6    | <0.5   |
| C5         | 4.9    | 1.3    | <0.5   | 0.7    |
| V5         | 7.7    | 0.6    | 0.7    | 0.7    |
| C5n        | 8.7    | 2.8    | 0.6    | <0.5   |
| V5n        | 6.7    | 0.5    | <0.5   | <0.5   |
| C6         | 3.7    | 1.4    | 0.5    | <0.5   |
| V6         | 2.2    | <0.5   | 0.5    | 0.7    |
| C7         | 1.6    | <0.5   | <0.5   | <0.5   |
| V7         | 6.6    | <0.5   | <0.5   | <0.5   |
| C8         | 2.0    | 0.5    | <0.5   | <0.5   |
| V8         | 1.1    | 1.8    | <0.5   | 1.6    |
| C22        | 1.5    | 0.9    | 0.5    | 0.8    |
| V22        | 1.9    | 1.8    | 0.7    | 1.1    |
Methyl-directed Small Loop Repair in Vitro

FIG. 3. Dependence of methyl-directed small heterologies on the loop size and mutS gene product. Data are from Table III. Symbols at point 0 of unmatched extra bases represent base-base mismatches. Circles, wild-type extract reactions; triangles, mutS extract reactions. Open symbols, heteroduplexes with insertion on unmethylated viral strand (V-substrates); closed symbols, heteroduplexes with deletion on unmethylated viral strand (C-substrates).

TABLE IV
In vitro complementation of mutator extracts in mismatch correction, small heterology repair, and large heterology reaction

Heteroduplex repair was determined as described under “Experimental Procedures,” and the reaction (10 μl) contained 75 μg of indicated cell extracts and 20 fmol of circular heteroduplex DNA with the indicated heterology and methylation state. Activity is expressed as repair occurring on either the viral strand (V) or the complementary strand (C). The state of dam methylation of viral and complementary strands is indicated by + and −, respectively.

| Source of extract | Repair V/C | Repair V/C | Repair V/C |
|-------------------|------------|------------|------------|
|                   | G-T        | V2gg       | C8         | V22        |
| Wild type (NM522) | 218<7      | 173<7      | 40/9       | 22/11      |
| mutH              | 16<7       | 22<7       | 11/11      | 9/15       |
| mutL              | 32<7       | 30<7       | 7/<7       | 3/<7       |
| mutS              | 8<7        | 12<7       | 7/10       | 11/17      |
| uvrD              | 22<7       | 24<7       | 8/<7       | 7/19       |
| mutH + mutL       | 249<7      | 253<7      | 26<7       | 16/11      |
| mutH + mutS       | 254<7      | 253<7      | 39/10      | 28/13      |
| mutH + uvrD       | 211<7      | 198<7      | 22/7       | 19/11      |
| mutL + mutS       | 249<7      | 247<7      | 35/<7      | 21/10      |
| mutL + uvrD       | 191<7      | 178<7      | 16/<7      | 15/7       |
| mutS + uvrD       | 224<7      | 230<7      | 27/<7      | 22/14      |

match and surrounding sequence, we compared the repair of different dinucleotide insertion and deletion heteroduplexes with various sequence contexts (Table III). The only difference, other than identities of the insertion/deletion is the absence of an A-T base pair adjacent to the bulky heterologies of C2ac and V2gt. Nevertheless, these heteroduplexes were subject to differential repair in E. coli extracts. Analysis of individual heteroduplexes in Table III suggested that V2gg and C2ac are better substrates (with repair from 12.6 to 13.5 fmol/nl) than V2gt and C2ca (with repair from 9.0 to 9.4 fmol/nl). Given the highly homologous nature of this set of substrates, their differential repair implies that heterology recognition is associated with the correction events as observed in base pair mismatch repair.

Low Level in Vitro Correction of Large DNA Heterologies Is MutHLS- and Methylation State-dependent—Although functions derived from mutH, mutL, mutS, and uvrD are deficient in heteroduplex correction in the in vitro system, activity is recovered upon mixing of extracts. Results of this type of analysis are shown in Table IV. With well repaired G-T mispair and V2gg dinucleotide insertion heteroduplex, complementation was observed with all possible pairs of extracts and in many cases resulted in a level of repair comparable with that observed with extracts derived from wild-type cells. As in the case of wild type reactions, the repair in mixed extracts was also dependent on the state of DNA methylation, repair being highly biased to unmethylated strands (Table IV). Interestingly, while the poorly repaired C8 deletion and V22 insertion heteroduplexes were subject to this type of analysis, a low level of complementation was also observed. Methyl-directed reactions in combined mutant extracts were 2–4-fold higher than reactions in individual extracts, while reaction levels on the methylated strand remained unchanged (Table IV).

With increasing length of heterology, we have occasionally observed low levels of reaction on heteroduplexes that are independent of MutS and methylation state (Fig. 2C, lanes 7 and 9; Table III, V5, V6, V8, C22, and V22). Repair in these cases may be mediated by alternative pathway such as the methyl-independent pathway described by Fishel et al. (26). To ascertain that part of large loop repair was truly methyl-directed, we tested a subset of heteroduplex substrates in all four possible methylation states (V−/C+, V+/C+, V+/C−, V−/C−). The combined extracts of mutH471::Tn5 and mutS201::Tn5 were used, since the extracts showed higher repair activity (Table IV). Results of this type of analysis are shown in Table V.

Repair of either hemimethylated configuration in extracts was directed to the unmethylated strand with different biases. For V−/C− heteroduplexes, the biases were >20:1 for V2gg, 15:1 for V5, and 2:5:1 for C22. In the case of V+/C− heteroduplexes, repair was also biased to the unmethylated strand but with reduced values of 2.8:1 for V2gg, 1.3:1 for V5, and 2:1 for C22. This may be due in part to the fact that methylation on viral strands of V+/C− heteroduplexes was incomplete (see “Experimental Procedures”).

The Dam-dependent repair pathway loses its ability to discriminate the wild-type strand from the deletion strand when DNA is devoid of adenine methylation in d(GATC) sequences. Table V shows the results of reactions with unmethylated heteroduplexes. The results of V−/C− heteroduplexes repair were consistent with the loss of strand bias, since reaction gave rise to approximately equal levels of repair on both strands.

Since methylated d(GATC) sequences are resistant to nicking by MutH endonuclease, correction of mismatched bases and insertion and deletion heterologies should be greatly reduced. The results of V+/C+ heteroduplex in Table V are consistent with this prediction. The repair levels on methylated complementary strands were reduced to background levels. However, the reduced repair levels on methylated viral strands were 2–5-fold higher than background levels. The basis for this difference may be incomplete methylation on viral strands.
of V+C+ heteroduplex substrates (see “Experimental Procedures”).

DISCUSSION

Genetic studies provided initial evidence supporting the correction of small insertion and deletion heterologies by E. coli mismatch repair. Dohet et al. (15) demonstrated that one-base insertions and deletions are efficiently repaired in vivo (13). Learn and Graffstrom, using an in vitro system, have demonstrated that one-, two-, and three-base deletions can be repaired as efficiently as G-T mismatches in a mutHLS-dependent, methyl-directed process (14). Parker and Marinus (15) and Carraway and Marinus (16) using in vivo assay found that heterologies larger than four bases were not recognized and processed by Dam-dependent mismatch repair. Mismatch repair systems in other cell types have also been implicated in the repair of small insertion/deletion mutations. The Hex system of Streptococcus pneumoniae efficiently corrects one- and two-base heterologies (27), while mismatch repair in Saccharomyces cerevisiae corrects one-base deletion in vivo (28, 29) and corrects four- and seven-base heterologies in vitro (30). The human strand break-directed mismatch repair system also repairs one-, two-, and three- and four-base insertions/deletions efficiently in vitro (32).

The results present in this report confirm the earlier findings in in vivo and in vitro studies that the methyl-directed DNA repair pathway in repair of 1–3-base insertions is as efficient as in vivo. Our results showing that insertions/deletions up to four or five bases were subject to methyl-directed repair at levels of 35–70% of that observed for G-T mispair contradict an earlier report showing that heterologies larger than four bases were not repaired (15). To reconcile the differences between their results and ours, we propose that the heteroduplex sequences within overlapping restriction endonuclease recognition sites adopt configurations suitable for repair, whereas the insertions of previous studies do not. Previous studies used only a single set of four- and five-base deletion heteroduplexes (15). We tested several four- or five-base insertion/deletion heteroduplexes, and the repair levels showed as much as a 2-fold difference (Table III). Thus, the different conclusions of the studies could be explained by sample size differences.

As shown in Fig. 3 and Table III, as the base number of the insertion/deletion increased, generally the repair levels decreased. However, this tendency is not tightly followed by all heterologies; some exceptions do exist. The hemimethylated seven-base insertion heteroduplex was significantly repaired by the methyl-directed mismatch repair pathway (see V7 in Table III). It had been suggested that larger heterology might induce a secondary structure within the loop that recognized by MutS (15). The other possibility is that the realignment of the loop in the duplex may produce transient base pair mismatch. This transient mismatch may be recognized by MutS, and it provoke repair reaction. This is consistent with the finding that low level repair of large loop is MutHLS-dependent and methyl-directed (Tables IV and V). Thus, the MutHLS system may process large heterologies through partial mismatch recognition as well as a co-repair mechanism (16).

With increasing length of heterology, our system detected a low level of Dam-independent reactions on heteroduplexes (Table III). This activity may be mediated by alternative pathway such as methyl-independent RecF pathway (26). The efficiency of correction of these substrates is so near background levels that we have difficulty in determining the exact contributions of Dam-dependent and Dam-independent activities in these large loop reactions.

The MutS protein has been shown to bind DNA at the site of mispaired bases (4, 11). Since crystallographic structure of MutS bind mismatched DNAs is not available at present, the elements of heterologous structure that the repair system recognizes and that subsequently determine repair efficiency are not understood. Since base-base mismatches such as T-G, G-G, C-A, A-A, and A-G can assume intrahelical conformations and since several studies suggest that it is this conformation that is recognized (33–35), the enzymatic system responsible for correction must be capable of detecting subtle perturbations in helix structure associated with the presence of the different mispairs. Heteroduplexes with large unpaired nucleotide sequences may form extrahelical loops. Such loop structure may not be readily recognized by MutS. Pursuit of this line of reasoning would imply that a small insertion/deletion accommodated into the double-helix structure can be recognized by MutS. It remains to be seen whether or not this is the case. If it is, then this could be the basis of specificity for the repair of small heterologies. It also can be expected that neighboring sequences would affect the distortion caused by an insertion/deletion mismatch. Griffith and colleagues (36, 37) had shown kinking of heteroduplexes with looped mispair, presumably due to intercalation of extra nucleotides into duplex. Rosen et al. (38, 39) had shown duplex containing an extra 1–3 unpaired bases adopt intrahelical conformations, stacking within the duplex. Heteroduplexes can easily accommodate the bulge loop with little structural perturbation beyond the immediate vicinity of the Duplex itself (38, 39). Therefore, it is conceivable that partially stacked bases would still be accessible to bases of other strand to form mismatched base pairs.

Mutations of the E. coli mutS and mutL genes as well as their analogs in yeast, mammalian, and human cell lines result in several hundred-fold enhancement in the frequency of mutations. Our demonstration that the bacterial mismatch repair system can repair insertion or deletion heterologies with different efficiencies suggests that eukaryotic mismatch repair proteins, which are highly homologous to the E. coli enzymes (18), may play an analogous role. Several endometrial and colorectal carcinoma cell lines that are defective in mismatch repair show instability of simple DNA microsatellites (32, 40). Since most of the genetic instability found in human diseases is at microsatellite tandem repeats, it is very interesting to know the repair efficiency for small heterologies within tandem repeats. The in vitro assay described here can be modified to test the repair of slippage in simple repeat heterology. The present study provides a basis for further investigation.

Acknowledgments—We thank Dr. Paul Modrich for strains and critical evaluation of the manuscript and Dr. Klim King for strains and continuing encouragement during the course of this work.

REFERENCES

1. Modrich, P. (1991) Annu. Rev. Genet. 25, 229–253
2. Lahue, R. S., Au, K. G., and Modrich, P. (1989) Science 245, 160–164
3. Cooper, D. L., Lahue, R. S., and Modrich, P. (1993) J. Biol. Chem. 268, 11823–11829
4. Lu, A. L., Clark, S., and Modrich, P. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 3395–3397
5. Grilley, M., Welsh, K. M., Su, S.-S., and Modrich, P. (1989) J. Biol. Chem. 264, 1099–1104
6. Au, K. G., Welsh, K., and Modrich, P. (1992) J. Biol. Chem. 267, 12142–12148
7. Grilley, M., Griffith, J., and Modrich, P. (1993) J. Biol. Chem. 268, 11830–11837
8. Lahue, R. S., Su, S.-S., and Modrich, P. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 1482–1486
9. Lu, A. L., Clark, S., and Modrich, P. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 4659–4663
10. Pukkila, J. P., Peterson, J., Herman, G., Modrich, P., and Meselson, M. (1983) Genetics 104, 571–582
11. Su, S. S., Lahue, R. S., Au, K. G., and Modrich, P. (1988) J. Biol. Chem. 263, 6829–6835
12. Claverys, J. P., Lacks, S. A. (1986) Microbiol. Rev. 50, 133–165
13. Dohet, C., Wagner, R., and Radman, M. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 3385–3397
22720  

Methyl-directed Small Loop Repair in Vitro

14. Learn, B. A., and Graffstrom, R. H. (1989) J. Bacteriol. 171, 6473–6481
15. Parker, B. O., and Marinus, M. G. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1730–1734
16. Carraway, M., and Marinus, M. G. (1993) J. Bacteriol. 175, 3972–3980
17. Fishel, R. A., Siegel, E. C., and Kolodner, R. (1986) J. Mol. Biol. 188, 147–157
18. Modrich, P., and Lahue, R. S. (1996) Annu. Rev. Biochem 65, 101–133
19. Umar, A., Boyer, J. C., and Kunkel, T. A. (1994) Science 266, 814–816
20. Fishel, R., Ewel, A., Lee, S., Lescoe, M. K., and Griffith, J. (1994) Science 266, 1403–1405
21. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103–119
22. Kramer, W., Drutsa, V., Jansen, H. W., Kramer, B., Pflugfelder, M., and Fritz, H. J. (1984) Nucleic Acids Res. 12, 9441–9456
23. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
24. Holmes, J., Jr., Clark, S., and Modrich, P. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 5837–5841
25. Fang, W., and Modrich, P. (1993) J. Biol. Chem. 268, 11838–11844
26. Fishel, R., and Kolodner, R. (1989) J. Bacteriol. 171, 3046–3052
27. Gasc, A. M., Sicard, A. M., and Claverys, J. P. (1989) Genetics 121, 29–36
28. Bishop, D. K., Andersen, J., and Kolodner, R. D. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 3713–3717
29. Kramer, B., Kramer, W., and Fritz, H. J. (1984) Cell 38, 879–887
30. Muster-Nassal, C., and Kolodner, R. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 7618–7622
31. Thomas, D. C., Roberts, J. D., and Kunkel, T. A. (1991) J. Biol. Chem. 266, 3744–3751
32. Parsons, R., Li, G. M., Longley, M. J., Fang, W., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K. W., Vogelstein, B., and Modrich, P. (1993) Cell 75, 1227–1236
33. Werntges, H., Steger, G., Riesner, D., and Fritz, H. J. (1986) Nucleic Acids Res. 14, 3773–3790
34. Kenward, O. (1985) J. Biomol. Struct. & Dyn. 3, 205–226
35. Hunter, W. N., Brown, T., Anand, N. N., and Kenward, O. (1986) Nature 320, 552–555
36. Wang, Y. H., and Griffith, J. (1991) Biochemistry 30, 1358–1363
37. Hsieh, C. H., and Griffith, J. D. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4833–4837
38. Rosen, M. A., Shapiro, L., and Patel, D. J. (1992) Biochemistry 31, 4015–4026
39. Rosen, M. A., Live, D., and Patel, D. J. (1992) Biochemistry 31, 4004–4014
40. Umar, A., Boyer, J. C., Thomas, D. C., Nguyen, D. C., Risinger, J. I., Boyd, J., Ionov, Y., Perucho, M., and Kunkel, T. A. (1994) J. Biol. Chem. 269, 14367–14370