Requirement for d(GATC) sequences in *Escherichia coli* *mutHLS* mismatch correction

(DNA repair)

ROBERT S. LAHUE, SHIN-SAN SU, AND PAUL MODRICH

Department of Biochemistry, Duke University Medical Center, Durham, NC 27710

Communicated by I. Robert Lehman, November 3, 1986

ABSTRACT The involvement of d(GATC) sequences in *Escherichia coli* DNA mismatch correction was ascertained by analyzing *in vitro* repair efficiencies of a series of related, covalently closed circular DNA heteroduplexes that contained from zero to four d(GATC) sites. A heteroduplex with four d(GATC) sites was repaired with high efficiency by extracts of *E. coli*, whereas no significant correction occurred on a closely related molecule lacking such sequences. Heteroduplexes containing one or two d(GATC) sites were corrected at rates between 10% and 93% of that observed for the four-site molecule, but repair efficiency did not correlate in a simple way with the number of sites present. The methylation state at a single d(GATC) sequence was sufficient to direct strandedness of repair, and correction of heteroduplexes containing one or more d(GATC) sites required functional *mutH*, *mutL*, and *mutS* gene products. In addition, DNA repair synthesis dependent on *mutH* and *mutS* also required the presence of at least one d(GATC) site. Although mismatch correction was not observed on a covalently closed circular heteroduplex lacking a d(GATC) sequence, such molecules were subject to strand-specific repair if they contained a strand-specific single-strand break. However, this correction reaction did not require *mutH*, *mutL*, *mutS*, or *uvrD* gene products. Consequently, we have concluded that d(GATC) sequences are directly involved in mismatch correction mediated by the *mutHLS* system.

Wagner and Meselson (1) postulated in 1976 that mismatch correction directed to a newly synthesized DNA strand might function to eliminate replication errors. They also suggested that DNA methylation could provide the basis for the strand discrimination in such a system. Since polynucleotide methylation occurs postsynthetically, newly synthesized DNA is expected to be transiently undermethylated relative to the parental template strands.

The occurrence of methyl-directed mismatch correction in *Escherichia coli* has been demonstrated by transfection experiments using artificially constructed heteroduplex DNAs. Such experiments showed that the state of strand methylation at d(GATC) sequences can determine the pattern of mismatch repair (2-7). Correction of hemimethylated heteroduplexes has been found to be highly biased to the unmethylated strand, whereas unmethylated molecules are repaired with little strand bias, and fully methylated heteroduplexes are repaired with reduced efficiency.

Repair of hemimethylated or unmethylated heteroduplexes in *E. coli* requires function of *mutH*, *mutL*, *mutS*, and *uvrD* loci (3, 5, 8, 9), mutations in any of which confer high spontaneous mutability. Furthermore, genetic studies have implicated *mutH*, *mutL*, and *mutS* loci in a pathway involving function of the *dam* gene, which encodes the DNA methylase that modifies d(GATC) sequences (10, 11). Protein products

of *mutH*, *mutL*, and *mutS* genes have recently been isolated in near-pure, biologically active forms, and the *mutS* protein has been found to recognize at least some mismatched base pairs (refs. 12 and 13; K. Welsh, A.-L. Lu, and P.M., unpublished data).

Although it is clear that methylation at d(GATC) sites can modulate mismatch correction, it has not been shown that d(GATC) sequences are directly involved in the correction process. In the study described in this paper, we exploited a cell-free system that supports methyl-directed mismatch correction *in vitro* (6, 14) to address this question. The results show that d(GATC) sequences are required for the reaction and that the state of methylation of a single d(GATC) sequence can govern strand bias of mismatch correction 1000 base pairs distant.

MATERIALS AND METHODS

***E. coli* Strains.** W3110 was provided by C. Richardson (Harvard University), JJ119 (*metJam185*, HfrH) was from R. Greene of this department, and D6432 [*argE*, *metB*, *Nal*^R, *Rif*^R, Δ (*lac-pro*), *sulI*⁺, *mutL::Tn10*] was from M. Meselson (Harvard University). Strains KMBL 3752 (*endA101*, *thyA306*, *lysA65*, *argA103*, *bio-87*, *metE72*, *pheA97*, *purA*, *aroB*, *cysC*), KMBL 3773 (as 3752 but *mutH101*, *thyA*⁺, *lysA*⁺), and KMBL 3775 (as 3752 but *mutS101*, *cysC*⁺) were from B. Glickman (York University).

DNA Preparations and Mismatch Repair Assays. Replicative form (RF) and single-stranded DNAs of bacteriophage ϕ 1 were isolated as described (6) except that RFI DNA was isolated from strain JJ119 grown in medium supplemented with methionine (15 μ g/ml). RFI preparations were efficiently modified at d(GATC) sequences in this host. In some experiments, fully methylated viral single-stranded DNA was prepared in JJ119 harboring pTP166, which carries the *dam* gene under *tac* control (15). In these cases, host cells were induced with 1 mM isopropyl β -D-thiogalactopyranoside prior to ϕ 1 infection. Unless specified otherwise, heteroduplex substrates were prepared in a covalently closed, negatively supercoiled form, as described previously (6). A heteroduplex hemimethylated at the single *Cla* I restriction site was prepared in a similar manner except that the RFI DNA used in heteroduplex preparation was modified *in vitro* with *Cla* I methylase (New England Biolabs). A variant heteroduplex containing a site-specific single-strand break in the complementary strand at the unique *HincII* site was generated by omission of the ligation step after heteroduplex formation (6).

Recombinant manipulations were by standard methods (16), and oligonucleotide mutagenesis was according to Kramer *et al.* (17). Mutants generated in this manner were identified by plaque hybridization (18) using the appropriate oligonucleotide as a probe, and mutant sequences were confirmed by the dideoxy sequencing procedure (19).

Mismatch correction in cell lysates was assayed as described (6).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

FIG. 1. Construction of hybrid phage containing from zero to four d(GATC) sites. The circular substrates utilized in this study are shown, for purposes of comparison, in linear form starting from the unique *HincII* site. fl sequences are represented by a thin line, sequences from M13mp10 as a solid thick line, and those from M13 DNA as an open thick line. d(GATC) sites are designated 1 through 4 from left to right in the orientation shown. Also shown are sequence changes that eliminate these sites. As described in the text, the numerical designation is used in subscript notation to indicate the presence of particular d(GATC) sequences in mismatch heteroduplexes. The asterisk indicates that d(GATC) sequence 4 in fl_{1-4} is also a *Bam*HI site. All heteroduplexes contained a G-T mismatch at position 5621 within the *Eco*RI site:

Viral GAATTT (*Eco*RI-resistant)
Complementary CTTAAG (*Eco*RI-sensitive)

Those bearing d(GATC) sites were methylated on the complementary strand and unmethylated on the viral strand. The substrates shown were constructed in the order of presentation. $fl_{h_{2,3}}$ was prepared in two steps. In the first, the 2-kilobase (kb) *Ban*I fragment from M13 was ligated with the 4.4-kb *Ban*I fragment from M13mp10 (21, 22). In the second step, a 4.8-kb DNA fragment from the intermediate construction was isolated by cleavage with *HincII* and *Hgi*AI and ligated to a 1.6-kb fragment from $flR229$ DNA (containing the single *Eco*RI site) prepared by treatment with *HincII* and partial reaction with *Hgi*AI. The product of this construction is $fl_{h_{2,3}}$, which was mutagenized with the oligonucleotide d(GTAACGGTCTAAAG) to create fl_{h_2} . Mutagenesis of fl_{h_2} with d(TGCGGGGTCGTCAC) yielded fl_{h_0} . fl_{h_1} was prepared by exchange of a 357-base-pair *HincII*-*Xmn*I fragment spanning d(GATC) site 1 between $flR229$ and fl_{h_0} . An *Eco*RI-resistant form of each construction was prepared by exchanging the 1.6-kb *HincII*-*Hgi*AI DNA fragment with the corresponding fragment from phage $flM28$, a derivative of $flR229$ that contains a C→T transition within the *Eco*RI site at position 5621 (6). Heteroduplex DNA was prepared from *Eco*RI-sensitive and *Eco*RI-resistant derivatives of each phage as described (6).

RESULTS

Construction of DNAs Containing Different Numbers of d(GATC) Sites. Previous experiments (6, 14) have utilized heteroduplexes constructed from phage $flR229$ DNA (20), which contains four d(GATC) sites at positions 216, 1382, 1714, and 2221. Fig. 1 illustrates the structure of this molecule, as well as those of four derivatives constructed for this study, which contain two, one, or no d(GATC) sequences.* These DNAs are designated using subscript notation to indicate

*As a precursor of M13mp18, phage M13mp10 was expected to contain a d(GATC) sequence at position 2221 (22). However, our sequence analysis of M13mp10 indicates the sequence d(GATT) at this position. In addition, we have been unable to detect a site for *Sau*3A at this position, indicating the absence of a d(GATC) at position 2221.

FIG. 2. Kinetic analysis of mismatch correction as a function of d(GATC) site density. DNA substrates utilized are those shown in Fig. 1, each containing a G-T mispair within the single *Eco*RI site. Reactivity of each substrate alone (A) or in competition with fl_{1-4} (B) was measured by the method of Lu *et al.* (6). For A, extract from strain W3110 was present at 28 mg/ml and the DNA concentration was 10 μ g/ml. In B, the extract concentration was 20 mg/ml and each DNA was present at 7.5 μ g/ml. Samples were removed as indicated and reactions were quenched with EDTA. After phenol extraction and ethanol precipitation, DNA samples were hydrolyzed with *Eco*RI, *HincII*, and *Bam*HI endonucleases and analyzed by electrophoresis in 0.8% agarose gels. Since only fl_{1-4} contains a *Bam*HI site (Fig. 1), correction of the other substrates in competition with fl_{1-4} can be monitored by using this set of three enzymes. These endonucleases cleave the mismatch-corrected form of fl_{1-4} to yield 3.4-, 0.8-, and 2.2-kb products, while uncorrected substrate yields 4.2- and 2.2-kb products. Correction products generated from the other substrates yield 5.6- and 0.8-kb fragments reflecting cleavage by *Eco*RI and *HincII*, while uncorrected DNA yields a 6.4-kb linear species due to cleavage at the *HincII* site. Error bars in B indicate the range of activity observed on fl_{1-4} in the four competition experiments.

particular d(GATC) sites present. For example, $flR229$ will be referred to as fl_{1-4} , indicating the presence of all four d(GATC) sequences, and the hybrid derivative containing only the sites at 1382 and 1714 is called $fl_{h_{2,3}}$.

Although the hybrid viral molecules are composed of DNA from phage fl , M13, and M13mp10, the number of nucleotide differences is limited. In addition to the single nucleotide change in the d(GATC) site at position 219, there are a total of 40 nucleotide differences between fl_{1-4} and $fl_{h_{2,3}}$ localized to the region between coordinates 875 and 4681 (Fig. 1 and refs. 21 and 23). However, each of the remaining constructs shown in Fig. 1 differs from its precursor by only a single nucleotide change within a d(GATC) site.

Efficiency of Mismatch Correction Depends on d(GATC) Sequences. Each of the phage DNAs shown in Fig. 1 contains an *Eco*RI sequence at coordinates 5616–5621. For purposes of mismatch heteroduplex preparation, a derivative of each

Table 1. Relative repair efficiencies of DNA substrates

Substrate	Relative rate of repair, %	
	No fl ₁₋₄ DNA	With fl ₁₋₄ DNA
fl ₁₋₄	100	100
flh _{2,3}	18	35
flh ₂	10	28
flh ₀	<5	6
flh ₁	93	79

Results shown are based on initial rate measurements from the experiments shown in Fig. 2 and represent the average of several determinations. Values of 100% correspond to 120 fmol per hr per mg of protein for repair efficiencies measured on individual heteroduplexes and 94 ± 14 fmol per hr per mg of protein for rate measurements performed in competition with fl₁₋₄. All heteroduplexes containing one or more d(GATC) sites were hemimethylated with modification on the complementary strand.

phage was constructed that contained a G-C→A-T transition mutation at position 5621 within the recognition sequence for this enzyme (6). *EcoRI*-sensitive and *EcoRI*-resistant forms of each phage were used to construct a set of heteroduplexes containing a variable number of d(GATC) sites, each member of which contained a single mismatch (G-T) at position 5621.

As shown in Fig. 2A and summarized in Table 1, members of this set were corrected *in vitro* with different efficiencies. Hemimethylated heteroduplexes of fl₁₋₄ and flh₁ were efficiently corrected, hemimethylated substrates derived from flh_{2,3} and flh₂ were corrected at reduced rates, and the flh₀ heteroduplex lacking a d(GATC) sequence was not subject to detectable repair (<5% repair in 60 min). Although flh₁ and flh₀ heteroduplexes differ only by a single base pair within d(GATC) site 1, their correction efficiencies differ markedly.

Variation of repair efficiency within this set of heteroduplexes was not due to the presence of inhibitors in substrate molecules, as demonstrated by competition experiments in which flh₁, flh₂, flh_{2,3}, and flh₀ were each incubated with fl₁₋₄ (Fig. 2B and Table 1). These experiments yielded a pattern of repair efficiencies qualitatively similar to that obtained for the individual heteroduplexes, with reactivities of the molecules being fl₁₋₄ ≥ flh₁ > flh_{2,3} > flh₂ > flh₀. In this set of experiments the reactivity of flh₀ was only slightly above background levels. Correction in each case where the heteroduplex contained one or more d(GATC) sites was *mut*-dependent. Correction was not observed in extracts derived from strains deficient in *mutH*, *mutL*, or *mutS*, but mixtures of these extracts did complement one another (Table 2).

Repair-Associated DNA Synthesis Is Dependent on the Presence of d(GATC) Sites. Previous work showed that methyl-directed mismatch correction of hemimethylated fl₁₋₄ is accompanied by repair synthesis on the unmethylated DNA strand (6, 14). Occurrence of this synthesis was found to depend on the presence of a mismatch, to be suppressed

FIG. 3. DNA synthesis associated with mismatch correction. DNA synthesis was measured according to Lu *et al.* (14), except that dTTP was present at 100 μ M. The wild-type extract was from strain KMBL 3752, *mutH* extract was from KMBL 3773, and *mutS* extract was from KMBL 3775. Filled circles, open circles, and triangles represent data obtained when wild-type extract was used with fl₁₋₄, flh₁, and flh₀, respectively. Squares are data obtained using fl₁₋₄, flh₁, or flh₀ with *mutH* or *mutS* extracts, and error bars indicate the range of results observed with *mut*⁻ extracts. All substrates contained the G-T mispair described in the legend to Fig. 1, and substrates containing one or more d(GATC) sites were methylated on the complementary strand.

in symmetrically methylated heteroduplexes, and to require products of *mutH*, *mutL*, and *mutS* genes.

To determine whether *mut*-dependent repair synthesis occurs on a mismatch heteroduplex lacking a d(GATC) sequence, we compared repair synthesis on the flh₀ heteroduplex with that on hemimethylated fl₁₋₄ and flh₁ heteroduplexes in extracts derived from *mut*⁺, *mutH*, and *mutS* strains. Fig. 3 shows that the majority of DNA synthesis on both fl₁₋₄ and flh₁ is dependent on *mutH* and *mutS* function. In addition, *mut*-dependent synthesis on flh₁ is comparable to that on fl₁₋₄ and is in good agreement with their similar repair efficiencies (Table 1). In contrast, *mut*-dependent DNA synthesis was not observed with the flh₀ heteroduplex.

A Single d(GATC) Site Is Sufficient to Direct Mismatch Correction in a 6-Kilobase Heteroduplex. The G-T mismatch employed in the experiments described above and in previous studies (6, 14) is limited in the sense that correction on one DNA strand can be monitored by *EcoRI* cleavage (legend to Fig. 1) but correction on the other strand is not directly detectable. This problem can be circumvented by the use of a heteroduplex in which the two nucleotides comprising the mismatch each reside in the recognition site for a different type II endonuclease (24). We have prepared a G-T mismatch of this form in the flh₁ and flh₀ sequence backgrounds (Tables 3 and 4). In these substrates (flhGT₁ and flhGT₁₀, respectively),[†] repair on the complementary strand of the G-T heteroduplex renders the DNA sensitive to *HindIII*, whereas correction on the viral strand results in sensitivity to *Xho I* cleavage.

The minimum distance between the G-T mismatch and the single d(GATC) site within flhGT₁ is 1024 base pairs. Nevertheless, the state of methylation of this sequence

Table 2. *mut*-dependence of mismatch repair of fl₁₋₄ and flh₁

Extract(s)	Repair, fmol per hr per mg of protein	
	fl ₁₋₄	flh ₁
<i>mut</i> ⁺	35	34
<i>mut</i> ⁻	<4	<4
<i>mutH</i> + <i>mutS</i>	25	23
<i>mutH</i> + <i>mutL</i>	19	17
<i>mutS</i> + <i>mutL</i>	32	31

Extracts: *mut*⁺, KMBL 3752; *mutH*, KMBL 3773; *mutS*, KMBL 3775; *mutL*, D6432. No detectable repair was observed with individual *mutH*, *mutL*, or *mutS* extracts, and these results are summarized under the single entry *mut*⁻.

[†]In this designation "flh" refers to the hybrid phage from which the DNA used to make the heteroduplex was derived, "GT1" designates a particular G-T mismatch with the structure shown in the legend to Table 3, and the subscript denotes the location of d(GATC) sites.

Table 3. Strand bias of mismatch correction in a substrate with one d(GATC) site

Methylation state (v/c)*	Repair, fmol per hr per mg of protein					
	<i>mut</i> ⁺		<i>mutH</i>		<i>mutS</i>	
	v	c	v	c	v	c
+/+	<5	5	<7	8	<4	5
-/+	39	4	<7	7	<4	4
+/-	<5	35	<7	10	<4	6
-/-	14	5	<7	7	<4	4

Repair assays were performed as described (6), using extracts derived from KMBL 3752 (*mut*⁺), KMBL 3773 (*mutH*), or KMBL 3775 (*mutS*), which were present at 17–27 mg/ml. Correction was analyzed by restriction analysis with the enzymes *Hind*III and *Xho* I. The substrate flhGT₁ was derived from flh₁ and contained a G–T mispair, the correction of which could be monitored on either strand. It was prepared by first eliminating the *Cla* I site at position 6040 by mutagenesis with the oligonucleotide d(GAATCAATGAACGG). A *Hind*III site was introduced into flh₁ by insertion of the synthetic duplex

5' AATTGCTAGCAAGCTTTTCGAGTCTAGA
CGATCGTTTCAAAGCTCAGATCTTTAA 5'

into the *Eco*RI site at position 5616. Plaque hybridization and sequence determination were used to identify a phage containing a single insert with the upper sequence present in the viral strand. This phage was mutagenized with d(ACTCGAGAGCTTGC) to introduce an A→G point mutation within the *Hind*III site, to generate an *Xho* I sequence. A G–T mismatch heteroduplex was then prepared from viral DNA containing the *Hind*III sequence and complementary strands containing the *Xho* I site:

Viral strand 5' AAGCTTTTCGAG (*Hind*III-sensitive)
Complementary strand TTCGAGAGCTC (*Xho* I-sensitive)

Methylated viral DNAs used in heteroduplex construction were prepared from a JJ119(pIP166) host (see *Materials and Methods*). *v, Viral strand; c, complementary strand.

modulates correction of the mismatch (Table 3). Mismatch repair of either hemimethylated configuration in extracts derived from the *mut*⁺ strain KMBL 3752 was directed to the unmethylated strand with a bias of about 8:1. No detectable repair was observed in extracts from isogenic *mutH* or *mutS* strains (Table 3), but repair with methyl-directed bias was restored upon mixing of mutant extracts (data not shown).

As shown in Table 3, symmetric methylation of the d(GATC) site of flhGT₁ reduced the level of repair in the *mut*⁺ extract to near background levels. The unmethylated form of this heteroduplex was also corrected, but at a somewhat lower level than the hemimethylated forms, with repair occurring predominantly on the viral strand. Extracts of other *mut*⁺ strains (W3110 and AB1157; data not shown) yielded results similar to those obtained with KMBL 3752 (Table 3), with two exceptions. First, correction of both hemimethylated and unmethylated DNA by such extracts was more efficient than with KMBL 3752. Second, correction of the unmethylated heteroduplex by W3110 and AB1157 extracts partitioned approximately equally between the two DNA strands, with the sum of repair on the two strands being approximately equal to the level of correction observed on the unmethylated strand of the hemimethylated forms. The basis for this strain difference is unknown. We have used KMBL 3752 for most of our work due to the availability of an isogenic set of *mut* strains (see *Materials and Methods*).

In contrast to the hemimethylated or unmethylated flhGT₁ heteroduplexes, the analogous heteroduplex lacking a d(GATC) site (flhGT₀) was not subject to detectable repair on either DNA strand (Table 4), results identical to those obtained with the substrate series shown in Table 1. In addition, no detectable repair was observed with an flhGT₁ heteroduplex hemimethylated at the single *Cla* I site at

Table 4. Mismatch correction on nicked and covalently closed substrates

	Repair, fmol per hr per mg of protein					
	flhGT ₁		flhGT ₀		Nicked flhGT ₀	
	v	c	v	c	v	c
<i>mut</i> ⁺	32	<4	<4	<4	<4	12
<i>mutH</i>	<7	<7	<7	<7	<7	18
<i>mutL</i>	<5	<5	<5	<5	<5	16
<i>mutS</i>	<4	<4	<4	<4	<4	12
<i>uvrD</i>	<7	<7	<7	<7	<7	21

The flhGT₁ heteroduplex contained a G–T mismatch within overlapping *Hind*III and *Xho* I sites and was methylated at the single d(GATC) site on the complementary strand. It is identical to the substrate described for Table 3. flhGT₀ and nicked flhGT₀ heteroduplexes were identical to the flhGT₁ substrate except for a single base change within d(GATC) site 1 (Fig. 1) and the presence of a single-strand break at the *Hinc*II site (position 1; *Materials and Methods*) in the complementary strand of nicked flhGT₀. Assays were performed as described in the legend to Table 3. Activity is expressed as repair occurring on either the viral strand (v) or the complementary strand (c).

position 2527 [modification of the 3'-penultimate adenine of 5' d(ATCGAT) 3'; data not shown]. Thus, the failure to repair heteroduplexes lacking a d(GATC) site cannot be attributed to a deficiency of *N*⁶-methyladenine in a hemimethylated configuration.

DNA Termini Can Direct Mismatch Correction by a *mut*-Independent Process. DNA heteroduplexes used in the studies described above were all covalently closed duplex DNA circles, and the findings presented indicate that d(GATC) sequences have a major role in the *mutHLS*-dependent correction on such molecules. Based on the inference that DNA ends direct mismatch correction in *Streptococcus pneumoniae*, Lacks and colleagues (25, 26) have suggested that strand discrimination based on DNA methylation is of only minor importance in *E. coli*, with the major pathway of strand targeting in this organism also being dependent on the presence of single-strand breaks. Indeed, experiments with mammalian cells have led Hare and Taylor (24) to suggest that mismatch correction in higher cells can be directed either by the presence of a strand-specific nick or by the strand methylation pattern.

We have explored the possible significance of strand breaks in directing *E. coli* mismatch correction by use of an open circular variant of the flhGT₀ heteroduplex described above. This molecule lacked a d(GATC) sequence but contained a site-specific strand break within the *Hinc*II recognition sequence of the complementary strand (see *Materials and Methods*). As shown in Table 4, this nicked heteroduplex was preferentially repaired on the strand containing the break, whereas the corresponding covalently closed form was not subject to detectable correction. However, in contrast to correction of the hemimethylated form of flhGT₁ heteroduplex, repair occurring on the nicked flhGT₀ molecule was not dependent on *mutH*, *mutL*, *mutS*, or *uvrD* gene products, proteins that are required for postsynthetic elimination of biosynthetic errors (Table 4). The nature of repair occurring on the nicked heteroduplex is unknown, although it cannot represent nick-translation by DNA polymerase I, since the reaction was also observed in extracts prepared from a *polA* deletion strain (data not shown; ref. 27).

DISCUSSION

In vitro correction of heteroduplexes containing from zero to four d(GATC) sites varied over a range of at least 20-fold, in the order fl₁₋₄ ≥ fl₁ > fl_{2,3} > fl₂ > fl₀, where the

zero-site molecule was almost inert as a substrate. Inspection of this hierarchy of reactivity indicates that repair efficiency is a complex function of the d(GATC) sites present. Reactivity decreases with the number of d(GATC) sites for the substrate series flh_{2,3}, flh₂, and flh₀, suggesting that the density of such sequences may affect the correction process. However, the high reactivity of the flh₁ heteroduplex is an exception to this observation. Thus, if density effects are important in determination of reactivity, the behavior of flh₁ indicates that other factors, such as the sequence environment of a d(GATC) site or its placement relative to a mismatch, can also contribute in a major way to the efficiency of the correction process.

Although the efficiency with which mismatch correction occurs is highly dependent on the presence of d(GATC) sequences and their state of methylation, we have occasionally observed low levels of correction on heteroduplexes that lack a d(GATC) site or in which the d(GATC) site is symmetrically methylated (Tables 1 and 3). The efficiency of correction of these substrates is so near background levels that we have not attempted to determine the *mut*-dependence of the reaction, but repair in these cases may be mediated by the low efficiency methyl-independent pathway described by Fishel *et al.* (28). If this is the case, then our results indicate that the methyl-directed pathway is more efficient by a factor of at least 10.

Our finding that single-strand breaks can target correction *in vitro* is of interest in view of the role of strand breaks in directing mismatch repair in *S. pneumoniae* and in terms of their proposed role in the *E. coli* system (25, 26). The fact that the reaction in *E. coli* extracts is independent of known *mut* gene products raises questions concerning its biological significance in this organism. While this reaction may represent an as yet undescribed *bona fide* pathway of mismatch correction, it is evident that strand breaks cannot in general suffice for strand targeting of the *mutHLS*-dependent correction system. However, our findings concerning the involvement of d(GATC) sequences in *E. coli* mismatch correction are consistent with the suggestion that methyl-directed mismatch correction may involve scission of the unmethylated DNA strand in the vicinity of a d(GATC) site and the idea that this cleavage event is dependent on the presence of a mismatched base pair elsewhere on the substrate (14).

Radman and colleagues (29) have recently described *in vivo* experiments that utilized a transfection assay to assess dependence of G-T and G-A mismatch correction on d(GATC) sequences. Heteroduplexes used in their study were derived from ϕ X174 and contained zero or two d(GATC) sites. Although the two classes of heteroduplex differed at a number of nucleotide positions in addition to the variation at d(GATC) sites, the conclusions of their study are essentially identical to those drawn here.

This work was supported by Grant GM23719 from the National Institute of General Medical Sciences and by funds from the North Carolina Biotechnology Center Small Grants Program. R.S.L. was

supported by Damon Runyon-Walter Winchell Cancer Fund Fellowship DRG-877.

1. Wagner, R. & Meselson, M. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4135-4139.
2. Radman, M., Wagner, R. E., Glickman, B. W. & Meselson, M. (1980) in *Progress in Environmental Mutagenesis*, ed. Alacevic, M. (Elsevier/North-Holland, Amsterdam), pp. 121-130.
3. Radman, M., Dohet, C., Bourguignon, M.-F., Doubleday, O. P. & Lecomte, P. (1981) in *Chromosome Damage and Repair*, eds. Seeberg, E. & Kleppe, K. (Plenum, New York), pp. 431-445.
4. Kramer, W., Schughart, K. & Fritz, H.-J. (1982) *Nucleic Acids Res.* **10**, 6475-6485.
5. Pukkila, P. J., Peterson, J., Herman, G., Modrich, P. & Meselson, M. (1983) *Genetics* **104**, 571-582.
6. Lu, A.-L., Clark, S. & Modrich, P. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4639-4643.
7. Wagner, R., Dohet, C., Jones, M., Doutriaux, M.-P., Hutchinson, F. & Radman, M. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 611-615.
8. Nevers, P. & Spatz, H. (1975) *Mol. Gen. Genet.* **139**, 233-243.
9. Bauer, J., Krammer, G. & Knippers, R. (1981) *Mol. Gen. Genet.* **181**, 541-547.
10. Glickman, B. W. & Radman, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1063-1067.
11. McGraw, B. R. & Marinus, M. G. (1980) *Mol. Gen. Genet.* **178**, 309-315.
12. Su, S.-S. & Modrich, P. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5057-5061.
13. Lahue, R. S., Su, S.-S., Welsh, K. & Modrich, P. (1986) *UCLA Symp. Mol. Cell. Biol.* **47**, in press.
14. Lu, A.-L., Welsh, K., Su, S.-S. & Modrich, P. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 589-596.
15. Marinus, M. G., Poteete, A. & Arraj, J. A. (1984) *Gene* **28**, 123-125.
16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
17. Kramer, B., Kramer, W. & Fritz, H.-J. (1984) *Cell* **38**, 879-887.
18. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180-182.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
20. Boeke, J. D. (1981) *Mol. Gen. Genet.* **181**, 288-291.
21. van Wezenbeek, P. M. G. F., Hulsebos, T. J. M. & Schoenmakers, J. G. G. (1980) *Gene* **11**, 129-148.
22. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103-119.
23. Hill, D. F. & Petersen, G. B. (1982) *J. Virol.* **44**, 32-46.
24. Hare, J. T. & Taylor, J. H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7350-7354.
25. Lacks, S. A., Dunn, J. & Greenberg, B. (1982) *Cell* **31**, 327-336.
26. Claverys, J.-P. & Lacks, S. A. (1986) *Microbiol. Rev.* **50**, 133-165.
27. Joyce, C. M. & Grindley, N. F. (1984) *J. Bacteriol.* **158**, 636-643.
28. Fishel, R. A., Siegel, E. C. & Kolodner, R. (1986) *J. Mol. Biol.* **188**, 147-157.
29. Laengle-Rouault, F., Maenhaut-Michel, G. & Radman, M. (1986) *EMBO J.* **5**, 2009-2013.