Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3395–3397, May 1986
Genetics

Methyl-directed repair of frameshift mutations in heteroduplex DNA

(bacteriophage λ/*Escherichia coli*/mismatch repair/mutagenesis/gene conversion)

CHRISTIANE DOHET, ROBERT WAGNER, AND MIROSLAV RADMAN

Institut Jacques Monod, Centre National de La Recherche Scientifique, Université Paris 7, Tour 43, 2 Place Jussieu, 75251 Paris Cedex 05, France

Communicated by Franklin W. Stahl, January 2, 1986

ABSTRACT **DNA heteroduplexes with single unpaired bases of the four different kinds were prepared by annealing separated strands of bacteriophage λ DNA and used to transfect *Escherichia coli*. Genetic analysis of the progeny phages obtained from transfected bacteria indicates that the *E. coli* mismatch repair system can recognize and repair heteroduplexes with single unpaired bases—i.e., frameshift/wild-type heteroduplexes. The repair of a particular strand of the heteroduplex is inhibited by full methylation of the adenines in the GATC sequences of that strand. Thus, it appears that the *E. coli* mismatch repair system can act on newly synthesized DNA strands to remove replication errors involving the insertion or deletion of a single base.**

The *Escherichia coli* mismatch repair system is able to recognize noncomplementary base pairs in DNA and acts, apparently via localized excision and resynthesis, to replace mispaired bases (for review, see ref. 1). Regions of DNA in which GATC sequences are fully adenine-methylated appear to be refractory to mismatch repair (2, 3), and it appears to be the transient undermethylation of newly synthesized 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 (1–5).

The *E. coli* mismatch repair system does not recognize and/or repair all mismatches with equal efficiency (6, 7). Both transition mismatches (G·T and C·A) are readily recognized and repaired, whereas three of the six transversion mismatches are not (6). This pattern can account, in part, for the findings that the mutator effects observed in *E. coli mutH*, *mutL*, *mutS*, and *mutU* mutants, which are deficient in mismatch repair (refs. 2–8; for review, see ref. 1), and *dam* mutants, which have undirected mismatch repair (2, 6), are due primarily to an increase in transition and frameshift mutations (1).

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—i.e., frameshift/wild-type heteroduplexes. This hypothesis was tested. The results indicate that heteroduplexes with one unpaired base can be recognized and repaired by the *E. coli* mismatch repair system.

MATERIALS AND METHODS

λ phages with sequenced frameshift mutations in the *c*I gene were obtained from Franklin Hutchinson (Yale University). Strand preparation and annealing, transfection conditions, and scoring procedures have been 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.

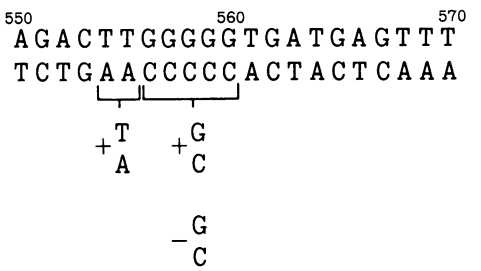

FIG. 1. DNA sequence of the *c*I gene in the region of the three sequenced frameshift mutations used in this study. Numbers indicate the number of bases from the start of the *c*I promoter for rightward transcription (11).

Unmethylated DNA is prepared from phages grown in *dam* (deficient in adenine methylation) bacteria, GM33 (9). Fully methylated DNA is prepared from phages grown in a methylase overproducer strain (10) [the GATC sequences in λ DNA prepared from phages grown in wild-type *E. coli* are only about 75% methylated (2)].

RESULTS AND DISCUSSION

When single strands of DNA from wild-type and frameshift mutant phages are annealed, the resulting heteroduplexes contain one or more unpaired bases. If the frameshift is of the +1 addition type, a single unpaired base will be present on the mutant strand. If the frameshift is of the −1 deletion type, a single unpaired base will be present on the wild-type strand. In other words, +1 addition frameshift/wild-type heteroduplexes and −1 deletion/wild-type heteroduplexes are structurally equivalent and differ only with respect to the genotype of the strand with an unpaired base. Each frameshift/wild-type pair of DNAs will form two heteroduplexes that differ only with respect to the unpaired base. Thus, by annealing separated strands of DNA from wild-type bacteriophage λ with separated strands of DNA from phages with addition frameshift mutations, one +1 G·C and one +1 A·T, it has been possible to form individual heteroduplexes with single unpaired bases of the four different kinds (Fig. 1; Table 1, lines 1–4).

Repair of unmethylated frameshift/wild-type heteroduplexes is relatively efficient, as judged by the small fraction of mixed infective centers—i.e., those yielding phages of both clear (*c*I) and turbid (c^+) genotypes, obtained from transfections of wild-type bacteria relative to transfections of mismatch repair-deficient (*mutL*) bacteria (Table 1, lines 1–4; see ref. 6). The extent of the repair is comparable to that observed for heteroduplexes with single base-pair mismatches (6) and much greater than that observed for heteroduplexes with a large (800-base-pair) single-stranded loop (ref. 3; unpublished data). As was observed for single base-pair mismatches (6), there is a consistent, but unexplained, bias in favor of L-strand genotypes among pure

Table 1. Genetic analysis of infective centers derived from heteroduplexes with single unpaired bases

Heteroduplex*	Wild-type				*MutL*			
	Mixed (c^+/c), %	Clear (c), %	Turbid (c^+), %	Total	Mixed (c^+/c), %	Clear c, %	Turbid (c^+), %	Total
G c ⊥ l c^+ — r	12.1†	56.1	31.8	280	78.1	19.9	2.0	256
c^+ — l c ⊤ r C	8.8†	43.5	47.7	262	90.5	2.0	7.5	254
C‡ c^+ ⊥ l c — r	17.7	37.6	44.7	599	90.2	3.0	6.8	592
c — l c^+ ⊤ r G‡	12.4	57.0	30.6	509	77.6	18.6	3.8	505
T c ⊥ l c^+ — r	6.5	33.0	60.5	306	81.4	1.1	17.5	274
c^+ — l c ⊤ r A	5.1	65.3	29.6	274	78.3	15.4	6.3	253
T c ⊥ l c^+ — r (me$^+$)	4.4	3.2	92.4	251	82.8	3.4	13.8	261
c^+ — l (me$^+$) c ⊤ r A	6.8	12.9	80.2	278	78.3	13.8	7.9	254
T c ⊥ l (me$^+$) c^+ — r	3.6	86.8	9.6	281	82.2	3.5	14.3	343
c^+ — l c ⊤ r (me$^+$) A	6.6	92.3	1.1	274	83.1	12.6	4.3	326
T c ⊥ l (me$^+$) c^+ — r (me$^+$)	72.7	12.0	15.5	267	84.8	3.2	11.9	277
c^+ — l (me$^+$) c ⊤ r (me$^+$) A	77.2	14.9	8.0	276	83.1	12.1	4.8	248

Wild-type cells are *E. coli* C600 (see ref. 6), and *MutL* cells are *E. coli* C600 *mutL* (*mut-211*):Tn*5* constructed by P1 transduction from ES 1293 (12). DNA strands are unmethylated unless otherwise indicated (me$^+$).

*The representation of the heteroduplexes in this table is not meant to reflect actual heteroduplex structure. NMR studies of short heteroduplexes have shown that single unpaired bases can be either intra- or extrahelical (13, 14).

†Preliminary results obtained from transfections with these heteroduplexes (3) could not be confirmed.

‡Mutant strands of these heteroduplexes have a (G·C) deletion frameshift.

infective centers—i.e., those infective centers yielding only cI or c^+ phages.

To confirm that the genotype of the strand with the unpaired base does not somehow influence mismatch repair, frameshift/wild-type heteroduplexes were prepared using −1 deletion frameshift DNA. The results of transfections with these heteroduplexes (Table 1, lines 3 and 4) are indistinguishable from those obtained from transfections with +1 addition frameshift/wild-type heteroduplexes.

To mimic more closely the situation in which mismatch repair is believed to operate *in vivo*—i.e., in hemimethylated DNA immediately behind the replication fork (1)—heteroduplexes were prepared in which one strand was methylated and the other was not (Table 1, lines 7–10). Repair of frameshift/wild-type heteroduplexes occurs in hemimethylated DNA and occurs preferentially, if not exclusively, on unmethylated strands. Repair is greatly reduced when both strands are methylated (Table 1, lines 11 and 12).

The *E. coli* mismatch repair system appears to recognize and repair heteroduplexes with single unpaired bases no less efficiently than it recognizes and repairs heteroduplexes with repairable single base-pair mismatches. Because the repair is methyl-directed, the repair system can presumably act on newly synthesized DNA strands to remove replication errors

involving the insertion or deletion of a single base. It may also be possible for the *E. coli* mismatch repair system to recognize and repair heteroduplexes with somewhat larger unpaired regions (2 or 3 bases). However, the finding that a large unpaired region (800 base pairs) is not repaired by the *E. coli* mismatch repair system (ref. 3; unpublished data) suggests that there is a maximum-sized unpaired region that the system can recognize as a mismatch. Further experiments are required to determine what that size is.

The sequenced mutants used in this study were provided by Dr. F. Hutchinson. R.W. was supported by senior fellowships from the Association pour la Recherche sur le Cancer (A.R.C.) and the Ligue Nationale Française contre le Cancer. This work was supported by the Centre National de la Recherche Scientifique and by grants from the Ligue Nationale Française contre le Cancer and the A.R.C.

1. Radman, R. & Wagner, R. (1984) *Curr. Top. Microbiol. Immunol.* **108,** 23–28.
2. Pukkila, P. J., Peterson, J., Herman, G., Modrich, P. & Meselson, M. (1983) *Genetics* **104,** 571–582.
3. Wagner, R., Dohet, C., Jones, M., Doutriaux, M.-P., Hutchinson, F. & Radman, M. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **49,** 611–615.
4. Wagner, R. & Meselson, M. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 4135–4139.
5. Radman, M., Wagner, R., Glickman, B. W. & Meselson, M. (1980) in *Progress in Environmental Mutagenesis*, ed. Alacevic, M. (Elsevier, Amsterdam), pp. 121–130.
6. Dohet, C., Wagner, R. & Radman, M. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 503–505.
7. Kramer, B., Kramer, W. & Fritz, H.-J. (1984) *Cell* **38,** 879–887.
8. Glickman, B. W. & Radman, M. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 1063–1067.
9. Marinus, M. G. & Morris, N. R. (1975) *Mutat. Res.* **28,** 15–26.
10. Herman, G. E. & Modrich, P. (1981) *J. Bacteriol.* **145,** 644–646.
11. Sauer, R. T. (1978) *Nature (London)* **276,** 301–302.
12. Siegel, E. C., Wain, S. L., Meltzer, S. F., Binion, M. L. & Steinberg, J. L. (1982) *Mutat. Res.* **93,** 25–33.
13. Morden, K. M., Chu, Y. G., Martin, F. H. & Tinoco, I., Jr. (1983) *Biochemistry* **22,** 5557–5563.
14. Pardi, A., Morden, K. M., Patel, D. J. & Tinoco, I., Jr. (1982) *Biochemistry* **21,** 6567–6574.