The EMBO Journal vol.6 no.6 pp.1809–1815, 1987

DNA mismatch-repair in *Escherichia coli* counteracting the hydrolytic deamination of 5-methyl-cytosine residues

Roland Zell and Hans-Joachim Fritz

Max-Planck-Institut für Biochemie, Abteilung Zellbiologie, D-8033 Martinsried bei München, FRG

Communicated by G.Gerisch

Derivatives of phage M13 were constructed and used for the *in vitro* preparation of heteroduplex DNA molecules containing base/base mismatches that mimick DNA lesions caused by hydrolytic deamination of 5-meC residues in *Escherichia coli* DNA (i.e. they carry a T/G mismatch in the special sequence context provided by the recognition site -CC^{A}/$_{T}$GG- of the Dcm-methyltransferase). Upon introduction of these heteroduplex DNAs into $CaCl_2$-treated *E. coli* cells, the mismatches are efficiently repaired with high bias in favour of the DNA strand containing the mismatched guanine residue. This special DNA mismatch-repair operates on fully *dam*-methylated DNA and is independent of gene *mutH*. It thus fulfills the salient requirements of a repair pathway responsible for counteracting the spontaneous hydrolytic deamination of 5-meC *in vivo*. The repair efficiency is boosted by a 5-methyl group present on the cytosine residue at the next-nearest position to the 5′ side of the mismatched guanine. The repair is severely impaired in host strains carrying a mutation in any of the three loci *dcm*, *mutL* and *mutS*.

Key words: DNA methylation/DNA mismatch-repair/hydrolytic DNA deamination/mutational hot-spots/very short patch repair

Introduction

Margaret Lieb has carried out an extensive study of genetic recombination in the repressor gene of bacteriophage lambda (Lieb, 1985, and literature quoted therein). From the extraordinary properties of *cI* mutations am6 and 1-1 she was able to infer the presence in *Escherichia coli* of a hitherto unknown DNA mismatch-repair pathway she called 'very short patch' (VSP) mismatch correction because of its ability to cause excess recombination between am6 and very closely spaced markers (Lieb, 1983, 1985).

According to Lieb's analysis, this repair pathway appears to be specialized to process T/G mismatches in the context of the recognition sequence of the *E. coli* Dcm-methyltransferase. Its biological function seems to lie in restoring such sites that have undergone spontaneous hydrolytic deamination at the inner (5-methylated) cytosine residue (see Figure 1). This is the only known spontaneous intracellular reaction that converts a naturally occurring DNA component (5-meC) into another (T) and thus leads to a base/base mismatch in quiescent (i.e. non-replicating and non-recombining) DNA molecules. The resultant pre-mutagenic DNA lesion was initially thought to be non-repairable, a conclusion based on the observed hypermutability of certain *dcm* methylation sites (Coulondre *et al.*, 1978; Duncan and Miller, 1980).

There are several other DNA mismatch-repair mechanisms acting in *E. coli*, the best-studied of which is the *dam/mutH* pathway ('methylation-instructed' DNA mismatch-repair). This repair system contributes to the overall fidelity of DNA replication by post-replicatively correcting nucleotide misincorporations that have escaped proof-reading (for a recent review, see Claverys and Lacks, 1986). The strand choice in this repair process is directed by transient undermethylation of GATC sites in the newly synthesized DNA strand (Pukkila *et al.*, 1983); repair tract

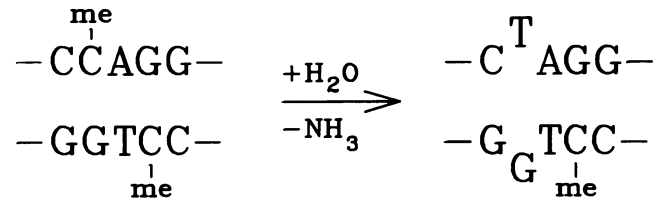

Fig. 1. Hydrolytic deamination of 5-methyl-cytosine in DNA. Spontaneous deamination of a 5-meC residue, as illustrated for the recognition site of the *E. coli* Dcm-methyltransferase, yields thymine and thus creates a T/G mismatch.

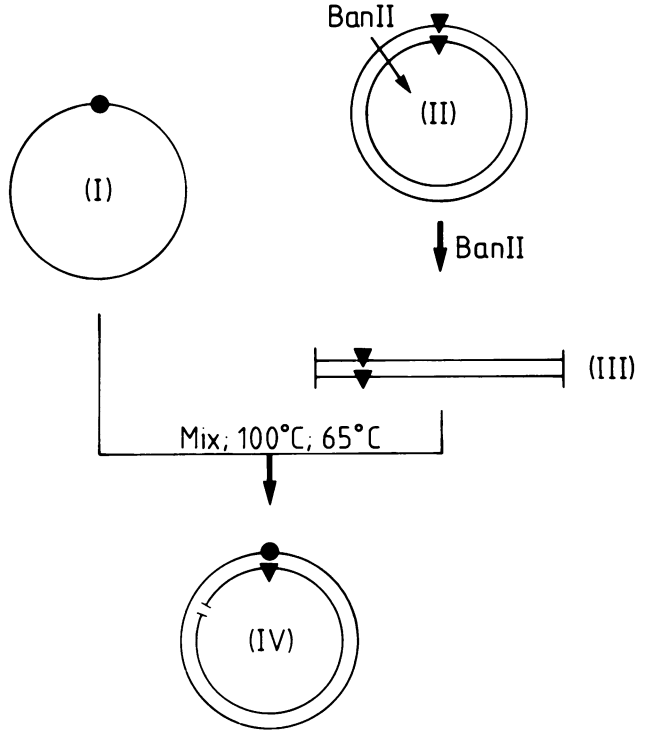

Fig. 2. Construction of nicked duplex DNA (ndDNA). Replicative form DNA (RF-DNA) (II) is linearized with *Ban*II. The *Ban*II linearized DNA (III) is mixed with single-stranded (virion) DNA (I) in a molar ratio of 1:5. The symbols ▲ and ● represent different genetic markers. Denaturation was achieved by heating for 5 min to 100°C, renaturation at 65°C for 10 min in 1 × SSC. By this procedure, four different DNAs are obtained: ndDNA (IV), reformed double-stranded linear DNA (III), remaining circular ssDNA (I), and excess single-stranded linear (+) strand DNA (not indicated). The ndDNA is purified by sucrose gradient centrifugation.

lengths are in the order of a few thousand nucleotides (Wagner and Meselson, 1976).

We embarked on a project aimed at elucidating mechanistic and evolutionary relationships between different DNA mismatch-repair systems of *E. coli* and of other organisms. In this context, a repair pathway with short DNA synthesis tracts is of particular interest, since it may provide a mechanistic model case for certain gene conversion phenomena in higher organisms (for an overview, see Kourilsky, 1986).

Here we describe application of molecular genetics to characterize the very short patch DNA mismatch-repair pathway (VSP repair) with respect to (i) substrate recognition properties, (ii) choice of template strand, (iii) influence of gene loci *dcm*, *mutH*, *mutL* and *mutS* and (iv) competition with other DNA mismatch-repair systems of *E. coli*.

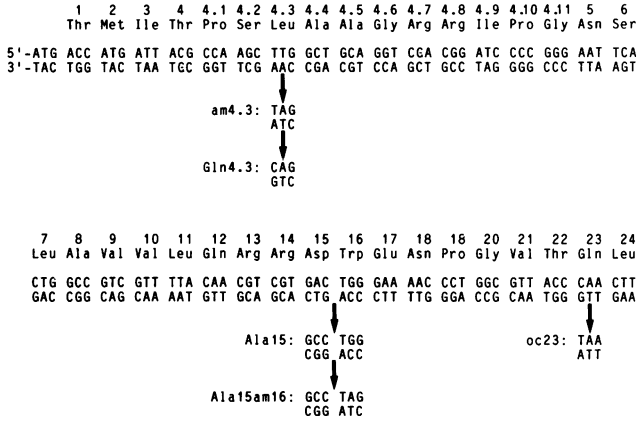

Fig. 3. Constructed mutations in the N-terminal region of the *lacZ* gene. Codon numbering is derived from the wild-type *lacZ* gene (Kalnins *et al.*, 1983). Numbers 4.1–4.11 designate additional codons corresponding to the poly-linker region of phage M13mp9 (Messing and Vieira, 1982). Mutation *lacZ*oc23 was described earlier (Kramer *et al.*, 1984a). For construction of the mutations *lacZ* am4.3, *lacZ*Gln4.3, *lacZ*Ala15 and *lacZ*Ala15am16 refer to Materials and methods.

Results

The genetic assay, construction of mutant M13 phage genomes and of nicked duplex DNA molecules

Previously we devised a two-step genetic assay to determine the efficiency and the directional bias of DNA mismatch-repair acting *in vivo* on heteroduplex DNA molecules of the M13 phage genome (Kramer *et al.*, 1984a). These heteroduplex DNA molecules are constructed *in vitro* by thermal denaturation/renaturation of mixtures of single-stranded (virion) DNA of a particular M13 phage and the double-stranded genome (RF-DNA) of an allelic derivative linearized by cleavage with a restriction enzyme. A nicked duplex DNA molecule (ndDNA) results, containing a single site of nucleotide mispairing (Figure 2). The ndDNA is used to transfect $CaCl_2$-treated host cells. The two allelic phages display a blue or colourless plaque phenotype on 'x-gal' plates respectively, due to their ability or inability to complement a *lacZ*ΔM15 mutation of the host (Langley *et al.*, 1975). Marker distribution is measured twice: firstly by plating directly after transfection and secondly by titrating the different phage types produced in the infected culture (re-infection step). The efficiency of DNA mismatch-repair acting on the transfecting ndDNA can then be estimated by the excess of coloured (i.e. homogeneously blue plus mottled) plaques in the transfection step relative to the amount of blue plaques in the re-infection step. The template choice bias of repair is directly reflected in the distribution of blue and colourless plaques (Kramer *et al.*, 1984a).

To make this assay applicable to the study of VSP DNA mismatch-repair, several new mutants of phage M13 were constructed: two contain in the plus strand the sequences C(5-meC)AGG or C(5-meC)TGG and display a blue plaque phenotype, two others contain, instead, the sequences CTAGG and CCTAG and display a colourless phenotype. This was achieved by oligonucleotide-directed mutation construction (Kramer *et al.*, 1984b) with the results indicated in Figure 3. These phages M13mp9-revGln4.3, M13mp9revAla15, M13mp9revam4.3 and M13mp9-revAla15am16 could then be used to construct ndDNAs with heteroduplex regions that mimick the product of the deamination reaction illustrated in Figure 1 (see below).

Table I. Repair of two T/G mismatches correlated to the hydrolytic deamination of 5-methyl-cytosine

(A) Mismatch in codons 4.3 and 16 respectively					(B) Mismatch in codon 23[a]	
Mismatch	Plaque phenotype of (−)strand marker	State of adenine methylation	WK-6/BMH 71-17 Yield of (−)strand marker (%)		BMH 71-18 Yield of (−)strand marker (%)	Mismatch
(+)Strand/ (−)strand		(+)Strand/ (−)strand	After transfection	After re-infection	After re-infection	(+)Strand/ (−)strand
T/G	Blue	me^+/me^+	98 (4/0)	97 (4/0)	78	T/G
codon 4.3		me^-/me^+	98 (4/1)	98 (4/1)	90	
(set I)		me^+/me^-	70 (8/8)	71 (8/14)	30	
		me^-/me^-	87 (8/7)	87 (8/8)	51	
G/T	Colourless	me^+/me^+	8 (3/1)	9 (3/2)		
codon 16		me^-/me^+	51 (3/8)	52 (3/5)		
(set II)		me^+/me^-	6 (3/1)	7 (3/2)		
		me^-/me^-	34 (3/5)	34 (3/2)		

Nicked duplex DNAs containing T/G (G/T) mismatches at codons 4.3 and 16 respectively, were constructed as outlined in Figures 2 and 4. Strain WK-6 was used as the transfection host with ndDNA set I, BMH 71-18 with set II. The two strains are identical in all relevant genetic properties with the exception of *supE* allele present only in BMH 71-18. This difference is important because mutation *lacZ*am4.3 is suppressed by *supE*, whereas mutation *lacZ*Ala15am16 is not. Each percentage value represents an average derived from several experiments rounded off to the nearest integer. Pairs of numbers in parentheses (part A) document the number of individual experiments in each case (left number) and the maximum deviation of a percentage value obtained in an individual experiment from the mean (right number). Typically, between 500 and 2000 plaques were scored in each individual experiment in both the transfection and the re-infection step.

[a]Data taken from Kramer *et al.* (1984a).

Phage M13 heteroduplex genomes that contain a T/G mismatch mimicking the deamination product of a 5-meC/G base pair show unusual DNA mismatch-repair

Two sets of ndDNAs each were constructed, one from phages M13mp9revGln4.3 and M13mp9revam4.3 (set I), the other from M13mp9revAla15 and M13mp9revAla15am16 (set II). Each set contains a T/G (G/T) mismatch. They are illustrated in Figure 4. Note that the mismatches have opposite orientation with respect to the plus and minus DNA strands. Within each set, all four possible states of *dam*-methylation were realized. These ndDNAs were used together with host strains WK-6 and BMH 71-18 in the transfection/re-infection assay outlined in the preceding paragraph. Results are summarized in Table I. For comparison, Table I also contains the results obtained with an 'ordinary' T/G mismatch described earlier (Kramer *et al.*, 1984a).

Essentially identical marker yields in the transfection and the re-infection step within each of the eight experiments listed indicate efficient DNA mismatch-repair. This result is expected for T/G mismatches. The response of the marker yield to the state of adenine methylation, however, is strikingly different from the normal pattern (see Table I, part B). In all cases, yields of phage displaying the blue plaque phenotype are considerably increased. This is particularly pronounced with hemimethylated DNA in which the methylated adenine residues are present in the DNA strand encoding the colourless plaque phenotype. The 71% yield of minus strand marker obtained in this case with set I constitutes the first observed example of DNA mismatch-repair with a bias in favour of the marker encoded by the unmethylated strand of a heteroduplex DNA molecule hemimethylated in its GATC sites. The large experimental variation in that particular case is discussed below (see last paragraph of Results). In summary, the data presented in Table I support Lieb's original notion of a special repair mechanism acting on the particular DNA mismatch illustrated in Figure 1. This support is strengthened by the variation of three fundamental experimental parameters: (i) replacing phage lambda by M13, (ii) placing the *dcm* recognition site containing the mismatch into a different DNA sequence environment and, most importantly, (iii) by studying the phenomenon by transfection with heteroduplex DNA constructed *in vitro* rather than by classical phage crosses in which *in vivo* formation of heteroduplex DNA can only be inferred indirectly. In the following, we therefore refer to the special repair acting on these base/base mismatches as VSP DNA mismatch-repair as termed by Lieb (1983) although the question of repair tract length is not addressed in the present study.

Very short patch DNA mismatch-repair is strongly influenced by gene loci mutL and mutS but is independent of mutH

DNA mismatch-repair in *E. coli* as characterized previously is completely dependent on intact genes *mutL* and *mutS* and requires intact *mutH* for adenine-methylation-dependent DNA strand discrimination (Kramer *et al.*, 1984a; W.Kramer, B.Kramer and H.-J.Fritz, in preparation). To elucidate possible relationships between the different pathways, the eight different ndDNAs introduced in the preceding paragraph were tested for repair in derivatives of host bacteria WK-6 and BMH 71-18 that carry mutations in the loci *mutH*, *mutL* or *mutS*. Results are compiled in Table II.

For *mutS* and *mutL*, the interpretation is straightforward and identical for both T/G mismatches. The excess of colourless plaques in the transfection as compared with the re-infection steps points to inefficient or no repair. The yields of minus strand marker observed in the re-infection step are clustered closely around a mean value of 73.5% for set I and 69.5% for set II. A similar value, 66.5%, was demonstrated before (Kramer *et al.*, 1984a) to result from complete lack of DNA mismatch-repair. These results demonstrate that VSP repair is severely impaired in a *mutL* or *mutS* background; they do not rule out, however, a small residual VSP repair activity. Indeed, preliminary evidence for such activity was found in double mismatch experiments (R. Zell and H.-J.Fritz, unpublished).

In the *mutH* strain, the T/G mismatch is efficiently repaired. The bias for the 'blue marker' is very high and unaffected by adenine methylation (on average 97% with set I and 91% with set II). The data summarized in Table II lend strong support to the assumption that *E. coli* possesses a *mutH*-independent DNA mismatch-repair mechanism which is capable of amending the hydrolytic deamination of 5-meC residues in DNA sites methylated by Dcm-methyltransferase. This repair pathway is strongly influenced by the *mutL* and *mutS* genes and thus (partly) shares an important property with other DNA mismatch-repair pathways of *E. coli* (also see Discussion).

Different repair mechanisms compete for the same base/base mismatch

Transition-type mismatches (i.e. wrong pyrimidine/purine oppositions) are generally good substrates for the *dam/mutH* repair (Kramer *et al.*, 1984a; Dohet *et al.*, 1985). Furthermore, this

Table II. Influence of loci *mutH*, *mutS* and *mutL* on VSP DNA mismatch-repair

Mismatch	State of adenine methylation	*mutH* Yield of (−)strand marker (%)		*mutS* Yield of (−)strand marker (%)		*mutL* Yield of (−)strand marker (%)	
(+)Strand/ (−)strand	(+)Strand/ (−)strand	After transfection	After re-infection	After transfection	After re-infection	After transfection	After re-infection
T/G (set I)	me^+/me^+	98 (4/1)	98 (4/1)	90 (3/3)	76 (3/3)	87 (3/0)	74 (3/3)
	me^-/me^+	97 (4/1)	98 (4/1)	90 (4/3)	74 (4/5)	89 (3/1)	75 (3/6)
	me^+/me^-	95 (4/2)	96 (4/3)	89 (4/2)	76 (4/2)	88 (3/1)	73 (3/4)
	me^-/me^-	96 (4/2)	97 (4/2)	88 (4/3)	76 (4/3)	87 (3/3)	74 (3/4)
G/T (set II)	me^+/me^+	6 (2/1)	11 (2/1)	39 (2/3)	71 (2/1)	40 (2/0)	71 (2/2)
	me^-/me^+	6 (2/1)	10 (2/2)	36 (2/4)	68 (2/2)	36 (2/4)	69 (2/2)
	me^+/me^-	5 (2/1)	9 (2/0)	39 (2/3)	71 (2/0)	36 (2/3)	69 (2/2)
	me^-/me^-	5 (2/1)	7 (2/0)	37 (2/3)	69 (2/1)	35 (2/3)	69 (2/3)

Mean percentage values were derived and displayed as indicated in the footnotes to Table I. For the re-infection step, wild-type parental strains WK-6 or BMH 71-18 were used throughout. The plaque phenotype encoded by the minus strand is blue for the mismatch in codon 4.3 and colourless for the one in codon 16 (see Figure 4).

Table III. Competition of VSP DNA mismatch-repair with other DNA mismatch-repair systems of *E. coli*

Strain[a]	State of adenine methylation (+)Strand/(−)strand	T/G (G/T) mismatch at codon 23 Yield of (−)strand marker (%)[b] After re-infection	codon 4.3 (set I) Yield of (−)strand marker (%)[c] After re-infection	codon 16 (set II) Yield of (−)strand marker (%)[c] After re-infection	$\bar{w}$[d]
mutH	me^+/me^+	74	98 (0.9)	11 (0.9)	
	me^-/me^+	73	97 (0.9)	10 (0.9)	
	me^+/me^-	73	95 (0.8)	9 (0.9)	0.9
	me^-/me^-	74	96 (0.9)	7 (0.9)	
Wild-type	me^+/me^+	78	98 (0.9)	9 (0.9)	
	me^-/me^+	90	98 (0.8)	52 (0.4)	
	me^+/me^-	30	71 (0.6)	7 (0.8)	0.6[e]
	me^-/me^-	51	87 (0.7)	34 (0.3)	

Mathematical simulation of competition of several DNA mismatch-repair pathways for the same base/base mismatch. Values of w (numbers given in parentheses) represent the probability with which a mismatch is processed by the VSP pathway. For details of matching procedure refer to text.
[a]BMH 71-18 and BMH 71-18 *mutH* for experiments involving mismatches at codon 23, WK-6 and WK-6 *mutH* for experiments with mismatches at codon 4.3.
[b]Data taken from Kramer *et al.* (1984a).
[c]Data taken from Tables I and II.
[d]Average value of w taken from set I and set II.
[e]Average taken from the six experiments involving *dam*-unmethylated or hemimethylated DNA.

type of mismatch was found to be efficiently repaired in MutH$^-$ strains, yet without the usual bias in favour of the *dam*-methylated DNA strand (Kramer *et al.*, 1984a). These findings suggest the *mutH* gene product to be responsible for strand discrimination and also the existence of a DNA mismatch-repair mechanism in *E. coli* that is independent of *dam/mutH* but shares substrate recognition properties with the latter pathway. We have recently verified the existence of such a methylation-independent repair mechanism (W.Kramer, B.Kramer and H.-J.Fritz, in preparation).

From the overlapping substrate recognition properties it was to be expected that in the transfection assay the VSP repair has to compete with the *dam/mutH* and the methylation-independent pathways for processing of its cognate base/base mismatch. We have tried to quantitate roughly this competition by matching the results of mathematical simulations with experimental marker yields.

The rationale underlying the matching procedure is as follows. Consider n different repair mechanisms competing to process a given base/base mismatch, each with a characteristic strand bias b. Each bias factor is by convention expressed as the percentage of repair using the minus strand as template. After the bias factors are fixed, the simulated marker yield Y_s can be adjusted to the experimental value Y_E by varying assumed probabilities w with which a base/base mismatch in a transfecting heteroduplex DNA molecule is processed by one or the other repair mechanism. The following equations apply:

$$Y_s = \sum_{i=1}^{n} w_i \times b_i, \quad \sum_{i=1}^{n} w_i = 1 \text{ and } Y_S = Y_E.$$

Starting point of the simulation is the set of data obtained with a T/G mismatch in codon 23 (see Figure 3) as a model case of an efficient repair substrate and BMH 71-18 *mutH* as the transfection host (upper left part of Table III). This mismatch is not subject to VSP repair. Since the *dam/mutH* mechanism is abolished in the MutH$^-$ background, methylation-independent repair is the only mechanism to be considered. Evidently, the bias of template choice is unaffected by adenine methylation and is 73–74% in favour of the minus strand. It was shown separately that this value reflects DNA mismatch-repair with the given strand bias and neither lack of repair (which fortuitously leads to a similar marker distribution after re-infection) nor increased strand loss (Kramer *et al.*, 1984a; W.Kramer, B.Kramer and H.-J.Fritz, in preparation).

If VSP DNA mismatch-repair has evolved to protect *dcm*-recognition sites from mutation by hydrolytic deamination of 5-meC, it seems plausible to assume that it has high intrinsic selectivity to choose the strand containing the intact *dcm*-recognition sequence as the repair template. With this strand selection bias set at 100% one can calculate the probability w, with which a T/G mismatch at codon 4.3 or codon 16 is processed by the VSP mechanism. For example, in the case of the mismatch at codon 4.3 and fully *dam*-methylated DNA, the experimental yield in percentage of minus strand marker equals 100 times the probability of processing by the VSP pathway plus 74 times the probability of processing by the methylation-independent pathway or:

$$98 = 100w + 74\,(1 - w).$$

For the mismatch at codon 16 the corresponding equation is:

$$11 = 0w + 74\,(1 - w).$$

(Note the opposite strand bias of VSP repair in the two cases considered.)

Evidently, the probability w roughly equals 0.9 in all eight cases measured with the MutH$^-$ strains (Table III). The corresponding probability of processing by the methylation-independent pathway is thus 0.1. In other words, VSP repair is about nine times more efficient than the methylation-independent pathway in competing for a T/G mismatch at codon 4.3 or codon 16. (Again, lack of repair need not be considered because of the extremely asymmetrical marker distribution which, furthermore, is similar in the transfection and the re-infection step — compare Table II.) Strand loss as the source of the asymmetric marker distribution can be excluded, because the effect is qualitatively and quantitatively equal for either strand orientation.

In the wild-type background, the situation is more complex.

Table IV. Influence of cytosine methylation and *dcm* gene

Mismatch in codon 4.3	State of adenine and cytosine methylation	WK-6 Yield of (−)strand marker (%)		RHZ 2-2 Yield of (−)strand marker (%)	
(+)Strand/ (−)strand	(+)Strand/ (−)strand	After transfection	After re-infection	After transfection	After re-infection
T/G	A me^+/A me^- C me^+/C me^+	70 (9/8)[a]	71 (9/14)[a]	–	36 (2/3)
	A me^+/A me^- C me^+/C me^-	49 (3/6)	47 (3/7)	–	14 (2/3)

Mean percentage values were derived and displayed as indicated in the footnote to Table I.
[a]Data taken from Table I.

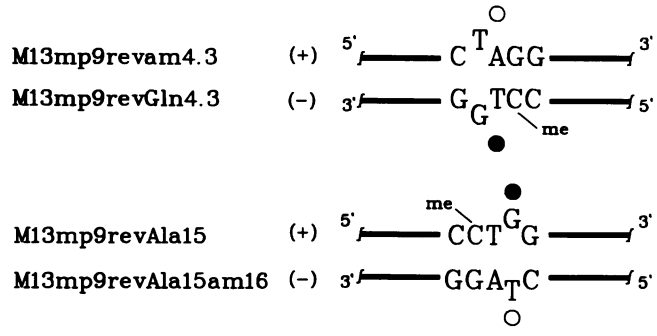

Fig. 4. Heteroduplex DNA molecules containing defined T/G mismatches. For codon numbers and nucleotide sequences refer to Figure 2. For construction of the upper heteroduplex DNA, the minus strand of the genome of phage M13mp9revGln4.3 was annealed to the plus strand of M13mp9revam4.3 (see Materials and methods). Likewise, the lower heteroduplex DNA was constructed from the plus strand of phage M13mp9revAla15 and the minus DNA strand derived from phage M13mp9revAla15am16. Filled and open circles designate relevant codons causing a blue of colourless plaque phenotype.

Mismatch-repair according to the *dam/mutH* mechanism has to be considered additionally, the exception being fully *dam*-methylated DNA which is refractory to *dam/mutH* repair (Pukkila *et al.*, 1983; W.Kramer, B.Kramer and H.-J.Fritz, in preparation). For codon 23, the marker yield is in this case strongly dependent on adenine methylation and reflects the combined action of both the *dam/mutH* and the methyl-independent repair mechanisms (lower left part of Table III). The results obtained with fully *dam*-methylated DNA reflect the ones from the experiments with $MutH^-$ host bacteria — the probability of processing the T/G mismatch by the VSP pathway again equals 0.9. In the other six cases (unmethylated or hemimethylated DNA) this probability is clustered around a mean value of 0.6. The variation around the mean is substantial, especially for the mismatch at codon 16 (lower right corner of Table III). This is not too surprising since recognition of base/base mismatches by the *dam/mutH* system is somewhat dependent on the sequence context (Fazakerley *et al.*, 1986) and may not be identical for mismatches of opposite strand orientation (T/G versus G/T). Even taking into account that the T/G mismatch at codon 23 may not be an ideal point of reference for every other T/G (G/T) mismatch considered, the data strongly suggest that VSP repair has a competition efficiency of roughly one or two times that of the *dam/mutH* and the methylation-independent pathways together.

Cytosine methylation of substrate DNA and wild-type dcm gene are important for VSP repair

The results outlined above demonstrate that VSP DNA mismatch-repair is an efficient process. One of the remaining questions is how the substrate T/G mismatches are recognized and channelled into the VSP repair pathway. The signal could conceivably lie in the special DNA sequence context provided by the *dcm*-recognition site and/or in the presence of a 5-meC residue at the next-nearest position on the 5'-side of the mismatched guanosine nucleotide in the template strand of DNA repair synthesis (see Figure 1).

We have investigated this question by varying the state of cytosine methylation of the transfecting heteroduplex DNA. Furthermore, we have tested for a possible direct involvement of the *dcm* gene in the process.

First, heteroduplex DNA containing the codon 4.3 T/G mismatch was constructed from RF-DNA grown in a Dcm^- host. As a net result, this removes the methyl group from the cytosine residue close to the mismatch. With the idea in mind that different efficiencies of processing the codon 4.3 T/G mismatch by the VSP pathway would most clearly show in competition with *dam/mutH* repair, the heteroduplex DNA was hemimethylated in its GATC sites with the A(6)-methyl groups residing in the plus strand. The two pathways would thus tend to work with opposite strand bias and differences in efficiency would be magnified. We think that this effect is also responsible for the large yield variation of up to 14% observed with the similarly hemimethylated ndDNAs documented in Table I. Preparation of *E. coli* cells for transfection is a procedure of limited reproducibility. Apparently, cells are in rather poor physiological condition after $CaCl_2$ treatment and significant variations in the relative activity of competing repair pathways may be expected from batch to batch. With the appropriate DNA methylation state, this would result in correspondingly magnified variations of marker yields.

As is evident from Table IV, removal of the cytosine methyl group results in a drop of minus strand marker yield from 71 to 47% which, however, still represents an unusually high yield of minus strand marker. From this result, we conclude that the immediate DNA sequence context of the mismatch is sufficient for recognition but the added presence of C(5)-methyl group has a stimulatory effect on repair efficiency.

We have initiated experiments using the same heteroduplex DNAs to transfect strain RHZ 2-2 which carries a *dcm* mutation. In such a genetic background of the transfection host, VSP repair is further reduced — even with fully *dcm*-methylated heteroduplex DNA. With the cytosine C(5)-methyl group removed from the mismatch region it seems to be completely abolished. These results point to a direct participation of the *dcm* gene product in the repair process.

Discussion

Substrate requirements and choice of repair template

The high excess recombination of phage lambda mutation *c*Iam6 with very close markers prompted Lieb (1983) to postulate the existence of a DNA mismatch-repair pathway that processes the special class of T/G mismatches illustrated in Figure 1. The study presented here demonstrates that this repair is a general phenomenon. Mutations *lacZ*am4.3, *lacZ*Gln4.3, *lacZ*Ala15 and *lacZ*Ala15am16 were constructed by oligonucleotide-directed mutagenesis, with no bias imposed on any particular recombination/repair properties of the respective heteroduplex DNAs, yet the corresponding T/G mismatches (Figure 4) display the high bias in favour of the intact *dcm*-recognition sequence which was

to be predicted by extrapolation of Lieb's results. Measurement of competition with *dam/mutH* repair and methylation-independent repair confirm that VSP repair has indeed a template choice bias of close to 100%, which is unaffected by the state of *dam* methylation.

The 5-methyl group present on the inner cytosine residue of the template *dcm*-recognition sequence (see Figure 1) has significant influence on the process: its removal results in less efficient VSP repair (Table IV). This may be due to less efficient competition for the mismatch or less biased choice of repair template or both.

Influence of gene loci dcm, mutH, mutL, mutS and relationship to other pathways

Besides its special substrate requirements and its short repair tracts VSP DNA mismatch-repair differs from other pathways in a number of criteria. (i) VSP repair is completely independent of *mutH*; if the substrate properties are such that competition by *dam/mutH* repair can play a role, its effects are even more pronounced in a MutH$^-$ strain. (ii) While gene loci *mutL* and *mutS* have a strong influence on VSP repair, their respective products may not strictly be required (R.Zell and H.-J.Fritz, unpublished). The nature of their (auxiliary?) role in the VSP pathway remains obscure. In *dam/mutH* DNA mismatch-repair the MutS protein binds to DNA regions containing base/base mismatches (Su and Modrich, 1986). If one accepts the MutS protein to be the mismatch recognition factor in the *dam/mutH* pathway, its involvement in VSP repair appears paradoxical since one would predict the latter pathway to have its own, specialized recognition factor. (iii) The *dcm* gene of the transfected bacterium has a strong influence on VSP repair. This confirms an earlier observation of Lieb (1983) made in genetic crosses of phage lambda with a Dcm$^-$ strain as the host. At present, it cannot be decided if the *dcm* gene product is involved in site recognition or in a subsequent part of the overall process.

Biological relevance for E. coli

As measured by our two-step transfection/re-infection assay, T/G mismatches as they arise within *E. coli* from spontaneous hydrolytic deamination of 5-meC residues in DNA are readily repaired. Thus, the findings presented here confirm and extend the conclusions reached on the basis of results from an entirely different type of experiment (Lieb, 1983).

We showed VSP repair and *dam/mutH* repair to compete with comparable rates for the two T/G mismatches illustrated in Figure 4. Hence, VSP repair must be considered to be a very efficient process which could well have evolved for the function to counteract the mutagenic effect of 5-meC deamination. This competition was measured here to obtain a point of reference for VSP repair efficiency. *In vivo* competition of VSP repair with the *dam/mutH* pathway may be irrelevant since hydrolytic deamination of 5-meC is expected to occur mainly in quiescent, i.e. fully *dam*-methylated, DNA on which no *dam/mutH* repair should occur. In any case, the question as to the molecular nature of the mutational hot-spots described by Miller and colleagues is thrown open again (Coulondre *et al.*, 1978; Duncan and Miller, 1980; see Introduction). Several alternative explanations now seem to be at least as likely as the one originally offered, namely, non-repairability of the DNA lesion caused by spontaneous hydrolytic deamination of 5-methyl cytosine in DNA (Coulondre *et al.*, 1978; Duncan and Miller, 1980). Reconcilitation of the conflicting views has to await further experiments.

Materials and methods

Bacterial strains

The following *E. coli* K-12 strains were described previously (Kramer *et al.*, 1984a,b): BMH 71-18 [Δ(*lac−proAB*), *supE*, *thi*; F′ *lacI*^{q} Z Δ M15, *proA*$^+$*B*$^+$)] (source B.Müller-Hill); BMH 71-18 *mutH*, *-mutS*, *-mutL* carrying the mutations *mutH*::Tn*5*, *mutS*::Tn*10* and *mutL*::Tn*10* respectively (source W.Kramer, this laboratory); WK 2-6-2 [(*mal* 354, *tsx*, *dam*-4 Tn*10* (tet^r), F′ *galT*::Tn*5* (kan^r)] (source W.Kramer, this laboratory); MK 30-3 [Δ(lac−*proAB*), *recA*, *galE*, *strA*, F′ *lacI*^{q}, *lacZ*ΔM15, *proAB*] (source M.Koenen). WK-6 [Δ(*lac−pro*), *galE*, *strA*, *nal*^{r}; F′ *lacI*^{q} Z ΔM15, *pro*$^+$] is a spontaneously arisen RecA$^+$ revertant of MK 30-3 (W.Kramer, this laboratory). WK-6 *mutH*, *-mutL*, *-mutS* were constructed from WK-6 by P1 transduction (Miller, 1972). P1 lysates were prepared from strains AB 1157 *mutH*::Tn*5*, *-mutL*::Tn*10* or *-mutS*::Tn*10* (source G.C. Walker) respectively. Transductants were selected by their resistance to tetracyclin (Tn*10*) or kanamycin (Tn*5*) respectively, and were tested for repair deficiency. RHZ 2-2 [*strA*, *dcm*-6, *dam*-3, *hsdS*-21 (r_k^- m_k^-), *lacY*-1, *tsx*-78, *supE*-44, *galK*-2, *galT*-22 *mtl*-1, *metB*-1; F′ *galT*::Tn*5* (kan^r)] was constructed from spontaneously arisen streptomycin-resistant derivatives of GM 272 (F$^-$) (source M. Marinus) crossed wtih GM 99 (F′) (source M.Marinus) (Miller, 1972). F-ductants were selected by resistance to streptomycin and kanamycin, and were checked for susceptibility to infection by M13 phages.

Phage strains

M13mp9rev and M13mp9revam16 (Kramer *et al.*, 1984b) were used to construct the derivatives M13mp9revGln4.3, M13mp9revam4.3, M13mp9revAla15 and M13mp9revAla15am16 using the gapped duplex DNA approach to oligonucleotide-directed mutagenesis (see below). In M13mp9revGln4.3 the original *lacZ* codon 4.3 (TTG, leu) was replaced by CAG (gln), in M13mp9revam4.3 by TAG (amber). Likewise, in M13mp9revAla15 wild-type codon 15 (GAC, asp) was replaced by GCC (ala) (see Figure 3). In the double mutant M13mp9revAla15am16 the wild-type codons 15 and 16 GAC TGG (asp trp) were replaced by GCC TAG (ala amber).

Preparation of nicked duplex DNAs, transfection and re-infection

Host bacteria proficient or deficient in *dam-* and/or *dcm*-methylation were infected with different M13 phages. Single-stranded and double-stranded phage genomes were prepared as described (Kramer *et al.*, 1984a). Their methylation state was checked at the stage of RF-DNA or ndDNA by treatment with restriction endonucleases *Mbo*I, *Dpn*I, *Sau*3A, *Eco*RII and *Bst*NI. To construct nicked duplex DNA, RF-DNA was digested with *Ban*II, mixed with a 5-fold molar excess of single stranded DNA and incubated at 100°C for 5 min in 1 × SSC (150 mM NaCl, 15 mM Na-citrate) to denature double-stranded linear DNA. The DNA was allowed to renature for 10 min at 65°C. The resulting nicked duplex DNA was purified from single-stranded DNA by sucrose gradient centrifugation (7−30% sucrose in 100 mM Tris−HCl pH 8.0/100 mM NaCl/10 mM EDTA/0.002% ethidium bromide, rotor Beckamn SW41, 14 h, 15°C, 27 000 r.p.m.). Transfection and re-infection have been described previously (Kramer *et al.*, 1984a).

Synthesis of 2′-deoxyoligonucleotides and oligonucleotide-directed mutation construction

Oligonucleotides were synthesized by the phosphoramidite method (Beaucage and Caruthers, 1981) using a DNA synthesizer model 380 A (Applied Biosystems). Work-up, purification and enzymatic 5′-phosphorylation of synthetic oligonucleotides were carried out as described (Kramer *et al.*, 1984a). Mutations were constructed by the gapped duplex DNA method following published procedures (Kramer *et al.*, 1984b).

For construction of M13mp9revam4.3 the mutagenic primer d(pCCTGCAGCC-TAGCTTGGCG) was annealed to a gapped duplex DNA (gdDNA) made of ssDNA of M13mp9 and the large *Pvu*II fragment of the M13mp9rev genome (gap of 268 nucleotides length). After gap-filling reaction with DNA polymerase I (Klenow fragment) and ligation with T4 DNA ligase, the heteroduplex DNA was used for transfection of BMH 71-18 *mutS*. The yield of phage producing colourless plaques was 60%.

A simplified protocol was used for construction of M13mp9revGln4.3. The mutagenic primer d(pCCTGCAGCCTGGCTTGGCG) was annealed to gdDNA constructed from ssDNA of M13mp9revam4.3 and the large *Pvu*II fragment of M13mp9rev (see above). This ternary DNA was directly used to transfect BMH 71-18 *mutS*. Fifteen per cent of the resulting progeny phage displayed a blue plaque phenotype.

For construction of M13mp9revAla15am16, gdDNA was constructed from ssDNA of M13mp9 and the large *Eco*RI/*Pvu*I fragment of M13mp9rev. To this gdDNA the mutagenic primer d(pGTTTTCCTAGGCACGACGTTG) was hybridized. Gap-filling reaction and ligation were performed with T4 DNA polymerase and *E. coli* ligase (Kramer and Fritz, 1987). The heteroduplex DNA was used to transfect WK-6 *mutS*. The yield of phage producing colourless plaques was

77%. The same procedure was applied to construct M13mp9revAla15. Single-stranded DNA of M13mp9revAla15am16 and the large *Eco*RI/*Pvu*II fragment of M13mp9revam16 were used to hybridize the gapped-duplex DNA. The mutagenic primer was d(pGTTTTCCCAGGCACGAC).

All mutations were verified by DNA sequence analysis (Bankier and Barrell, 1983; data not shown).

Acknowledgements

We thank Martin Marinus for the gift of strains, Harald Kolmar for construction of M13mp9revam4.3, Nicole Gerwin for construction of M13mp9revAla15, Wilfried Kramer for construction of strain WK-6, and Judith Graff and Elisa Ferrando for other valuable experimental contributions. Furthermore, we thank Ruth Ehring, Günther Gerisch, Börries Kemper, Harald Kolmar, Ulrich Krawinkel, Wolfgang Nellen and Beate Stern for criticizing drafts of the manuscript. M.Jones, R.Wagner and M.Radman (personal communication) have independently reached similar conclusions about VSP repair from transfection experiments with lambda phage heteroduplex DNA. This work was supported by Deutsche Forschungsgemeinschaft through Forschungsschwerpunkt Gentechnologie.

References

Bankier,A.T. and Barrell,B.G. (1983) In Flavell,R.A. (ed.), *Techniques in the Life Sciences.* Vol. B5. *Nucleic Acid Biochemistry.* Elsevier Scientific Publishers Ireland, pp. 1–34.

Beaucage,S.L. and Caruthers,M.H. (1981) *Tetrahedron Lett.*, **22**, 1859–1862.

Claverys,J.-P. and Lacks,S.A. (1986) *Microbiol. Rev.*, **50**, 133–165.

Coulondre,C., Miller,J.H., Farabaugh,P.J. and Gilbert,W. (1978) *Nature*, **274**, 775–780.

Dohet,C., Wagner,R. and Radman,M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 503–505.

Duncan,B.K. and Miller,J.H. (1980) *Nature*, **287**, 560–561.

Fazakerley,G.V., Quignard,E., Woisard,A., Guschlbauer,W., van der Marel, G.A., van Boom,J.H., Jones,M. and Radman,M. (1986) *EMBO J.*, **5**, 3697–3706.

Kalnins,A., Otto,K., Rüther,U. and Müller-Hill,B. (1983) *EMBO J.*, **2**, 593–597.

Kourilsky,P. (1986) *Trends Genet.*, **2**, 60–63.

Kramer,W. and Fritz,H.-J. (1987) *Methods Enzymol.*, in press.

Kramer,B., Kramer,W. and Fritz,H.-J. (1984a) *Cell*, **38**, 879–887.

Kramer,W., Drutsa,V., Jansen,H.-W., Kramer,B., Pflugfelder,M. and Fritz,H.-J. (1984b) *Nucleic Acids Res.*, **12**, 9441–9456.

Langley,K.E., Villarejo,M.R., Fowler,A.V., Zamenhof,P.J. and Zabin,I. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 1254–1257.

Lieb,M. (1983) *Mol. Gen. Genet.*, **191**, 118–125.

Lieb,M. (1985) *Mol. Gen. Genet.*, **199**, 465–470.

Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269–276.

Miller,J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press, New York.

Pukkila,P.J., Peterson,J., Herman,G., Modrich,P. and Meselson,M. (1983) *Genetics*, **104**, 571–582.

Su,S.-S. and Modrich,P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5057–5061.

Wagner,R., Jr and Meselson,M. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 4135–4139.

Received on February 18, 1987