Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 10003–10008, August 1998
Genetics

The role of SOS and flap processing in microsatellite instability in *Escherichia coli*

(short repeats/polymerase slippage)

PATRICE MOREL*, CÉLINE REVERDY†, BÉNÉDICTE MICHEL, S. DUSKO EHRLICH, AND ERA CASSUTO‡

Génétique Microbienne, Institut National de la Recherche Agronomique, Domaine de Vilvert, 78352 Jouy en Josas Cedex, France

Communicated by I. Robert Lehman, Stanford University School of Medicine, Stanford, CA, June 22, 1998 (received for review February 3, 1998)

ABSTRACT **Mutations affecting mismatch repair result in elevated frequencies of microsatellite length alteration in prokaryotes and eukaryotes. However, the finding that microsatellite instability is found often in cells with a functional mismatch repair system prompted a search for other factors of tract alteration. In the present report, we show that, in *Escherichia coli*, poly(AC/TG) tracts are destabilized by mutations that induce SOS. These observations may have implications for eukaryotic cells because recent results suggest the existence of a mammalian SOS response analogous to that in prokaryotes. In addition, a defect in the 5′–3′ exonuclease domain of DNA polymerase I, homologous to the mammalian FEN1 and the yeast RAD27 nucleases, leads to a marked increase in repeat expansions characteristic of several genetic disorders. Finally, we found that the combination of a proofreading defect with mismatch repair deficiency results in extreme microsatellite instability.**

Microsatellites are tandemly repeated sequence motifs of <6 nt, which represent a substantial part of the eukaryotic genome. Length variation of trinucleotide repeats is known to be associated with a number of genetic diseases (1, 2), but recent work has demonstrated an association with somatic diseases as well, which is not restricted to triplets. For example, mono-, di-, and trinucleotide repeats are unstable in colon cancer cells (3–5).

Although several mechanisms have been proposed to account for the dramatic length variation of simple repetitive DNA, there is presently a consensus on a major role for polymerase slippage (6–8). Numerous studies in prokaryotes and eukaryotes have shown that cells carrying mutations affecting mismatch repair are associated with elevated frequencies of tract instability (6, 8, 9). Conversely, human cells that exhibit microsatellite instability are frequently deficient in mismatch repair (10). However, a large proportion of sporadic cancers displaying microsatellite instability have no defect in any of the known mismatch repair genes (11). Mutations in the *RAD27* (*RTH1*) *Saccharomyces cerevisiae* gene also destabilize microsatellites (12) by a mechanism distinct from mismatch repair (13). Moreover, it is becoming increasingly clear that, aside from proteins, a number of other factors are involved, such as the nature of the repeated motif, the length of the tract, the nature of the flanking sequences, and the orientation of the microsatellite with respect to the direction of replication.

Studies in *Escherichia coli* as well as in yeast generally have used microsatellites cloned in phages or plasmids. To identify other potential factors of microsatellite instability without the complications introduced by the mode of replication and the copy number of the vectors, we inserted the most common dinucleotide tracts, poly(AC) and poly(TG), into the *E. coli* chromosome and examined length variation as a function of: (*i*) the original length, (*ii*) the location of the tract on the leading or lagging strand, and (*iii*) the influence of mutations in several genes involved in recombination, replication, or repair. In accord with previous studies, we found that instability increased exponentially with the number of repeats in the insert (6, 14). We also observed that poly(TG) tracts placed in the leading strand were consistently more unstable than poly(AC) tracts of the same length and location.

Mutations affecting replication and/or excision repair (but not homologous recombination) substantially increased microsatellite instability. In some strains, these effects appear to be caused by constitutive induction of the SOS repair pathway, which therefore ranks as a novel factor of microsatellite instability in *E. coli*. In addition, deletion of the 5′ exonuclease domain of DNA polymerase I (homologous to the nuclease encoded by *RAD27* in yeast and to the mammalian FEN1) led to a marked increase in the proportion of expansions by a mechanism independent of SOS. Finally, a proofreading defect associated with mismatch repair deficiency resulted in an extreme level of tract alteration.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Phages. The *E. coli* strains used in this study are listed in Table 1, and plasmids and phages are listed in Table 2. Mutants were constructed by P1 transduction in strain MC1061 containing prophage λ ZAP III $(AC)_{18}$, $(AC)_{51}$, $(TG)_{21}$, or $(TG)_{42}$ (Fig. 1) and were selected by resistance to antibiotics or by UV sensitivity. Mutator activity was determined by measuring the frequency of generation of cells resistant to rifampicin (15).

Standard DNA Procedures. Restriction enzymes, alkaline phosphatase, and ligase were purchased from commercial sources and were used according to the suppliers' recommendations. Phage λ DNA was extracted as described (16). The Amersham λ DNA *in vitro* packaging module was used for encapsidation of phage λ derivatives. Phage and plasmid preparations and analysis, preparation, and transformation of *E. coli* competent cells were carried out as described (16).

Construction of an Integrative λ Vector: λ ZAP III. The left arm of λ ZAP II was obtained by digestion of the phage DNA with *Eco*RI and *Hin*dIII, and the right arm was obtained by digestion with *Pvu*II. Both arms were purified from contaminating fragments by electrophoresis on agarose gel and were electroeluted. In parallel, a fragment carrying the 3′ part of the *lacZ* gene was obtained by digestion of pFR97 with *Eco*RI and *Dra*I and inserted by ligation between the left (*Eco*RI) and

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.

*Present address: RDP Ecole Normale Supérieure, 46 Allée d'Italie, 69364 Lyon Cedex, 07 France.
†Present address: Biologie Cellulaire et Moléculaire, INRA, 78352 Jouy en Josas Cedex, France.
‡To whom reprint requests should be addressed. e-mail: cassuto@biotec.jouy.inra.fr.

Table 1. *E. coli* strains

Strains	Relevant genotype	Origin
MC1061	*araD139* Δ(*ara-leu*)*7696* Δ(*lac*)*l74 galU galK hsdR*2(r_{K-} m_{K+})*mcrB1 rpsL* (Str^r)	Laboratory collection
JM109	E14$^-$ (McrA$^-$) *recA1 endA1 gyrA96 thi-1 hsdR17* (r_{K-} m_{K+}) *supE44 relA1* Δ(*lac-proAB*) [F′ *traD36 proAB lacI^q ZΔM15*]	Laboratory collection
AQ2178	*polA1 zig*::Tn*10*	T. Kogoma*
CJ233	Δ*polA*::Kan^r [F′ Klenow Cm^r]	21
FG252	*sbcD300*::Kan^r	45
FR98	*mutD*::Tn*10*	M. Radman†
FR559	*mutS*::Tn*5*	M. Radman
JJC16	Δ(*recA-srl*)::Tn*10*	Laboratory collection
JJC18	*uvrD*::Tn*5*	Laboratory collection
JJC213	Δ*rep*::Kan^r	Laboratory collection
GM2199	*dam13*::Tn*9*	M. G. Marinus‡
GY4588	*sfiA*::Kan^r *pyrD*	R. Devoret§
GY4587	*sfiA11*	R. Devoret
GY5425	*lexA71*(Def)::Tn*5*	R. Devoret
GY6781	*lexA1* (Ind$^-$)*mal*::Tn*9*	R. Devoret
JJC443	*lexA3* (Ind$^-$)*mal*::Tn*10*	Laboratory collection
DE582	*recA730 srlC*::Tn*10 lexA57*(Def)	M. A. Petit¶
PM1	MC1061 λ ZAP III $(AC)_{39}$	This work
PM2	MC1061 λ ZAP III $(TG)_{33}$	This work
PM3	MC1061 λ ZAP III $(AC \text{ or } TG)_n$	This work
PM4	MC1061 *polAl zig*::Tn*10* λ ZAP III $(AC \text{ or } TG)_n$	This work
PM5	MC1061 Δ*polA*::Kan^r [F′ Klenow Cm^r] λ ZAP III $(AC \text{ or } TG)_n$	This work
PM6	MC1061 *sbcD300*::Kan^r λ ZAP III $(AC \text{ or } TG)_n$	This work
PM7	MC1061 *mutD*::Tn*10* λ ZAP III $(AC \text{ or } TG)_n$	This work
PM8	MC1061 *mutS*::Tn*5* λ ZAP III $(AC \text{ or } TG)_n$	This work
PM9	MC1061 Δ(*recA-srl*)::Tn*10* λ ZAP III $(AC \text{ or } TG)_n$	This work
PM10	MC1061 *uvrD*::Tn*5* λ ZAP III $(AC \text{ or } TG)_n$	This work
PM11	MC1061 Δ*rep*::Kan^r λ ZAP III $(AC \text{ or } TG)_n$	This work
PM12	MC1061 *dam13*::Tn*9* λ ZAP III $(AC \text{ or } TG)_n$	This work
PM13	MC1061 *sfiA11 lexA71*(Def)::Tn*5* λ ZAP $(AC \text{ or } TG)_n$	This work
PM14	MC1061 *sfiA11 lexA71*(Def)::Tn*5 recA730 srlC*::Tn*10* λ ZAP III $(AC)_{51}$	This work
PM15	MC1061 *polA1 zig*::Tn*10 lexA1*(Ind$^-$)*mal*::Tn*9* λ ZAP III $(AC)_{51}$	This work
PM16	MC1061 Δ*polA*::Kan^r[F′ Klenow Cm^r]*lexA3*(Ind$^-$) *mal*::Tn*10* λ ZAP III $(AC)_{51}$	This work
PM17	MC1061 *mutS*::Tn*5 lexA1* (Ind$^-$)*mal*::Tn*9* λ ZAP III $(AC)_{51}$	This work
PM18	MC1061 *uvrD*::Tn*5 lexA1*(Ind$^-$)*mal*::Tn*9* λ ZAP III $(AC)_{51}$	This work

*University of New Mexico, Albuquerque.
†Institut Jacques Monod, Paris.
‡University of Massachusetts, Worcester.
§Institut Curie, Orsay.
¶Institut National de la Recherche Agronomique, Jouy en Josas.

right (*Pvu*II) arms of λ ZAP II (Fig. 1). After *in vitro* encapsidation and transduction into *E. coli* MC1061 (Δ*lac*), cells were grown at 28°C to allow lysogenization on Luria–Bertani medium plates containing 5-bromo-4-chloro-3-indolyl-D-galactoside (X-Gal), isopropyl-D-thiogalactoside (IPTG), and ampicillin at 20 μg/ml. Blue clones were picked, and the structure of the new vector (λ ZAP III) was verified by restriction digestions after lytic growth.

Integration of CA and TG Repeats into the *E. coli* Chromosome. When cloned into the viral strand of M13 mp18, $(AC)_{39}$ and $(TG)_{33}$ repeats are referred to as M13 mp18-$(AC)_{39}$ and M13 mp18-$(TG)_{33}$ (a gift from M. Dutreix, Institut Curie, Paris). They were obtained by ligating hybridized $(AC)_8$ and $(TG)_8$ oligonucleotides synthesized on a Beckman synthesizer. The ligation mixture was used to transform *E. coli* JM109. Recombinant phages were analyzed by sequencing.

Table 2. Plasmids and phages

	Relevant genotype	Origin
Plasmids		
pFR97	*lacZ*	46
pBS II SK^+	Amp^r	Stratagene
Phages		
M13pm18 $(AC)_{39}$		M. Dutreix*
M13mp18 $(TG)_{33}$		M. Dutreix
λ ZAP II	Amp^r	Stratagene
λ ZAP III	*lacZ*$^+$ Amp^r	This work
λ ZAP III $(AC)_{39}$	*lacZ*$^+$ Amp^r	This work
λ ZAP III $(TG)_{33}$	*lacZ*$^+$ Amp^r	This work

*Institut Curie, Paris.

The replicative form DNAs from M13 mp18-$(AC)_{39}$ and M13 mp18-$(TG)_{33}$ were digested with *Xba*I and *Pst*I, and the fragments carrying the repeats were purified on 6% polyacrylamide gels, electroeluted, and ligated into pBS II SK^+ linearized by the same enzymes. The ligation mixture was used to transform *E. coli* JM109. Recombinant plasmids were analyzed by sequencing and were purified and digested with *Not*I and *Eco*RI. Fragments carrying $(AC)_{39}$ and $(TG)_{33}$ were purified as above. The vector λ ZAP III was digested with the same enzymes, and the left and right arms were purified on 0.8% agarose gel to eliminate the *Not*I–*Eco*RI fragment present in the polylinker (Fig. 1). The *Not*I–$(AC)_{39}$–*Eco*RI and *Not*I–$(TG)_{33}$–*Eco*RI fragments then were ligated to the arms of λ ZAP III. After *in vitro* packaging, transduction into MC1061, and growth at 28°C, as above, blue colonies containing prophages λ ZAP III $(AC)_{39}$ (strain PM1) and λ ZAP III $(TG)_{33}$ (strain PM2) were isolated. The structure of the clones was verified by sequencing of PCR fragments directly amplified from the bacterial chromosome.

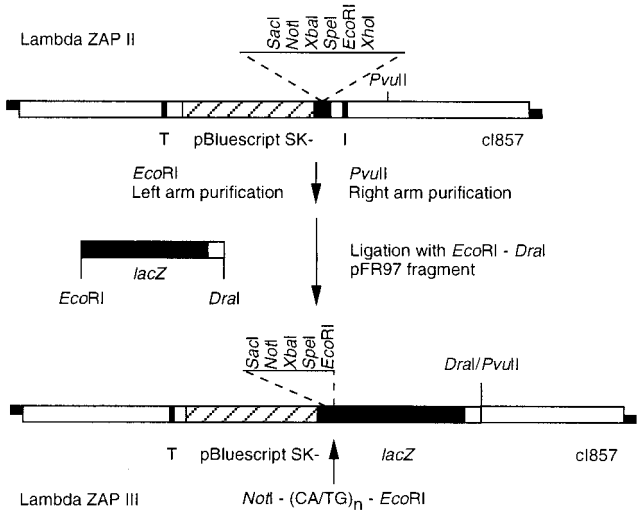

FIG. 1. Construction of λ ZAP III and cloning of dinucleotide repeats. λ Zap III was constructed by cloning the entire *lac*Z gene into λ ZAP II. $(AC)_{39}$ and $(TG)_{33}$ repeats purified from pBS II SK^+ $(AC)_{39}$ and pBS II SK^+ $(TG)_{33}$ were introduced into λ ZAP III by using two restriction enzymes, *Not*I and *Eco*RI, to ensure the directionality of the insertion. Integration of λ ZAP III at the λ attachment site directs the insertion of the (AC) or (TG) sequences on the leading or lagging strand of the *E. coli* chromosome. The position of the two kinds of repeats with respect to the direction of replication was verified by sequencing, by using primers specific for the leading strand. See *Materials and Methods* for more details.

Oligonucleotide PCR Primers. Oligonucleotide primers complementary to chromosomal stretches of dinucleotide repeats were synthesized in the DNA synthesizer Oligo 1000 from Beckman. Primer 1 (5′-AGCTCACTCATTAGGCACC) and primer 2 (5′-TCTTCGCTATTACGCCAGC) were used to amplify DNA fragments containing $(AC)_n$ or $(TG)_n$ short direct repeats. Primer 3 (5′-AAGTTGGGTAACGCCAG) was used to sequence previously obtained amplified DNA fragments.

PCR Amplification. Amplification was performed with a Perkin–Elmer Cetus 9600 thermal cycler. Isolated colonies resuspended in distilled water were used as starting material. The reaction mixtures containing 2 mM each deoxynucleotide triphosphate, 0.3 mM primers, and 2.5 units of *Taq* DNA polymerase (Boehringer Mannheim) were adjusted to a final volume of 20 μl. They were processed through a 30-cycle program consisting of a 30-sec denaturation period at 94°C, a 30-sec annealing period at 55°C, and a 30-sec elongation period at 72°C. PCR fragments were electrophoresed on 6% polyacrylamide gels and stained with ethidium bromide.

DNA Sequencing. PCR sequencing was performed with the Applied Biosystems PRISM dye-terminator sequencing kit on the Perkin–Elmer Cetus 9600 thermal cycler or the Applied Biosystems Catalyst lab station.

Tract Alteration Frequency in *lac*$^+$ Clones with $(AC)_{15-51}$ and $(TG)_{18-42}$ Inserts. Blue colonies were picked on Petri dishes after 1-night incubation at 28°C, were resuspended in Luria–Bertani medium, and were diluted and plated in the presence of X-Gal and IPTG to give 1,000–2,000 colonies per dish. The frequency of tract alteration was estimated by dividing the proportion of white colonies derived from the original bacterium by the number of generations.

RESULTS

Frameshift Mutation Assay. The designation $(AC)_n$ or $(TG)_n$ refers to AC or TG repeats located on the leading strand of the *E. coli* chromosome (Fig. 1). To follow instability of dinucleotide repeats, we inserted the *lacZ* gene preceded by a tract of $(AC)_n$ or $(TG)_n$ into the chromosome of *E. coli* Δ(*lac*) by means of a λ ZAP vector (Fig. 1). Cells carrying this construct display a blue phenotype when plated on medium containing X-Gal and IPTG if the tract comprises 3n bp because active β-galactosidase is synthesized. Any insertion or deletion of 1 or 2 bp (as well as 3n + 1 or 3n + 2) will disrupt the reading frame and will cause the colony to be white, whereas alterations of 3 bp will not be detected.

Frequency of Tract Alteration as a Function of the Number of Repeats. To carry out the following experiments, we constructed strains PM1 and PM2, which carried, respectively, $(AC)_{39}$ and $(TG)_{33}$. After growth on Petri dishes for several generations, isolated colonies were resuspended and plated in the presence of X-Gal and IPTG. Each original clone produced a minority of white colonies of which 50 were analyzed by PCR amplification and electrophoresis on polyacrylamide gel. Colonies carrying noticeably amplified or deleted tracts were picked and grown as in *Materials and Methods* and yielded rare blue colonies from which the insert was sequenced. We isolated *lac*$^+$ clones with the following inserts: $(AC)_{15,\ 18,\ 24,\ 39,\ 45,\ 51}$ and $(TG)_{18,\ 21,\ 30,\ 33,\ 36,\ 42}$, located on the leading strand (Fig. 1). For each, three independent clones were analyzed by PCR amplification and sequencing and were used to study instability, as described in *Materials and Methods*. As shown in Fig. 2, the frequency of tract alteration increased exponentially with the number of repeats in the insert. In addition, when placed on the leading strand, (TG) repeats were reproducibly more unstable than (AC) repeats at the same location.

The nature of the changes in tract length (amplification or deletion) was determined for each of the 12 inserts from a limited number of colonies (up to 24) by PCR amplification of the insert and analysis of the amplified fragments by electrophoresis on polyacrylamide gel, followed by sequencing in case of ambiguity (data not shown). Regardless of the size and composition of the insert, deletions always represented a large majority (≈90%) of the events. We then examined the effect of mutations affecting the major DNA transactions on the instability of repeated doublets. In all of the experiments below, mutations were introduced in the MC1061 strain containing prophage λ ZAP III carrying either a short $[(AC)_{18}, (TG)_{21}]$ or a long $[(AC)_{51}, (TG)_{42}]$ insert, referred to as (AC) or $(TG)_n$ in Table 1.

Influence of Mutations Affecting Homologous Recombination, Replication, and Repair. For each type of mutation, the frequency of tract alteration was determined on three to eight independent clones. The results are presented in Tables 3 and 4. In *E. coli*, inactivation of the *recA* gene reduces homologous recombination 10^4-fold. A *recA* mutation slightly increased

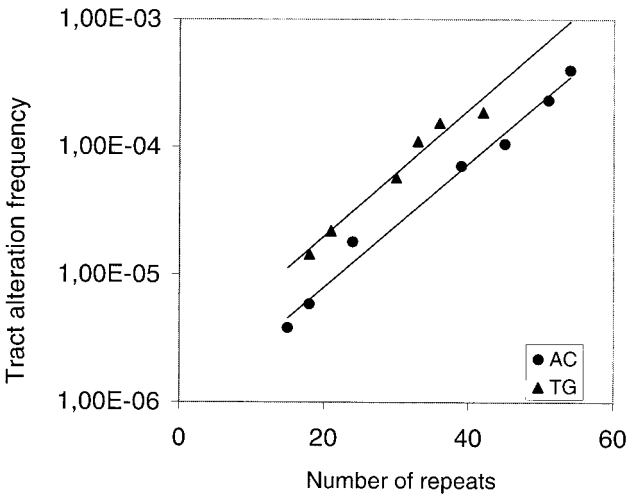

FIG. 2. Frequency of tract alteration in $(AC)_n$ and $(TG)_n$ microsatellites as a function of tract length. Results correspond to inserts placed on the leading strand.

Table 3. Effect of mutations on the frequency of tract alteration in $(AC)_n$ and $(TG)_n$ inserts

Relevant genotype	$(AC)_{18}$	$(AC)_{51}$	$(TG)_{21}$	$(TG)_{42}$
Wild type	5.9×10^{-6} (1.0)	2.2×10^{-4} (1.0)	2.2×10^{-5} (1.0)	1.9×10^{-4} (1.0)
recA	2.5×10^{-5} (4.2)	1.9×10^{-4} (0.9)	4.4×10^{-5} (2.0)	1.9×10^{-4} (1.0)
sbcD	8.1×10^{-6} (1.4)	2.3×10^{-4} (1.0)	7.4×10^{-6} (0.3)	2.8×10^{-4} (1.5)
rep	nd	8.3×10^{-4} (3.8)	6.7×10^{-5} (3.0)	8.3×10^{-4} (4.4)
polA1	1.3×10^{-4} (22)	8.7×10^{-4} (4.0)	1.9×10^{-4} (8.6)	6.5×10^{-4} (3.4)
Δ*polA* (F′ Klenow)	1.8×10^{-4} (30)	2.7×10^{-3} (13)	3.7×10^{-4} (17)	3.3×10^{-3} (17)
uvrD	nd	3.2×10^{-3} (15)	5.3×10^{-4} (24)	3.7×10^{-3} (19)
mutS	5.7×10^{-4} (97)	5.9×10^{-3} (27)	9.6×10^{-4} (44)	3.0×10^{-2} (158)
mutD	(≥1,000)	(≥1,000)	(≥1,000)	(≥1,000)
dam	nd	2.8×10^{-3} (13)	8.5×10^{-4} (39)	3.4×10^{-3} (18)
lexA (Def)	1.3×10^{-4} (22)	1.7×10^{-3} (7.7)	2.4×10^{-4} (11)	1.7×10^{-3} (8.8)
lexA (Def) *recA730*	nd	1.5×10^{-3} (6.8)	nd	nd

Frequencies of tract alteration represent the average of at least three independent determinations. Numbers in parentheses represent the rates of tract length alteration (fold) relative to the rate in the wild-type strain for each insert. nd, not determined.

tract instability in short inserts and had no detectable effect on long ones.

Mutations in *sbcC* originally were isolated as cosuppressors of the recombination deficiency of *recBC sbcB* mutants (17). The *sbcC* gene was found to be cotranscribed with another gene, *sbcD*. *sbcCD* codes for a nuclease that was proposed to cleave secondary structures formed at replication forks, thus initiating recombination events by allowing the entry of RecBCD at broken ends (18). Although secondary structures could be expected to occur frequently when replication forks encounter microsatellite sequences, the *sbcD* mutation did not affect the rate of instability for any of the inserts.

DNA polymerase I, encoded by the *polA* gene, functions in both excision repair and in chromosomal replication. In the absence of exogenous DNA damaging agents, the processing of Okasaki fragments is probably its main function. Two mutations in *polA* were used in the present experiments. *polA1* eliminates most of the polymerizing activity of the enzyme, whereas ΔpolA (F′ Klenow) lacks the 5′–3′ exonuclease. The latter mutation is of particular interest because of the functional and structural homology between the *E. coli* 5′–3′ exonuclease and the mammalian FEN1 or the yeast RAD27 nucleases (19) that are thought to increase tract instability by a mechanism contributing to repeat expansion (12, 13, 20). Both *E. coli polA* mutants grow normally on rich medium because either the Klenow fragment or the 5′–3′ exonuclease is sufficient for viability (21). Both mutations increased the instability of the inserts. The effect of *polA1* was moderate in long inserts and was more important in short ones, whereas in the strain lacking the 5′–3′ exonuclease, alteration was observed for all inserts with a frequency 13- to 30-fold higher than in the wild-type strain.

The Rep helicase is required for growth of certain phages. *E. coli rep* mutants are characterized by a slower propagation of the replication fork compared with wild-type strains (22–24). On the assumption that a slow propagation of the fork would favor polymerase slippage between repeats, we introduced a *rep* mutation into the constructs, described above. A moderate increase (3- to 4-fold) in instability, apparently unrelated to insert size, was observed for the three inserts examined. As expected, a mutation in the *mutS* gene resulted in a substantial increase in instability (27- to 160-fold) for all insert sizes. We examined another mutation affecting mismatch repair, *uvrD,* the product of which, helicase II (UvrD), is known to be involved also in excision repair and recombination. Tract instability increased 15- to 24-fold in this strain.

In the *E. coli* mismatch repair pathway, strand discrimination is ensured by specific methylation of the parental strand by the Dam methylase at the replication fork. Mutations in *dam* confer to the cell a mutator phenotype and allow mismatch repair to take place on both strands (25). Instability increased 13- to 39-fold in the *dam* strain. A mutation (*mutD*) in the *dnaQ* gene encoding the proofreading exonuclease of DNA polymerase III leads to the most potent mutator phenotype known in *E. coli* (15, 26, 27). In a *mutD* strain carrying inserts, the frequency of tract alteration was so high that it could not be estimated from the proportion of white colonies. All colonies were of mixed color, rendering the blue–white test inoperative. The high instability phenotype was in large part caused by saturation of the mismatch repair system (15) because plating of the cells on minimal medium, where mismatch repair is fully active (15), lowered the rate of tract alteration to ≈10-fold that in the wild-type strain (data not shown).

To examine a possible role of SOS in microsatellite instability, a defective *lexA* allele was introduced in the wild-type strain carrying inserts. Constitutive derepression of the *lexA* operon increased instability 8- to 22-fold, and transduction of *lexA* (Def) strains back to $lexA^+$ restored the wild-type level of tract alteration (data not shown). No further increase was observed when the SOS pathway was fully induced by introduction of a *recA730* allele in a *lexA* (Def) strain. Derepression of the *lexA* operon without RecA activation is therefore sufficient to provoke the destabilization of simple repetitive DNA.

Mutations in *uvrD* and *rep* (28), *polA* (29), and *dam* (30) are known to induce the SOS pathway to various extents. A *lexA* noninducible allele then was introduced in the *polA1*, Δ*polA* (F′ Klenow), *uvrD*, and *mutS* strains carrying one of the inserts, $(AC)_{51}$. It should be noted that the noninducible *lexA* allele could not be introduced in the *dam* strain, because *dam lexA* double mutants are inviable (31). The *lexA* (Ind$^-$) mutation did not affect instability in the *mutS* strain. In contrast, instability decreased to the same level as that in the wild-type strain carrying the same insert for the *polA1* and *uvrD* strains and to an intermediate level for the Δ*polA* (F′ Klenow) strain (Table 4), indicating that all or part of the effects of these mutations results from SOS induction. The *lexA* (Ind$^-$) allele also reduced by 10-fold the spontaneous mutation rate of the *uvrD* strain (data not shown).

The nature of the tract alteration events was examined by sequencing clones derived from the wild-type strain and

Table 4. Effect of mutations on the frequency of tract alteration in the $(AC)_{51}$ insert in the absence of SOS induction

Relevant genotype	Tract alteration frequency*
lexA (Ind$^-$)	2.1×10^{-4} (1.0)
mutS lexA (Ind$^-$)	5.1×10^{-3} (24)
polA1 lexA (Ind$^-$)	2.8×10^{-4} (1.3)
Δ*polA* (F′ Klenow) *lexA* (Ind$^-$)	9.9×10^{-4} (4.7)
uvrD lexA (Ind$^-$)	3.2×10^{-4} (1.5)

*Tract alteration frequencies represent the average of at least three independent determinations. Numbers in parentheses represent the rates of tract length alteration (fold) relative to the rate in the *lexA* (Ind$^-$) strain.

Table 5. Types of alterations in $(AC)_{51}$ in wild-type and mutated strains

Relevant genotype	Tracts sequenced, *n*	Deletions*				Insertions*			
		>−4†	−4	−2	−1	+1	+2	+4	>+4†
Wild type	21	3	2	4	10	1	0	0	1
lexA(Def)*recA730*	23	1	1	2	15	2	0	2	0
mutS	21	0	0	2	13	5	1	0	0
uvrD	23	0	0	4	17	0	2	0	0
dam	20	0	0	0	18	1	1	0	0
Δ*polA* (F′ Klenow) *lexA* (Ind⁻)	20	7	1	0	4	7	0	1	0

*Expressed as the number of repeat units deleted (−) or added (+). The increase in insertions in the Δ*polA* (F′ Klenow) *lexA* (Ind⁻) strain compared to the wild-type strain is significant ($P = 0.027$), as determined by the hypergeometric test (47).
†Deletions larger than four repeat units were: −7, −8, and −25 in the wild-type strain; −11 in *lexA* (Def) *recA730*; −7 (2), −14, −16 (2), −19, and −28 in Δ*polA* (F′ Klenow) *lexA* (Ind⁻). The insertion larger than four repeats in the wild-type strain was +5.

from the relevant mutated strains, all carrying the $(AC)_{51}$ insert. In each case, 20–23 white clones derived from six independent blue colonies were analyzed (Table 5). Comparison with the sequence of the original insert led to the following observations: (*i*) In the wild-type, *lexA*(Def) *recA730*, *uvrD*, and *dam* strains, deletions represented close to 90% of the alteration events. (*ii*) In the *mutS* strain and in the strain lacking the 5′–3′ exonuclease of DNA polymerase I and carrying *lexA* (Ind⁻), the proportion of insertions increased to 28% and 40%, respectively. (*iii*) In the *mutS, uvrD*, and *dam* strains, alterations were remarkably homogeneous in size (one or two repeats), consistent with the inability of the mismatch repair system of *E. coli* to detect efficiently mismatches (and supposedly insertions and deletions) of >4 bp (32).

DISCUSSION

Most, if not all, studies of microsatellite stability in prokaryotes, as well as the majority of studies in yeast, have made use of a versatile plasmid or phage system (6–8, 14, 33–35). In the present work, we inserted poly(AC) and poly(TG) into the chromosome of a wild-type *E. coli* strain, by means of an appropriate λ ZAP vector carrying a thermosensitive repressor. Various mutations affecting replication, homologous recombination, excision, and postreplicative repair were introduced into the wild-type construct. We then could examine the variations in tract length without the added complication of phage or plasmid replication idiosyncrasies.

Our results (Fig. 2) confirm the existence of a direct relationship between the length of the insert and the frequency of tract alteration (6, 14), which also was reported recently in yeast (36). In addition, we observed a definite orientation effect because the instability of poly(TG) tracts on the leading strand is approximately twice as high as that of poly(AC) tracts of comparable length at the same location. In yeast, the rate of instability of poly(TG) tracts is independent of orientation (7). Mutations in *mutS* and *recA* (or their yeast homologs) have been examined previously for their effect on microsatellite sequences. Our results, reported in Tables 3 and 4, confirm the absence of effect of a *recA* mutation on tract instability and an important role of mismatch repair. The magnitude of the effect observed here for a *mutS* mutation varies from 27- to 160-fold, with no obvious relationship with the length or nature of the insert.

The introduction of a defective *lexA* allele (with or without concomitant RecA activation) into the wild-type strain carrying inserts substantially increased tract length alteration. As mentioned above, several mutations affecting DNA replication (*dam*, *rep, uvrD, polA*) are known to induce the SOS response in *E. coli* in the absence of DNA damage. We therefore examined microsatellite instability under conditions that precluded SOS induction. Two strains carrying a noninducible *lexA* gene in addition to the SOS-inducing mutation (*polA, uvrD*) displayed a level of tract instability close to that in the wild-type strain. These results show that these mutations cause microsatellite instability by an indirect mechanism dependent on SOS induction and suggest that instability in the *rep* and *dam* strains also depends on SOS. In the strain lacking the 5′–3′ exonuclease domain of polymerase I, introduction of the *lexA* (Ind⁻) allele reduced instability to an intermediate level, indicating that tract alteration is, in part, independent of SOS in this strain. Mutations in the *mutS* gene do not induce SOS (37). As expected, introduction of *lexA* (Ind⁻) in the *mutS* strain had no effect on tract instability.

The product of the *uvrD* gene is required for long patch mismatch repair, and the length of the alterations observed in the *uvrD* strain (Table 5) does reflect a mismatch repair defective phenotype. In contrast, the finding that the *uvrD* strain carrying *lexA* (Ind⁻) behaves like a wild-type strain with respect to microsatellite instability is difficult to reconcile with an essential role for UvrD in mismatch repair. However, UvrD is involved in a number of different processes, and the complexity of *uvrD* mutant phenotypes and genetic properties is notorious. As one example among many others, methyl-directed mismatch repair acts as barrier to interspecies recombination by efficiently destroying recombination intermediates between diverged sequences. Whereas mutations in *mutL* and *mutS* increase ≈1,000-fold, recombination between *E. coli* and *S. tyhimurium* mutations in *uvrD* have very little effect (38).

The extreme mutator effect in *mutD* cells is known to result for the most part from saturation of postreplicational mismatch repair by a high level of replication errors under rapid growth conditions (15). In slowly dividing *mutD* cells, full mismatch repair capacity is retained (15), as confirmed by our results with *mutD* cells carrying inserts plated on minimum medium. The level of instability in these cells (≈10-fold that of the wild-type strain) is probably a consequence of the proofreading defect. Therefore, the extremely high microsatellite instability in rapidly growing *mutD* cells (Table 3) appears to be because of a proofreading defect associated with a defect in mismatch repair.

RAD27 mutants of *S. cerevisiae* display rates of microsatellite instability comparable to those observed in mismatch repair mutants. The *RAD27* gene encodes a 5′–3′ exonuclease with structural and functional homology to the mammalian DNase IV (FEN1) and to the 5′ nuclease domain of *E. coli* DNA polymerase I, both known to function in lagging-strand synthesis. These nucleases remove by endonucleolytic cleavage a 5′ flap generated by displacement synthesis (19), for example when DNA polymerase reaches the 5′ end of an Okasaki fragment. We have shown in the present work that deletion of the 5′ nuclease domain of DNA polymerase I increases the rate of tract alteration and that this increase is in part independent of SOS induction. Because DNA polymerase I is not required for methyl-directed mismatch repair (39), the SOS-independent increase in tract instability must be because of the absence of a repair pathway distinct from mismatch repair. In the absence of SOS induction, deletion of the

5′–3′ exonuclease (strain PM16) led to a marked increase in expansions that, of interest, also was found in *RAD27* mutants (12). These results are consistent with a role for this class of 5′–3′ exonucleases in removing single-stranded flaps that, according to recent models (13, 20), could mediate expansions, for example by realignment out of register on the template. Thus, at least two factors of tract stability, mismatch repair and flap removal, appear to be conserved from prokaryotes to eukaryotes.

The identification of SOS as a factor raises questions concerning the mechanisms of destabilization. SOS induction could modify one or several subunits of DNA polymerase III holoenzyme in a way that favors slippage upon encounter of microsatellite sequences. Modification of the β subunit (sliding clamp), the γ complex (clamp loader), or the τ subunit (holoenzyme organizer) would be expected to affect the high processivity of the polymerase (40). SOS induction by UV irradiation has been shown to produce a smaller form of the β subunit (41). Also, it has been proposed that, on induction, UmuCD might serve as an alternate sliding clamp (42), less tightly bound to the template–primer than the bona fide β clamp. Whether the processivity of the polymerase is affected similarly when the SOS response is induced genetically in the absence of DNA damage remains to be determined.

Numerous genes are inducible by stress or DNA damage in yeast and mammals, and little is known about their regulation. However, recent findings suggest the existence in mammalian cells of a p53-mediated SOS response that, like the prokaryotic response, can be elicited in the absence of DNA damage (43). Also, *hMSH2* tumor cell lines were shown to accumulate tract alterations when maintained at high density (44). The authors point out that such growth conditions may resemble those encountered by cells in a tumor. They suggest that the conditional dramatic level of microsatellite instablity in *hMSH2* cells reflects a stress-induced, error-prone repair pathway as a consequence of replication under restrictive culture conditions in the absence of mismatch repair. Whether our results are relevant for tumor development will become clearer as the search for replication and repair mutants develops.

We are grateful to M. Dutreix for a generous gift of M13 mp18-$(AC)_{39}$ and M13 mp18-$(TG)_{33,}$ to M. Seigneur and H. Bierne for help with P1 transduction, to P. Bellenand for help with the statistical analysis, and to P. Noirot, T. Petes, and M. A. Petit for helpful comments and discussions. This work was supported by Grant 39/94 from the Groupement de Recherches et d'Etudes sur les Génomes (GREG) to E.C.

1. Richards, R. I. & Sutherland, G. R. (1994) *Nat. Genet.* **6,** 114–116.
2. Ashley, C. T., Jr., & Warren, S. T. (1995) *Annu. Rev. Genet.* **29,** 703–728.
3. Aaltonen, L. A., Peltomäki, P., Leach, F. S., Sistonen, P., Pylkkänen, L., Mecklin, J., Järvinen, H., Powell, S. M., Jen, J., Hamilton, S. R., *et al.* (1993) *Science* **260,** 812–816.
4. Thibodeau, S. N., Bren, G. & Schaid, D. (1993) *Science* **260,** 816–819.
5. Ionov, Y., Peinado, M. A., Malkhosian, S., Shibata, D. & Perucho, M. (1993) *Nature (London)* **363,** 558–561.
6. Levinson, G. & Gutman, G. A. (1987) *Nucleic Acids Res.* **15,** 5323–5338.
7. Henderson, S. T. & Petes, T. D. (1992) *Mol. Cell. Biol.* **12,** 2749–2757.
8. Strand, M., Prolla, T. A., Liskay, R. M. & Petes, T. D. (1993) *Nature (London)* **365,** 274–276.
9. Sia, E. A., Kokoska, R. J., Dominska, M., Greenwell, P. & Petes, T. D. (1997) *Mol. Cell. Biol.* **17,** 2851–2858.
10. Kroutil, L. C., Register, K., Bebenek, K. & Kunkel, T. A. (1996) *Biochemistry* **35,** 1046–1053.
11. Liu, B., Nicolaides, M. C., Markowitz, S., Wilson, J. K. W., Parsons, J. R., Jen, J., Papadopoulos, N., Peltomäki, P., de la Chapelle, A., Hamilton, S. R., *et al.* (1995) *Nat. Genet.* **9,** 48–55.
12. Johnson, R. E., Kovvali, G. K., Prakash, L. & Prakash, S. (1995) *Science* **269,** 238–240.
13. Tishkoff, D. X., Filosi, N., Gaida, G. M. & Kolodner, R. D. (1997) *Cell* **88,** 253–263.
14. Freund, A., Bichara, M. & Fuchs, R. P. P. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7465–7469.
15. Schaaper, R. M. & Radman, M. (1989) *EMBO J.* **8,** 3511–3516.
16. Sambrook, J., Fritsch, E. & Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
17. Lloyd, R. G. & Buchman, C. (1985) *J. Bacteriol.* **164,** 836–844.
18. Connelly, J. C. & Leach, D. R. F. (1996) *Genes Cells* **1,** 285–291.
19. Robins, P., Pappin, D. J. C., Wood, R. D. & Lindahl, T. (1994) *J. Biol. Chem.* **269,** 28535–28538.
20. Gordenin, D. A., Kunkel, T. A. & Resnick, M. A. (1997) *Nat. Genet.* **16,** 116–118.
21. Joyce, C. M. & Grindley, N. D. F. (1984) *J. Bacteriol.* **158,** 636–643.
22. Lane, H. E. D. & Denhardt, D. T. (1974) *J. Bacteriol.* **120,** 805–814.
23. Lane, H. E. D. & Denhardt, D. T. (1975) *J. Mol. Biol.* **97,** 99–112.
24. Cosalanti, J. & Denhardt, D. T. (1987) *Mol. Gen. Genet.* **209,** 382–390.
25. Modrich, P. (1991) *Annu. Rev. Genet.* **25,** 229–253.
26. Schaaper, R. M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8126–8130.
27. Fijalkowska, I. J. & Schaaper, R. M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2856–2861.
28. Ossanna, N. & Mount, D. W. (1989) *J. Bacteriol.* **171,** 303–307.
29. Bates, H., Randall, S. K., Rayssiguier, C., Bridges, B. A., Goodman, M. F. & Radman, M. (1989) *J. Bacteriol.* **171,** 2480–2484.
30. Marinus, M. G. & Norris, N. R. (1975) *Mutat. Res.* **28,** 15–26.
31. Peterson, K. R. & Mount, D. W. (1993) *J. Bacteriol.* **175,** 7505–7508.
32. Parker, B. O. & Marinus, M. G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1730–1734.
33. Bichara, M., Schumacher, S. & Fuchs, R. P. P. (1995) *Genetics* **40,** 897–907.
34. Bowater, R. P., Rosche, W. A., Jaworski, A., Sinden, R. R. & Wells, R. D. (1996) *J. Mol. Biol.* **264,** 82–96.
35. Ohshima, K., Kang, S. & Wells, R. D. (1996) *J. Biol. Chem.* **271,** 1853–1856.
36. Wierdl, M., Dominska, M. & Petes, T. D. (1997) *Genetics* **146,** 769–779.
37. Matic, I., Rayssiguier, C. & Radman, M. (1995) *Cell* **80,** 507–515.
38. Rayssiguier, C., Thaler, D. S. & Radman, M. (1989) *Nature (London)* **342,** 396–401.
39. Dzidic, S. & Radman, M. (1989) *Mol. Gen. Genet.* **217,** 254–256.
40. Yuzhakov, A., Turner, J. & O'Donnell, M. (1996) *Cell* **86,** 877–886.
41. Skaliter, R., Paz-Elizyr, T., Livneh, Z. (1996) *J. Bacteriol.* **271,** 2478–2481.
42. Battista, J. R., Nohmi, T., Donnelly, C. E. & Walker, G. C. (1988) in *Mechanisms and Consequences of DNA Damage Processing,* eds. Friedberg, E. C. & Hanawalt, P. C. (Liss, New York), pp. 455–459.
43. Eller, M. S., Maeda, T., Magnoni, C., Atwal, D. & Gilchrest, B. A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12627–12632.
44. Richards, B., Zhang, H., Phear, G. & Meuth, M. (1997) *Science* **277,** 1523–1526.
45. Gibson, F. P., Leach, D. R. F. & Lloyd, R. G. (1992) *J. Bacteriol.* **174,** 1222–1228.
46. Shapira, S. K., Chou, J., Richaud, F. V. & Casabadan, M. J. (1983) *Gene* **25,** 71–82.
47. CETAMA (1986) *Statistique appliquée à l'exploitation des mesures* (Masson, Paris), p. 29.