High-accuracy lagging-strand DNA replication mediated by DNA polymerase dissociation

Katarzyna H. Maslowska[a,b], Karolina Makiela-Dzbenska[b], Jin-Yao Mo[a], Iwona J. Fijalkowska[b,1], and Roel M. Schaaper[a,1]

[a]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709; and [b]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-106 Warsaw, Poland

Edited by Philip C. Hanawalt, Stanford University, Stanford, CA, and approved March 2, 2018 (received for review November 27, 2017)

The fidelity of DNA replication is a critical factor in the rate at which cells incur mutations. Due to the antiparallel orientation of the two chromosomal DNA strands, one strand (leading strand) is replicated in a mostly processive manner, while the other (lagging strand) is synthesized in short sections called Okazaki fragments. A fundamental question that remains to be answered is whether the two strands are copied with the same intrinsic fidelity. In most experimental systems, this question is difficult to answer, as the replication complex contains a different DNA polymerase for each strand, such as, for example, DNA polymerases δ and ε in eukaryotes. Here we have investigated this question in the bacterium *Escherichia coli*, in which the replicase (DNA polymerase III holoenzyme) contains two copies of the same polymerase (Pol III, the *dnaE* gene product), and hence the two strands are copied by the same polymerase. Our in vivo mutagenesis data indicate that the two DNA strands are not copied with the same accuracy, and that, remarkably, the lagging strand has the highest fidelity. We postulate that this effect results from the greater dissociative character of the lagging-strand polymerase, which provides additional options for error removal. Our conclusion is strongly supported by results with *dnaE* antimutator polymerases characterized by increased dissociation rates.

DNA replication fidelity | leading and lagging strands | DNA polymerase III holoenzyme | *dnaE* antimutators | DNA polymerase dissociation

The accuracy by which organisms are able to duplicate their chromosomal DNA is generally high, producing on average only one error for every 10^9–10^{11} copied bases. This high fidelity is not achieved in a single step, but instead is produced via the operation of several sequential error-avoidance and editing steps. These steps include selection of the correct DNA base by the DNA polymerase (i.e., nucleotide insertion step), the editing (i.e., removal) of polymerase misinsertion errors by exonucleolytic proofreading, and finally postreplicative DNA mismatch repair, which detects and corrects DNA mismatches in newly replicated DNA (1).

Chromosomal DNA replication is generally performed by multisubunit DNA polymerase complexes (replicases) that conduct the simultaneous, coordinated synthesis of the two new DNA strands at the replication fork. The replication speed can be very high, up to 500–1,000 nucleotides per second for the bacterium *Escherichia coli* (2). In this organism, replication is performed by the multisubunit DNA polymerase III holoenzyme (HE) complex (2), which contains two copies of a DNA polymerase core, one for each strand, with composition αεθ, in which the α subunit is the DNA polymerase and ε is the associated exonucleolytic proofreader. The two core polymerases are joined together through interactions with the τ subunit of the $DnaX_5$ complex. The latter also serves as a loader/unloader complex for two donut-shaped (β_2) processivity clamps that tether the two polymerases to the DNA. The speed of the fork is controlled through interactions with the DNA helicase (DnaB gene product) that unzips the DNA ahead of the moving fork (2). How such high-speed replication machines can copy DNA with impressive fidelity is an important question.

Detailed mutational studies in *E. coli* have suggested that the in vivo polymerase misinsertion rate is approximately 10^{-4}–10^{-5} per base copied, and that proofreading by the ε subunit reduces the error rate by approximately 10^2-fold, yielding an overall polymerase error rate of 10^{-6}–10^{-7} per replicated base (1). DNA mismatch repair (via the *mutHLS* system), which follows the replication fork, then reduces the observed error rate by another 10^2- to 10^3-fold, accounting for the overall mutation rate of approximately 10^{-10} per replicated base (1, 3). However, many aspects of the replisome and its fidelity remain incompletely understood, including its precise composition (4, 5), the fidelity role of other HE subunits (6), the interference of the accessory DNA polymerases (7), and the efficiency of exonucleolytic proofreading (8–10).

In the present work, we have addressed another fundamental question: whether the two DNA strands—leading and lagging—are subject to the same fidelity rules. Insight into this matter may also help understand the origin of DNA strand biases observed in studies of mutagenesis, evolution, and cancer (11–13).

Results and Discussion

A System to Investigate in Vivo Chromosomal DNA Replication Fidelity. To study any differential effects of leading and lagging-strand replication on chromosomal replication fidelity, we developed the system shown schematically in Fig. 1*A*. Mutations are scored in the 1,100-bp *lacI* gene, encoding the repressor of the *lacZYA* operon. Forward mutations occurring throughout the *lacI* gene, inactivating the repressor function, lead to constitutive expression

Significance

The accuracy (fidelity) by which cells are able to duplicate their chromosomal DNA before cell division is an important factor in the frequency at which they accumulate mutations. Because mutations are generally harmful, organisms have developed various mechanisms to minimize the frequency of errors during DNA replication. Replication is generally performed by large multiple subunit complexes (replicases), which simultaneously and in a coordinated fashion copy the two DNA strands. Due to the antiparallel nature of the two DNA strands, the replication enzymology of the two individual strands is slightly different, and our study demonstrates that the two strands are copied with different accuracies. Specifically, the discontinuous lagging strand is significantly more accurate than the more processive leading strand.

Author contributions: I.J.F. and R.M.S. designed research; K.H.M., K.M.-D., and J.-Y.M. performed research; K.H.M., K.M.-D., I.J.F., and R.M.S. analyzed data; and R.M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: iwonaf@ibb.waw.pl or schaaper@niehs.nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1720353115/-/DCSupplemental.

Published online April 2, 2018.

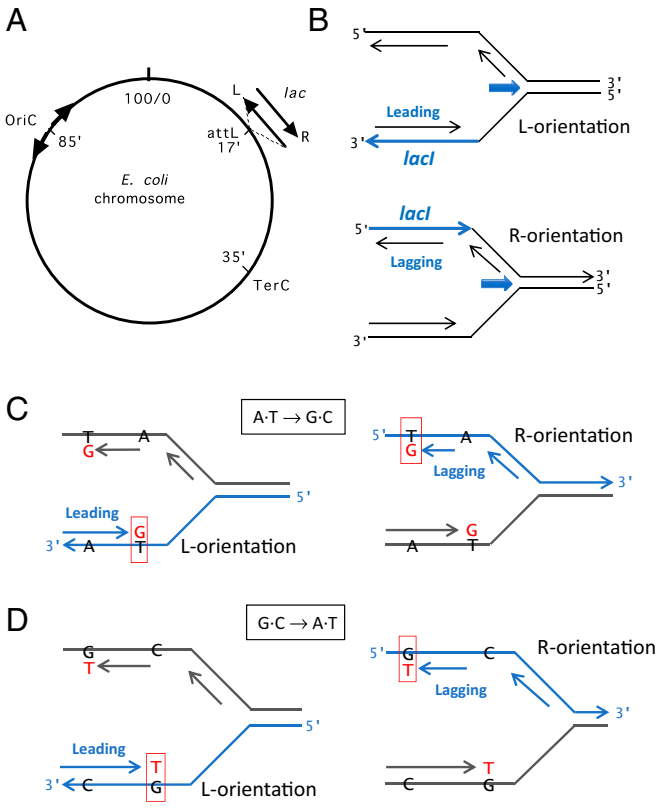


Fig. 1. A system for assaying differential replication fidelity in leading and lagging strands. (*A*) *E. coli* chromosome with the *lac* operon inserted at the *attL* locus in either the R or the L orientation. Also shown are the OriC origin, where bidirectional replication begins (the two arrowheads), and the TerC terminator region, where the two forks eventually meet. As shown, the *lac* operon will be copied by the rightward (clockwise) fork. (*B*) Diagrams showing how the replication fork copies the *lac* genes in the two gene orientations. The coding (or reference) strand of the *lacI* gene (in blue) is copied by the leading strand machinery when in the L orientation and by the lagging-strand machinery when in the R orientation. (*C*) Analysis of A·T → G·C mutations resulting from T·G replication errors ($T_{template}$·G mispairings). When annotated in the *lacI* coding strand, A·T → G·C can be scored at T sites (T → C) or A sites (A → G). However, as shown on the left for the L orientation, mutations at T sites result from leading strand T·G errors, and those at A sites result from lagging-strand T·G errors (red boxes). For the R orientation, the reverse applies. Thus, the comparison of A·T → G·C mutations at coding strand T (T → C) and A sites (A → G) is a convenient measure of the strand-dependent occurrence of T·G errors. (*D*) As in *C*, but for G·C → A·T transitions mediated by G·T errors. Here the comparison between mutations at coding strand G vs. C sites is a measure of the strand dependence of the G·T errors.

of the operon. Such constitutive mutants can be readily selected by their ability to grow on solid media containing the sugar phenyl-β-D-galactoside (P-gal) as the sole carbon source (14, 15). The *lacI* gene has been used as a mutational target in many previous studies, and considerable knowledge about the types of *lacI* mutations leading to inactivation has been accumulated (9, 15–22). In the system shown in Fig. 1*A*, the *lac* genes (*lacI* and the immediately adjacent *lacZYA*) are deleted from the normal location (at ∼8′ of the *E. coli* map), but are reintroduced at the phage lambda attachment site (*attL* at ∼17′) in the two possible orientations, which we arbitrarily label R and L (Fig. 1*A*).

Replication of the *E. coli* chromosome initiates at the single OriC origin (∼85′) producing two replication forks traveling in opposite directions until ultimately meeting at the chromosomal terminus (∼35′). Within each fork, one DNA strand, called the leading strand, is replicated continuously in the same direction as the moving fork, while the other (lagging) strand is replicated in the opposite direction in the form of short Okazaki fragments. Fig. 1*B* illustrates the consequences of the inversion in the way that the operon is replicated. For the R orientation, the *lacI* coding sequence (indicated in blue) is copied by the lagging-strand mechanism, while for the L orientation, it is copied by the leading strand replication.

Our experimental system to determine possible differences in replication fidelity between the leading and lagging strands is based on a detailed analysis of the frequencies and types of *lacI* mutations occurring for the two gene orientations. For example, the most prominent replication error made by DNA polymerases is the misincorporation of dGTP opposite a template T, denoted here as T·G (template strand underlined). This event will ultimately be observed as an A·T → G·C transition mutation. Using the *lacI* coding sequence as a reference, these mutations can be scored at both T sites (read as T → C) and A sites (read as A → G) within this sequence, but their genesis at the two sites is different. For example, in the L orientation (Fig. 1*C*), the mutations scored at T sites (T → C) result from T·G errors in the leading strand, while those scored at A sites (A → G) result from T·G errors in the lagging strand, and the reverse is true for the R orientation (Fig. 1*C*). Thus, insight into the relative frequency of T·G errors during leading strand and lagging-strand replication can be obtained by simply comparing the ratio of these transitions at coding strand T and A sites.

Likewise, for the case of the reciprocal G·C → A·T transitions, analysis in terms of the preferred G·T mispairings can be performed simply by comparing events at reference strand G and strand C. Referring to Fig. 1*D*, for the L orientation, G·C → A·T observed at coding strand G sites (G → A) correspond to G·T errors in the leading strand, while those observed at C sites (C → T) correspond to G·T in the lagging strand, and vice versa for the R orientation. A caveat for this overall approach is that A·T → G·C mutations can alternatively result from A·C mispairings, while G·C → A·T can result from C·A errors. The logic and resulting strand assignments would lead to exactly the opposite conclusion. However, while A·C and C·A errors can undoubtedly occur, their frequency, based on DNA polymerase studies in vitro (23–28), is generally significantly lower than that of T·G or G·T errors.

Differential Replication Fidelity in Leading and Lagging Strands. Here we describe the DNA sequence analysis of 1,366 *lacI* mutations throughout the 1,100-bp *lacI* gene target. The strains used were defective in postreplicative DNA mismatch repair (carrying a *mutL* defect), so that the observed mutations may be analyzed straightforwardly in terms of uncorrected DNA replication errors. The data in Table 1 show that in the *mutL* strain, mutations occur at essentially identical frequency for the R and L orientations (14.7 and 14.3 × 10^{-6}, respectively). Unchanged frequencies are also observed for the various subclasses of mutations: base substitutions, transitions (i.e., A·T → G·C or G·C → A·T), transversions, and −1 frameshift mutations. This finding is as expected for a large mutational target in which the DNA sequences of the two DNA strands (coding and noncoding) have similar intrinsic mutational potential. Overall, transitions (A·T → G·C or G·C → A·T) significantly outnumber transversions, a well-established feature of mismatch repair-defective strains (18) indicative of the predominant nature of primary replication errors.

Despite the overall similarity noted for R- and L-oriented *lacI* genes (Table 1), significant differences become apparent when analyzing the spectra of individual A·T → G·C or G·C → A·T mutations. The results of this analysis are summarized in Fig. 2*A*, and the complete sequence data are provided in Fig. S1 and Dataset S1. Fig. 2*A* shows that A·T → G·C mutations are unequally distributed over coding strand T sites (T → C) and A sites (A → G). More remarkably, inversion of the gene orientation switches the bias to the opposite direction. Thus, for the L orientation, T → C mutations are more frequent than A → G (64 vs.

GENETICS

Table 1. Number and frequency of sequenced *lacI* mutations in L and R chromosomal *lac* orientations in *mutL* and *mutL dnaE915* backgrounds

	mutL		*mutL dnaE915*	
Mutation	L orientation	R orientation	L orientation	R orientation
Total	349 (14.3)	357 (14.7)	343 (7.23)	317 (7.23)
Base substitutions	296 (12.1)	292 (12.0)	161 (3.39)	198 (4.52)
Transitions	283 (11.7)	286 (11.8)	141 (2.97)	182 (4.15)
A·T → G·C	100 (4.10)	118 (4.86)	73 (1.54)	67 (1.53)
G·C → A·T	183 (7.50)	168 (6.92)	68 (1.43)	115 (2.62)
Transversions	13 (0.53)	6 (0.25)	20 (0.42)	16 (0.36)
Frameshifts (Fs)	52 (2.13)	64 (2.64)	181 (3.82)	118 (2.69)
Fs (−1)	38 (1.46)	43 (1.77)	125 (2.63)	83 (1.89)
Fs (+1)	14 (0.57)	21 (0.86)	56 (1.18)	35 (0.80)

Calculations are based on the DNA sequencing results of a total of 1,366 independent *lacI* mutants in the indicated strains. The full listing of the sequenced mutations is provided in Dataset S1 and shown graphically in Figs. S1 and S2. The frequency (in parentheses, $\times 10^{-6}$) for each subcategory is derived by multiplying the overall mutant frequency by the observed fraction of each subcategory.

36), but in the R orientation they are less frequent (33 vs. 85). A similar reversal occurs for the G·C → A·T mutations; in the L orientation, G → A mutations are more frequent than C → T mutations (112 vs. 71), while in the R orientation, this is reversed (43 vs. 125).

In both cases, the observed switches are fully consistent with a differential fidelity of leading and lagging-strand DNA replication, with the latter being more accurate. For A·T → G·C mutations, changing the orientation from L to R moves the coding strand T·G errors (red boxes) from the leading strand to the lagging strand (Fig. 1*C*), which is associated with a decrease in mutant frequency. Likewise, for the G·C → A·T mutations, reversal of the orientation from L to R moves the coding strand G·T errors (red box) from the leading strand to the lagging strand (Fig. 1*D*), which is also associated with decreased mutant frequency. Thus, for both T·G and G·T errors, lagging-strand replication is more accurate. The higher accuracy of lagging-strand replication had been suggested earlier based on a limited number of reversion systems (29, 30), but the present data prove that this higher accuracy is observed for a large number of sites throughout a large gene and can be assumed to apply genome-wide.

A Mechanism for High Accuracy of Lagging-Strand Replication. What is the possible basis for the greater accuracy of lagging-strand DNA synthesis on the *E. coli* chromosome? We suggest that it is related to the greater "dissociability" of the lagging strand half of the polymerase complex. The lagging-strand DNA polymerase dissociates from its primer-terminus repeatedly every ∼500 bp when reaching the end of the Okazaki fragment. From there, it rapidly recycles to the next available 3′ primer terminus approximately every 1 s (31–33). The signal that causes the lagging-strand polymerase to dissociate has been shown to be the presence of a newly produced upstream primer (34), and delays in the progress of the lagging-strand polymerase lead to polymerase dissociation before the Okazaki fragment is completed (34). As it is well established that 3′-terminal mispairs are much more difficult to extend than correct base pairs (27, 28, 35–37), and this has been well demonstrated for *E. coli* DNA polymerase III (10, 21, 36, 37), it is reasonable to assume that delays in lagging-strand synthesis at polymerase errors may lead to dissociation from the mismatch. These dissociation events must be considered clear fidelity events, because any abandoned terminal mispairs are unlikely to survive as a mutation and they will be ready prey for any 3′-exonuclease, either free or DNA polymerase-associated (7), including, for example, DNA polymerase II, which can act as back-up proofreader for HE (38). In contrast, in the more processive leading strand, terminal mismatches, when not removed by the exonucleolytic proofreading, may have no other fate than being extended. This will fix the mismatch as a potential mutation and account for the higher error rate for this strand. Recent studies have suggested that leading strand synthesis might not be as continuous as previously thought (39, 40), but our proposed mechanism would apply as long as a sufficient processivity difference exists between the two strands.

As an alternative to the proposed dissociative mechanism, we have considered the possibility that other, strand-specific factors or proteins influence the intrinsic polymerase accuracy, and we cannot fully exclude such alternative possibilities. Importantly, we note that the fidelity effect of strand inversion applies equally to transcribed and nontranscribed *lacI* strands (Fig. 2*A*), arguing against a possible role for transcription/replication encounters. In addition, the *lacI* gene is generally poorly expressed, and such replication/transcriptions encounters are likely to be rare. Below, we describe several additional experiments in further support of the dissociative mechanism.

***dnaE* Antimutators as Tools to Investigate Replication Fidelity.** We sought further evidence for a dissociative lagging-strand fidelity mechanism by analyzing the strandedness of mutagenesis in the *E. coli dnaE915* antimutator strain. Antimutators are mutants with a lower mutation rate than the WT strain (41). Our laboratory has been successful in isolating many *E. coli* antimutator mutants (42), and we have found the responsible defects to map to the *dnaE* gene (43), encoding the alpha (polymerase) subunit of DNA polymerase III (44). As assayed by a number of reversion assays, these *dnaE* alleles are able to reduce replication errors by several-fold (42, 45, 46). Fig. 3 shows the locations of 28 sequenced *dnaE* antimutator mutations. The *dnaE915* antimutator carries the Ala498Thr mutation (43). Importantly, the underlying *dnaE* amino acid substitutions are widely distributed throughout the central part of the polymerase, including the palm, thumb, and finger domains that compose the catalytic portion of the polymerase (47). It is unlikely that any of these mutations exert their effect directly through improvement of the polymerase insertion fidelity, as such base-selection specific effects are restricted to a limited number of residues within the catalytic pocket responsible for ensuring the proper geometry of the nascent terminal base pair (35, 48). Instead, the altered residues are thought to affect the catalytic performance or stability of the enzyme. In other words, they are impaired polymerases that achieve reduced error rates indirectly. For example, reduced polymerase stability could directly lead to an enzyme with enhanced dissociation probability (i.e., reduced processivity). A reduced catalytic forward rate from a terminal mismatch would also enhance, indirectly, the

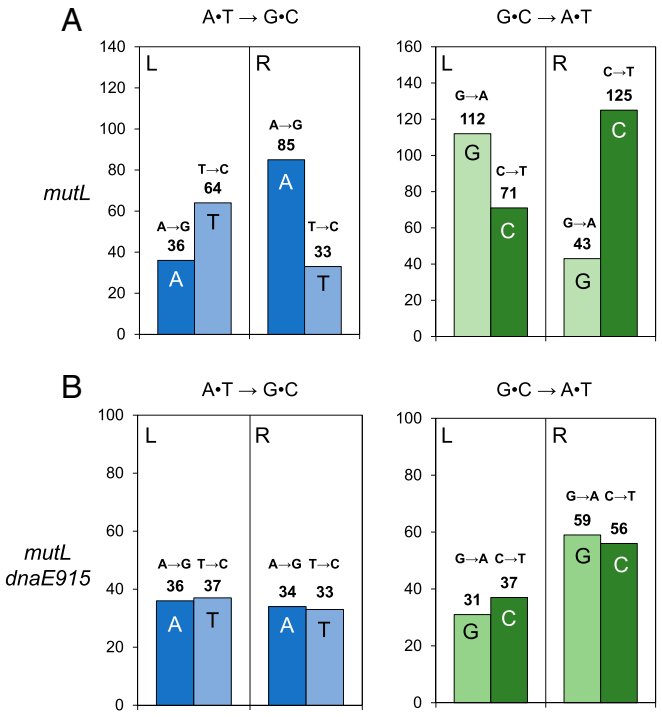


Fig. 2. Effect of gene reversal on the distribution of A·T → G·C and G·C → A·T mutations across the *lacI* gene, indicative of greater accuracy of lagging-strand replication. (*A*) In the *mutL* background, A·T → G·C (*Left*) occur at unequal frequencies at coding strand A sites (A → G) and T sites (T → C), and their ratio inverts with gene orientation ($P = 1.48 \times 10^{-7}$, Fisher's exact test for L vs. R orientation). Likewise, G·C → A·T mutations (*Right*) occur at unequal frequencies at C and G sites, and their ratio inverts with gene orientation ($P = 1.63 \times 10^{-11}$). As explained in the text (see also the diagrams in Fig. 1 *C* and *D*), this orientation dependence is consistent with the highest fidelity for the lagging-strand replication. Regarding the possible role for transcription, A·T → G·C transitions (resulting from T·G errors) are represented by A → G events in the transcribed strand and by T → C events in the nontranscribed (reference) strand. Combined with the frequency data of Table 1, the lagging-strand fidelity advantage is similar for nontranscribed and transcribed strands (1.9- and 2.4-fold, respectively). Similarly, for G·C → A·T mutations caused by G·T errors, the lagging-strand advantage is seen for both nontranscribed and transcribed strands (2.6- and 1.8-fold, respectively). (*B*) In the *mutL dnaE915* (antimutator) background, the orientation bias is no longer observed, suggesting that the fidelity difference between leading and lagging strands has been diminished or lost (see text for details).

likelihood of dissociation and proofreading, in all cases leading to lower error rates.

We have purified alpha subunits from WT and antimutator isolates and confirmed several of these assumptions biochemically. These data are provided in Tables S1–S3. The data confirm that the DnaE915 (Ala498Thr) alpha subunit has reduced specific activity, indicative of impaired polymerase activity, while processivity measurements, reflecting the probability of the polymerase terminating synthesis at various template positions, show that the DnaE915 polymerase has reduced processivity and thus is more dissociative. In complex with the ε (proofreading) subunit, the DnaE915 polymerase displays a significantly enhanced turnover ratio, again reflective of catalytically impaired polymerase, and so more frequent dissociation and enhanced proofreading likely contribute to the antimutator effect. Overall, it follows that the increased dissociability of the *dnaE915* allele is a promising tool for testing our model.

The mutant frequency results presented in Table 1 show that the *dnaE915* allele is indeed a clear antimutator in the *lacI* forward system. The effect is approximately twofold for either *lac* orientation [7.23×10^{-6} vs. $(14.3–14.7) \times 10^{-6}$]. For the group of base pair substitutions, the antimutagenic effect is even slightly larger, 3.6-fold for the L orientation (12.9 vs. 3.39) and 2.7-fold for the R orientation (12.0 vs. 4.52) (Table 1). Interestingly, analysis of the strand dependence of the base pair substitutions (using the complete spectra presented in Fig. S2) shows that the leading/lagging-strand bias for the base pair substitutions has essentially disappeared, with equal numbers of mutations now occurring in each strand (Fig. 2*B*). Mutant frequency calculations (Table 2) indicate that in fact mutations in both strands are reduced, but this effect is larger for the more error-prone leading strand, resulting in a loss of the difference in fidelity between the two strands. Table 2 shows an average 4.8-fold reduction for leading strand events (average of four leading strand entries), compared with an average 2.2-fold effect for the lagging-strand events. In work to be published elsewhere, we report that this is a general feature of the *dnaE* antimutator alleles, i.e., they reduce leading strand errors more than lagging-strand errors. As one example, using a set of *lacZ* reversion assays, the *dnaE941* antimutator (L611F, Fig. 3), which was initially isolated as a suppressor of the *mutT* mutator (49) but is also a strong antimutator for DNA replication errors, reduces *lacZ* G·C → A·T transitions by approximately 25-fold for leading strand vs. 7-fold for lagging strand, and reduces *lacZ* A·T → G·C transitions by approximately 23-fold for leading strand vs. 6-fold for lagging strand.

The present data are highly consistent with our model in which dissociative DNA polymerases improve fidelity, and point to the differential effects of dissociative features on the two strands. While polymerase dissociation may occur in both strands, the relative importance of this step is clearly different between the strands. In the leading strand, dissociation may be a minor step, but increased dissociation has a disproportionately large effect, perhaps because dissociation, in addition to proofreading, has now become a rate-contributing fidelity step. In the lagging strand, dissociation may already be a rate-determining fidelity

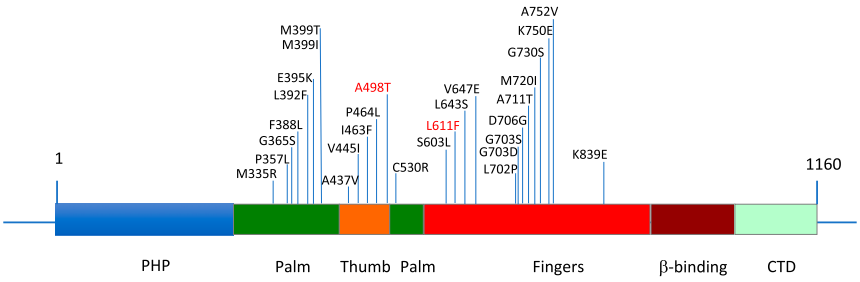


Fig. 3. Location of *dnaE* antimutator mutations across the *dnaE* gene, encoding the alpha (polymerase) subunit of DNA polymerase III. The indicated domains of the polymerase are as described by Lamers et al. (47). The amino acid substitutions of 9 of the 28 indicated *dnaE* alleles have been reported previously (43, 46); the remainder are described here. The *dnaE915* (A498T) and *dnaE941*(L611F) alleles used in this study are in red.

Table 2. *dnaE915*-mediated antimutator effects for leading and lagging-strand replication

Mutation	Strand	L orientation	R orientation
A·T → G·C	Leading (T → C)	3.4	4.5
	Lagging (A → G)	1.95	1.8
G·C → A·T	Leading (G → A)	7.1	4.0
	Lagging (C → T)	3.7	1.3

Fold antimutator effects are calculated from the overall *dnaE915* antimutator effects obtained from the frequencies of Table 1 and adjusted by the leading/lagging-strand distributions detailed in Fig. 1. For example, for A·T → G·C in the L orientation, the overall *dnaE915* antimutator effect is 2.66 (4.10/1.54) (Table 1). Of these, leading strand mutations (T → C) were 64/100 in the *mutL* strains and 37/74 in the *mutL dnaE915* strains (Fig. 1). Thus, the specific leading strand antimutator effect was 2.66 × (64/100)/(37/74) = 3.4-fold. For the lagging-strand mutations in this orientation (A → G), the specific effect was 2.66 × (36/100)/(36/73) = 1.95. All other numbers in the table are calculated similarly.

step, and only a modest effect occurs. It is important to note that the antimutator strains display generally normal viability, and thus their in vivo replication defect must be a subtle one.

***dnaE* Antimutators Become Mutators in the Presence of Error-Prone DNA Polymerases.** Further evidence for the dissociative model can be derived from the seemingly counterintuitive mutator effect displayed by *dnaE* antimutators, at least in certain special cases. For example, we noted a slight mutator effect for frameshift mutations in the *dnaE915* strain (Table 1). To account for this effect, it may be argued that terminal mispairs, when abandoned or in the process of being abandoned, may be subject to potentially mutagenic primer–template misalignments in suitable DNA sequences (50), and this propensity to misalign and create more easily extendable substrates has been well documented for Pol III (21, 37). Terminal mispairs resulting from dissociation could also become a prey for error-prone accessory DNA polymerases, such as *E. coli* Pol IV or Pol V (51–53). These enzymes are normally kept at low (Pol IV) or undetectable (Pol V) levels, but can become induced as part of the bacterial SOS system. Because they lack 3′-exonuclease activity, their access to and extension of a mismatched primer is likely to be mutagenic. Thus, a clear further prediction can be made. While the *dnaE* alleles like *dnaE915* are antimutators under normal conditions, they can be predicted to behave as mutators in the presence of increased amounts of Pol IV or Pol V. We have tested and confirmed this prediction.

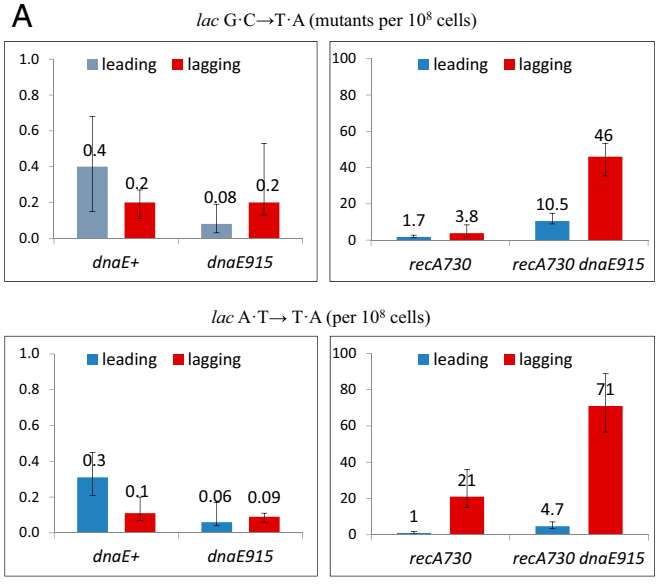
In the experiment of Fig. 4, we induced the constitutive presence of Pol V using the regulatory *recA730* allele (45, 54), and measured mutagenesis using two separate *lac* reversion assays that score specifically a G·C → T·A or an A·T → T·A base pair substitution (55). For both cases, it is clear that while in the rec^+ background, *dnaE915* causes substantial antimutator effects (Fig. 4*A*, *Left*), in the *recA730* background, it produces substantial mutator effects (Fig. 4*A*, *Right*). This experiment allows for several distinct and internally consistent conclusions: (*i*) in the $dnaE^+$ background, *dnaE915* acts as a clear antimutator for both *lac* reversions; (*ii*) this antimutator effect is largely specific for the leading strand; (*iii*) the *recA730* allele causes a strong mutator effect due to the constitutive presence of Pol V; (*iv*) this *recA730* mutator effect is largely specific for the lagging strand (Fig. 4*B*), consistent with the preferred presence of abandoned terminal mispairs in this strand; (*v*) in the *dnaE915* background, the Pol V mutator effect is greatly enhanced, consistent with the now-increased number of available terminal mismatches; and (*vi*) in the *dnaE915* background, the *recA730* mutator effect is now broadly seen for both leading and lagging strands (Fig. 4*B*).

Conclusion. The results presented here strongly indicate that on the *E. coli* bacterial chromosome, the lagging strand is synthesized with several-fold greater accuracy than the leading strand, and that this phenomenon is related to and caused by the increased dissociability of the lagging-strand polymerase. This increased dissociability can have both positive and negative fidelity consequences and thus is naturally limited in extent. Our conclusion is based on an analysis of mutations occurring at >100 detectably mutable sites throughout the 1,100-bp *lacI* gene, and can be safely assumed to be valid chromosome-wide. Because in essence, all nonviral chromosomes are subject to separate leading and lagging-strand synthesis, this result has broad significance. Even in cases where different DNA polymerases are responsible for synthesis of the two strands, an intrinsic strand-dependent fidelity affect should contribute.

Materials and Methods

Bacterial Strains and Culture Conditions. The *E. coli* strains and the media used in this study are described in *SI Materials and Methods*. All strains are derivatives of strain MC4100 and carry the *lacIZYA* genes inserted in the phage lambda attachment site (*attB*) in the two possible orientations (*lac*R and *lac*L) (30). For recording *lacI* spectra, we used the WT *lac* operon derived from strain NR9102 (17), while for analyzing *lacZ* reversion frequencies, we used the *lac* operon from strains CC104 or CC105 carrying specific *lacZ* missense mutations (55). All strains were also mismatch repair-deficient (*mutL*::Tn*5*) to facilitate scoring of DNA replication errors.

A



B

Fold-mutator effect of *recA730* in *dnaE*⁺ and *dnaE915* backgrounds

Mutation	Strand	*dnaE*⁺	*dnaE915*
lac G·C→T·A	leading	4.2	137.5
	lagging	19	230
lac A·T→T·A	leading	3.3	78
	lagging	210	789
Rif^r	-	0.8	7.3

Fig. 4. *dnaE* antimutators reveal themselves as mutators in the presence of an error-prone DNA polymerase. (*A*) *lac* reversion assays for G·C → T·A (*Upper charts*) and A·T → T·A (*Lower charts*) transversions (revertant frequency per 10^8 cells). All strains are *mutL*::Tn*5*. The leading/lagging-strand distinction is based on C·T (for G·C → T·A) or T·T mispairings (A·T → T·A), as described previously (29). The results show that *dnaE915* acts as an antimutator in the $dnaE^+$ background (*Left*), but as a mutator in the *recA730* background (*Right*). Note the different scales in the two panels. Mutant frequencies are median frequencies for 10–15 independent cultures, with error bars representing 95% confidence intervals. (*B*) Calculated *recA730* mutator effects (i.e., fold increases in mutant frequency relative to the corresponding rec^+ control), along with results for the frequency of rifampicin-resistant mutants (rif^r), which are not subject to gene inversion. As shown previously, in the $dnaE^+$ background, the *recA730* mutator effect has a strong preference for the lagging strand, but in the *dnaE915* background, both strands show strong mutator effects.

Mutant Frequency Determinations and Mutant Isolation. In brief, for recording *lacI* mutant spectra 384 independent LB cultures were used for each strain and *lac* orientation. Aliquots of the grown cultures were spread on plates containing P-gal (phenyl-β-D-galactoside) as carbon source selecting for mutants with constitutive *lac* expression (*lacI*). One mutant was picked randomly from each culture for DNA sequencing. *LacZ* revertant frequencies were determined from 10 to 15 independent LB cultures by selection of lac^+ mutants on minimal lactose plates. More details are provided in *SI Materials and Methods*.

DNA Sequencing. DNA sequencing of the entire *lacI* gene was performed as described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Kevin Gerrish [National Institute of Environmental Health Sciences (NIEHS) Molecular Genomics Core Laboratory] and John Otstot and Greg Solomon (NIEHS Epigenomics Core Laboratory) for assistance with the sequencing of the *lacI* mutants, Dr. Charles McHenry for providing plasmids and advice on protein purification; Dr. Scott Lujan for his expert help with the mutant database handling; and Drs. Katie Glenn and Bradley Klemm (NIEHS) for their helpful comments on the manuscript. This research was supported by the NIEHS Intramural Research Program (Project Z01 ES065086) and the Foundation for Polish Science (International PhD Projects Program MPD/2009-3/2, "Studies of Nucleic Acids and Proteins–from Basic to Applied Research," and Grant TEAM/2011-8/1, "New Players Involved in the Maintenance of Genomic Stability," cofunded by the European Union Regional Development Fund).

1. Schaaper RM (1993) Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem* 268:23762–23765.
2. Lewis JS, Jergic S, Dixon NE (2016) The *E. coli* DNA replication fork. *Enzymes* 39:31–88.
3. Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA* 88:7160–7164.
4. Dohrmann PR, Correa R, Frisch RL, Rosenberg SM, McHenry CS (2016) The DNA polymerase III holoenzyme contains γ and is not a trimeric polymerase. *Nucleic Acids Res* 44:1285–1297.
5. Reyes-Lamothe R, Sherratt DJ, Leake MC (2010) Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* 328:498–501.
6. Gawel D, Jonczyk P, Fijalkowska IJ, Schaaper RM (2011) *dnaX36* mutator of *Escherichia coli*: Effects of the tau subunit of the DNA polymerase III holoenzyme on chromosomal DNA replication fidelity. *J Bacteriol* 193:296–300.
7. Fijalkowska IJ, Schaaper RM, Jonczyk P (2012) DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair. *FEMS Microbiol Rev* 36:1105–1121.
8. Fernandez-Leiro R, et al. (2017) Self-correcting mismatches during high-fidelity DNA replication. *Nat Struct Mol Biol* 24:140–143.
9. Pham PT, Zhao W, Schaaper RM (2006) Mutator mutants of *Escherichia coli* carrying a defect in the DNA polymerase III tau subunit. *Mol Microbiol* 59:1149–1161.
10. Pham PT, Olson MW, McHenry CS, Schaaper RM (1998) The base substitution and frameshift fidelity of *Escherichia coli* DNA polymerase III holoenzyme in vitro. *J Biol Chem* 273:23575–23584.
11. Lujan SA, et al. (2014) Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res* 24:1751–1764.
12. Rocha EP, Danchin A, Viari A (1999) Universal replication biases in bacteria. *Mol Microbiol* 32:11–16.
13. Shah SN, Hile SE, Eckert KA (2010) Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes. *Cancer Res* 70:431–435.
14. Miller JH, Ganem D, Lu P, Schmitz A (1977) Genetic studies of the *lac* repressor. I. Correlation of mutational sites with specific amino acid residues: Construction of a colinear gene-protein map. *J Mol Biol* 109:275–298.
15. Schaaper RM, Danforth BN, Glickman BW (1986) Mechanisms of spontaneous mutagenesis: An analysis of the spectrum of spontaneous mutation in the *Escherichia coli lacI* gene. *J Mol Biol* 189:273–284.
16. Schaaper RM, Dunn RL, Glickman BW (1987) Mechanisms of ultraviolet-induced mutation: Mutational spectra in the *Escherichia coli lacI* gene for a wild-type and an excision repair-deficient strain. *J Mol Biol* 198:187–202.
17. Schaaper RM, Dunn RL (1991) Spontaneous mutation in the *Escherichia coli lacI* gene. *Genetics* 129:317–326.
18. Schaaper RM, Dunn RL (1987) Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: The nature of in vivo DNA replication errors. *Proc Natl Acad Sci USA* 84:6220–6224.
19. Schaaper RM (1993) The mutational specificity of two *Escherichia coli dnaE* antimutator alleles as determined from *lacI* mutation spectra. *Genetics* 134:1031–1038.
20. Fowler RG, Schaaper RM (1997) The role of the *mutT* gene of *Escherichia coli* in maintaining replication fidelity. *FEMS Microbiol Rev* 21:43–54.
21. Mo JY, Schaaper RM (1996) Fidelity and error specificity of the alpha catalytic subunit of *Escherichia coli* DNA polymerase III. *J Biol Chem* 271:18947–18953.
22. Oller AR, Fijalkowska IJ, Dunn RL, Schaaper RM (1992) Transcription-repair coupling determines the strandedness of ultraviolet mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci USA* 89:11036–11040.
23. Bebenek K, Joyce CM, Fitzgerald MP, Kunkel TA (1990) The fidelity of DNA synthesis catalyzed by derivatives of *Escherichia coli* DNA polymerase I. *J Biol Chem* 265: 13878–13887.
24. Joyce CM, Sun XC, Grindley ND (1992) Reactions at the polymerase active site that contribute to the fidelity of *Escherichia coli* DNA polymerase I (Klenow fragment). *J Biol Chem* 267:24485–24500.
25. Kunkel TA, Hamatake RK, Motto-Fox J, Fitzgerald MP, Sugino A (1989) Fidelity of DNA polymerase I and the DNA polymerase I-DNA primase complex from *Saccharomyces cerevisiae*. *Mol Cell Biol* 9:4447–4458.
26. Mendelman LV, Boosalis MS, Petruska J, Goodman MF (1989) Nearest- neighbor influences on DNA polymerase insertion fidelity. *J Biol Chem* 264:14415–14423.
27. Mendelman LV, Petruska J, Goodman MF (1990) Base mispair extension kinetics: Comparison of DNA polymerase alpha and reverse transcriptase. *J Biol Chem* 265: 2338–2346.
28. Perrino FW, Loeb LA (1989) Differential extension of 3′ mispairs is a major contribution to the high fidelity of calf thymus DNA polymerase-alpha. *J Biol Chem* 264: 2898–2905.
29. Maliszewska-Tkaczyk M, Jonczyk P, Bialoskorska M, Schaaper RM, Fijalkowska IJ (2000) SOS mutator activity: Unequal mutagenesis on leading and lagging strands. *Proc Natl Acad Sci USA* 97:12678–12683.
30. Fijalkowska IJ, Jonczyk P, Tkaczyk MM, Bialoskorska M, Schaaper RM (1998) Unequal fidelity of leading strand and lagging strand DNA replication on the *Escherichia coli* chromosome. *Proc Natl Acad Sci USA* 95:10020–10025.
31. Wu CA, Zechner EL, Marians KJ (1992) Coordinated leading and lagging strand synthesis at the *Escherichia coli* DNA replication fork, I: Multiple effectors act to modulate Okazaki fragment size. *J Biol Chem* 267:4030–4044.
32. Wu CA, Zechner EL, Reems JA, McHenry CS, Marians KJ (1992) Coordinated leading and lagging strand synthesis at the *Escherichia coli* DNA replication fork, V: Primase action regulates the cycle of Okazaki fragment synthesis. *J Biol Chem* 267:4074–4083.
33. Zechner EL, Wu CA, Marians KJ (1992) Coordinated leading and lagging strand synthesis at the *Escherichia coli* DNA replication fork, II: Frequency of primer synthesis and efficiency of primer utilization control Okazaki fragment size. *J Biol Chem* 267: 4045–4053.
34. Yuan Q, McHenry CS (2014) Cycling of the *E. coli* lagging strand polymerase is triggered exclusively by the availability of a new primer at the replication fork. *Nucleic Acids Res* 42:1747–1756.
35. Beard WA, Wilson SH (2003) Structural insights into the origins of DNA polymerase fidelity. *Structure* 11:489–496.
36. Kim DR, McHenry CS (1996) In vivo assembly of overproduced DNA polymerase III: Overproduction, purification, and characterization of the alpha, alpha-epsilon, and alpha-epsilon-theta subunits. *J Biol Chem* 271:20681–20689.
37. Pham PT, Olson MW, McHenry CS, Schaaper RM (1999) Mismatch extension by *Escherichia coli* DNA polymerase III holoenzyme. *J Biol Chem* 274:3705–3710.
38. Banach-Orlowska M, Fijalkowska IJ, Schaaper RM, Jonczyk P (2005) DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in *Escherichia coli*. *Mol Microbiol* 58: 61–70.
39. Beattie TR, et al. (2017) Frequent exchange of the DNA polymerase during bacterial chromosome replication. *eLife* 6:e21763.
40. Lewis JS, et al. (2017) Single-molecule visualization of fast polymerase turnover in the bacterial replisome. *eLife* 6:e23932.
41. Schaaper RM (1998) Antimutator mutants in bacteriophage T4 and *Escherichia coli*. *Genetics* 148:1579–1585.
42. Fijalkowska IJ, Dunn RL, Schaaper RM (1993) Mutants of *Escherichia coli* with increased fidelity of DNA replication. *Genetics* 134:1023–1030.
43. Fijalkowska IJ, Schaaper RM (1993) Antimutator mutations in the alpha subunit of *Escherichia coli* DNA polymerase III: Identification of the responsible mutations and alignment with other DNA polymerases. *Genetics* 134:1039–1044.
44. Welch MM, McHenry CS (1982) Cloning and identification of the product of the *dnaE* gene of *Escherichia coli*. *J Bacteriol* 152:351–356.
45. Fijalkowska IJ, Dunn RL, Schaaper RM (1997) Genetic requirements and mutational specificity of the *Escherichia coli* SOS mutator activity. *J Bacteriol* 179:7435–7445.
46. Fijalkowska IJ, Schaaper RM (1995) Effects of *Escherichia coli dnaE* antimutator alleles in a proofreading-deficient *mutD5* strain. *J Bacteriol* 177:5979–5986.
47. Lamers MH, Georgescu RE, Lee SG, O'Donnell M, Kuriyan J (2006) Crystal structure of the catalytic alpha subunit of *E. coli* replicative DNA polymerase III. *Cell* 126:881–892.
48. Beard WA, et al. (1998) Vertical-scanning mutagenesis of a critical tryptophan in the minor groove binding track of HIV-1 reverse transcriptase: Molecular nature of polymerase-nucleic acid interactions. *J Biol Chem* 273:30435–30442.
49. Schaaper RM (1996) Suppressors of *Escherichia coli mutT*: Antimutators for DNA replication errors. *Mutat Res* 350:17–23.
50. Bebenek K, Kunkel TA (1990) Frameshift errors initiated by nucleotide misincorporation. *Proc Natl Acad Sci USA* 87:4946–4950.
51. Goodman MF (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* 71:17–50.
52. Goodman MF, Woodgate R (2013) Translesion DNA polymerases. *Cold Spring Harb Perspect Biol* 5:a010363.
53. Jarosz DF, Beuning PJ, Cohen SE, Walker GC (2007) Y-family DNA polymerases in *Escherichia coli*. *Trends Microbiol* 15:70–77.
54. Witkin EM, McCall JO, Volkert MR, Wermundsen IE (1982) Constitutive expression of SOS functions and modulation of mutagenesis resulting from resolution of genetic instability at or near the *recA* locus of *Escherichia coli*. *Mol Gen Genet* 185:43–50.
55. Cupples CG, Miller JH (1989) A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci USA* 86: 5345–5349.