Published in final edited form as:

Mutat Res. 2014 January ; 759: 22–28. doi:10.1016/j.mrfmmm.2013.11.003.

Effect of dNTP pool alterations on fidelity of leading and lagging strand DNA replication in *E. coli*

Damian Gawel^{a,b,*}, Iwona J. Fijalkowska^a, Piotr Jonczyk^a, and Roel M. Schaaper^b

^aInstitute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02-106, Poland

^bLaboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Abstract

The fidelity with which organisms replicate their chromosomal DNA is of considerable interest. Detailed studies in the bacterium *Escherichia coli* have indicated that the fidelity of leading- and lagging-strand DNA replication is not the same, based on experiments in which the orientation of certain mutational targets on the chromosome was inverted relative to the movement of the replication fork: different mutation rates for several base-pair substitutions were observed depending on this orientation. While these experiments are indicative of differential replication fidelity in the two strands, a conclusion whether leading or lagging strand is the more accurate depends on knowledge of the primary mispairing error responsible for the base substitutions in question. A broad analysis of *in vitro* base-pairing preferences of DNA polymerases led us to propose that lagging-strand is the more accurate strand. In the present work, we present more direct *in vivo* evidence in support of this proposal. We determine the orientation dependence of mutant frequencies in *ndk* and *dcd* strains, which carry defined dNTP pool alterations. As these pool alterations lead to predictable effects on the array of possible mispairing errors, they mark the strands in which the observed errors occur. The combined results support the proposed higher accuracy of lagging-strand replication in *E. coli*.

Keywords

DNA replication fidelity; leading and lagging strands; base-base mispairs; dNTP precursors; base selection; exonucleolytic proofreading

*Corresponding author at current address: Department of Biochemistry and Molecular Biology, Medical Centre for Postgraduate Education, Warsaw, Ul. Marymoncka 99, 02-813 Warsaw, Poland. Tel: +48 22 834 6847, Fax: +48 22 834 0470, gawel@cmkp.edu.pl.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest statement The authors declare that there are no conflicts of interest.

1. Introduction

High accuracy DNA replication is a fundamental process that allows the transfer of genetic information to the next generation with minimal information loss. Although the fidelity of DNA replication is high, errors during this process are a significant source of spontaneous mutations. High fidelity DNA replication is achieved by a combination of activities, among which are (i) the discrimination between correct and incorrect nucleotides during the insertion step of DNA polymerization, (ii) removal of incorrectly inserted nucleotides by an exonucleolytic proofreading step, and (iii) postreplicative DNA mismatch repair, in which newly synthesized DNA is surveyed for mismatches, which are then corrected using the sequence of the (correct) parental strand as template [1–4]. These three processes, which are broadly conserved among all forms of life, are capable of reducing the error rate per base pair replicated to values 10^{-9} to 10^{-11} [4, 5].

While these processes comprise a broad outline, several important aspects still need to be understood. One particular issue is whether DNA replication occurs with similar fidelity in the leading and lagging strands of the replication fork. This issue is pertinent because of the slightly different enzymology of leading and lagging strand replication: due to the antiparallel nature of the two DNA strands, synthesis of the leading strand proceeds, continuously, in the same direction as the moving replication fork, while the lagging strand is synthesized discontinuously in the form of small sections called Okazaki fragments [6]. Because of this enzymatic dichotomy, fidelity differences may occur. Also, in eukaryotic organisms the two strands are replicated by different DNA polymerases, like polymerases δ and ϵ that replicate the eukaryotic lagging and leading strands, respectively [7, 8].

Studies from our laboratories have focused on understanding DNA replication fidelity in the bacterium *Escherichia coli*. In this model system, both strands are replicated by DNA polymerase III [6]. Specifically, the chromosome is replicated by a dimeric DNA polymerase complex termed DNA polymerase III holoenzyme (HE), whose composition is $(\alpha\epsilon\theta)_2\beta_2\tau_2(\gamma\delta\delta'\chi\psi)$ [6]. The complex contains two core polymerases ($\alpha\epsilon\theta$), one for each strand. Within the core complex, α is the polymerase subunit, ϵ the associated proofreading subunit, and θ a stabilizing subunit for ϵ [9, 10]. Importantly, despite the fact that in *E. coli* the leading and lagging strand are replicated by an identical polymerase (Pol III core), our studies have revealed that the two strands are replicated with differential fidelity [11–16], thus revealing an intrinsic strand-dependent phenomenon, the basis for which we would like to understand.

Our results are based on measurements of *lac* reversion frequencies in pairs of strains in which the *lac* operon is inserted in the two opposite orientations relative to the replication fork [11]. In this manner, a DNA sequence of interest will be synthesized, for example, as part of the leading strand in one orientation, but as part of the lagging strand in the other orientation. The results using a series of four *lac* base substitution alleles revealed in each case a different *lac* reversion frequency between the two orientations [11–16]. These experiments were performed in a strain background lacking the postreplicative DNA mismatch repair system (*mutHLS* system) [3], facilitating interpretation of the mutant frequencies in terms of DNA replication error rates. It was concluded that on the *E. coli*

chromosome leading and lagging strands have a different fidelity [11]. The differences varied by the type of base substitution investigated, but were in the 2- to 4-fold range [11–16].

While these experiments indicate differential replication fidelity for the two strands, they do not yet provide direct information about which of the strands is replicated more accurately. For example, when scoring A·T → T·A transversion mutations, these could be arising from A·A mispairings during copying of the A-containing strand or from T·T mispairings during copying of the T-containing strand, and unless a decision is made which of the two mispairs (A·A or T·T) is the main contributor to the observed mutations, a decision which strand is more accurate cannot be made. In a previous report we analyzed the properties of a variety of DNA polymerases that had been studied *in vitro* with regard to their misinsertion preferences and ability to extend a variety of base-base mismatches (extension of mismatches is needed for replication errors to be expressed as mutations), and made predictions for the corresponding properties of *E. coli* DNA polymerase III *in vivo*. On this basis, we postulated that A·T→T·A transversions would derive primarily from T·T errors (template base stated first and underlined), G·C→T·A transversions from C·T errors, G·C→A·T transitions from G·T errors, and A·T→G·C transitions from T·G errors [11]. Using these assumptions it emerged that, in four out of four cases examined, the lagging strand is the more accurate strand. A model as to why lagging-strand replication might be more accurate was discussed [11].

In the present work, we have further investigated this question of strand assignment. We make use of certain mutant of *E. coli* that possess alterations of their dNTP pools. The dNTPs (5'-deoxynucleoside triphosphates) are the DNA precursors used by the polymerases, and they have been shown to be critical factors in the ability of DNA polymerases to synthesize DNA accurately. First, at any given DNA template position, the one correct dNTP has to compete for insertion with the three incorrect ones [17, 18]. While polymerases perform this discrimination fairly well, the actual rate of misinsertion is directly proportional to the ratio of incorrect to correct dNTPs. Hence, mutants containing altered dNTP pools are expected to be mutators for certain base substitution mutations, as is indeed the case [19–21]. A second reason why dNTP pool mutants could display a mutator phenotype is because of the so-called “next nucleotide effect” [1, 22]. Here, elevated dNTP levels can enhance extension of base-base (or other) mismatches by promoting insertion of the next nucleotides following the mismatch. As the mismatch extension step competes directly with the removal of the erroneous base by proofreading, elevated dNTPs can be mutagenic through this mechanism. When the specific changes in the dNTP pool mutants are known, clear predictions can be made as to the mispairings that will be promoted [19, 21]. These predictions can then be used to “mark” the strands and to verify the earlier conclusion that lagging strand replication is more accurate than leading-strand replication. For this purpose, we use *ndk* and *dcd* mutants of *E. coli*, defective in nucleoside diphosphate kinase (*ndk* gene) and dCTP deaminase (*dcd* gene), respectively, both of which are characterized by defined dNTP pool changes and a corresponding mutator phenotype [21]. The strains were also defective in the *mutHLS* DNA mismatch-repair system to facilitate the interpretation of the mutation rates in terms of replication error rates. Our analysis of the mutational

responses for several *lac* alleles in the two orientations on the chromosome fully supports our previously deduced higher accuracy of lagging-strand replication.

2. Materials and methods

2.1 Media

Solid and liquid media were as described [23]. Solid minimal medium (MM) contained 0.5% glucose or 0.4% lactose as a carbon source, 50 µg/ml of L-proline, and 5 µg/ml of thiamine. MM solid medium was further supplemented with 50 µg/ml of thymidine for experiments involving *dcd* strains. Antibiotics, where required, were added at the following final concentrations: ampicillin, 25 µg/ml; chloramphenicol, 20 µg/ml; kanamycin, 25 µg/ml; rifampicin, 100 µg/ml; and tetracycline, 12.5 µg/ml.

2.2 Bacterial strains

The *E. coli* strains used in the mutagenesis experiments are all derivatives of strains EC3114, EC3120, EC3126, EC3132, EC3138, EC3144, EC3150, and EC3156, which have been described [11]. They contain a *lac* deletion at the normal *lac* location near 8 min on the chromosome, but carry a reinserted *lac* operon at the bacteriophage lambda attachment site (*attB*) near 17 min. This *lac* operon was derived from strains CC102, CC104, CC105 and CC106 [24], which each carry a defined *lacZ* missense mutation that allows scoring of mutagenesis *via* reversion to Lac⁺ by G·C→A·T, G·C→T·A, A·T→T·A or A·T→G·C base substitution [24]. Each *lac* operon is inserted in the two possible chromosomal orientations (termed L and R, arbitrarily) [11] to permit investigation of the effects of gene orientation on mutagenesis and DNA replication fidelity. To more directly on focus replication fidelity, all strains used were also defective in DNA mismatch repair (*mutL*::Tn5 or *mutL*::Tn10) [11]. The *dcd* allele (*dcd-12*::Tn10*dkan*) was inserted by P1 transduction from strain BW1040 [25] using kanamycin resistance as selective marker. The *ndk* allele (*ndk*::*cam*) was inserted by P1 transduction from NR11814 [21] using chloramphenicol resistance as selective marker.

2.3 Lac and rif^r mutant frequency measurements

For each strain, a total of 20 independent LB cultures (2 ml each) were initiated from single colonies (one colony per tube). The colonies were taken from three to nine independent isolates (transductants) for each strain. The cultures were grown to saturation at 37°C on a rotator wheel. One hundred microliters of the cultures, diluted to 10⁻⁶, were plated on MM glucose plates to determine the total cell count, while 100 µl of the undiluted cultures were plated on MM-lactose plates to determine the number of Lac⁺ mutants. To measure the frequency of rifampicin-resistant cells, 100 µl of undiluted or 10⁻¹ diluted cultures were plated on LB plates containing rifampicin. Plates were incubated for 20 h (LB) or 40 h (MM) at 37°C. To calculate mutant frequencies, the average number of mutants per plate was divided by the number of total cells plated. Statistical analysis was performed using the software program Prism (GraphPad).

3. Results

3.1 The chromosomal *lac* reversion system

In Fig. 1, we present a set of diagrams outlining the DNA polymerization events that can lead to reversion of each of the *lacZ* alleles used in this study. The alleles are specific in the sense that, for each allele, the indicated missense codon can be reverted by only a single, defined base substitution to the GAG codon encoding the critical Glu461 residue of β -galactosidase (*lacZ* gene product) [24]. These *lac* alleles were first created on an *F'prolac* [24], but were transferred to the chromosome by insertion of the corresponding *lac* operons into the *attB* insertion site of bacteriophage lambda in the two possible orientations [11]. The two orientations, arbitrarily termed L and R [11], are shown in Fig. 1 along with the specific mispairing events that would revert the missense codons to the wild-type GAG codon. For example, in Fig. 1A, the *lac* allele originally present in strain CC104 [24] contains the GCG missense codon in either leading (L-orientation, left side of diagram) or lagging strand (R-orientation, right). Reversion of the GCG to GAG can occur by either $\underline{C}\cdot T$ or $\underline{G}\cdot A$ mispairing (both yielding the necessary G·C $\rightarrow$ T·A transversion). The $\underline{C}\cdot T$ mispairing would be promoted by increased levels of dTTP (incorrect nucleotide) relative to dGTP (correct nucleotide), and/or increased levels of dCTP, the next (correct) nucleotide. The requirements for promoting the $\underline{G}\cdot A$ mispairing are different: enhanced dATP (incorrect nucleotide) relative to dCTP (correct nucleotide) and by enhanced dGTP (next correct nucleotide). In view of these distinguishing requirements for promoting the possible mispairs, alterations of the dNTP pool are likely to be a useful tool for deciding which particular mispair is primarily responsible for the *lac* reversions and, hence, for deciding which strand is replicated more accurately. Below, we describe results with the *dcd* and *ndk* pool-bias mutants.

3.2 Effect of *dcd* on *lac* reversions

The *E. coli dcd* mutant is defective in dCTP deaminase, which performs the deamination of dCTP to produce dUTP, an important step in *de novo* dTTP biosynthesis [25]. The dUTP is converted to dUMP by *E. coli* dUTPase, which is then converted to dTMP by thymidylate synthase, followed by conversion to dTDP and dTTP. Due to their lack of deaminase activity, *dcd* mutants accumulate dCTP [21, 26, 27], which is one contributor to the elevated mutation level that they display (mutator phenotype) [21]. Despite their deaminase defect, *dcd* strains still retain some ability to synthesize dTTP through alternative pathways [28], although the dTTP concentrations are significantly reduced when cells are grown in minimal medium [21, 26, 27]. In the present series of experiments, cells were grown in LB medium, and dNTP measurement have shown the dTTP concentration to be essentially normal, presumably due to the presence of thymine and/or thymidine in the medium, which permit synthesis of dTTP through salvage pathways [28]. The observed dNTP levels in *dcd* strains are shown in Table 1 (from [21]). It is seen that the dCTP concentration is increased nearly four-fold, while the dGTP concentration is decreased by about 2-fold.

In Table 2, we present the mutant frequencies of the four studied *lac* reversions in *dcd*⁺ (wt) and *dcd* strains. The strains are also mismatch-repair deficient to facilitate interpretation of the mutant frequencies in terms of replication errors. For the wt strain the results reiterate

the previous findings that the two orientations yield different reversion frequencies. For G·C→T·A, A·T→T·A, and G·C→T·A, the L-orientation yields the higher frequency, while for the A·T→G·C the R-orientation is higher. Typically, the frequency differences are 2- to 4-fold, as observed in multiple repeated experiments [11–16].

Table 2 also clearly shows the mutator effect of the *dcd* strain, most prominently for the G·C→T·A and A·T→T·A transversions (3.1- to 5.6-fold). Overall, these *dcd* mutator effects are smaller than those reported recently for the same *lac* alleles residing on F'*prolac* [21], likely reflecting certain features specific to the F' replication. Nevertheless, the *dcd* mutator effects observed here for the chromosomal *lac* location should be informative with regard to the preferences of *in vivo* mispairings and, hence, the strand bias of DNA replication fidelity. We also note that the mutator effects are similar or near-identical regardless of the *lac* orientation. A full analysis will be presented in the Discussion.

3.3 Effect of *ndk* on *lac* reversions

The *ndk* gene encodes the enzyme nucleoside diphosphate kinase (Ndk), a non-specific enzyme responsible for phosphorylating all cellular nucleoside diphosphates (NDPs and dNDPs) to the corresponding triphosphates [29, 30]. Initially thought to be essential, an *ndk* deletion mutant was found to be viable [31]. However, not unexpectedly, the strain showed altered dNTP pools as well as increased mutability (mutator phenotype) [32]. It was assumed that another cellular kinase could replace Ndk, although not with the same efficiency and specificity. This alternative activity was shown to be adenylate kinase [32].

The dNTP levels observed in *ndk* strains are presented in Table 1: the dATP level is reduced by about four-fold, while there are increases for dCTP (2.0-fold) and, possibly, dTTP (1.5-fold). Table 3 also clearly shows the mutator effect of the *ndk* strain. Like is the case on the F', the chromosomal mutator effect is strongly specific for the A·T→T·A transversion (23- to 26-fold), with only small effects for the G·C→T·A transversion [21]. The two transitions were not monitored in these experiments. As in the case of the *dcd* mutator, we note that the increases are occurring similarly for both orientations.

3.4 Frequency of rifampicin-resistant mutants in *lac*-inverted strains

As a control in the above-described experiments, we also routinely determined the frequency of rifampicin-resistant (Rif^r) mutants, which result from mutations in the *rpoB* gene [33]. As shown in Table 4, the *dcd* and *ndk* defects result in a mutator effect for Rif^r mutations (1.5-fold and 20-fold, respectively). However, as the *rpoB* gene is not subject to inversion, the frequency of rifampicin-resistant mutants is not significantly changed between strains with L- and R-oriented *lac* operons (Table 4), as expected.

4. Discussion

The purpose of the current series of experiments was to provide support for our previous conclusion that on the *E. coli* chromosome the fidelity of lagging-strand replication is higher than that of leading-strand replication. We used the tool of alterations in the dNTP DNA precursor concentrations to make predictions as to which specific mispairings will be promoted or disfavored. The responsiveness of the tested *lac* base substitutions to the known

pool changes should then be an indication whether the choice of preferred mispair is correct., page

In the Figs. 1 and 2 we present diagrams of replication forks traversing the relevant section of the *lac* operon in the two (L or R) orientations. Shown is the missense codon for each of the four monitored *lac* missense alleles: GCG, GTG, GGG, and AAG [24], as well as their complement in the opposite strand. For each case, we have also indicated the particular dNTPs that are relevant for the fidelity of the copying step. These include the correct dNTP competing with the creation of each mispair (its elevation will be antimutagenic at the misinsertion step) and the dNTP that follows the misinsertion (its elevation is predicted to be mutagenic). The mutagenic action of increased levels of the next nucleotide, generally known as the next-nucleotide effect, is well established from *in vitro* DNA polymerase reactions and has also been deduced to occur *in vivo* [19, 20]. Mechanistically, high levels of the next nucleotide promote extension of a mismatch and reduce the probability of removal of the misinserted base by proofreading [1, 22].

We will first analyze the results obtained with the *dcd* strain (Fig. 1). Diagram 1A concerns the production of the *lac* G·C→T·A transversions, which were enhanced by 5- to 6-fold in the *dcd* strain (Table 2). In our previous studies we have proposed that the C·T mismatch was the primary mismatch responsible for these events [11]. In the diagrams we have indicated those dNTPs whose changes are predicted to be mutagenic in RED color, while the antimutagenic dNTP effects are indicated in BLUE. In the *dcd* strain the dCTP level is enhanced while the dGTP level is reduced (Table 1). Hence, it is easy to see how the production of C·T errors would be enhanced: dGTP is the correct nucleotide and its reduction is predicted to be mutagenic, while dCTP, the next nucleotide, is enhanced, which should also be mutagenic. In contrast, the reciprocal G·A error is not expected to be enhanced. In fact, it is expected to be reduced: the dCTP elevation will be antimutagenic, as it is the correct nucleotide, while the decrease in dGTP, the next nucleotide, will disfavor extension of the mismatch. Thus, these data are fully consistent with our earlier assumption that the primary cause of the G·C→T·A transversions is C·T mispairing (boxed). The similar enhancement (mutator effect) for the two gene orientations is further consistent with the implicit assumption that the same mispair underlies the mutation in either strand.

The *dcd* mutator also enhances the *lac* A·T→T·A transversion (Table 2). Here, the T·T mismatch (boxed) was proposed to be the major contributor [11]. The diagram Fig. 1B indicates that there is no predicted effect for T·T mispairing at the misinsertion step (both the correct dATP and incorrect dTTP are unchanged), but extension of the T·T will still be promoted by the increase in dCTP, the next nucleotide. In contrast, any A·A mismatches that might occur are not predicted to be promoted. In fact, the lowering of dGTP would disfavor any extension. Thus, the previous choice of T·T [11] is supported by the *dcd* mutator data.

Modest mutator effects are observed for the two transitions in the *dcd* strain. Fig. 1C for the G·C→A·T transition, for which we favor the G·T mismatch (boxed) [11], shows that the elevation of dCTP would have two opposing effects on the G·T mispair, as dCTP is both the correct nucleotide (antimutagenic effect) and the next nucleotide (mutagenic effect). An identical effect is predicted for the reciprocal C·A error. Here, dGTP is both the correct

nucleotide and the next nucleotide, and the effects of its reduction in the mutant may cancel out. Thus, no strong conclusions regarding this base-pair substitution can be drawn.

For the A·T→G·C transitions of Fig. 1D, where we favor the T·G mismatch (boxed) [11], somewhat similar arguments apply. For the T·G events, the dGTP decrease of the *dcd* mutant would be antimutagenic (it is the incorrect nucleotide). For the reciprocal A·C mismatch, the increase in dCTP is predicted to be mutagenic (it is the incorrect nucleotide). Thus, while no direct positive evidence is obtained for the prevalence of the T·G mismatch, indirect evidence still supports this notion. Indeed, if the A·C mismatch were the major contributor to the A·T→G·C transitions, then the strongly elevated dCTP concentration (incorrect nucleotide) would be predicted to have a strong mutagenic effect, which is not observed.

The data obtained for the *ndk* strain are also informative. Between the two tested transversions, the A·T→T·A transversion is strongly enhanced (Table 3). The analysis of the *ndk* effects is shown in the diagrams of Fig. 2. As the *ndk* strain is characterized by elevated dCTP and lowered dATP (Table 1), it can be clearly seen that for the A·T→T·A transversion (Fig. 2B) the T·T mismatch is favored by both the higher dCTP (next nucleotide) and the lowered dATP (correct nucleotide). In contrast, the reciprocal A·A mismatch would be disfavored by the low dATP (incorrect nucleotide). Thus, these data strongly support the proposed T·T mismatch as the main contributor to the A·T→T·A transversions.

With regard to G·C→T·A transversions, only a small *ndk* mutator effect is observed (Table 3). From the diagram of Fig. 2A, one might argue that extension of the C·T mismatch would be promoted by the elevated dCTP level (next nucleotide), predicting a possible mutator effect. No mutator effect is predicted for the reciprocal G·A mismatch, as high dCTP (correct nucleotide) and low dATP (incorrect nucleotide) would both disfavor the event. It is possible that the modest dCTP elevation in the *ndk* strain is not substantial enough for a strong enhancement of this event. We note that the G·C→T·A is significantly enhanced in the *dcd* strain, which is characterized by a larger dCTP increase than in the *ndk* strain and a decreased dGTP level (both favorable).

In summary, the current data provide solid support for several of our assumptions about the predominant mispair responsible for the *lacZ* base substitutions used in our leading/lagging strand system. We also note that the *dcd* or *ndk* mutator effects are found to be very similar or near identical for the *lac* L or R orientation, consistent with the notion that the same mismatch is responsible for the mutational events in leading and lagging strands. These data strengthen our conclusion that on the *E. coli* chromosome the fidelity of lagging-strand replication is higher than that of the leading strand. From Tables 2 and 3 we see that for the G·C→T·A, A·T→T·A and G·C→T·A substitutions the L orientation has the highest mutant frequency, while for the A·T→G·C the R orientation has the highest frequency. In all four cases (see Figs. 1 and 2), this corresponds with the preferred (boxed) mispair residing in the leading strand, thus indicating that leading-strand synthesis is more error prone. Although this strand-dependent fidelity difference is modest, about 2-fold in the present study, it has been observed consistently, at the 2- to 4-fold level, in a series of extended studies [11–16].

It is our contention that this fidelity bias does not reflect a difference in the intrinsic polymerase misinsertion error rate in the two strands, but rather a difference in the subsequent processing of the error [11, 12]. How there can be more effective error processing in the lagging strand is a question that remains to be answered. It is attractive to speculate that improved lagging-strand fidelity results from the more dissociative properties of the polymerase in this strand, permitting more efficient proofreading (note that proofreading is essentially a dissociative process in which the primer terminus is moved from the polymerase active site to the distant exonuclease site). Alternatively, the dissociative character of the lagging-strand polymerase may lead to increased removal of the terminal mismatch by the exonuclease of Pol II, which has been shown to proofread for Pol III under some conditions [12, 13, 34, 35] or by a third copy of Pol III that has been proposed to be present at the replication fork [36].

Acknowledgements

We thank Drs. Jordan Saint-Charles and Kasia Maslowska of the NIEHS for their careful reading of the manuscript for this paper. This work was supported by project number Z01 ES065086 of the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. IJF acknowledges support from project TEAM/2011-8/1 (New players involved in the maintenance of genomic stability) from the Foundation for Polish Science TEAM program.

References

- [1]. Kunkel TA. DNA replication fidelity. *J Biol Chem.* 2004; 279:16895–16898. [PubMed: 14988392]
- [2]. Kunkel TA, Bebenek K. DNA replication fidelity. *Annu Rev Biochem.* 2000; 69:497–529. [PubMed: 10966467]
- [3]. Kunkel TA, Erie DA. DNA mismatch repair. *Annu Rev Biochem.* 2005; 74:681–710. [PubMed: 15952900]
- [4]. Schaaper RM. Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem.* 1993; 268:23762–23765. [PubMed: 8226906]
- [5]. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics.* 1998; 148:1667–1686. [PubMed: 9560386]
- [6]. McHenry CS. Bacterial replicases and related polymerases. *Curr Opin Chem Biol.* 2011; 15:587–594. [PubMed: 21855395]
- [7]. Burgers PM. Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem.* 2009; 284:4041–4045. [PubMed: 18835809]
- [8]. Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PM, Kunkel TA. Division of labor at the eukaryotic replication fork. *Mol Cell.* 2008; 30:137–144. [PubMed: 18439893]
- [9]. Chikova AK, Schaaper RM. Mutator and antimutator effects of the bacteriophage P1 *hot* gene product. *J Bacteriol.* 2006; 188:5831–5838. [PubMed: 16885451]
- [10]. Taft-Benz SA, Schaaper RM. The theta subunit of *Escherichia coli* DNA polymerase III: a role in stabilizing the epsilon proofreading subunit. *J Bacteriol.* 2004; 186:2774–2780. [PubMed: 15090519]
- [11]. Fijalkowska IJ, Jonczyk P, Tkaczyk MM, Bialoskorska M, Schaaper RM. Unequal fidelity of leading strand and lagging strand DNA replication on the *Escherichia coli* chromosome. *Proc Natl Acad Sci U S A.* 1998; 95:10020–10025. [PubMed: 9707593]
- [12]. Fijalkowska IJ, Schaaper RM, Jonczyk P. DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol Rev.* 2012; 36:1105–1121. [PubMed: 22404288]
- [13]. Gawel D, Jonczyk P, Fijalkowska IJ, Schaaper RM. *dnaX36* Mutator of *Escherichia coli*: effects of the τ subunit of the DNA polymerase III holoenzyme on chromosomal DNA replication fidelity. *J Bacteriol.* 2011; 193:296–300. [PubMed: 21036999]

- [14]. Kuban W, Banach-Orlowska M, Bialoskorska M, Lipowska A, Schaaper RM, Jonczyk P, Fijalkowska IJ. Mutator phenotype resulting from DNA polymerase IV overproduction in *Escherichia coli*: preferential mutagenesis on the lagging strand. *J Bacteriol.* 2005; 187:6862–6866. [PubMed: 16166552]
- [15]. Makiela-Dzibenska K, Jaszczur M, Banach-Orlowska M, Jonczyk P, Schaaper RM, Fijalkowska IJ. Role of *Escherichia coli* DNA polymerase I in chromosomal DNA replication fidelity. *Mol Microbiol.* 2009; 74:1114–1127. [PubMed: 19843230]
- [16]. Maliszewska-Tkaczyk M, Jonczyk P, Bialoskorska M, Schaaper RM, Fijalkowska IJ. SOS mutator activity: unequal mutagenesis on leading and lagging strands. *Proc Natl Acad Sci U S A.* 2000; 97:12678–12683. [PubMed: 11050167]
- [17]. Goodman MF, Creighton S, Bloom LB, Petruska J. Biochemical basis of DNA replication fidelity. *Crit Rev Biochem Mol Biol.* 1993; 28:83–126. [PubMed: 8485987]
- [18]. Pham PT, Olson MW, McHenry CS, Schaaper RM. The base substitution and frameshift fidelity of *Escherichia coli* DNA polymerase III holoenzyme *in vitro*. *J Biol Chem.* 1998; 273:23575–23584. [PubMed: 9722597]
- [19]. Ahluwalia D, Bienstock RJ, Schaaper RM. Novel mutator mutants of *E. coli nrdAB* ribonucleotide reductase: insight into allosteric regulation and control of mutation rates. *DNA Repair (Amst).* 2012; 11:480–487. [PubMed: 22417940]
- [20]. Kumar D, Abdulovic AL, Viberg J, Nilsson AK, Kunkel TA, Chabes A. Mechanisms of mutagenesis *in vivo* due to imbalanced dNTP pools. *Nucleic Acids Res.* 2011; 39:1360–1371. [PubMed: 20961955]
- [21]. Schaaper RM, Mathews CK. Mutational consequences of dNTP pool imbalances in *E. coli*. *DNA Repair (Amst).* 2013; 12:73–79. [PubMed: 23218950]
- [22]. Kunkel TA, Schaaper RM, Beckman RA, Loeb LA. On the fidelity of DNA replication. Effect of the next nucleotide on proofreading. *J Biol Chem.* 1981; 256:9883–9889. [PubMed: 6456268]
- [23]. Fijalkowska IJ, Schaaper RM. Effects of *Escherichia coli dnaE* antimutator alleles in a proofreading-deficient *mutD5* strain. *J Bacteriol.* 1995; 177:5979–5986. [PubMed: 7592352]
- [24]. Cupples CG, Miller JH. A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci U S A.* 1989; 86:5345–5349. [PubMed: 2501784]
- [25]. Wang L, Weiss B. *dcd* (dCTP deaminase) gene of *Escherichia coli*: mapping, cloning, sequencing, and identification as a locus of suppressors of lethal *dut* (dUTPase) mutations. *J Bacteriol.* 1992; 174:5647–5653. [PubMed: 1324907]
- [26]. Neuhaud J, Thomassen E. Altered deoxyribonucleotide pools in P2 eductants of *Escherichia coli* K-12 due to deletion of the *dcd* gene. *J Bacteriol.* 1976; 126:999–1001. [PubMed: 177407]
- [27]. O'Donovan GA, Edlin G, Fuchs JA, Neuhaud J, Thomassen E. Deoxycytidine triphosphate deaminase: characterization of an *Escherichia coli* mutant deficient in the enzyme. *J Bacteriol.* 1971; 105:666–672. [PubMed: 5541540]
- [28]. Weiss B, Wang L. *De novo* synthesis of thymidylate via deoxycytidine in *dcd* (dCTP deaminase) mutants of *Escherichia coli*. *J Bacteriol.* 1994; 176:2194–2199. [PubMed: 8157589]
- [29]. Hama H, Almaula N, Lerner CG, Inouye S, Inouye M. Nucleoside diphosphate kinase from *Escherichia coli*; its overproduction and sequence comparison with eukaryotic enzymes. *Gene.* 1991; 105:31–36. [PubMed: 1657712]
- [30]. Hama H, Lerner C, Inouye S, Inouye M. Location of the gene (*ndk*) for nucleoside diphosphate kinase on the physical map of the *Escherichia coli* chromosome. *J Bacteriol.* 1991; 173:3276. [PubMed: 1646199]
- [31]. Lu Q, Zhang X, Almaula N, Mathews CK, Inouye M. The gene for nucleoside diphosphate kinase functions as a mutator gene in *Escherichia coli*. *J Mol Biol.* 1995; 254:337–341. [PubMed: 7490752]
- [32]. Lu Q, Inouye M. Adenylate kinase complements nucleoside diphosphate kinase deficiency in nucleotide metabolism. *Proc Natl Acad Sci U S A.* 1996; 93:5720–5725. [PubMed: 8650159]
- [33]. Garibyan L, Huang T, Kim M, Wolff E, Nguyen A, Nguyen T, Diep A, Hu K, Iverson A, Yang H, Miller JH. Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst).* 2003; 2:593–608. [PubMed: 12713816]

- [34]. Banach-Orlowska M, Fijalkowska IJ, Schaaper RM, Jonczyk P. DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in *Escherichia coli*. Mol Microbiol. 2005; 58:61–70. [PubMed: 16164549]
- [35]. Gawel D, Pham PT, Fijalkowska IJ, Jonczyk P, Schaaper RM. Role of accessory DNA polymerases in DNA replication in *Escherichia coli*: analysis of the *dnaX36* mutator mutant. J Bacteriol. 2008; 190:1730–1742. [PubMed: 18156258]
- [36]. Georgescu RE, Kurth I, O'Donnell ME. Single-molecule studies reveal the function of a third polymerase in the replisome. Nat Struct Mol Biol. 2012; 19:113–116. [PubMed: 22157955]

Highlights

- In *E. coli* leading and lagging strand are replicated with different fidelity
- Lagging strand replication is more accurate
- dNTP pool alterations can be used to mark the strands
- The fidelity difference may relate to differential removal of terminal mismatches

Fig 1.

Diagram for analyzing the effect of dNTP alterations in the *dcd* strain of *E. coli*. Panels A, B, C, and D show, respectively, the potential mispairing events responsible for the indicated $G \cdot C \rightarrow T \cdot A$, $A \cdot T \rightarrow T \cdot A$, $G \cdot C \rightarrow A \cdot T$ or $A \cdot T \rightarrow G \cdot C$ base-pair substitutions in the *lacZ* gene. These events revert the indicated *lacZ* missense codon (GCG, GTG, GGG, and AAG, respectively) to GAG, restoring the wild-type genotype. The *lacZ* missense alleles were first described residing on an *F'prolac* episome by Cupples and Miller [24], but in the current study they are present on the *E. coli* chromosome in the two possibly orientations (indicated as L or R) relative to the replication fork, as described [11]. For each base substitution, we present the two possible mispairs that can give rise to the required base substitution. It can be seen that, when comparing the two *lac* orientations, a given mispair occurring in the leading strand in one orientation occurs in the lagging strand in the other orientation, and *vice versa*. The mispair that we have proposed in a previous study [11] to be the primary contributor to the base substitution is indicated in bold and surrounded by a rectangle (for example, the C·T mispair for the $G \cdot C \rightarrow T \cdot A$ transversion of panel A). With arrows, we indicate the correct dNTP for this particular position as well as the next dNTP to be incorporated following the mismatch. The analysis focuses on the effects of the changes in dCTP and dGTP levels, which are primarily affected in the *dcd* strain (Table 1). When the dCTP and dGTP changes are predicted to be mutagenic, the dNTP is represented in RED color, while dCTP or dGTP changes predicted to be antimutagenic are presented in BLUE. As described in the text, both the competition between the incorrect and the correct

nucleotide at the misinsertion step, and the concentration of the next dNTP are determining factors for the mutational outcome. In panel D, we also show the 5' T that precedes the AAG missense codon, because it is the template base that needs to be copied following an A-C mispairing event. See text for more details.

Fig 2.

Diagram for analyzing the effect of dNTP alterations in the *ndk* strain of *E. coli*. This analysis focuses on the effects of changes in dATP and dCTP levels, which are primarily affected in the *ndk* strain (Table 1). RED colors for dATP or dCTP indicate that the change in this dNTP is predicted to be mutagenic. BLUE colors indicate an antimutagenic effect. See the Legend to Fig. 1 and the main text for more details.

Table 1dNTP pools in *dcd* and *ndk* strains (pmoles/10⁸ cells)*

Strain	dATP	dTTP	dCTP	dGTP
wild-type	16 ± 7	15 ± 6	75 ± 19	8.0 ± 2.4
<i>dcd</i>	11 ± 5	13 ± 7	277 ± 39	3.8 ± 5
<i>ndk</i>	4.0 ± 4.8	23 ± 7	150 ± 18	7.5 ± 1.7

* Data from Schaaper and Mathews [21]

Table 2

Mutator effects of the *dcd* deficiency for *lac* reversion of two chromosomal orientations (L or R) of the *lac* operon

		<i>lac</i> ⁺ per 10 ⁸ cells in indicated strain		Mutator effect
<i>lac</i> reversion	<i>lac</i> orientation	wt	<i>dcd</i>	
G·C→T·A	L	1.5 ± 0.2	8.2 ± 1	5.5
	R	0.75 ± 0.2	4.2 ± 1	5.6
A·T→T·A	L	2.2 ± 0.2	8.5 ± 0.4	3.9
	R	1.6 ± 0.2	5.0 ± 0.1	3.1
G·C→A·T	L	78 ± 4	113 ± 51	1.5
	R	31 ± 2	20 ± 21	0.6
A·T→G·C	L	34 ± 3	69 ± 6	2.2
	R	83 ± 4	147 ± 8	1.8

All strains are *mutL*. The mutator effect indicates the increase in mutant frequency between *dcd* and wild-type (wt) strains.

Table 3

Mutator effects of the *ndk* deficiency on *lac* reversion for two chromosomal orientations (L and R) of the *lac* operon

		<i>lac</i> ⁺ per 10 ⁸ cells in indicated strain		Mutator effect
<i>lac</i> reversion	orientation	wt	<i>ndk</i>	
G·C→T·A	L	1.1 ± 0.1	1.7 ± 0.3 (1.5)	1.5
	R	0.33 ± 0.1	1.1 ± 0.2 (3.3)	3.3
A·T→T·A	L	1.4 ± 0.2	36 ± 1 (26)	26
	R	0.78 ± 0.1	18 ± 0.5 (23)	23

All strains are *mutL*. The mutator effect indicates the increase in mutant frequency between *ndk* and wild-type (wt) strains.

Table 4

Unchanged mutant frequencies towards rifampicin resistance in strains carrying the *lac* operon in L or R orientation on the chromosome

<i>lac</i>		Rif ^r per 10 ⁸ cells in indicated strain		
allele	orientation	<i>wt</i>	<i>dcd</i>	<i>ndk</i>
G·C→T·A	L	184 ± 5	240 ± 10	nd
	R	170 ± 4	240 ± 12	nd
A·T→T·A	L	190 ± 15	nd	5900 ± 800
	R	180 ± 14	nd	6200 ± 1000

nd = not determined. All strains are also *mutL*.