

Transcription regulatory elements are punctuation marks for DNA replication

Ekaterina V. Mirkin*, Daniel Castro Roa†, Evgeny Nudler†, and Sergei M. Mirkin**

*Department of Biochemistry and Molecular Genetics, University of Illinois, Chicago, IL 60607; and †Department of Biochemistry, New York University Medical Center, New York, NY 10016

Edited by Philip C. Hanawalt, Stanford University, Stanford, CA, and approved March 27, 2006 (received for review February 9, 2006)

Collisions between DNA replication and transcription significantly affect genome organization, regulation, and stability. Previous studies have described collisions between replication forks and elongating RNA polymerases. Although replication collisions with the transcription-initiation or -termination complexes are potentially even more important because most genes are not actively transcribed during DNA replication, their existence and mechanisms remained unproven. To address this matter, we have designed a bacterial promoter that binds RNA polymerase and maintains it in the initiating mode by precluding the transition into the elongation mode. By using electrophoretic analysis of replication intermediates, we have found that this steadfast transcription-initiation complex inhibits replication fork progression in an orientation-dependent manner during head-on collisions. Transcription terminators also appeared to attenuate DNA replication, but in the opposite, codirectional orientation. Thus, transcription regulatory signals may serve as “punctuation marks” for DNA replication *in vivo*.

collisions | promoter | terminator

Impairment of DNA replication is believed to be a major factor in genomic instability (1–13). Because transcription and replication share the same template, occasional collisions between the two machineries are inevitable and can interfere with replication fork progression. Collisions between the elongating RNA polymerase and the replication fork have been well documented *in vitro* (14–17) and *in vivo* in both *Escherichia coli* (18–20) and *Saccharomyces cerevisiae* (6, 21, 22). The consensus from those studies was that head-on collisions with elongating RNA polymerase are much more detrimental for replication fork progression than codirectional collisions. Although it was suggested that replication stalling during the head-on collisions with transcription was caused by topological stress in the DNA separating the two machineries (18, 19, 22, 23), we have recently shown that it is caused by their direct, physical interaction (20). These results, combined with the data on preferred codirectional alignment of transcription units with the direction of replication in prokaryotes (23–27), have led to the suggestion that the main disadvantage of the head-on collisions could be their inhibitory effect on DNA replication.

All of the experimental studies cited in the preceding paragraph evaluated the effects of elongating RNA polymerase on the progression of the replication fork. Is there an interplay between the replication machinery and the transcription-initiation complex? To the best of our knowledge, there have been few studies on this matter. One intriguing example was the detection of a polar replication fork pause site at the tRNA locus of *S. cerevisiae*, which depended on the functionality of both the promoter and the RNA polymerase III (pol III) (22). It was believed that the replication fork was attenuated during the encounter with the elongating RNA polymerase (22). A later study, however, suggested that the pol III-initiation complex was responsible for the replication slowing in this system (6). In our recent study of transcription–replication collisions in *E. coli*, the synthetic *trc* promoter was shown to stall the replication fork in

an orientation-dependent manner even in the repressed state (20). Unfortunately, we could not distinguish whether this replication stalling was caused by the lactose repressor, RNA polymerase, or both. Thus, a thorough study of the interplay between the replication fork and the transcription-initiation complex seemed to be warranted.

Initiation of transcription in *E. coli* is a multistep process (for review, see ref. 28). First, RNA polymerase holoenzyme binds to the promoter, forming the closed complex. The second step is the isomerization of the closed complex into the open complex, accompanied by DNA unwinding. Third, the ternary complex is formed during binding of the first NTP. The fourth step is the synthesis of the first ≈ 10 RNA bases without the movement of RNA polymerase along DNA, resulting in the formation of the so-called stressed complex. This short RNA may be released from the ternary complex, leading to abortive initiation, or it may become the 5' end of the RNA transcript. Finally, promoter clearance is accompanied by the dissociation of the σ factor and engagement of the core RNA polymerase in the stable elongation complex.

Although the principles of the initial promoter recognition are well understood, the sequence requirements for the further steps in the initiation process are much less clear. Their importance, however, is illustrated by the results of the computational analysis of bacterial genomic DNA, which reveals many false-positive promoters, whereas the true promoters may not even have the highest computational scores (ref. 29 and references therein).

To study collisions between the replication machinery and transcription-initiation complex carefully, we needed a system in which a transcription-initiation complex without additional proteins would stably exist inside the cell. Such a system could be achieved by shifting the equilibrium from the promoter clearance toward abortive initiation. In this case, the majority of short RNA products would be released while the RNA polymerase would remain in its initiating mode. To ensure efficient binding of the RNA polymerase to the promoter, we chose a strong promoter, the bacteriophage T7 early promoter A1 (30). To shift the equilibrium toward abortive initiation, we decided to modify the initial transcribed sequence (ITS) of the promoter, the importance of which for the strength of the promoters and the transition to the productive elongation mode was demonstrated in refs. 31 and 32. Specifically, the original promoter ITS, positioned between +1 and +20, was converted into a 90% AT-rich element, carrying multiple Ts on the nontemplate strand. Our idea was based on the fact that the weak RNA·DNA

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CD, codirectional orientation of the promoter to the direction of replication in the plasmid; HO, head-on orientation; ITS, initial transcribed sequence; pA, promoter A in a plasmid; pB, promoter B in a plasmid; pol I, II, and III, RNA polymerase I, II, and III, respectively; RNA I, origin-specific negative regulator of ColE1-like plasmid replication; TR, A/B promoter-specific transcript probe.

†To whom correspondence should be addressed. E-mail: mirkin@uic.edu.

© 2006 by The National Academy of Sciences of the USA

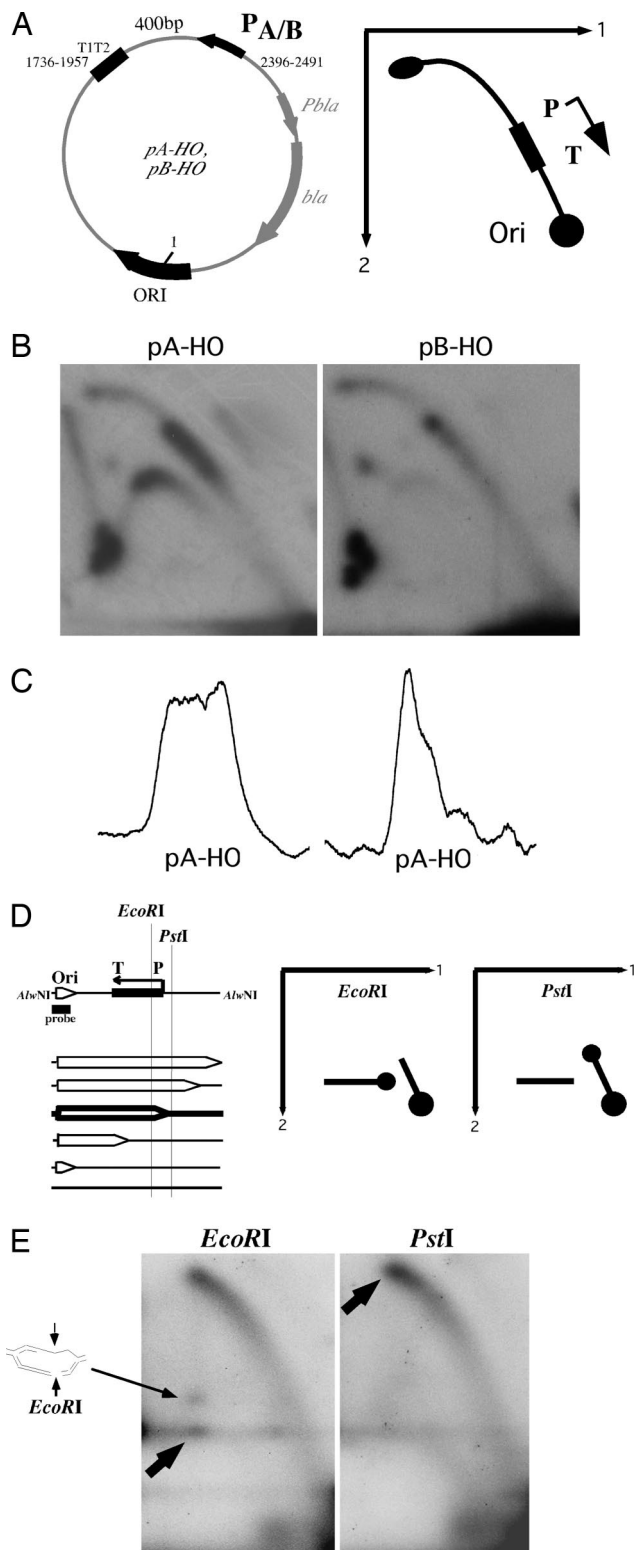


Fig. 2. Replication stalling by the head-on oriented transcription-initiation complex. (A Left) Map of the plasmids pA-HO and pB-HO. Transcription cassettes with promoters A and B, respectively, are oriented head-on to the direction of replication; nucleotide position 1 in the plasmids corresponds to the replication start site; positions of the promoters and T1T2 terminators are shown. (A Right) Schematic representation of the two-dimensional gel electrophoresis of replication intermediates. The bubble arc consists of replication intermediates, starting from the small bubble at the origin of replication (at the bottom) and going up to the biggest, fully replicated bubble (at the top). Replication stalling is detected as a thickening of a particular segment of

the arc. Consequently, the drastically diminished activity of promoter B is not the result of inefficient RNA polymerase binding and open-complex formation. We believe, therefore, that transcriptional impairment, caused by the AT-rich ITS, is the result of the inability of the RNA polymerase to undergo the transition to the productive elongation mode. Mutant promoter B was thus chosen to study the interplay between the steady-state transcription-initiation complex and DNA replication *in vivo*.

Replication Fork Progression Is Inhibited During Head-On Collision with the Transcription-Initiation Complex *in Vivo*. We evaluated replication fork progression through a transcription cassette by using the two-dimensional electrophoretic analysis of replication intermediates (37). Briefly, replication intermediates of ColE1-derived, unidirectionally replicating plasmids from *E. coli* cells were cleaved immediately upstream from the origin, and the resultant bubble-shaped intermediates were resolved in a two-dimensional gel, forming the so-called bubble arc (38). This arc is smooth if the replication fork progressed with the same speed throughout the plasmid; if replication was slowed down at a particular position, a distinct bulge appears on the arc because of the accumulation of replication intermediates of the defined size and shape. We have previously applied this approach successfully to detecting transcription–replication collision in *E. coli* cells (20).

The general map of our plasmids, in which transcription cassettes face the replication origin head-on, is presented in Fig. 2A alongside the schematic drawing of the anticipated bubble arc. The experimental data in Fig. 2B (Left) show that the replication is slowed down throughout the whole area transcribed from the wild-type promoter, in accord with our previous results (20). [Note that the Y arc underneath the bubble arc comes from replicating plasmid dimers (39) and also shows extensive transcription–replication collisions, albeit with a worse resolution. The strength of this Y arc depends on the proportion of plasmid dimerization and varies substantially among various constructs and experiments.] For the mutant promoter B, the results are strikingly different (Fig. 2B Right); replication stalling is mostly evident at a defined spot in the whole plasmid, which roughly corresponds to the position of the promoter. Fig. 2C shows the quantitative analysis of the replication data, obtained from a different experiment. Clearly, most of the replication stalling occurs at the promoter in the pB-HO plasmid, whereas the stalling is evenly spread throughout the transcribed region in promoter A.

To confirm that replication stalling indeed occurs at the mutant promoter, we have undertaken fine mapping of the pause site. To this end, we have used a modification of the two-

the arc. P, promoter; T, terminator; arrow, direction of transcription. (B) Two-dimensional gel electrophoresis of replication intermediates. Replication stalling is evident as a long segment that corresponds to the whole transcribed area in the pA-HO plasmid (Left) or as a bulge that corresponds to the promoter in the pB-HO plasmid (Right). (C) Quantitative analysis of the replication arcs around the transcribed units in plasmids pA-HO and pB-HO. The experiment shown here is independent of the one shown in Fig. 4B. (D) Scheme for mapping replication pause sites by in-gel digestion of replication intermediates (for details, see Results). (Left) The vertical lines show the positions of the restriction sites immediately upstream (PstI) and downstream (EcoRI) from the promoter (P). Replication intermediates are shown as bubbles, and the bold bubble corresponds to stalled replication intermediates. (Center and Right) During EcoRI digestion (Center), stalled intermediates become Y-shaped and move to the line, whereas after PstI digestion (Right), they remain bubble-shaped and stay on the arc. (E) Experimental mapping of stalled replication intermediates by in-gel digestion with EcoRI (Left) or PstI (Right). Stalled replication intermediates are shown by arrows. Note the underreplicated stalled intermediates migrating between the arc and the line (for details, see Results).

dimensional gel-electrophoretic analysis of the replication intermediates, where an extra in-gel restriction digestion was performed during separation in the first dimension. The schematic representation of this approach is illustrated in Fig. 2*D*. During in-gel digestion, a fraction of replication intermediates converts into identical Y-shaped molecules that migrate as a horizontal line in the second dimension. When digestion occurs immediately downstream from the promoter (EcoRI), replication intermediates stalled at the promoter must shift from the bubble arc to the horizontal line. In contrast, when the digestion occurs immediately upstream from the promoter (PstI), the same stalled intermediates must remain on the bubble arc (Fig. 2*E*). The majority of stalled intermediates remain on the bubble arc during the PstI digestion. After EcoRI digestion, one fraction of stalled intermediates moves to the line, and another fraction migrates between the bubble arc and the line. The latter intermediates apparently correspond to DNA molecules in which the lagging strand around the EcoRI site was under-replicated, resulting in incomplete digestion and “butterfly-like” structures (11, 40). The third fraction of the stalled intermediates remained on the arc. Those intermediates reflect the residual activity of the mutant promoter evident from Fig. 1*C*. Notwithstanding the latter fraction, the fact that a noticeable amount of stalled intermediates moved away from the bubble arc during EcoRI digestion is sufficient to conclude that replication stalling occurred at the promoter during collision with the transcription-initiation complex.

We conclude, therefore, that the replication pause site is located between the EcoRI and PstI sites, i.e., at the mutant promoter. These data show that a replication fork can stall when encountering a transcription-initiation complex head-on in *E. coli* cells.

Codirectionally Positioned Transcription Terminators Inhibit the Replication Fork Progression. To study the effect of the transcription-initiation complex on replication fork progression in the codirectional orientation, we have inverted both the wild-type and mutant transcription cassettes, generating the plasmids shown in Fig. 3*A*. On an anticipated bubble arc, the transcription cassette would face up, with the promoter positioned proximally and terminators positioned distally to the replication origin.

Fig. 3*B Left* and *Center* shows that, in contrast with the head-on orientation, there is no evident replication stalling by either of the codirectionally oriented promoters A or B. One can see a slight thickening of the replication arc in the whole area transcribed from the wild-type promoter, which reflects a modest replication slow-down during collisions with the codirectionally moving RNA polymerase (17). Unexpectedly, however, there is a more prominent replication stall at the transcription terminators in the *greA,greB* knockout strain (which is otherwise isogenic to the wild-type strain used in this study). The GreA and GreB proteins (41) are known to assist RNA polymerase in escaping from backtracked conformation by activating its endonucleolytic activity, which leads to RNA cleavage, formation of the new 3' end in the active center, and resumption of transcription elongation (42). The replication stalling at terminators was indeed increased in the *greA,greB* mutant (Fig. 3*B Right*).

To confirm additionally that the latter replication pausing was caused by the transcription terminators, we have mapped the pause site by following the same logic described in the previous section (Fig. 3*C*). The experimental data in Fig. 3*D* show that stalled intermediates moved to the line during in-gel digestion with XbaI but remained on the bubble arc if treated

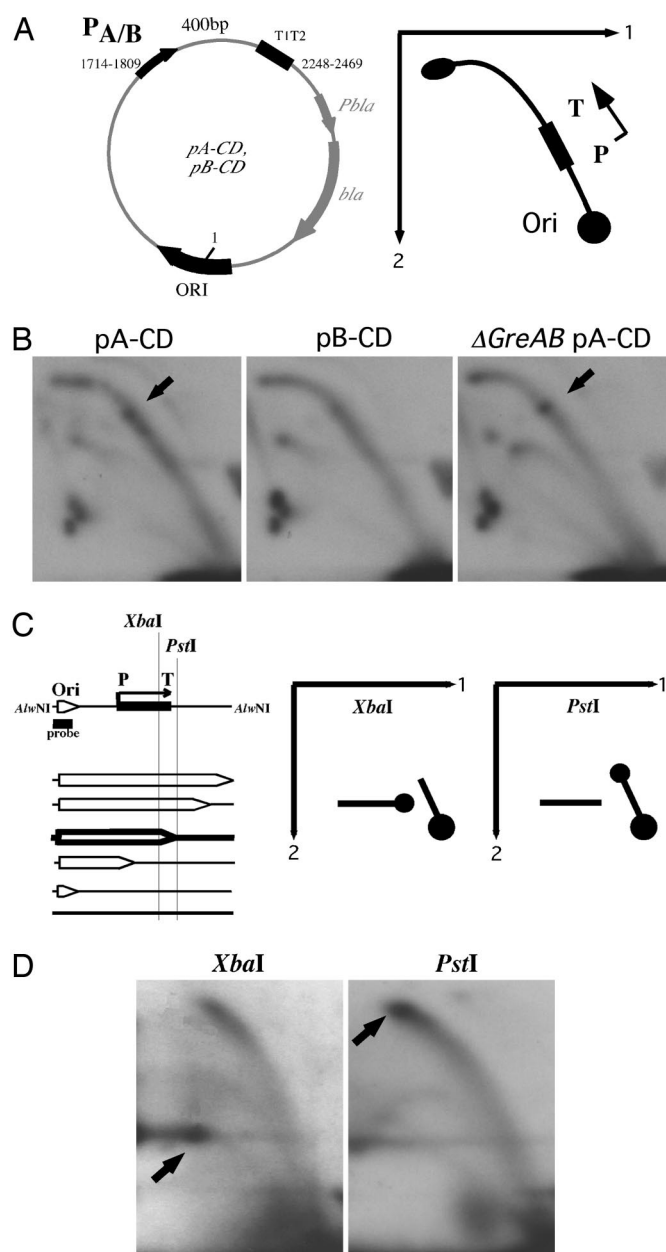


Fig. 3. Replication stalling by the codirectionally oriented transcription terminators. (*A Left*) Map of plasmids pA-CD and pB-CD. Transcription cassettes with promoters A or B, respectively, are oriented codirectionally with the direction of replication; nucleotide position 1 corresponds to the replication start site; positions of the promoters and T1T2 terminators are indicated. (*A Right*) Schematic representation of the two-dimensional gel electrophoresis of replication intermediates. Unlike Fig. 2, a transcription cassette faces up, i.e., away from the origin. (*B*) Two-dimensional gel electrophoresis of replication intermediates. Replication stalling is evident in the pA-CD plasmid at the position that corresponds to the position of the transcription terminators (arrow). (*C*) A scheme of mapping stalled intermediates by in-gel digestion is shown. Replication intermediates were separated in the first dimension followed by digestion with the XbaI and PstI enzymes, flanking the T1T2 terminators (T). (*D*) Experimental mapping of stalled intermediates. During XbaI digestion (*Left*), stalled intermediates become Y-shaped and move to the line, whereas during PstI digestion (*Right*), they remain bubble-shaped and stay on the arc.

with PstI. Consequently, the replication stall zone is located between these two restriction sites, coinciding with the transcription terminators.

Discussion

To look at the interplay between transcription-initiation complexes and DNA replication, we have constructed a promoter that supports efficient formation of closed and open complexes with RNA polymerase but prevents promoter clearance. This end was achieved by converting the ITS (+1 to +20 region) of the very strong A1 promoter of the bacteriophage T7 into the 90% AT-rich sequence without altering any other position. This change led to a dramatic (up to 10-fold) decrease in transcriptional activity both *in vitro* (Fig. 1*B*) and *in vivo* (Fig. 1*C*), without decreasing the efficiency of open-complex formation (Fig. 1*D*). Interestingly, the activity of the mutant promoter was not decreased further in the *greA,greB* knockout *E. coli* strain, and the addition of the GreA and GreB proteins did not improve the efficiency of this promoter *in vitro* (data not shown). Thus, GreAB proteins fail to facilitate promoter clearance in our case, indicating that either the abortive initiation is responsible for the impairment of our mutant promoter or that the RNA polymerase backtracking is so overwhelming that Gre factors do not make any difference.

The mutated promoter seemed, therefore, ideal for studying collisions between transcription-initiation complexes and the replication fork *in vivo*. For the replication studies, we have chosen a bacterial plasmid (43) that contains the entire pBR322 replication origin, including *pasL* and *pasH* sites required for the efficient switching from pol I- to pol III-mediated DNA replication (44). Our transcription cassettes in various orientations (Figs. 2*A* and 3*A*) were positioned $\approx 1,700$ bp downstream from the replication start site and $\approx 1,300$ bp beyond the *pasH* site, i.e., in the area solidly replicated by the DNA polymerase III holoenzyme. Thus, the data obtained in this plasmid system should be perfectly applicable to *E. coli* chromosomal replication.

We observed potent inhibition of DNA replication during head-on collision of the fork with the mutant promoter (Fig. 2*A–C*). Mapping of this replication pause site by in-gel digestion of replication intermediates immediately upstream and downstream from the promoter confirmed that the replication fork was indeed halted at the promoter (Fig. 2*D* and *E*).

When wild-type promoter A faced replication head-on, we observed a severe replication stalling throughout the whole transcribed area (Fig. 2*B Left*), in accord with our previous study (20). That study was carried out in the *E. coli* DH5 α strain (20), which carries mutations in some important components of DNA metabolism. Because our current studies performed in the wild-type, MG1655 Rph⁺ strain gave identical results, head-on transcription–replication collisions could significantly influence nucleic acid metabolism in wild-type *E. coli*.

Unlike what we saw in our previous study (20), however, we saw a trace of replication stalling during codirectional collisions with elongating RNA (Fig. 3*B Left*). Although it was diminutive compared with that for the head-on collisions (compare Figs. 2*B Left* and 3*B Left*), this study demonstrates replication stalling by codirectional collisions with transcription *in vivo*. Such collisions between phage replisomes and bacterial RNA polymerases were, so far, detected only *in vitro* (14, 17). Clearly, however, the head-on collisions with either elongating or initiating RNA polymerase have much more dire consequences for DNA replication. It is tempting to speculate, therefore, that the front edge of the RNA polymerase could be a potent contrahelicase (45).

The amount of full-length transcript generated at the wild-type promoter is approximately twice as high in its head-on orientation as in its codirectional orientation (Fig. 1*C*). This difference was not the result of a variation in copy number for different plasmids because promoter-specific transcripts were normalized to the amount of RNA I in all cases. RNA I is the origin-specific negative regulator of ColE1-like plasmid replication (for review, see ref. 46), and its abundance directly reflects the plasmid copy number. We

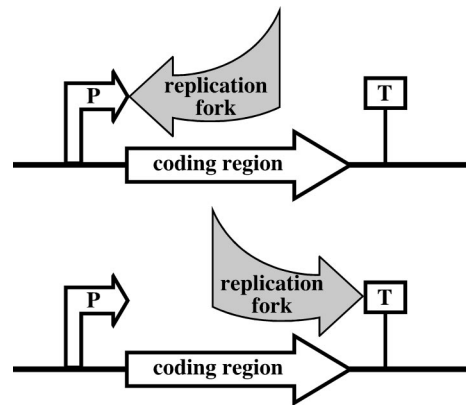


Fig. 4. A model of transcription regulatory elements, serving as punctuation marks for DNA replication. P, promoter; T, terminator. For details, see Discussion.

believe, therefore, that the difference in the mRNA levels depending on the promoter orientation could be caused by different outcomes of transcription–replication collisions.

Most essential genes in bacteria are oriented codirectionally with replication. It was therefore suggested that the deleterious consequences of head-on transcription–replication collisions could be the result of the formation of truncated transcripts and, consequently, truncated proteins, serving as dominant–negative forms of essential proteins (27). Our transcriptional data are in disagreement with this hypothesis because we see more full-length transcripts produced from head-on, rather than codirectionally oriented, promoters.

While studying codirectional collisions, we unexpectedly observed replication stalling in the plasmid region that corresponded to the transcription terminators T1T2 (Fig. 3*A* and *B*). Mapping of this replication pause site has indeed confirmed that it coincides with the terminators (Fig. 3*C* and *D*). Why would the replication fork pause at this regulatory element? A possible explanation could be that a fraction of RNA polymerase molecules may not dissociate from the template at a terminator, but remain bound in some form of a trapped or backtracked complex (47). If this explanation is correct, one would expect a stronger terminator-caused replication stalling in the *greA,greB* mutant (Fig. 3*B Right*). This explanation is also supported by the footprinting, attributed to RNA polymerase, observed at some terminators (48). Further experiments are needed to address the existence and structure of such complexes both *in vitro* and *in vivo*.

Altogether, our data show the replication fork stalling during its head-on encounter with the transcription-initiation complex or its codirectional encounter with the transcription-terminator complex. Notably in both instances, the replication fork is stalled after passing the coding region. It is plausible, therefore, that transcription-initiation and -termination elements could serve as polar punctuation marks for DNA replication, i.e., attenuate the replication fork progression as it traverses the coding areas (Fig. 4). This attenuation could provide extra room for the repair or gene conversion machineries to clear the coding areas off the newly acquired mutations, thus helping to maintain the integrity of the coding regions of sparsely transcribed bacterial genes.

Methods

Construction of Promoters. Both templates were generated by PCR-directed mutagenesis from T7A1 promoter template (49) with DNA polymerase (Phusion; Fermentas, Hanover, MD) and DNA oligonucleotides (IDT, Coralville, IA): 5'-AAAACCTGCAGTC-CAGATCCCCGAAAATTTATCAA-3' (sense primer for both templates) 5'-CTGTTGAATTCGGTTGGCGGAAAGAATA-AATTA AAAAGATGGCTGTAAGTATCCTATAGG-3' (anti-

