3400–3411 *Nucleic Acids Research, 2018, Vol. 46, No. 7* *Published online 21 February 2018*
doi: 10.1093/nar/gky118

Genome-wide relationship between R-loop formation and antisense transcription in *Escherichia coli*

Nalini Raghunathan[1,2], Rajvardhan M. Kapshikar[1,2], Jakku K. Leela[1], Jillella Mallikarjun[1,2], Philippe Bouloc[3] and Jayaraman Gowrishankar[1,*]

[1]Laboratory of Bacterial Genetics, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, Telangana 500039, India, [2]Graduate Studies, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India and [3]Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, F-91198, Gif-sur-Yvette cedex, France

Received November 24, 2017; Revised January 30, 2018; Editorial Decision January 31, 2018; Accepted February 09, 2018

ABSTRACT

Transcription termination by Rho is essential for viability in various bacteria, including some major pathogens. Since Rho acts by targeting nascent RNAs that are not simultaneously translated, it also regulates antisense transcription. Here we show that RNase H-deficient mutants of *Escherichia coli* exhibit heightened sensitivity to the Rho inhibitor bicyclomycin, and that Rho deficiency provokes increased formation of RNA–DNA hybrids (R-loops) which is ameliorated by expression of the phage T4-derived R-loop helicase UvsW. We also provide evidence that in Rho-deficient cells, R-loop formation blocks subsequent rounds of antisense transcription at more than 500 chromosomal loci. Hence these antisense transcripts, which can extend beyond 10 kb in their length, are only detected when Rho function is absent or compromised and the UvsW helicase is concurrently expressed. Thus the potential for antisense transcription in bacteria is much greater than hitherto recognized; and the cells are able to retain viability even when nearly one-quarter of their total non-rRNA abundance is accounted for by antisense transcripts, provided that R-loop formation from them is curtailed.

INTRODUCTION

Transcription termination at the ends of genes and operons in bacteria occurs by two processes, factor independent (or intrinsic) and factor dependent, whose respective contributions are believed to be approximately equal in *Escherichia coli* (1,2). The latter process is also known as Rho-dependent transcription termination (RDTT).

The molecular mechanism of RDTT is reasonably well characterized (reviewed in 1,2,3,4,5,6). Briefly stated, it is mediated by the binding to a nascent transcript of Rho protein, whose subsequent interaction with RNA polymerase (RNAP) in the transcription elongation complex leads to dissociation of the enzyme from the DNA template. The coupling of transcription with translation (which is the norm in bacteria) protects against RDTT within the open-reading frame (ORF) regions, since translating ribosomes sterically prevent Rho's access to the nascent transcript. A second protein NusG is also required for RDTT at some termination sites, and both Rho and NusG are essential for viability in several bacteria including *E. coli* (1–6).

RDTT has been suggested to participate, directly or indirectly, in several functions (that may not be mutually exclusive). These include: the silencing of horizontally transferred genes (7); maintenance of chromosomal integrity (8); prevention of gratuitous excision of prophages (9); regulation of gene expression by attenuation, small RNAs or riboswitches (10–15); suppression of pervasive antisense transcription (16–20); and avoidance of formation of excessive RNA–DNA hybrids or R-loops (21–24).

Antisense transcripts are those that are templated from the 'wrong' strand of ORFs in the genome. Although early studies had identified and characterized a limited number of such RNAs as regulators of gene product abundance, more recent data from next-generation-sequencing experiments have revealed an unexpected and substantial level of antisense transcription in both prokaryotes (25–31) and eukaryotes (32–34), which may be designated as the 'constitutive antisense transcriptome'.

Peters *et al*. (16) have subsequently shown that the potential for antisense transcription in *E. coli* is much higher than that suggested by the constitutive antisense data (35,36), and that this potential is indeed kept in check by RDTT; in their study, following growth in the presence of sublethal concentrations of the Rho inhibitor bicyclomycin (BCM), there was a substantial increase in extent of antisense transcription in the cells. Their findings are consistent with the discovery of numerous intragenic promoters in *E. coli* (37–

*To whom correspondence should be addressed. Tel: +91 40 2721 6000; Fax: +91 40 2721 6006; Email: shankar@cdfd.org.in

42) as well as with the concept that nascent untranslated transcripts are the target of RDTT (1,22), since antisense transcripts are by definition not translated. RDTT has also been shown to inhibit antisense transcription in other bacteria (17,19).

With respect to RDTT and R-loops, Leela *et al*. (23) have shown that the lethality conferred by deletion of *rho* or *nusG* in wild-type (WT) *E. coli* can be rescued by ectopic expression of UvsW, an R-loop helicase of T4 phage (43,44). The model is that, in the absence of RDTT, nascent transcripts that are not being translated are prone to re-annealing with upstream template DNA to generate R-loops which are toxic (22,24). By exploiting the property that C residues in the displaced single-stranded DNA of an R-loop suffer modification upon treatment with bisulfite, Leela *et al*. (23) were also able to infer the genome-wide locations of R-loops purportedly from both sense and antisense transcripts in *E. coli*, and to demonstrate their increased prevalence in a mutant deficient for RDTT.

In this work, we identify more than 500 sites on the *E. coli* chromosome from which antisense transcription is elevated in RDTT-deficient derivatives only when they are also expressing the R-loop helicase UvsW. These loci are well correlated with the antisense regions that were shown to be high-bisulfite-reactive by Leela *et al*. (23), suggesting that, for this subset, it is R-loop formation that precludes their detection following Rho inhibition alone.

MATERIALS AND METHODS

Growth media, bacterial strains, plasmids and primers

Unless otherwise indicated, LB and minimal A (with 0.2% glucose) were used as rich and defined media (45), respectively, and the growth temperature was 37°C. Supplementation with ampicillin (Amp), kanamycin (Kan) and trimethoprim (Tp) were at concentrations described previously (23). Xgal was added at 25 μg/ml, and isopropyl β-D-thiogalactoside (IPTG) at the indicated concentrations in different experiments.

Escherichia coli strain SA1751 (λ *int*$^+$ *xis439 cI857* [*cro-chlA*]Δ*H1*) used for preparation of supercoiled minicircle templates has been described earlier (46). All other strains are derivatives of the reference *E. coli* K-12 strain MG1655, and are listed in Supplementary Table S1.

The following plasmids have been described previously (salient features in parentheses): pSA508 (Amp^R, for generation of supercoiled minicircle templates with multiple-cloning-site and Rho-independent terminator) (46); and pHYD2411 (Tp^R, single-copy-number *rho*$^+$ *lacZ*$^+$) and pHYD2412 (Tp^R, single-copy-number *nusG*$^+$ *lacZ*$^+$) (23). Other plasmids constructed in this study, and the primers used therefor, are described in Supplementary Table S2.

Immunoblotting with S9.6 monoclonal antibody

Immunoblotting with S9.6 monoclonal antibody (Kerafast, USA) for detection of RNA–DNA hybrids was performed essentially as described (47–49). Total nucleic acids were prepared by spooling following chloroform isoamyl alcohol extraction as described in Ausubel *et al*. (50) for preparation of chromosomal DNA, with the modification that the step of RNase treatment was omitted. Cultures of the WT strain GJ13519 grown to mid-exponential phase without or with 25 μg/ml (sublethal) BCM were used, and the yield was ∼10 μg per ml of culture. For each culture, a pair of 10–20 μg aliquots of the preparations was immobilized, with the aid of vacuum suction through a Bio-Dot microfiltration apparatus (Bio-Rad, USA) followed by UV-crosslinking at 1200 J/m^2, on Hybond-N+ nylon membrane (Amersham Biosciences); one aliquot for each pair was treated with two units of RNase H for 1 h at 37°C before the immobilization. A 1:5000 dilution of the antibody preparation was used for immunoblotting overnight at 4°C, followed by reaction with enzyme-conjugated anti-mouse secondary antibody and detection with a chemiluminesence kit.

To determine the equivalence of immobilization of the different nucleic acid samples, the membrane was washed in water, soaked in 5% acetic acid with shaking for 15 min, and then stained with 0.05% methylene blue in 0.5 M sodium acetate buffer (pH5.2) (51).

For strain GJ13531 (Δ*rho* P_{tac}-UvsW), an overnight-grown culture in glucose-minimal A supplemented with 150 μM IPTG was washed twice in minimal A and then inoculated 1:50 in fresh glucose-minimal A medium (without IPTG). This IPTG-depleted cell suspension was incubated for 12 h before it was processed for S9.6 immunoblotting as above; an aliquot of the culture was also plated on appropriate media to confirm the absence of accumulation in it of suppressors or contaminants. Cells from an IPTG-supplemented log-phase culture of GJ13531 were used as the control for this experiment.

Protocols for RNA-Seq experiments

Viable clones of the UvsW-expressing strains GJ13531 (Δ*rho*) and GJ13507 (Δ*nusG*) were obtained as white colonies from their respective shelter plasmid-carrying derivatives GJ13531/pHYD2411 and GJ13507/pHYD2412 on glucose-minimal A plates supplemented with Xgal and IPTG at 200 μM (for Δ*rho*) or 3 μM (for Δ*nusG*), as previously described (23).

Starting from single colonies, the following cultures were set up in triplicate for overnight incubation: GJ13507, GJ13519, GJ13531 and GJ13533 in glucose-minimal A; and GJ13519 also in 0.2% glycerol-minimal A. All the cultures were supplemented with 200 μM IPTG, with the exception of the cultures of GJ13507 whose supplementation with IPTG was at 3 μM. The overnight-grown cultures were each subcultured into 20 ml of fresh medium of the same composition, with an inoculum of 1:50 for GJ13507 and GJ13531 and of 1:100 for the remainder, and grown to an A_{600} of 0.4–0.45, before the cells were harvested for making the RNA preparations as described below. Growth rates for the cultures were determined, and are given in Supplementary Table S3. Aliquots of the cultures were also plated on appropriate media to confirm the absence of accumulation of suppressors or contaminants in any of them. At the appropriate optical density, culture growth was instantaneously arrested by addition of an equal volume of chilled 100% ethanol, and the cells were stored at −80°C until they were processed for RNA extraction. The cells were lysed and total RNA was prepared by the hot phenol method essentially

as described (52,53) after chromosomal DNA has been digested with RNase-free DNase. The quality of RNA preparations was validated by microchannel electrophoresis (Agilent, USA). Strand-specific RNA-Seq data were generated, following rRNA depletion with a Ribo-Zero kit, with the aid of the di-tagged cDNA strategy (ScriptSeq) on an Illumina NextSeq platform.

The sequence data from these experiments, as also from the other publicly available datasets [Peters *et al*. (16); Larson *et al*. (54); and Sedlyarova *et al*. (55)], were then analyzed as described in the Supplementary Data.

In vitro transcription with supercoiled minicircle templates

Techniques for DNA manipulations and polymerase chain reaction were as described (50,56). Derivatives of plasmid pSA508 (46) each carrying the phage T7 A1 promoter and a specified *E. coli* genomic fragment were constructed as described in Supplementary Table S2, and all the insert regions were sequence verified. Transformants of strain SA1751 carrying these plasmids were temperature-induced in one litre-culture volumes as described (46), to enable the generation of supercoiled minicircles that were then purified from 1.5% agarose gels following electrophoresis.

In vitro transcription reactions were each performed essentially as described (57), with 0.45 pmol of supercoiled minicircle DNA as template, 3 pmol of RNAP holoenzyme and ^{32}P-α-CTP as radiolabel, in a total volume of 150 μl, with the modification that heparin was omitted from the mixture and the transcription reaction was stopped by heating to 65°C for 20 min. A total of 10 μl of the mix was retained as 'no-treatment' control; of the remainder, one-half was treated with two units of RNase H for 5 min and 0.5 μg/ml RNase A for 3 min ('A+H' sample) (58), while the other half was treated only with 0.5 μg/ml of RNase A for 3 min. ('A' sample). All three fractions were then treated with a stop-solution mix containing 70 μl of ice-cold precipitation buffer (40 mM ethylenediaminetetraacetic acid, 300 μg/ml salmon sperm DNA, 0.6 M sodium acetate), ethanol precipitated, and subjected to electrophoresis on denaturing 6% polyacrylamide— 8 M urea gels followed by autoradiography, as described (57).

RESULTS AND DISCUSSION

R-loop abundance is increased under RDTT-deficient conditions

To examine whether R-loops are more prevalent following Rho inhibition, we used S9.6 monoclonal antibody, which recognizes RNA–DNA hybrids (59) to perform immunoblotting experiments with total nucleic acid preparations from cultures grown without or with sublethal concentrations of BCM. The BCM-grown culture preparation displayed significantly more reactivity to S9.6 antibody than did the control; this reactivity was also abolished upon RNase H treatment, confirming that the antibody is specific to RNA–DNA hybrids (Figure 1A).

We have earlier shown (23) that lethality conferred by Δ*rho* in WT *E. coli* can be overcome by expression of UvsW, the R-loop helicase from phage T4. Reactivity to S9.6 antibody was low in nucleic acid preparations from cells of the

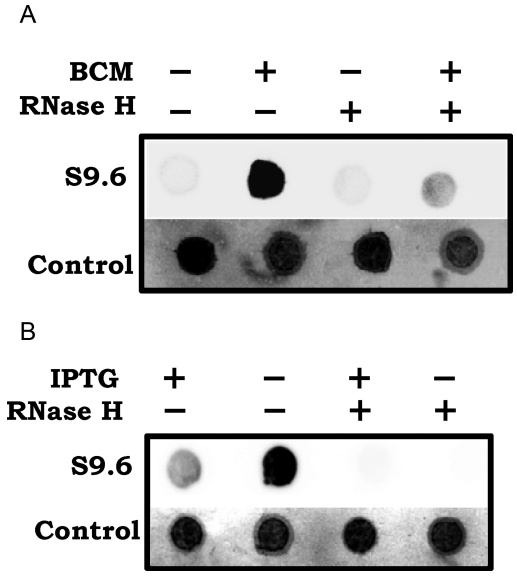

Figure 1. Immunoblot with S9.6 monoclonal antibody of total nucleic acid preparations from cultures (**A**) of WT strain (GJ13519) grown without or with sublethal BCM at 25 μg/ml, and (**B**) of Δ*rho*-UvsW strain (GJ13531) grown under IPTG-replete or -depleted conditions. Where indicated, the nucleic acid preparations were treated *in vitro* with RNase H. To serve as loading controls, methylene blue-stained images of the blotted membranes are also shown.

Δ*rho*-UvsW strain that had been grown in presence of the inducer IPTG needed for expression of UvsW, but it was strongly elevated in equivalent preparations from the culture that was growth-inhibited following IPTG withdrawal; once again, this signal was abolished upon treatment with RNase H (Figure 1B). We conclude that R-loops are increased in cells deficient for RDTT, and furthermore that UvsW expression under these conditions is correlated with both reduced R-loop prevalence and restoration of viability.

Heightened BCM sensitivity of RNase H-deficient mutants

In *E. coli*, RNA–DNA hybrids can be disrupted by the pair of endogenous RNase H enzymes I and II that are encoded, respectively, by *rnhA* and *rnhB*. BCM tolerance phenotypes for *rnhA* and *rnhB* mutants have been reported previously (60,61), but only in the context of studies on whole-genome single-gene knockout collections. In a more detailed and systematic characterization of these phenotypes in the *rnhA* and *rnhB* single and double mutants, we found that the *rnhA* derivative is more sensitive than is the WT strain to inhibition by BCM at both 37 and 42°C (Supplementary Figure S1A and B, respectively), and that the *rnhB* mutation confers no defect. The *rnhA rnhB* double mutant [which is known to be growth-sensitive at temperatures of 37°C and above (23,62)] was even more BCM-sensitive than *rnhA* alone at 30°C (Supplementary Figure S1C). These results indicate that the compromised ability in *rnhA* and *rnhA rnhB* mutants to remove R-loops is associated also with increased sensitivity to BCM.

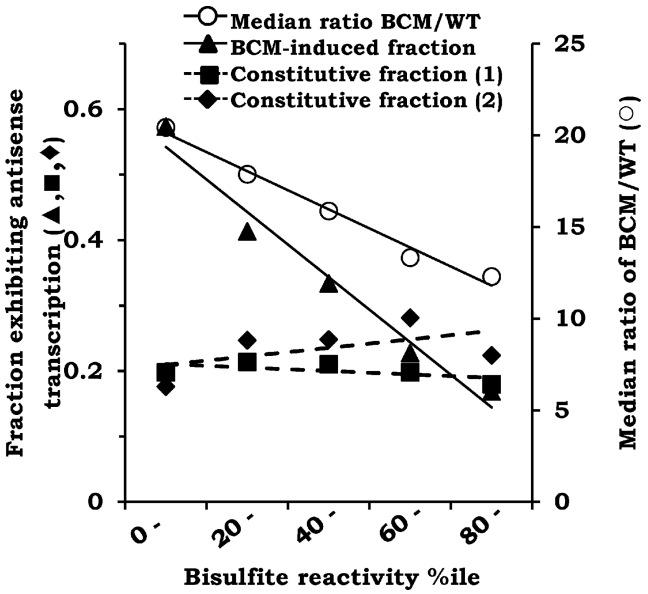

Figure 2. Inverse correlation between BCM-induced antisense transcription [Peters *et al.* (16)] and bisulfite reactivity. Across the different intervals of bisulfite reactivity percentiles are plotted both the fraction of BCM-induced antisense loci (filled triangles), and the median induction ratio of antisense transcripts for these loci (open circles), within each interval. Also shown (interrupted lines) are the fractions of the constitutively transcribed antisense loci (i.e. >80th percentile in the WT strain) for each of the intervals, the data for which have been taken from Peters *et al.* (16) (filled squares, constitutive fraction 1) and Dornenburg *et al.* (35) (filled diamonds, constitutive fraction 2).

Inverse correlation between Rho-inhibited antisense transcription and reactivity to bisulfite

Given both that Rho acts to terminate synthesis of nascent transcripts that are not simultaneously translated, and that such transcripts are also the perceived substrates for R-loop generation, we examined whether a correlation exists between those ORFs where antisense RNA synthesis had been induced following Rho inhibition by sublethal BCM [in the data of Peters *et al.* (16)], and the sites of presumed high-antisense R-loop prevalence [as had previously been determined in the bisulfite reactivity experiments (23)]; for the former, we applied a $\log_2$ induction ratio threshold of 3.

Against our initial expectation of a possible direct correlation between the two datasets, however, we in fact observed the opposite, that they were strongly inversely correlated across the entire range of bisulfite reactivity rankings (Figure 2; $R^2 = 0.97$, $P < 0.002$). Thus, whereas more than one-half of regions in the lowest-ranking interval of bisulfite reactivity exhibited BCM-induced antisense transcription, only one-sixth of those in the highest-ranking interval did so. This same counterintuitive pattern was observed even when the $\log_2$ BCM-induction ratio threshold was reduced to 2 (Supplementary Figure S2A; $R^2 = 0.99$, $P < 0.001$). The median value for the BCM induction ratio also exhibited a progressive decline with increasing bisulfite reactivity (Figure 2 and Supplementary Figure S2A; $R^2 = 0.98$, $P < 0.001$ and $R^2 = 0.92$, $P < 0.01$, respectively). On the other hand, there was little or no correlation between the constitutive antisense transcriptome [from each of two different datasets, namely those of Dornenburg *et al.* (35) and Peters *et al.* (16)], and the regions of varying bisulfite reactivity (Figure 2; $R^2 = 0.24$, $P = 0.39$ and $R^2 = 0.38$, $P = 0.26$, respectively, for the two datasets).

To account for this apparent paradox of inverse correlation between two phenomena that are individually both ostensibly upregulated upon Rho inhibition, we postulated that if a nascent antisense transcript synthesized following Rho inhibition were to form an R-loop, it would inhibit movement of succeeding RNAP molecules and hence lead to reduced transcript abundance for that locus (63–65). Thus, only those RNAs would be identified as BCM-induced that do not form R-loops when they are synthesized following Rho inhibition. Such a model could explain the progressive under-representation of BCM-induced antisense transcription loci with increasing bisulfite reactivity.

Antisense transcriptomes in RDTT-deficient strains expressing an R-loop helicase

One prediction of our model is that the antisense transcription sites above which are under-represented because of R-loop formation will be revealed upon concomitant expression of an R-loop helicase in the RDTT-deficient strain. (Of course, the transcripts previously identified as BCM-induced [that is, which do not form R-loops] would also be expected to be present in this strain).

Accordingly, we designed strand-specific RNA-Seq experiments to determine the transcription profiles in triplicate cultures of Δ*rho* or Δ*nusG* derivatives expressing UvsW (designated hereafter as Δ*rho*-UvsW and Δ*nusG*-UvsW, respectively). We also generated RNA-Seq data for cultures of the WT (*rho*$^+$ *nusG*$^+$) strain, and of the WT strain expressing UvsW (WT-UvsW), as controls.

After the sequence reads were aligned to the reference *E. coli* genome, base read counts (normalized to both ORF length and aggregate read counts) were determined for sense and antisense strands of each of 4091 ORFs that together comprise 84% of the chromosome length in WT *E. coli* (Supplementary Table S4, sheet 1). (In the process, we confirmed both the Δ*rho* and Δ*nusG* status of the test strains and that the Δ*nusG* derivative exhibited a 5-fold increase in transcription of the autoregulated *rho* gene, which is consistent with western blot data from an earlier study (66)). Normalized base read counts were similarly computed for these ORFs from the data of Peters *et al.* (16) for cultures that had been grown without or with BCM (Supplementary Table S4, sheet 1).

Antisense transcripts induced by BCM are also expressed in Δ*rho*-UvsW strain

We initially established that UvsW expression alone had negligible effect on global antisense transcription (Figure 3A, $\log_2$ median ratio of antisense expression for WT-UvsW relative to WT = 0.14, 95% confidence interval [CI] = 0.13–0.15), whereas there was substantial induction of antisense expression with either sublethal BCM exposure [from data of Peters *et al.* (16)] or in the Δ*rho*-UvsW derivative ($\log_2$ median ratios relative to WT, respectively, 2.26 [95% CI = 2.19–2.33] and 2.61 [95% CI = 2.55–2.67]) (Figure 3A).

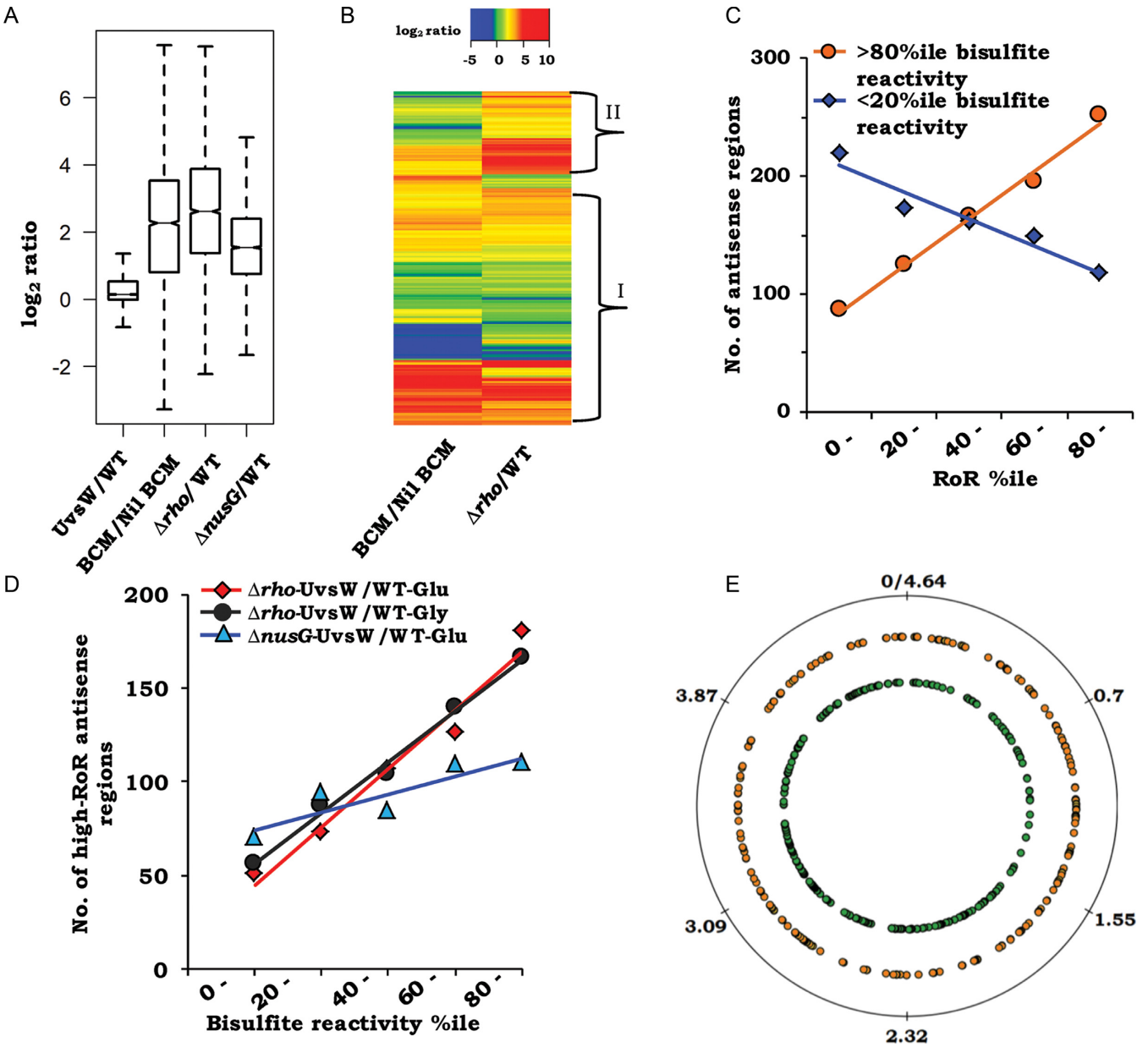

Figure 3. Analysis of antisense RNA-Seq data from RDTT-deficient strains expressing UvsW. (**A**) Box plot representations of the $\log_2$ antisense expression ratios (relative to glucose-grown WT) for the strains WT-UvsW, Δ*rho*-UvsW and Δ*nusG*-UvsW (labeled, respectively, as UvsW/WT, Δ*rho*/WT and Δ*nusG*/WT), as well as for the culture grown in presence of BCM to that in its absence (BCM/Nil BCM) [taken from the data of Peters *et al.* (16)]. (**B**) Heat map representation of the ratio data from panel A for the pair: BCM/Nil BCM and Δ*rho*/WT; hierarchical clusters were ordered by complete linkage method. Regions marked 'I' and 'II' are explained in the text. (**C**) Frequency distribution curves, across the range of Δ*rho*-UvsW RoR percentiles, of antisense loci that are (i) above the 80th percentile (orange circles) and (ii) below the 20th percentile (blue diamonds) of bisulfite reactivity. (**D**) Frequency distribution curves, across the range of bisulfite reactivity percentiles, of high-RoR antisense loci determined from the following comparisons: Δ*rho*-UvsW to WT grown in glucose (red diamonds); Δ*nusG*-UvsW to WT grown in glucose (blue triangles); and Δ*rho*-UvsW to WT grown in glycerol (black circles). (**E**) Representation, on map of circular *Escherichia coli* chromosome, of 535 high-RoR antisense regions in the Δ*rho*-UvsW strain, corresponding to ORFs oriented clockwise (orange, middle circle) and counterclockwise (green, inner circle). Genome coordinates (in Mb) are marked on the outer circle.

As mentioned above, the expectation from our model was that antisense transcripts that do not form R-loops [and thus had been detected earlier as BCM-induced (16)] would also be present in the Δ*rho*-UvsW strain, and that the latter would additionally contain the R-loop forming antisense RNAs as well. These expectations were validated upon inspection of the heat map representation of the $\log_2$ induction ratios for all 4091 antisense regions in the two sets of data (Figure 3B). For the subset of regions marked 'I' in Figure 3B, there was for the major part either no induction or approximately equivalent induction of antisense expression both upon BCM addition and in Δ*rho*-UvsW; in addition, there was also a subset (marked 'II' in Figure 3B) that was more prominently induced in Δ*rho*-UvsW than in the BCM-exposed cultures. These data also serve to confirm our previous finding (23), this time on a genome-wide scale,

that in rescuing Δ*rho* lethality, UvsW expression does not reverse the transcription termination defect in the strain.

Identification of novel antisense transcripts that are abundant only with combined RDTT deficiency and UvsW expression

To follow-up on the heat map data and to identify antisense regions that were differentially induced upon the combined influence of RDTT deficiency and R-loop helicase expression compared to that with RDTT deficiency alone, we determined the ratio of ratios (RoR) for the antisense strand of each of the ORFs. In the RoR calculation, the numerator and denominator ratios were the magnitudes of induction, respectively, (i) in Δ*rho*-UvsW (relative to WT) from our data, and (ii) by sublethal BCM (relative to the cultures without BCM) in the data of Peters *et al*. (16). The expectation was that high-RoR regions would represent those in which R-loop formation had precluded their induction by BCM addition alone.

For antisense regions which are high-ranking (>80th percentile) for bisulfite reactivity (that is, with presumed maximal R-loop prevalence), the frequency distribution plot showed a strong direct correlation with RoR for the Δ*rho*-UvsW derivative (Figure 3C; $R^2 = 0.99$, $P < 0.001$). The converse was also true, in that antisense regions that were the least ranking (<20th percentile) for bisulfite reactivity (minimal R-loop prevalence) exhibited a strong inverse correlation and were under-represented at high RoRs (Figure 3C; $R^2 = 0.93$, $P < 0.01$).

From the data, we then identified 535 sites of antisense transcription that fulfilled the following criteria to indicate that they are transcribed preferentially only under the combined conditions of Rho deficiency and UvsW expression (compared to that with Rho inhibition alone): $\log_2$ ratio of induction by BCM in data of Peters *et al*. (16) <3; $\log_2$ ratio of expression in Δ*rho*-UvsW strain relative to WT >2; and RoR >80th percentile (corresponding to $\log_2$ RoR of around two and above). Their distribution showed a strong positive correlation across the range of ranks of bisulfite reactivity, with a 3.5-fold difference in numbers between the highest and lowest quintile rank intervals (Figure 3D; $R^2 = 0.97$, $P < 0.003$). When the criterion relating to BCM induction was modified to $\log_2$ ratio <2, the identified antisense region numbers were decreased to 412 but the strong positive correlation across the range of ranks of bisulfite reactivity remained, this time with a 5-fold enrichment for the highest rank interval compared to the lowest (Supplementary Figure S2B; $R^2 = 0.91$, $P < 0.01$).

Taken together, these findings are strongly supportive of the model that a subset of antisense loci which are transcribed following Rho inhibition suffer R-loop formation and therefore need expression of an R-loop helicase as well for their detection. The 535 regions identified, which we refer to below as high-RoR antisense regions, are organized in 370 clusters and show a fairly uniform distribution across the genome on both strands (Supplementary Table S4, sheet 2; see also Figure 3E). Given that UvsW expression alone had no significant effect on antisense expression (Figure 3A), these results suggest that the majority of R-loops upon which the helicase acts are formed only after inhibition of RDTT.

Analysis of antisense transcription in representative high-RoR clusters

For two representative high-RoR clusters, a contig analysis of the RNA-Seq read data clearly established the occurrence of novel antisense transcripts that could span multiple ORF boundaries and extend even beyond 10 kb in their length (Figure 4). These RNAs were present only in the Δ*rho*-UvsW strain, but not in the WT strain nor in cultures exposed to just one of the two perturbations in isolation (BCM addition, or UvsW expression).

To obtain a more fine-grained insight from the RNA-Seq data, we determined the read counts for successive 100-base regions across the genome for both strands, and calculated $\log_2$ ratios (relative to WT) after normalization and thresholding, as described in the Supplementary Data (Supplementary Table S4, sheet 3). Plots for the representative examples of antisense regions that display unchanged transcription (nil or constitutive) under all conditions, transcription with RDTT deficiency, and transcription for multi-ORF clusters only upon combined RDTT deficiency and UvsW expression, are shown in Supplementary Figure S3. Our data indicate once again that at multiple loci, antisense transcripts are detected only with the last combination, and that the induction is quite marked (>100-fold) under these conditions. These high-RoR regions were also the ones that exhibited high bisulfite reactivity.

Sense transcription is not inhibited in Δ*rho*-UvsW strain

In cells of the Δ*rho*-UvsW strain, antisense RNAs comprised 22% of the total non-rRNA abundance (determined as the average of data for the three replicate cultures). Nevertheless, sense transcription in the Δ*rho*-UvsW strain was not inhibited (and in fact was moderately elevated), as determined both for the full set of 4091 ORFs ($\log_2$ median expression for WT and Δ*rho*-UvsW, respectively, 6.01 [95% CI = 5.90–6.11] and 6.45 [95% CI = 6.40–6.51]) and for the subset of 535 'high-RoR' ORFs (corresponding values, respectively, 5.71 [95% CI = 5.44–5.98] and 6.02 [95% CI = 5.87–6.18]) (see Supplementary Figure S4). Our findings are in agreement with the conclusions of Peters *et al*. (16) that increased antisense transcription following Rho inhibition is not associated with concomitant decrease in sense transcription.

Antisense transcription is similarly affected by UvsW expression in Δ*nusG* as in Δ*rho*, and is unrelated to growth rate changes

From the box plot representation (Figure 3A), it was evident that antisense expression in Δ*nusG*-UvsW was increased relative to that in WT ($\log_2$ median induction ratio = 1.54, 95% CI = 1.49–1.58), but to a less extent than in Δ*rho*-UvsW. It is known that RDTT is compromised more severely in absence of Rho than in absence of NusG (7,16,67,68).

From the Δ*nusG*-UvsW data, high-RoR antisense regions were determined by the approach similar to that described above for the Δ*rho*-UvsW strain, with the modification that a lower threshold of $\log_2$ induction ratio >1 (relative to WT) was employed. The frequency distribution of

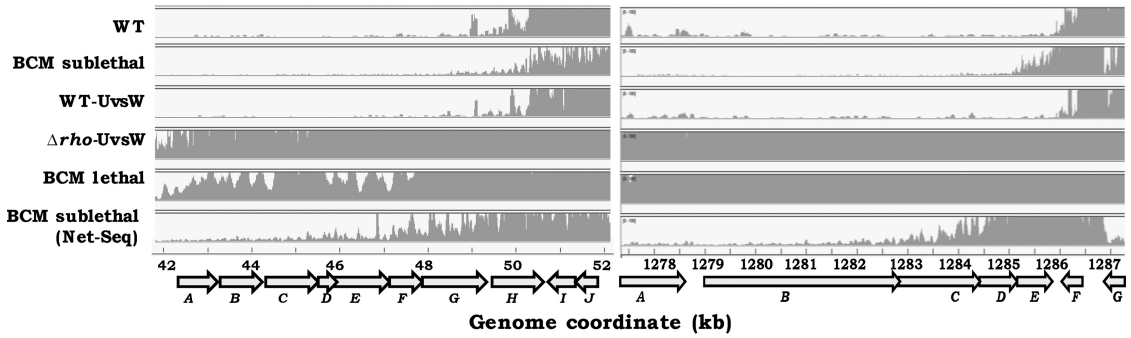

Figure 4. Contig analysis of sequence data from two representative genomic regions. For each panel, genome coordinates (in kb) and ORF annotations are given at the bottom. Plotted are the base-specific read numbers, with an upper cut-off limit at 100, in the data for RNA preparations from one replicate each of the following cultures: WT, glucose-grown GJ13519; (+) BCM sublethal, culture grown with sublethal BCM in study of Peters *et al*. (16); WT-UvsW, GJ13533; Δ*rho*-UvsW, GJ13531; (+) BCM lethal, culture harvested after lethal BCM exposure in study of Sedlyarova *et al*. (55); and (+) BCM sublethal (Net-Seq), culture grown with sublethal BCM in study of Larson *et al*. (54). In both the panels, data for only one strand (bottom) is shown which is antisense for the ORFs in clockwise orientation. ORFs have been alphabetically listed in the following order (from left to right): left panel, *fixA*, -*B*, -*C*, -*X*, *yaaU*, *kefF*, -*C*, *folA*, *apaH*, -*C*; and right panel, *narK*, -*G*, -*H*, -*J*, -*I*, *tpr*, *purU*.

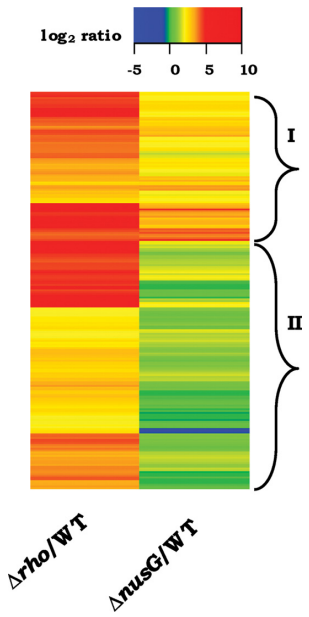

Figure 5. Heat map representation of $\log_2$ antisense expression ratios (relative to glucose-grown WT) for the strains Δ*rho*-UvsW and Δ*nusG*-UvsW (labeled, respectively, as Δ*rho*/WT, and Δ*nusG*/WT) with respect to the 535 high-RoR regions; hierarchical clusters were ordered by wardD2 linkage method. Regions marked 'I' and 'II' are explained in the text.

the high-RoR regions in Δ*nusG*-UvsW was moderately positively correlated across the range of ranks of bisulfite reactivity (Figure 3D; $R^2 = 0.78$, $P < 0.05$).

Peters *et al*. (16) had shown that antisense transcripts induced upon BCM addition can be divided into two categories depending upon whether they are, or are not, also induced in NusG-deficient cells; they had designated the cognate sites of Rho termination as, respectively, NusG-dependent and NusG-independent, whose relative proportion was ∼1:4 across the genome. The latter also exhibited stronger 'rho utilization' sites, with higher C/G ratios in the transcript sequences, than the former.

A comparative analysis of antisense transcription data for the Δ*rho*-UvsW and Δ*nus*G-UvsW derivatives was reasonably consistent with the findings above of Peters *et al*. (16). Thus, in the heat map representation of the 535 high-RoR antisense regions [which are completely non-overlapping with the BCM-induced regions of Peters *et al*. (16)], both NusG-dependent (that is, also induced in Δ*nus*G-UvsW albeit to less extent than in Δ*rho*-UvsW) and NusG-independent subsets could be identified, which are demarcated as 'I' and 'II', respectively, in Figure 5. The median C/G ratio of the deduced antisense transcript sequences for the NusG-independent subset of 353 regions ($\log_2$ induction ratio for Δ*nus*G-UvsW relative to WT <1) was 1.14, which was significantly higher than that for the NusG-dependent subset of 182 regions, at 1.07 ($P < 10^{-5}$, Mann–Whitney U test).

To exclude the possibility that slow growth of Δ*rho*-UvsW and Δ*nusG*-UvsW derivatives (relative to WT) may account for the differences in their transcriptomes, we also performed RNA-Seq analyses for triplicate cultures of the WT strain grown with glycerol, wherein the growth rate was about 60% of that for glucose-grown cultures [Supplementary Table S3, see also (69)]. A strong correlation was observed when the $\log_2$ antisense RNA induction ratios for the Δ*rho*-UvsW strain, calculated relative to the WT strain grown with glucose on the one hand or glycerol on the other, were compared (Supplementary Figure S5; $R^2 = 0.84$, $P < 10^{-15}$). The distribution of high-RoR regions determined from the data for the glycerol-grown cultures (using the same criteria as that for the glucose-grown cultures) was also positively correlated across the range of ranks of bisulfite reactivity (Figure 3D; $R^2 = 0.99$, $P < 0.0003$).

R-loops are generated in a high-RoR antisense region during *in vitro* transcription

R-loops generated during *in vitro* transcription can be identified by the property that the radiolabeled transcripts would be resistant to RNase A and sensitive to RNase H (58). We generated templates (Supplementary Table S2) in

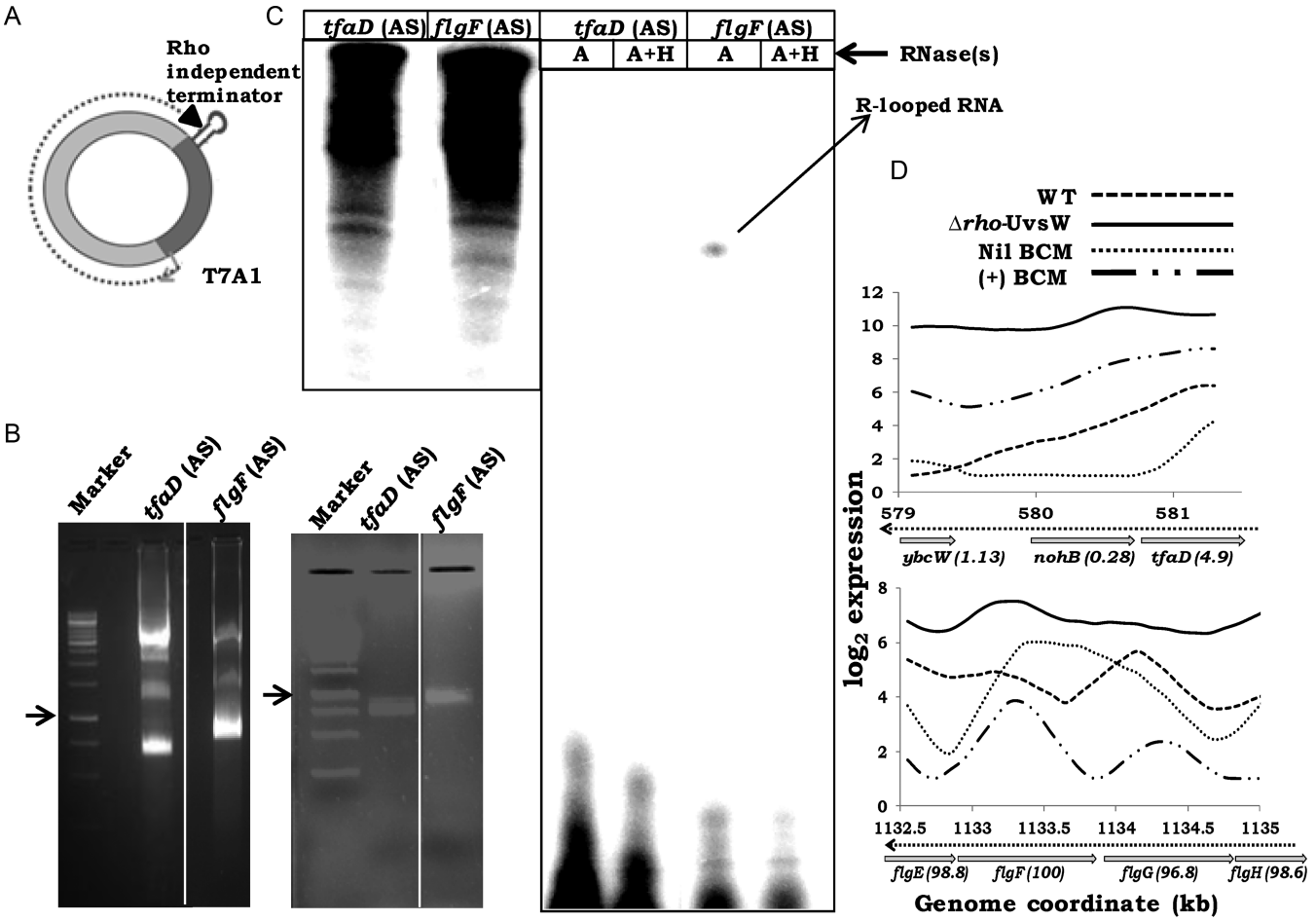

Figure 6. Detection of R-loops during *in vitro* transcription. (**A**) Schematic depiction of supercoiled minicircle template following excision from pSA508 derivative, with cloned genomic fragment between T7 A1 promoter and Rho-independent terminator. (**B**) Agarose gel electrophoregrams of pSA508 derivatives [bearing antisense (AS) orientations of fragments from *tfaD* and *flgF*] following temperature induction in SA1751 cultures (left), and after purification of the supercoiled minicircles (right). Arrows denote 1-kb band in each of the two marker lanes. (**C**) Autoradiographs following denaturing gel electrophoresis of *in vitro* transcription products from the supercoiled minicircles, before (left panel) and after (right panel) treatment with RNase(s) A or A with H (A+H), as indicated. Lanes for samples without RNase treatment were loaded with one-seventh of the amounts loaded on each of the lanes for RNase-treated samples. (**D**) Plots of the $\log_2$ antisense expression values for the *tfaD* and *flgF* loci for cultures of WT strain without (Nil BCM) and with sublethal BCM (+ BCM) [data from Peters *et al.* (16)], and from this study for WT and Δ*rho*-UvsW cultures. Genome coordinates (in kb) and ORF annotations are marked, with the antisense bisulfite reactivity percentile indicated in parentheses beside each ORF. Direction of antisense transcription is shown by the interrupted arrow.

antisense orientation from both a region (*flgF*) exhibiting high-antisense RoR, and a second control region (*tfaD*) whose antisense transcription was induced to equivalent extent either by BCM alone or in the Δ*rho*-UvsW derivative (bisulfite reactivity percentiles 100 and 4.9, respectively; see Figure 6D). The DNA fragments were cloned such as to reside downstream of the phage T7 A1 promoter and upstream of a Rho-independent terminator. Since R-loop formation is facilitated by negative supercoiling (70), the templates were prepared as supercoiled minicircles by a protocol of *in vivo* site-specific recombination from plasmids, as previously described (46) (Figure 6A and B).

Following *in vitro* transcription from the pair of supercoiled minicircle templates, RNase A-resistant transcripts were observed only for the *flgF* ORF region in its antisense orientation (Figure 6C); these transcripts were sensitive to digestion by RNase H. Transcripts from the control *tfaD* antisense template were fully RNase A-sensitive (Figure 6C). We conclude that transcripts from a high-RoR antisense region do form R-loops *in vitro* on a supercoiled template.

Antisense expression analyses in other transcriptome datasets of Rho inhibition

We analyzed two other public sets of *E. coli* transcriptome data related to RDTT (Supplementary Table S4, sheet 1): one was of nascent transcript sequencing (Net-Seq) (54) from cultures without or with sublethal BCM, that is, conditions that were similar to those employed earlier for RNA-Seq by Peters *et al.* (16); and the other was from RNA-Seq experiments of cultures without or with exposure for 20 min to lethal concentrations of BCM (55). As described below, the former set [Net-Seq with sublethal BCM (54)] resembled that of Peters *et al.* (16), whereas the latter [RNA-Seq with

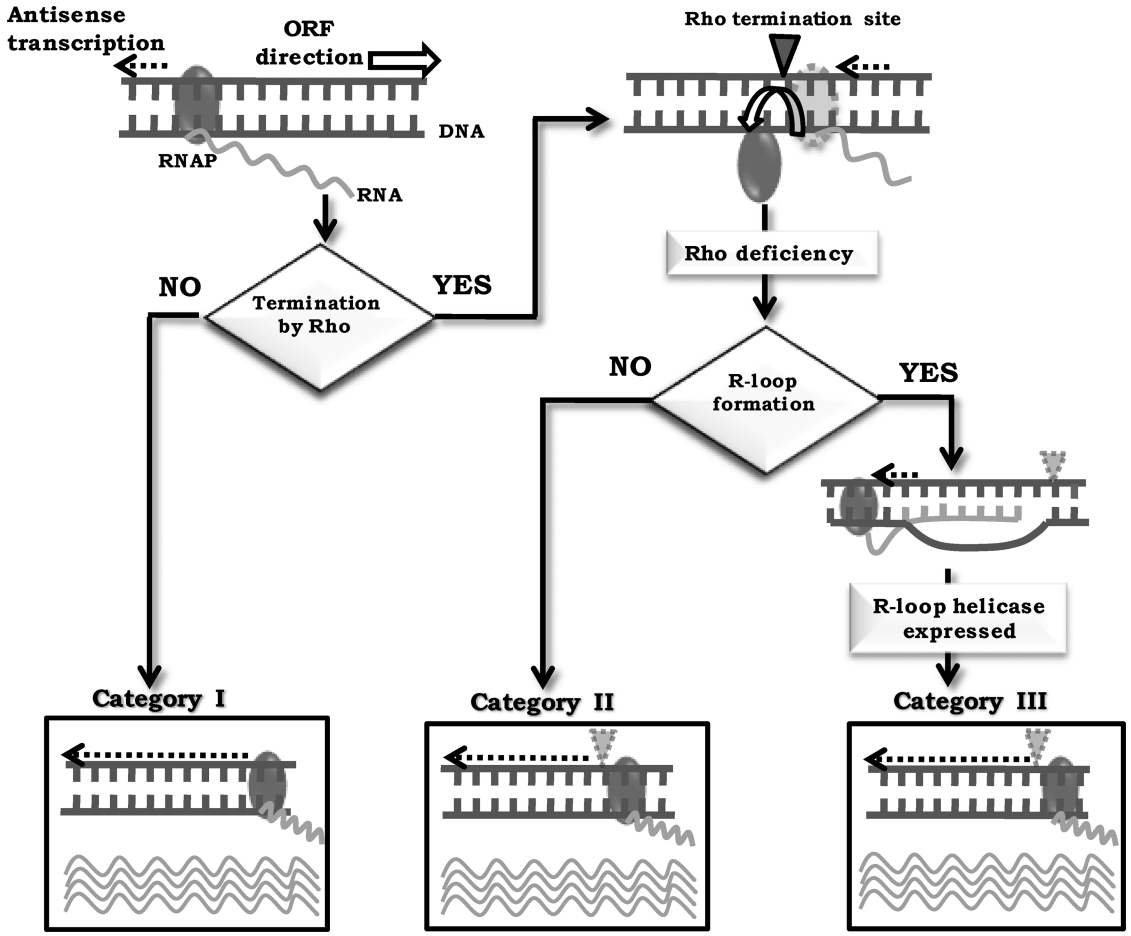

Figure 7. Model to explain three categories of bacterial antisense transcription; see text for details.

lethal BCM (54)] was similar to that from the Δ*rho*-UvsW derivative.

Thus, the BCM-induced antisense regions from the Net-Seq data exhibited an inverse correlation in their distribution across the range of ranks of bisulfite reactivity, whereas those from the lethal-BCM exposure experiment showed no significant correlation (Supplementary Figure S6A; $R^2 = 0.98$, $P < 0.001$ and $R^2 = 0.57$, $P = 0.13$, respectively). Furthermore as depicted in Supplementary Figure S6B for the 535 high-RoR antisense regions identified above, there was little induction following sublethal BCM exposure of synthesis of both free transcripts (data of Peters *et al.* (16)) and nascent transcripts (Net-Seq data) ($\log_2$ median induction ratios, respectively, 1.05 [95% CI = 0.92–1.18] and 1.32 [95% CI = 1.23–1.41]); on the other hand, these regions were induced by lethal BCM exposure ($\log_2$ median induction ratio = 4.89 [95% CI = 4.76–5.02]) to the same extent as in the Δ*rho*-UvsW derivative ($\log_2$ median induction ratio = 3.85 [95% CI = 3.72–3.97]). The same pattern was noted in the contig analysis of the sequence read data as well, wherein lethal BCM exposure was associated with strong antisense transcription of those regions that were not induced by sublethal BCM (Figure 4).

The finding that relative antisense transcript abundance in the high-RoR regions (in absence of UvsW expression) is low when the Rho deficiency is chronic (after sublethal BCM exposure for several hours), and high when it is acute (after lethal BCM exposure for 20 min), suggests that R-loops at these sites may accumulate with time, and furthermore that these R-loops then act to block progressive rounds of transcription elongation (63–65).

CONCLUSIONS

We have shown here that transcripts from more than 500 antisense loci are synthesized in RDTT-deficient strains predominantly only when UvsW is also expressed in them. These high-RoR regions were also correlated with the loci that had previously been inferred, based on their reactivity to bisulfite, to be R-looped *in vivo*. The simplest interpretations for these observations is that at a subset of regions where antisense transcription occurs following Rho inhibition, R-loops are generated that serve as road-blocks for further transcription (63–65); and that UvsW's action as an R-loop helicase relieves the road-blocks at these regions. The direct demonstration (i) in Rho-deficient cells of elevated R-loop prevalence that is alleviated upon UvsW expression, and (ii) of R-loop occurrence during *in vitro* transcription at a high-RoR antisense locus, supports this model. In an earlier study (23), we had also shown that

UvsW expression rescues the inviability of an *rnhA rnhB* double mutant at 42°C.

Apart from its R-loop helicase activity, UvsW has also been shown to possess activity *in vitro* as a junction DNA helicase that can catalyze branch migration and fork regression reactions (71–74). It has also been suggested that the mechanism of UvsW's action in rescuing Δ*rho* or Δ*nusG* lethality may perhaps be related to its ability to resolve problems related to DNA replication in the mutants (75). In the context of the present study, however, it is quite unlikely that UvsW's DNA junction helicase activity would explain the detection of novel antisense RNAs in RDTT-deficient cells from the loci of high bisulfite reactivity. Indeed, the strong and direct correlation established in this study, between bisulfite-reactive loci and the sites of novel antisense transcription, offers mutual reinforcement to the twin notions (i) that the former represent R-looped regions, and (ii) that the latter reflect UvsW's ability to unwind them.

Thus, three categories of bacterial antisense transcription may be envisaged (Figure 7). The first is constitutive antisense transcription whereas in the second and the third, transcripts are only synthesized following Rho inhibition. The second category are antisense RNAs that do not form R-loops whereas the third category (represented by the transcripts from high-RoR regions) form R-loops. The features that distinguish between the latter two categories remain to be determined; previous genome-scale studies in other organisms have variously implicated high levels of transcript expression, as well as different characteristics of the RNA sequences such as G-richness and G/C skew or poly-A tracts and A/T skew, for R-loop formation (76–80).

The fact that the Δ*rho*-UvsW derivative, containing high levels of antisense RNAs, is viable (while Δ*rho* is lethal) suggests that the antisense transcripts themselves are not toxic and that they become so only upon associating with DNA to form R-loops. Rho's essential role therefore is in curtailing antisense transcription and consequential R-loop formation. Rho is essential for viability in many other bacteria including major pathogens such as *Mycobacterium tuberculosis* (20), and here too it is possible that R-loops from untranslated antisense transcripts are the cause of lethality under the Rho-deficient conditions.

R-loop toxicity may be caused not only by transcriptional road-blocking as mentioned above (63–65), but also by arrest and titration of RNAP molecules (41), impedance of replication fork progression (81,82), and aberrant initiation of DNA replication (24,83,84). Washburn and Gottesman (8,75) have earlier suggested that Rho inhibition can lead to transcription-replication conflicts and thus to loss of genome integrity, which is consistent with the model for R-loop toxicity in bacterial cells. In eukaryotes as well, R-loops are now similarly recognized to be a major contributor to compromise of genomic integrity and thus to cancer pathogenesis (85–89), and action of an R-loop helicase has recently been shown to ameliorate R-loop toxicity in mammalian cells (90).

DATA AVAILABILITY

The RNA-Seq data generated in this study are available at GEO https://www.ncbi.nlm.nih.gov/geo/ under accession number GSE103937.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The RNA-Seq data for this study were generated at the Plateforme de Séquençage à Haut Débit facility of I2BC. BCM was a kind gift from Max Gottesman. We thank Saswat Mohapatra, Audrey Vingadassalon and especially Rohan Misra for assistance with experiments and data analysis; and COE team members for advice and discussions.

FUNDING

Centre of Excellence (COE) Project for Microbial Biology—Phase 2; Indo-French DST-ANR Project 2013–06; DST-INSPIRE Fellowships (to N.R., J.M.); ICMR Research Fellowship (to R.M.K.); J C Bose Fellowship (to J.G.); INSA Senior Scientist Award (to J.G.).
Conflicting interest statement. None declared.

REFERENCES

1. Adhya,S. and Gottesman,M.E. (1978) Control of transcription termination. *Annu. Rev. Biochem.*, **47**, 967–996.
2. Ray-Soni,A., Bellecourt,M.J. and Landick,R. (2016) Mechanisms of bacterial transcription termination: all good things must end. *Annu. Rev. Biochem.*, **85**, 319–347.
3. Boudvillain,M., Figueroa-Bossi,N. and Bossi,L. (2013) Terminator still moving forward: expanding roles for Rho factor. *Curr. Opin. Microbiol.*, **16**, 118–124.
4. Grylak-Mielnicka,A., Bidnenko,V., Bardowski,J. and Bidnenko,E. (2016) Transcription termination factor Rho: a hub linking diverse physiological processes in bacteria. *Microbiology*, **162**, 433–447.
5. Kriner,M.A., Sevostyanova,A. and Groisman,E.A. (2017) Learning from the leaders: gene regulation by the transcription termination factor Rho. *Trends Biochem. Sci.*, **41**, 690–699.
6. Mitra,P., Ghosh,G., Hafeezunnisa,M. and Sen,R. (2017) Rho protein: roles and mechanisms. *Annu. Rev. Microbiol.*, **71**, 687–709.
7. Cardinale,C.J., Washburn,R.S., Tadigotla,V.R., Brown,L.M., Gottesman,M.E. and Nudler,E. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science*, **320**, 935–938.
8. Washburn,R.S. and Gottesman,M.E. (2011) Transcription termination maintains chromosome integrity. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 792–797.
9. Menouni,R., Champ,S., Espinosa,L., Boudvillain,M. and Ansaldi,M. (2013) Transcription termination controls prophage maintenance in *Escherichia coli* genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 14414–14419.
10. Bossi,L., Schwartz,A., Guillemardet,B., Boudvillain,M. and Figueroa-Bossi,N. (2012) A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes Dev.*, **26**, 1864–1873.
11. Proshkin,S., Mironov,A. and Nudler,E. (2014) Riboswitches in regulation of Rho-dependent transcription termination. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1839**, 974–977.
12. Hollands,K., Sevostiyanova,A. and Groisman,E.A. (2014) Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E1999–E2007.

13. Kriner,M.A. and Groisman,E.A. (2015) The bacterial transcription termination factor Rho coordinates Mg^{2+} homeostasis with translational signals. *J. Mol. Biol.*, **427**, 3834–3849.
14. Gall,A.R., Datsenko,K.A., Figueroa-bossi,N., Bossi,L., Masuda,I., Hou,Y. and Csonka,L. (2016) Mg^{2+} regulates transcription of *mgtA* in *Salmonella typhimurium* via translation of proline codons during synthesis of the MgtL peptide. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 15096–15101.
15. Bastet,L., Chauvier,A., Singh,N., Lussier,A., Wade,J.T., Lamontagne,A., Prevost,K., Mass,E. and Lafontaine,A. (2017) Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation. *Nucleic Acids Res.*, **45**, 7474–7486.
16. Peters,J.M., Mooney,R.A., Grass,J.A., Jessen,E.D., Tran,F. and Landick,R. (2012) Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes Dev.*, **26**, 2621–2633.
17. Nicolas,P., Mäder,U., Dervyn,E., Rochat,T., Leduc,A., Pigeonneau,N., Bidnenko,E., Marchadier,E., Hoebeke,M., Aymerich,S. *et al.* (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science*, **335**, 1103–1106.
18. Mäder,U., Nicolas,P., Depke,M., Pané-Farré,J., Debarbouille,M., van der Kooi-Pol,M.M., Guérin,C., Dérozier,S., Hiron,A., Jarmer,H. *et al.* (2016) *Staphylococcus aureus* transcriptome architecture: from laboratory to infection-mimicking conditions. *PLOS Genet.*, **12**, e1005962.
19. Bidnenko,V., Nicolas,P., Grylak-Mielnicka,A., Delumeau,O., Auger,S., Aucouturier,A., Guerin,C., Francis,R., Bardowski,J., Aymerich,S. *et al.* (2017) Termination factor Rho: from the control of pervasive transcription to cell fate determination in *Bacillus subtilis*. *PLoS Genet.*, **13**, e1006909.
20. Botella,L., Vaubourgeix,J., Livny,J. and Schnappinger,D. (2017) Depleting *Mycobacterium tuberculosis* of the transcription termination factor Rho causes pervasive transcription and rapid death. *Nat. Commun.*, **8**, 14731.
21. Harinarayanan,R. and Gowrishankar,J. (2003) Host factor titration by chromosomal R-loops as a mechanism for runaway plasmid replication in transcription termination-defective mutants of *Escherichia coli*. *J. Mol. Biol.*, **332**, 31–46.
22. Gowrishankar,J. and Harinarayanan,R. (2004) Why is transcription coupled to translation in bacteria? *Mol. Microbiol.*, **54**, 598–603.
23. Leela,J.K., Syeda,A.H., Anupama,K. and Gowrishankar,J. (2013) Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 258–263.
24. Gowrishankar,J., Leela,J.K. and Anupama,K. (2013) R-loops in bacterial transcription: their causes and consequences. *Transcription*, **4**, 153–157.
25. Thomason,M.K. and Storz,G. (2011) Bacterial antisense RNAs: how many are there and what are they doing? *Annu. Rev. Genet.*, **44**, 167–188.
26. Sesto,N., Wurtzel,O., Archambaud,C., Sorek,R. and Cossart,P. (2013) The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nat. Rev. Microbiol.*, **11**, 75–82.
27. Lybecker,M., Bilusic,I. and Raghavan,R. (2014) Pervasive transcription: detecting functional RNAs in bacteria. *Transcription*, **5**, e944039.
28. Wade,J.T. and Grainger,D.C. (2014) Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat. Rev. Microbiol.*, **12**, 647–653.
29. Grainger,D.C. (2016) The unexpected complexity of bacterial genomes. *Microbiology*, **162**, 1167–1172.
30. Mars,R.A.T., Nicolas,P., Denham,E.L. and van Dijl,J.M. (2016) Regulatory RNAs in *Bacillus subtilis*: a Gram-positive perspective on bacterial RNA-mediated regulation of gene expression. *Microbiol. Mol. Biol. Rev.*, **80**, 1029–1057.
31. Lloréns-Rico,V., Cano,J., Kamminga,T., Gil,R., Latorre,A., Chen,W.-H., Bork,P., Glass,J.I., Serrano,L. and Lluch-Senar,M. (2016) Bacterial antisense RNAs are mainly the product of transcriptional noise. *Sci. Adv.*, **2**, e1501363.
32. Pelechano,V. and Steinmetz,L.M. (2013) Gene regulation by antisense transcription. *Nat. Genet.*, **14**, 880–893.
33. Khorkova,O., Myers,A.J., Hsiao,J. and Wahlestedt,C. (2014) Natural antisense transcripts. *Hum. Mol. Genet.*, **23**, 54–63.
34. Sun,Y., Li,D., Zhang,R., Peng,S., Zhang,G., Yang,T. and Qian,A. (2017) Strategies to identify natural antisense transcripts. *Biochimie*, **132**, 131–151.
35. Dornenburg,J.E., DeVita,A.M., Palumbo,M.J. and Wade,J.T. (2010) Widespread antisense transcription in *Escherichia coli*. *Mbio*, **1**, e00024-10.
36. Raghavan,R., Sloan,D.B. and Ochman,H. (2012) Antisense transcription is pervasive but rarely conserved in enteric bacteria. *Mbio*, **3**, e00156-12.
37. Kawano,M., Storz,G., Rao,B.S., Rosner,J.L. and Martin,R.G. (2005) Detection of low-level promoter activity within open reading frame sequences of *Escherichia coli*. *Nucleic Acids Res.*, **33**, 6268–6276.
38. Shimada,T., Yamazaki,Y., Tanaka,K. and Ishihama,A. (2014) The whole set of constitutive promoters recognized by RNA polymerase RpoD holoenzyme of *Escherichia coli*. *PLoS One*, **9**, e90447.
39. Wade,J.T. (2015) Where to begin? Mapping transcription start sites genome-wide in *Escherichia coli*. *J. Bacteriol.*, **197**, 4–6.
40. Thomason,M.K., Bischler,T., Eisenbart,S.K., Förstner,K.U., Zhang,A., Herbig,A., Nieselt,K., Sharma,C.M. and Storz,G. (2015) Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J. Bacteriol.*, **197**, 18–28.
41. Lamberte,L.E., Baniulyte,G., Singh,S.S., Stringer,A.M., Bonocora,R.P., Stracy,M., Kapanidis,A.N., Wade,J.T. and Grainger,D.C. (2017) Horizontally acquired AT-rich genes in *Escherichia coli* cause toxicity by sequestering RNA polymerase. *Nat. Microbiol.*, **2**, 16249.
42. Wade,J.T. and Grainger,D.C. (2018) Spurious transcription and its impact on cell function. *Transcription*, doi:10.1080/ 21541264. 2017.1381794.
43. Carles-Kinch,K., George,J.W. and Kreuzer,K.N. (1997) Bacteriophage T4 UvsW protein is a helicase involved in recombination, repair and the regulation of DNA replication origins. *EMBO J.*, **16**, 4142–4151.
44. Dudas,K.C. and Kreuzer,K.N. (2001) UvsW protein regulates bacteriophage T4 origin-dependent replication by unwinding R-loops. *Mol. Cell. Biol.*, **21**, 2706–2715.
45. Miller,JH (1992) *A Short Course in Bacterial Genetics: a Laboratory Manual and Handbook for* Escherichia coli *and Related Bacteria*. Cold Spring Harbor Lab Press, NY.
46. Choy,H.E. and Adhya,S. (1993) RNA polymerase idling and clearance in *gal* promoters: use of supercoiled minicircle DNA template made in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 472–476.
47. Wahba,L., Gore,S.K. and Koshland,D. (2013) The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *Elife*, **2**, e00505.
48. Schwab,R.A., Nieminuszczy,J., Shah,F., Langton,J., Martinez,D.L., Liang,C., Cohn,M.A., Gibbons,R.J., Deans,A.J. and Niedzwiedz,W. (2015) The Fanconi anemia pathway maintains genome stability by coordinating replication and transcription. *Mol. Cell*, **60**, 351–361.
49. Stork,C.T., Bocek,M., Crossley,M.P. and Sollier,J. (2016) Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *Elife*, **5**, e17548.
50. Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1995) *Short Protocols in Molecular Biology*, 3rd edn. Wiley, NY.
51. Ko,M., Huang,Y., Jankowska,A.M., Pape,U.J., Tahiliani,M., Bandukwala,H.S., An,J., Lamperti,E.D., Koh,K.P., Ganetzky,R. *et al.* (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, **468**, 839–843.
52. Aiba,H., Adhya,S. and de Crombrugghe,B. (1981) Evidence for two functional *gal* promoters in intact *Escherichiacoli* cells. *J. Biol. Chem.*, **256**, 11905–11910.
53. Bury-Moné,S., Nomane,Y., Reymond,N., Barbet,R., Jacquet,E., Imbeaud,S., Jacq,A. and Bouloc,P. (2009) Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS Genet.*, **5**, e1000651.
54. Larson,M.H., Mooney,R. a, Peters,J.M., Windgassen,T., Nayak,D., Gross,C.A., Block,S.M., Greenleaf,W.J., Landick,R. and Weissman,J.S. (2014) A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science*, **344**, 1042–1047.
55. Sedlyarova,N., Shamovsky,I., Bharati,B.K., Epshtein,V., Chen,J., Gottesman,S., Schroeder,R. and Nudler,E. (2016) sRNA-mediated control of transcription termination in *E. coli*. *Cell*, **167**, 111–121.

56. Sambrook,J. and Russell,D. (2001) *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor Lab Press, NY.
57. Rajkumari,K., Kusano,S., Ishihama,A., Mizuno,T. and Gowrishankar,J. (1996) Effects of H-NS and potassium glutamate on σ^S - and σ^{70} -directed transcription in vitro from osmotically regulated P1 and P2 promoters of *proU* in *Escherichia coli*. *J. Bacteriol.*, **178**, 4176–4181.
58. Li,X. and Manley,J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, **122**, 365–378.
59. Boguslawski,S., Smith,D., Michalak,M., Mickelson,K., Yehle,C., Patterson,W. and Carrico,R. (1986) Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids. *J. Immunol. Methods*, **89**, 123–130.
60. Tran,L., van Baarsel,J. A., Washburn,R.S., Gottesman,M.E. and Miller,J.H. (2011) Single-gene deletion mutants of *Escherichia coli* with altered sensitivity to bicyclomycin, an inhibitor of transcription termination factor Rho. *J. Bacteriol.*, **193**, 2229–2235.
61. Nichols,R.J., Sen,S., Choo,Y.J., Beltrao,P., Zietek,M., Chaba,R., Lee,S., Kazmierczak,K.M., Lee,K.J., Wong,A. *et al.* (2011) Phenotypic landscape of a bacterial cell. *Cell*, **144**, 143–156.
62. Itaya,M., Omori,A., Kanaya,S., Crouch,R.J., Tanaka,T. and Kondo,K. (1999) Isolation of RNase H genes that are essential for growth of *Bacillus subtilis*. *J. Bacteriol.*, **181**, 2118–2123.
63. Tous,C. and Aguilera,A. (2007) Impairment of transcription elongation by R-loops in vitro. *Biochem. Biophys. Res. Commun.*, **360**, 428–432.
64. Belotserkovskii,B.P. and Hanawalt,P.C. (2011) Anchoring nascent RNA to the DNA template could interfere with transcription. *Biophys. J.*, **100**, 675–684.
65. Belotserkovskii,B., Soo Shin,J. and Hanawalt,P.C. (2017) Strong transcription blockage mediated by R-loop formation within a G-rich homopurine – homopyrimidine sequence localized in the vicinity of the promoter. *Nucleic Acids Res.*, **45**, 6589–6599.
66. Saxena,S. and Gowrishankar,J. (2011) Modulation of Rho-dependent transcription termination in *Escherichiacoli* by the H-NS family of proteins. *J. Bacteriol.*, **193**, 3832–3841.
67. Shashni,R., Qayyum,M.Z., Vishalini,V., Dey,D. and Sen,R. (2014) Redundancy of primary RNA-binding functions of the bacterial transcription terminator Rho. *Nucleic Acids Res.*, **42**, 9677–9690.
68. Valabhoju,V., Agrawal,S. and Sen,R. (2016) Molecular basis of NusG-mediated regulation of Rho-dependent transcription termination in bacteria. *J. Biol. Chem.*, **291**, 22386–22403.
69. Hermsen,R., Okano,H., You,C., Werner,N. and Hwa,T. (2015) A growth-rate composition formula for the growth of *E. coli* on co-utilized carbon substrates. *Mol. Syst. Biol.*, **11**, 801.
70. Drolet,M. (2006) Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol. Microbiol.*, **59**, 723–730.
71. Webb,M.R., Plank,J.L., Long,D.T., Hsieh,T.S. and Kreuzer,K.N. (2007) The phage T4 protein UvsW drives Holliday junction branch migration. *J. Biol. Chem.*, **282**, 34401–34411.
72. Nelson,S.W. and Benkovic,S.J. (2007) The T4 phage UvsW protein contains both DNA unwinding and strand annealing activities. *J. Biol. Chem.*, **282**, 407–416.
73. Nelson,S.W. and Benkovic,S.J. (2011) Response of the bacteriophage T4 replisome to non-coding lesions and regression of a stalled replication fork. *J. Mol. Biol.*, **401**, 743–756.
74. Manosas,M., Perumal,S.K., Bianco,P., Ritort,F., Benkovic,S.J. and Croquette,V. (2013) RecG and UvsW catalyse robust DNA rewinding critical for stalled DNA replication fork rescue. *Nat. Commun.*, **4**, 2368.
75. Washburn,R. and Gottesman,M. (2015) Regulation of transcription elongation and termination. *Biomolecules*, **5**, 1063–1078.
76. Ginno,P.A., Lim,Y.W., Lott,P.L., Korf,I. and Chédin,F. (2013) GC skew at the 5′ and 3′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Res.*, **23**, 1590–1600.
77. Chan,Y.A., Aristizabal,M.J., Lu,P.Y.T., Luo,Z., Hamza,A., Kobor,M.S., Stirling,P.C. and Hieter,P. (2014) Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-Chip. *PLoS Genet.*, **10**, e1004288.
78. El Hage,A., Webb,S., Kerr,A. and Tollervey,D. (2014) Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.*, **10**, e1004716.
79. Wahba,L., Costantino,L., Tan,F.J., Zimmer,A. and Koshland,D. (2016) S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.*, **30**, 1327–1338.
80. Sanz,L.A., Hartono,S.R., Lim,Y.W., Steyaert,S., Rajpurkar,A., Ginno,P.A., Xu,X. and Chédin,F. (2017) Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. *Mol. Cell*, **63**, 167–178.
81. Gan,W., Guan,Z., Liu,J., Gui,T., Shen,K., Manley,J.L. and Li,X. (2011) R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.*, **25**, 2041–2056.
82. Kuzminov,A. (2018) When DNA topology turns deadly—RNA polymerases dig in their R-loops to stand their ground: new positive and negative (super)twists in the replication-transcription conflict. *Trends Genet.*, **34**, 111–120.
83. Wimberly,H., Shee,C., Thornton,P.C., Sivaramakrishnan,P., Rosenberg,S.M. and Hastings,P.J. (2013) R-loops and nicks initiate DNA breakage and genome instability in non-growing *Escherichia coli*. *Nat. Commun.*, **4**, 2115.
84. Gowrishankar,J. (2015) End of the beginning: elongation and termination features of alternative modes of chromosomal replication initiation in bacteria. *PLoS Genet.*, **11**, e1004909.
85. Groh,M. and Gromak,N. (2014) Out of balance: R-loops in human disease. *PLoS Genet.*, **10**, e1004630.
86. Sollier,J. and Cimprich,K.A. (2015) Breaking bad: R-loops and genome integrity. *Trends Cell Biol.*, **25**, 514–522.
87. Santos-Pereira,J.M. and Aguilera,A. (2015) R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.*, **16**, 583–597.
88. Richard,P. and Manley,J.L. (2017) R loops and links to human disease. *J. Mol. Biol.*, **429**, 3168–3180.
89. Stirling,P.C. and Hieter,P. (2017) Canonical DNA repair pathways influence R-loop driven genome instability. *J. Mol. Biol.*, **429**, 3132–3138.
90. Song,C., Hotz-Wagenblatt,A., Voit,R. and Grummt,I. (2017) SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev.*, **31**, 1370–1381.