Antisense transcripts enhanced by camptothecin at divergent CpG-island promoters associated with bursts of topoisomerase I-DNA cleavage complex and R-loop formation

Jessica Marinello[1], Giovanni Chillemi[2], Susana Bueno[2], Stefano G. Manzo[1] and Giovanni Capranico[1,*]

[1]Department of Pharmacy and Biotechnology, University of Bologna, via Irnerio 48, 40126 Bologna, Italy and [2]CASPUR, via dei Tizii 6, Rome 00185, Italy

ABSTRACT

DNA Topoisomerase I (Top1) is required to relax DNA supercoils generated by RNA polymerases (RNAPs). Top1 is inhibited with high specificity by camptothecin (CPT), an effective anticancer agent, and by oxidative base damage and ribonucleotides in DNA strands, resulting into Top1-DNA cleavage complexes (Top1ccs). To understand how Top1ccs affect genome stability, we have investigated the global transcriptional response to CPT-induced Top1ccs. Top1ccs trigger an accumulation of antisense RNAPII transcripts specifically at active divergent CpG-island promoters in a replication-independent and Top1-dependent manner. As CPT increases antisense transcript levels in the presence of 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole, a transcription inhibitor, Top1ccs likely impair antisense RNA degradation. Time-course data showed a burst of Top1ccs increased by CPT at promoter sites and along transcribed regions, causing a transient block of RNAPII at the promoter. Moreover, cell immunofluorescence analyses showed that Top1ccs induce a transient increase of R-loops specifically at highly transcribed regions such as nucleoli in a Top1-dependent manner. Thus, a specific and highly dynamic transcriptional response to Top1ccs occurs at divergent active CpG-island promoters, which may include a transient stabilization of R-loops. The results clarify molecular features of a response pathway leading to transcription-dependent genome instability and altered transcription regulation.

INTRODUCTION

Transcription by RNA polymerase II (RNAPII) is controlled by a regulated interplay of DNA transactions and factors controlling cell identity. The complexity of transcription regulation is increased by the recent knowledge that divergent transcription is a common feature of gene promoters in several organisms, generating short and long (>200 bp) non-coding RNAs (1–6). Moreover, DNA conformation and superhelicity can affect transcription and transcription regulation (1–6). On one hand, negative and positive supercoils of DNA template can impair transcription in living cells (4,6), whereas on the other, DNA supercoiling has a basic function in transcription regulation in prokaryotes (2) and during development and pathological states of eukaryotic cells (1,3,7). Interestingly, negatively supercoiled DNA can form non-B structures, such as R-loops, that can impact transcription-associated recombination and mutation, Ig gene class-switch recombination and transcription regulation (8–11).

DNA topoisomerase I (Top1) modulates DNA supercoiling during fundamental DNA metabolic pathways (12) and is the highly specific target of camptothecins (CPT), highly effective agents against human cancers (13,14). Moreover, drug therapeutic potential has recently been broadened, as CPT derivatives can unsilence the silent allele of Ube3a in a mouse model of human Angelman syndrome (15), suggesting a significant impact of Top1 inhibitors on neurological disorders. Top1 can remove DNA supercoils by cleaving a strand of a DNA duplex

and forming a transient DNA-Top1 cleavage complex (Top1cc) in which the cut strand is covalently linked to the protein. Then, Top1 allows the rotation of the cut strand around the intact one and eventually reseals the strand break (16,17). CPT interacts at the cleavage site with DNA and the enzyme impeding the break resealing reaction and stabilizing Top1ccs (18). Top1ccs are highly reversible both in vitro and in living cells, and CPT analogs act to increase significantly their half-life (2 min on average) resulting in a decrease of the DNA uncoiling rate by the enzyme (19). In living cells, CPT action results in an immediate increase of Top1-mediated DNA breaks, and inhibition of Top1 activity, transcription and replication (17,19–21). The latter drug effect is likely due to collisions of Top1ccs with advancing replication forks that can result into frank DNA double-stranded breaks. The replicative DNA breakage eventually triggers cell cycle arrest and apoptosis of cancer cells (17,21). However, the molecular bases of Top1cc interference with transcription regulation and induction of genome instability are not fully clarified (13).

According to the 'twin-supercoiling domain model', RNAP elongation can lead to local accumulation of positive supercoils ahead of the RNAP and negative supercoils behind (5,21). The bulk of cellular Top1 activity is present at transcribing regions, where its activity can be redundant with Top2, as both enzymes can relax either positive and negative supercoils (5). However, different solutions for supercoil removal and specialized roles of each enzyme can be operative in cells as indicated by findings in yeast where Top2 is mainly responsible for positive supercoil removal in front of RNAPs and Top1 removes negative supercoils behind the polymerase (22,23). Early works provided evidence that Top1 also functions at the transcription initiation step (24–26). Top1 is likely an enzyme that can more rapidly respond to local torsional tension in nucleosome-free regions (4,27), a molecular environment common to promoter regions where Top1 has been mapped (22,28). In addition, Top1 has a role in chromatin organization and remodeling, as the disassembly/assembly of nucleosomes is thought to be another main source of DNA supercoiling (29). Top1 has been shown to affect chromatin organization at yeast telomeric regions (30) and active promoters (22,28,31).

In recent years, we showed that CPT-stabilized Top1ccs lead to hyperphosphorylation of RNAPII and decreased recruitment at human HIF-1α, cMyc and GAPDH gene promoters (28,32). Unexpectedly, we found that Top1ccs favor RNAPII escape from promoter-proximal pausing sites and increase the levels of antisense transcripts at the 5′- and 3′-ends of the human HIF1α gene (31). The two antisense RNAs overlap with the primary sense transcript of the HIF-1α gene. Even though their functions are not fully known, we showed that the 5′ antisense HIF-1α RNA, corresponding to the first exon region, localizes mainly at the nuclear membrane close to the nuclear pore protein Nup62 suggesting a role in the transport and post-transcriptional regulation of mRNAs (33). Antisense transcripts accumulate for a 4-h period at the human HIF-1α locus, suggesting that Top1ccs can cause a specific transcriptional stress at active promoters.

Nevertheless, the nature and generality of increased antisense transcripts at gene 5′-ends were left unresolved.

Here, we investigated the molecular response to Top1ccs at human gene promoters by next-generation sequencing of cellular transcripts. We show that CPT-stabilized Top1ccs cause the accumulation of antisense RNAPII transcripts specifically at divergent CpG-island (CGI) promoters, mainly upstream of the transcriptional start site (TSS), in a replication-independent and Top1-dependent manner. In addition, Top1cc levels increase transiently at active promoters during CPT treatment, and parallel a transient RNAPII block and increase of R-loops in living cells. The findings highlight molecular events at specific active promoters that can trigger transcription-associated recombination and transcription-associated mutation and set the basis to understand the functions of antisense transcripts in gene regulation and genome instability.

MATERIALS AND METHODS

Cell lines

The cancer cell lines HCT116, PC3, U2OS and COLO205 were purchased from ATCC (LGC Standards S.r.l., Milan, Italy) and were grown in Dulbecco's modified Eagle's medium (HCT116) or RPMI (PC3, U2OS and COLO205) mediums with 10% fetal bovine serum (M-Medical S.r.l., Milan, Italy). HCT116-shRNATop1 cell line was gently provided by Y. Pommier (NCI, Bethesda, MD, USA) and was grown as HCT116 cells but in presence of 200 μg/ml Hygromycin B. The cell line stably expresses a short hairpin RNA targeting exon 17 in the TOP1 gene (34,35). Cells were maintained at 37°C in a humidified incubator containing 20% O_2 and 5% CO_2. Cell line identity was certified with Cell ID System (Promega) by BMR Genomics Srl (Padova, Italy).

Cell treatment and RNA extraction

Exponentially growing cells were exposed to 10 μM of CPT, Topotecan (TPT), Etoposide or carboxylate forms of CPT, for 4 h at 37°C, unless specified otherwise. In case of co-treatment, cells were incubated with aphidicolin (5 μM) or 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB) (50 μM) for 15 min, then CPT (10 μM) was added to the medium for further 4 h. For CDK9 silencing experiments, cells were transfected 24 h after plating (30% confluence) with RNAiMax Transfection reagent (Life Technologies) and 20 nM of scramble siRNA or CDK9 siRNA (Life Technologies) and exposed to drugs 72 h post-transfection. For transient Top1 silencing experiments, cells were transfected in the same way using 60 nM of scramble siRNA or Top1 siRNA (targeting exon 16 of the TOP1 gene; Life Technologies) and exposed to drugs 72 h post-transfection. Total cellular RNA was purified with the acid phenol method (33), quantified by ultraviolet absorbance and quality controlled by electrophoresis.

RNA bisulfite conversion and sequencing

RNA bisulfite treatment was performed according with He *et al.* (36), with minor modifications (Supplementary Information). RNA was converted to cDNA using

random octamers and Superscript III Reverse Transcriptase (Life Technologies). Libraries were prepared from cDNA with mRNA-Seq Sample Prep Kit (Illumina) following Illumina's standard protocols. RNA sequencing was performed on Illumina HiSeq2000 platform (pair-end 2 X100 bp, flowcell v3, two samples for each lane) at IGA Technology Services S.r.l. (Udine, Italy). The RNA-Seq data have been submitted to the SRA database (http://www.ncbi.nlm.nih.gov/sra) with identifier PRJNA186590.

Bioinformatics analysis

Two independent RNA-seq experiments for each cell line were performed with similar results, and here, we report the results of analyses of the pooled data. Low-quality sequenced bases were filtered out using DynamicTrim v.1.12 (37), which trims the reads to leave only the longest contiguous segment that surpasses the quality threshold ($P = 0.05$). The bisulfite-treated paired reads were then mapped twice to the GRCh37 human genome assembly in which all the cytosines were mutated in thymines (in the following named as CT-GRCh37) or all the guanine in adenine (GA-GRCh37) to identify the strand originating the sequence fragment. Sequences coming from positive strand transcripts align to the CT-GRCh37, whereas negative strand trascripts align to the GA-GRCh37. The read alignment was carried out with Bowtie v0.12.7 (38) and TopHat v1.4.1 (39) to identify known transcripts. For each of the four experiments, reads that aligned to both CT-GRCh37 and GA-GRCh37 were discarded. Cufflinks was used to assemble and identify novel transcripts, using reference genomes in which all the RefSeq genes in the + or − strands were alternatively masked. Gene expression levels were estimated in FPKM units (expected number of Fragments Per Kilobase of transcript sequence per Millions of sequenced nucleotides) using Cufflinks (40), and differential transcript expression between control and CPT-treated samples was then computed using Cuffdiff.

Plots of the RNA fragments distributions with respect to the TSS (both sense and antisense) were calculated on a subset of 9665 not overlapping RefSeq genes, composed by 6626 and 3039 genes with and without CpG islands, respectively, in the region −4000/+1000 bp as respect to the TSS. Cumulative plots were normalized for total aligned tags and number of genes in the respective class. The number of reads in the −4000/+1000 bp region for each not overlapping RefSeq gene is reported as box-plot of the difference between control and CPT-treated cells (in the following Δ antisense reads and Δ sense reads).

A single sample Wilcoxon signed rank test was used to estimate the central tendency of the Δ sense and Δ antisense reads groups (pseudo-median) with a 95% confidence interval (41). Moreover, two sample comparisons were performed using Wilcoxon rank sum tests to evaluate if couples of Δ antisense or Δ sense reads distributions were statistically different, one versus the other.

K⁺SDS precipitation assay

The K⁺SDS Precipitation protocol was described already (30). Briefly, cells were lysed with 1% SDS and DNA

sonicated (Bioruptor®, Diagenode) to an average fragment size of 200–300 bp. Proteins were then precipitated with 2.5 M KCl at 4°C and centrifugation at 1200g for 3 min. Pellets were washed five times with 0.25 M KCl, and then treated with Proteinase K. DNA was purified by phenol-extraction and ethanol precipitation and quantified by real-time PCR.

ChIP assay

ChIP protocol was as previously described (28) with some modifications (see Supplementary Information).

ICE assay

ICE assay was performed as previously reported (42) with minor modifications (see Supplementary Information).

Immunofluorescence

HCT116 and HCT116-siRNATop1 cells were seeded at low density and treated with 10 µM CPT 24 h after seeding. Cells were fixed and permeabilized in pure ice-cold methanol and incubated with standard protocols first with primary antibody (S9.6 antibody gently furnished by F. Chedin - University of California, Davis), and finally with secondary antibody for signal detection (Anti-mouse Ig, fluorescein linked whole antibody from sheep, Amersham Biosciences). Images acquisition by Nikon Ti-E microscopy equipped with A1R confocal system. Fluorescence quantitation analysis has been performed using ImageJ software with the following formula: Corrected Total Cell Fluorescence = Integrated Density— (Area of selected cell X Mean Fluorescence of Background Readings).

RESULTS

Top1ccs increase antisense sequence tags at active CGI promoters in human cells

To establish the genome-wide effects of CPT-stabilized Top1ccs at promoters, we mapped RNA-seq paired tags following 4-h CPT (10 µM) treatment of human colon cancer HCT116 cells. We also performed RNA-seq experiments in HCT116-shRNATop1 cells that have Top1 contents reduced to ~25% of wild-type cell levels (31,34) to directly determine the role of Top1. Paired tags were obtained from total cellular RNA depleted of ribosomal RNAs and treated with bisulfite to maintain the information of strand direction (Supplementary Data, 'Materials and Methods' section). A total of 2.18–2.89×10^{7} high-quality paired tags mapped uniquely to the human genome, with 43.4–45.0% of them mapping to the minus strand (Supplementary Table SI). Although a large portion of reads (43.8–49.3%) mapped to genomic regions without gene annotations both in control and treated cells (Supplementary Table SI), we have here focused on reads and transcripts mapping at gene promoters. First, we evaluated mRNA expression patterns altered by CPT, and determined 1054 and 15 Refseq transcripts with statistically significant variations (>2-fold) in HCT116 and HCT116-shRNATop1 cells, respectively.

The fraction of downregulated genes was higher than upregulated genes (69 against 31% in HCT116 cells). The results of Gene Set Enrichment analyses of CPT-altered genes with DAVID Bioinformatics Resources (43,44) were highly consistent with CPT mode of action (Supplementary Figure S1 and Supplementary Data Set 1) and previous findings (34,45), thus showing that we can readily measure significant variations in transcript levels.

As CPT-stabilized Top1ccs affected the overall ratio of sense/antisense tags (Supplementary Table SI), we then evaluated whether the balance of sense and antisense transcripts was changed at specific gene regions. By analyzing tag distributions along annotated Refseq genes, we found evidence that the number of antisense tags was specifically and markedly increased by Top1ccs in a region immediately upstream of TSS and in a manner dependent on gene expression levels (Supplementary Figure S2A). Sense tags showed some changes at the 5′- and 3′-gene regions while decreasing somewhat in the body of genes (Supplementary Figure S2A). As an increase of antisense reads upstream of TSS could be due to a modification of expression of an overlapping gene in a divergent orientation, we then determined antisense tag distributions at non-overlapping Refseq gene promoters and their dependence on promoter features, such as the presence of a CGI or its activity measured as FPKM of sense (mRNA) transcripts. The average number of antisense tags was increased up to 4000 bp upstream of TSS, with a peak at 450 bp from TSS, in active (0.3–100 FPKM) promoters containing CpG islands (Figure 1A). Top1 silencing significantly reduced CPT effects as antisense tags increased less in HCT116-shRNATop1 (Figure 1B), demonstrating that effects on antisense transcripts are strictly related to Top1 inhibition. Box-plot analyses show that antisense tag increase was specific for CGI promoters, as no increase was detected at CGI-less promoters (Figure 1C and Supplementary Figure S2C). Antisense tags were not increased at inactive CGI promoters (0.0–0.3 FPKM; Figure 1D and Supplementary Figure S2B) and at active promoters (>100 FPKM; Figure 1E and F, and Supplementary Figure S2B), indicating that CPT effects were restricted to CGI promoters of intermediate activity. As expected, sense tags show a peak downstream of TSS with slight changes by CPT in HCT116 cells (Figure 1A and Supplementary Figure S2), even though CPT increased sense tags in Top1-silenced cells (Figure 1B). No differences were found between protein-coding and non-protein-coding Refseq genes. Taken together, the data demonstrate that, even though CPT is a potent transcription inhibitor, it can markedly and specifically enhance the number of antisense sequence tags at human active CGI promoters.

Top1cc-enhanced antisense transcripts are found at divergent active CGI promoters

Next, to identify Top1cc-increased transcripts upstream to TSSs, we used the Cufflink package with stringent parameters to reduce false-positive instances, even though that may have reduced the number of true transcripts. We determined unknown (not annotated) tag clusters and their genome location relative to the closest annotated

Refseq gene and found 8964 clusters long <5000 nt in HCT116 cells, 809 of which were altered by Top1ccs (Supplementary Table SII). In all, 2464 tag clusters were located within 5000 bases from a Refseq gene, the majority (1328) of which were promoter-associated antisense transcripts consistently with antisense transcription being a significant process at human promoters (46,47). Of the 809 clusters altered by Top1ccs, a significant fraction (43.5%) was found within 5000 bp to a Refseq gene (Supplementary Table SII), suggesting that transcribed regions constitute a main target of CPT effects. Moreover, CPT-increased antisense tag clusters at promoters were the highest fraction (19.5%) of all CPT-altered clusters (Figure 2A and Supplementary Table SII). The analysis in HCT116-shRNATop1 cells showed similar general patterns but again CPT effects were less marked (Figure 2A and Supplementary Table SII).

At promoter and terminator gene regions, CPT-stabilized Top1ccs generally increased tag clusters, with a significantly major effect on promoter-associated antisense clusters (Figure 2A and Table 1). In particular, we identified 256 antisense clusters increased by Top1ccs in HCT116 cells, and almost all of them localized at CGI promoters (Table 1 and Supplementary Table SIII and Supplementary Data Set 2). Silencing of Top1 again decreased the CPT effect as we could detect fewer clusters (84) increased by Top1ccs at CGI promoters of HCT116-shRNATop1 cells (Figure 2A and Table 1 and Supplementary Data Set 3). A large majority of them are located up to 1500 bp upstream of TSS (Supplementary Figure S3), and 33 are in common between the two cell lines (Figure 2B). Their median size was larger in HCT116 than HCT116-shRNATop1 cells (853 versus 688, Supplementary Table SIV and Supplementary Figure S3), consistent with the observed antisense tag distributions upstream of TSS (Figure 1A and B). To determine whether the antisense transcript increase had an effect on sense transcription at the same promoter, we analyzed the expression levels of Refseq genes having increased antisense transcripts. These genes were expressed at moderate to high levels (from 0.2 to 100 FPKM), and their expression levels were on average unchanged after 4 h of CPT treatment (Figures 1E and 2E). Thus, in the studied time frame, the levels of antisense transcripts do not correlate with levels of sense RNAs.

Then, we confirmed RNA-seq results by randomly selecting 10 antisense transcripts increased by CPT at promoters and measuring their levels by rtPCR (Figure 2C). First, we confirmed the presence of these antisense transcripts and their CPT-dependent increase by primer-specific retrotranscription (Supplementary Figure S4). Then, we measured their fold-increases after CPT treatment demonstrating a good agreement with genomic sequencing analyses (Figure 2C). As a control, we also included one antisense transcript (5′ aTNIK) that was decreased by CPT based on RNA-seq data (not shown). As expected, CPT effects were dose dependent in HCT116 cells (Figure 2C). Antisense transcript increase was much lower in Top1-silenced cells showing a dependence on cellular Top1 levels (Figure 2C). However, Top1 silencing by itself did not change significantly antisense transcript levels (Supplementary Figure S5C).
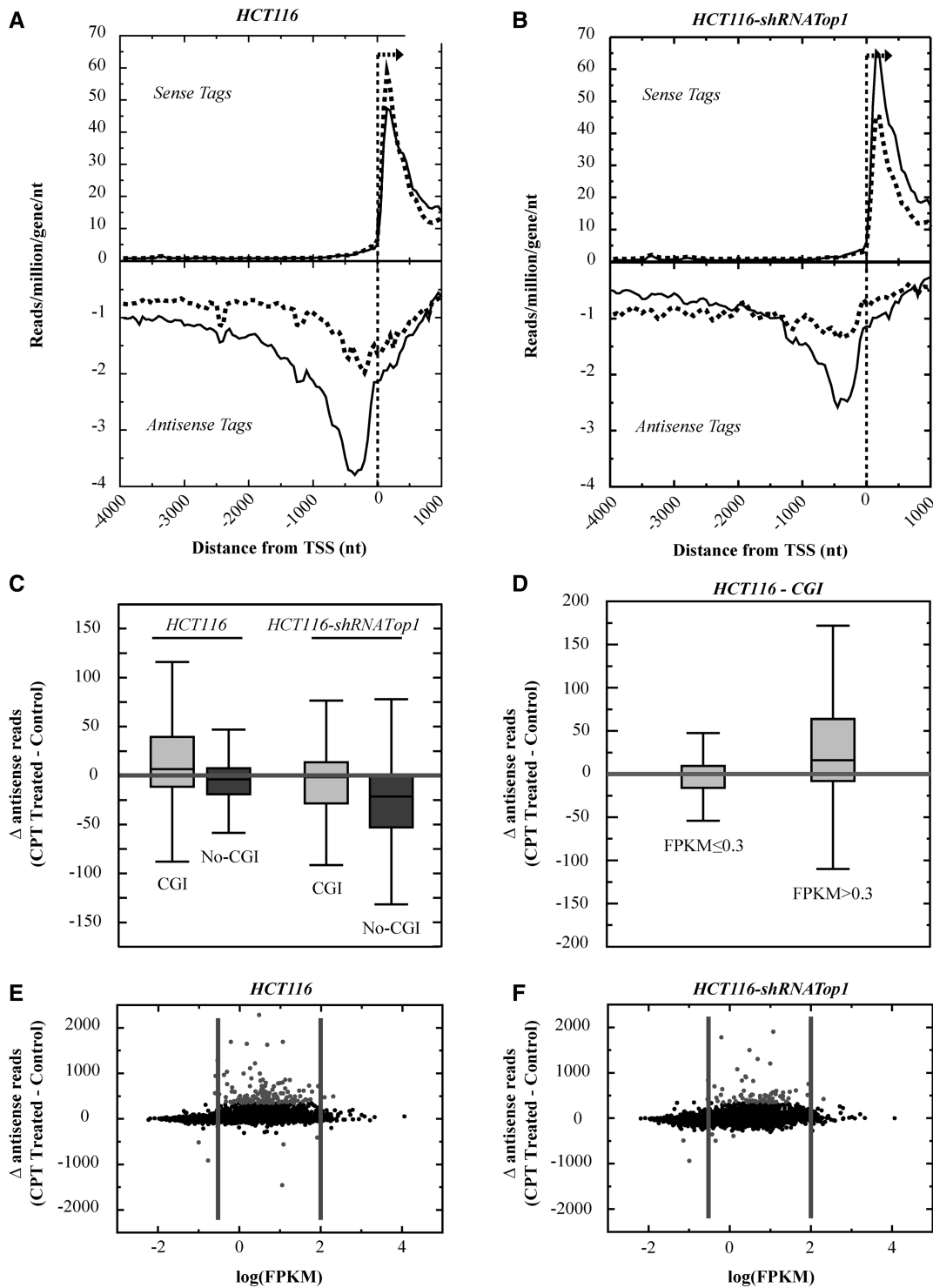
Figure 1. Antisense sequence tags are increased by CPT at active CGI promoters in a Top1-dependent manner. (**A**) Tag distributions along non-overlapping Refseq genes of HCT116 cells were analyzed in a region from 4000 bases upstream to 1000 bases downstream the TSS (dotted line and arrow). Control and CPT-treated cells are in dotted and straight line, respectively. Sense and antisense tags are above and below the *x*-axis, respectively. (**B**) Tag distributions in HCT116-shRNATop1 cells, in which Top1 content is reduced to about the 25% of wt levels. (**C**) Box-plots representing the variation of antisense sequence tags after CPT treatment in the indicated cell lines. Non-overlapping Refseq genes were split in two groups: genes with a CGI promoter (CGI, light gray), and genes without a CGI at the promoter (No-CGI, gray). (**D**) Box-plots representing the variation of antisense sequence tags after CPT treatment in HCT116 cells. Non-overlapping Refseq genes with a CGI promoter were split in two groups with the indicated level of gene expression (FPKM). (**E**) Plot of variation of antisense sequence tags after CPT treatment as a function of log(FPKM) in HCT116 cells, non-overlapping Refseq genes with CGI promoters. Vertical gray lines represent 0.3 and 100 FPKM value, respectively, from left to right. Gray dots are higher than the threshold: module (300). Genes with FPKM = 0 are not reported in the graph. All excluded genes had antisense variation below the threshold. (**F**) The same as in Panel (E), HCT116-shRNATop1 cells. Two sample comparisons using Wilcoxon rank sum tests indicate that all groups are statistically different, one versus the other (HCT116 CGI versus no CGI, P = 2.4e-119; HCT-shRNATop1 CGI versus no CGI, *P* = 8.6e-148; HCT116 CGI FPKM < 0.3 versus ≥0.3, *P* = 3.9e-115).
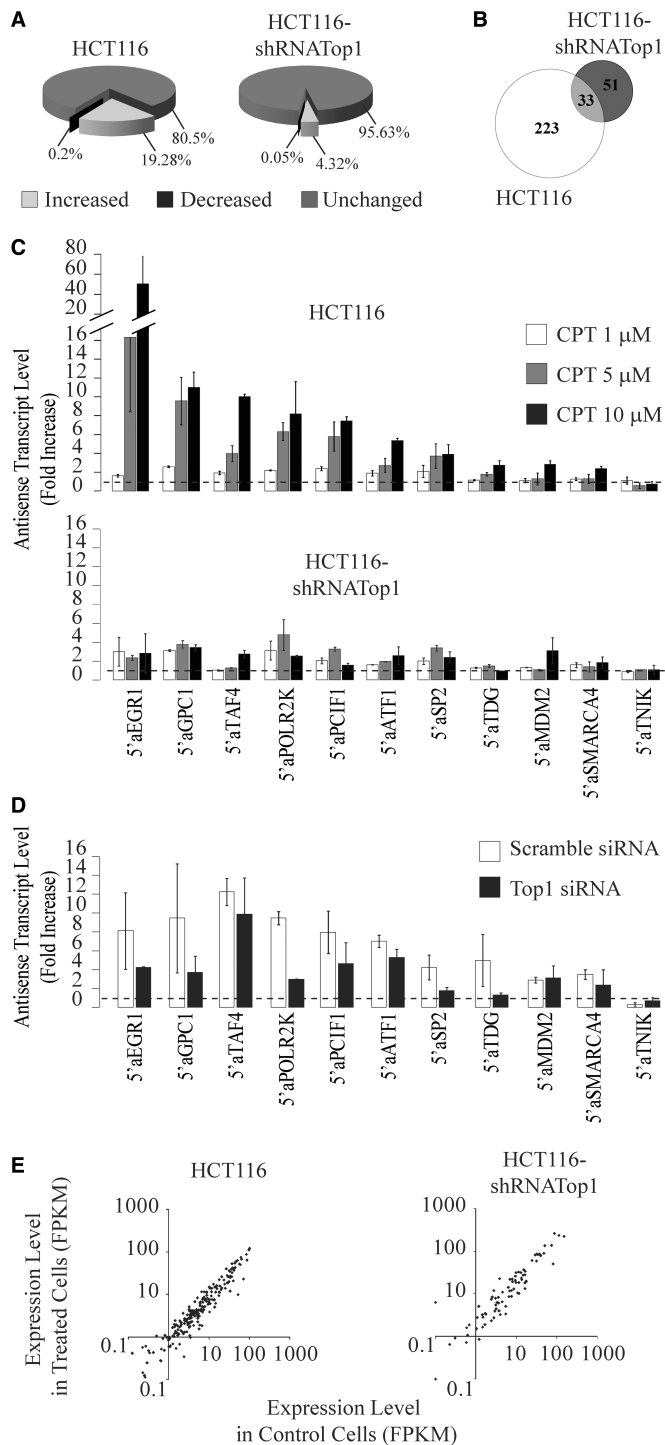
Figure 2. CPT induces increased levels of unknown tag clusters mainly at CGI promoters of genes in a Top1-dependent manner. (A) Pie charts representing the effect of CPT on unknown clusters at promoters of the indicated cell lines. Tag clusters were long <5000 nucleotides, with no annotation at the corresponding genomic region and localized from the TSS to 5000 bases upstream. Light gray, black and grey show unknown clusters that were increased, decreased and unchanged by CPT, respectively. (**B**) Common promoter-associated antisense transcripts (33) between HCT116 and HCT116-shRNATop1 cells. (C) Antisense transcripts were determined upstream the TSS of selected genes by rtPCR in the indicated cell lines. The selected genes showed CPT-increased tag clusters as determined with RNA-seq data, with the exception of TNIK gene that had an antisense transcript reduced by CPT. (**D**) Antisense transcripts were evaluated by rtPCR in HCT116 after transient

As HCT116-shRNATop1 cells are a stable cell line, we also tested the effect of a transient silencing of Top1 by siRNA in HCT116 cells (Supplementary Figure S6). We found that Top1-specific siRNAs suppressed CPT effects with different rates in HCT116 cells. In agreement with data on HCT116-shRNATop1 (Supplementary Figure S5C), transient Top1 silencing left substantially unchanged antisense transcript levels in untreated cells (Supplementary Figure S7). Thus, CPT effects at CGI promoters strictly depend on cellular Top1 levels. Moreover, CPT increased antisense transcripts at the studied CGI promoters in other human cell lines, including PC3, U2OS, COLO205 (Supplementary Figure S5A) and N-Tera cells (not shown), showing that the molecular response to Top1cc is not restricted to a specific cell type. Taken together, these findings confirmed Req-seq data and demonstrated that Top1ccs specifically increase antisense transcript levels at CGI promoters of intermediate activity in different human cell types.

To characterize the promoters with Top1cc-enhanced antisense transcripts, we have then investigated their chromatin features as determined by the ENCODE project in HCT116 cells (48). An analysis of gene promoter regions corresponding to the Top1cc-increased 256 antisense transcripts in HCT116 cells (Table 1 and Supplementary Table SIII) with WashU Epigenome Browser tools (http://epigenomegateway.wustl.edu) showed that these regions have chromatin features of divergent promoters (Figure 3A). RNAPII density, H3K4m3 and H3K27ac levels showed two peaks upstream and downstream of TSS, whereas gene promoter sets with no change or a decrease of antisense tags showed only one peak for each of the three chromatin marks downstream to TSS (Figure 3B and Supplementary Figure S8). In all cases, GC percentage was mainly increased around the TSS and downstream of it (Figure 3 and Supplementary Figure S8). Although the lower and upper quartiles of each single box-plot showed a significant variation of chromatin features among the selected promoters (Figure 3 and Supplementary Figure S8), nevertheless the findings clearly demonstrated a common symmetrical chromatin architecture signature with two peaks of paused RNAPII and two domains of histone marks at gene 5'-ends, likely related to a bi-directional (divergent) transcription activity in HCT116 cells (47,49).

Top1cc-enhanced promoter-associated antisense transcripts are regulated by Cdk9

As RNAPIIs at divergent promoters have been proposed to be regulated after P-TEFb recruitment (50), we have

Figure 2. Continued

silencing of Top1 protein (black bars). In (C) and (D), PCR determinations were normalized to cytochrome b mRNA and to untreated cells (dotted line). Values are means +/− SEM of at least four determinations of two independent experiments. Genomic coordinates of analyzed antisense transcripts are reported in Supplementary Data Set. Primers are reported in Supplementary Table SV. (**E**) Correlation between expression levels in CPT-treated and control cells of the genes with CPT-increased promoter-associated antisense transcripts in HCT116 (256) and HCT116-shRNATop1 (84) cell lines.

Table 1. Antisense transcripts are altered by CPT at CGI promoters of Refseq genes

CPT effect	Gene promoters				Gene terminators			
	Antisense transcripts		Sense transcripts		Antisense transcripts		Sense transcripts	
	HCT116	HCT116-shRNATop1	HCT116	HCT116-shRNATop1	HCT116	HCT116-shRNATop1	HCT116	HCT116-shRNATop1
Increased	256[a]	84[a]	30	17	14	21	41	51
Decreased	3[b]	1[b]	5	3	1	0	4	3

[a]248 out of the 256 (HCT116) and 80 out of the 84 (HCT116-shRNATop1) increased antisense transcripts at promoters are found at CGI promoters.
[b]Antisense transcripts decreased by CPT (4 in total) are found at CGI promoters.
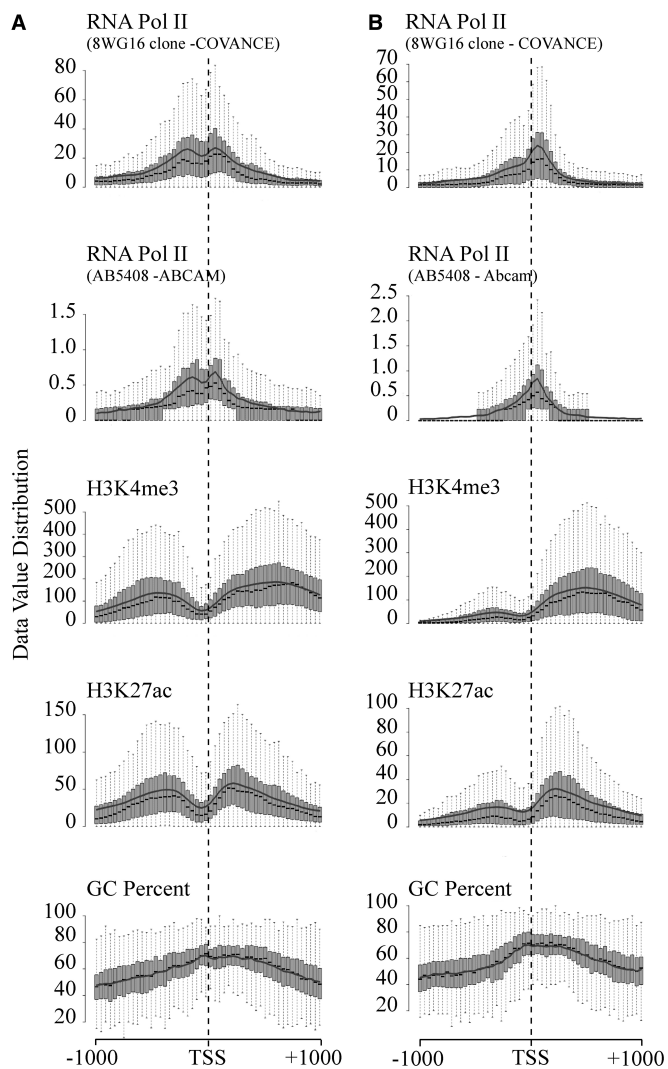

Figure 3. CPT exerts its effects at bidirectional promoters. (**A**) Analysis with WashU Epigenome Browser tools of the promoter region of the CPT-increased 256 antisense transcripts identified by Cufflinks. Shown chromatin features are respectively RNA Pol II density, H3K4me3 density, H3K27ac density and GC percentage in a region from 1000 bases upstream to 1000 bases downstream the TSS (dotted line). (**B**) Analysis with WashU Epigenome Browser tools of the promoter region of 249 randomly selected genes having FPKM between 1 and 100 and that did not demonstrate in our analysis a modification in antisense reads after CPT treatment.

then investigated the role of the transcriptional kinase Cdk9 (a subunit of P-TEFb) by directly downregulating it with siRNAs. Cdk9 silencing partially suppressed the Top1cc increase of antisense transcripts at almost all the tested promoters following 4 h of treatment (Figure 4A). The observed reduction of Top1cc-enhanced antisense transcripts was significant considering that kinase silencing was partial (Supplementary Figure S9A), and 4 h of drug treatment could allow transcript accumulation effects. We have then used DRB, a strong inhibitor of Cdk9 and RNAPII transcription elongation, which reduced the levels of antisense transcripts both in control and CPT-treated HCT116 cells (Figure 4B). The results therefore showed that Cdk9 activity is required for normal extents of antisense transcripts at the selected CGI promoters, likely indicating that antisense transcription is regulated by P-TEFb and Cdk9 at proximal-promoter RNAPII pausing step (50). Interestingly, even though DRB effectively reduced antisense transcripts with and without CPT treatment, we noticed that CPT could still increase antisense transcripts in the presence of DRB for almost all the studied promoters with similar fold change (Figure 4B and Supplementary Figure S10). We have then investigated whether CPT affected sense read levels downstream of TSS in the studied promoter set. However, we found no evidence of any common effect on sense transcripts (not shown), even though some increase of sense reads can be observed at some of the studied promoters (Figure 4C). Thus, the results show that CPT-stabilized Top1ccs specifically affect antisense versus sense transcripts at divergent CGI promoters, likely reducing RNA degradation.

Top1cc-induced increase of promoter-associated antisense transcripts is independent from DNA replication

Next, we determined whether ongoing DNA replication is required for the observed Top1cc effects at promoters by stopping replication with aphidicolin, a specific inhibitor of DNA polymerases. Aphidicolin could prevent CPT-induced H2AX phosphorylation, a marker of DNA damage (Supplementary Figure S9B), showing that replicative DNA damage promoted by CPT was markedly reduced by the compound. Nevertheless, CPT still increased antisense transcript levels in a dose-dependent manner at the selected CGI promoters even
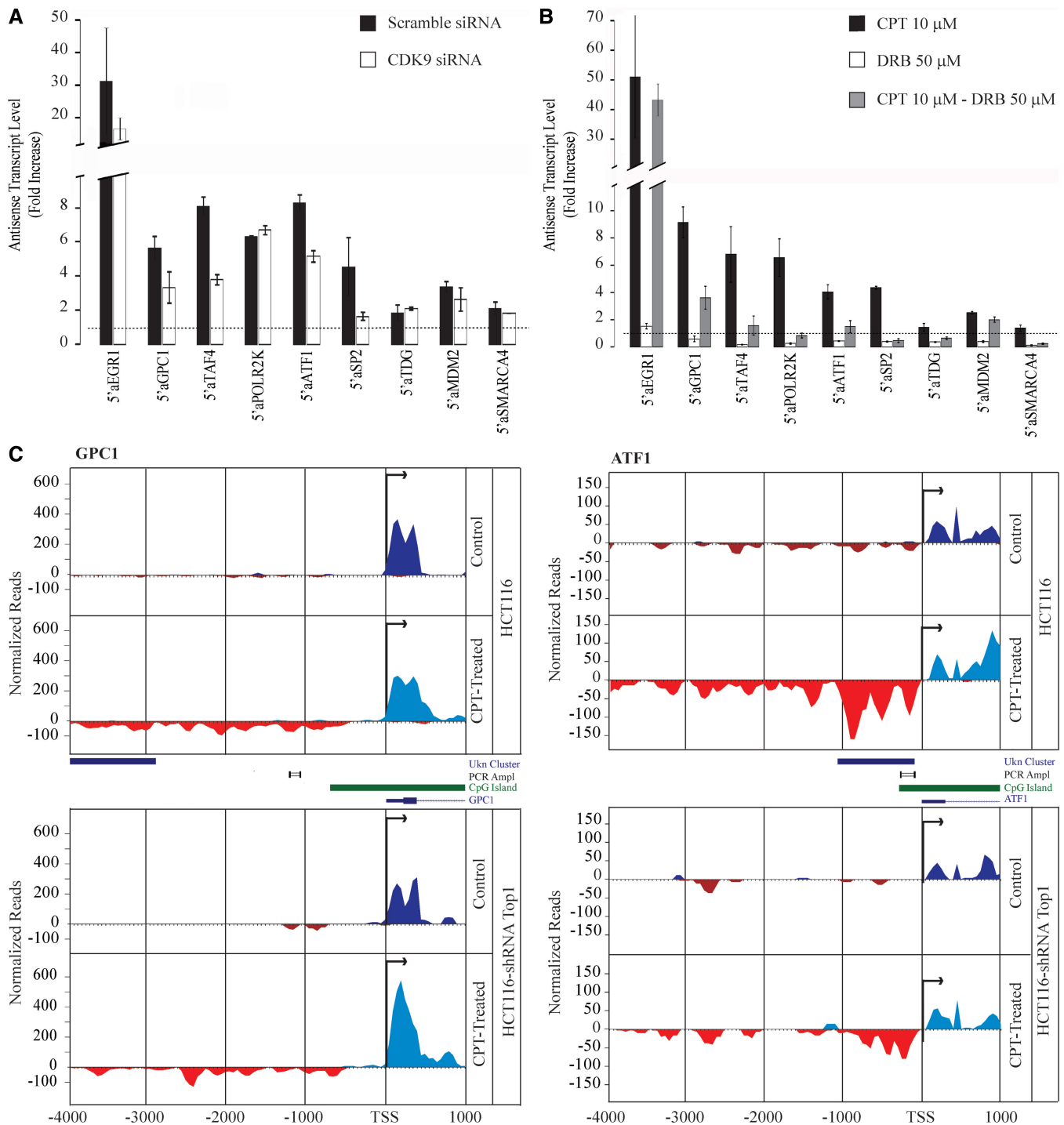
Figure 4. Dependence of CPT effects on antisense transcript levels on ongoing transcription. (**A**) The effects of 10 μM CPT for 4 h on antisense transcription have been studied after silencing of CDK9 (white bars) in HCT116 cells. Drug treatment was performed 72 h post-transfection of siRNAs. Scrambled siRNA samples are black bars. (**B**) CPT effects in HCT116 cells were studied at the selected locus with and without DRB. Co-treatment: cells were pretreated for 15 min with 50 μM DRB before the addition of CPT (10 μM) for 4 h. In all cases, PCR determinations were normalized to cytochrome b mRNA and to untreated cells (dotted line). Values are means ± SEM of at least four determinations of two independent experiments for each panel. (**C**) Genomic map of the region −4000 to +1000 bp around transcription start site (TSS, bold line) of two analyzed genes having increased antisense transcripts induced by CPT. In each case, the top and bottom graphs reports the read distributions in HCT116 and HCT116-shRNATop1 cells, respectively. In shades of blue the patterns of sense reads, in shades of red those of antisense reads (dark colors for control cells; light colors for CPT-treated cells). In the centre of each graph, the map from UCSC genomic browser, highlighting the chromosomal position of unknown cluster, PCR amplicon, CpG island and the gene 5′-end, respectively. The validation by qRT-PCR has been conducted or in a region overlapping the unknown cluster (i.e. ATF1) or in a position close to it (i.e. GPC1), where it is evident an increase of antisense transcription simply by reads mapping.

A

B
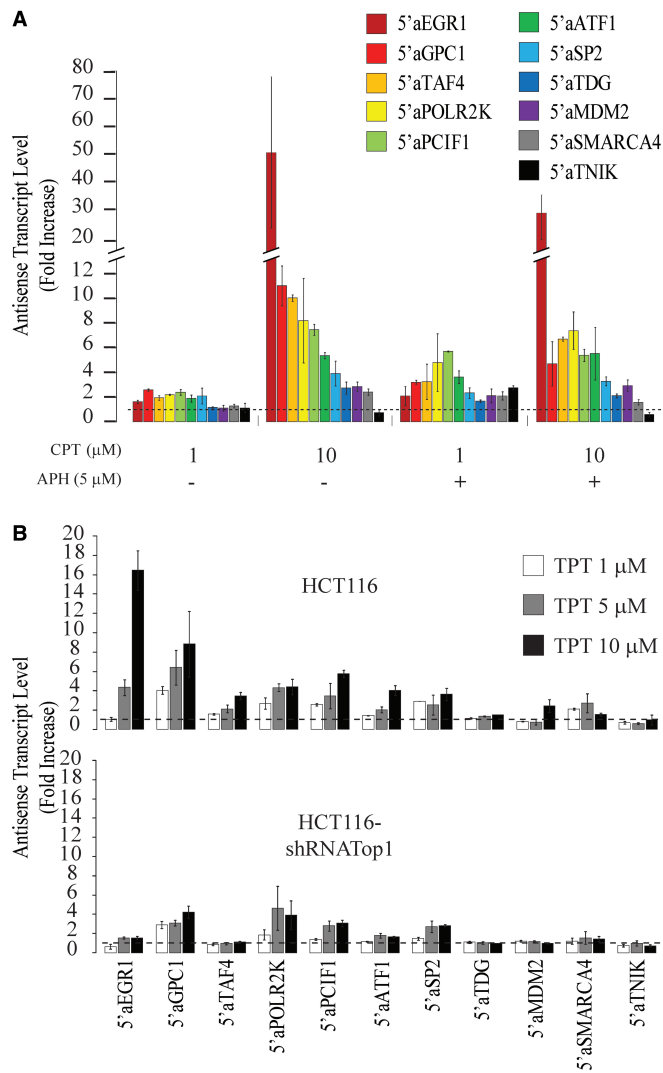
Figure 5. CPT effects on antisense transcript levels are independent from CPT-induced DNA damage at replication forks. Dose-response of HCT116 cells was analyzed at the selected locus with and without Aphidicolin. Co-treatment: cells were pretreated for 15 min with 5 µM Aphidicolin before the addition of CPT (1 or 10 µM) for 4 h. (**B**) Antisense transcripts were evaluated after treatment of the indicated cell lines with TPT 10 µM for 4 h. In (**A**) and (**B**), PCR determinations were normalized to cytochrome b mRNA and to untreated cells (dotted line). Values are means +/− SEM of at least four determinations of two independent experiments for each panel.

in the presence of aphidicolin (Figure 5A), showing that Top1ccs impacts promoter-associated antisense transcripts in a replication-independent manner. We also tested TPT, a CPT-derivative approved for clinical use and showed that it increased promoter-associated antisense transcripts in HCT116 cells in a dose-dependent manner, whereas the effect was reduced in HCT16-shRNATop1 cells (Figure 5B). Consistently, an inactive CPT analog and etoposide (VP-16, a TopII inhibitor) did not increase antisense RNAs in the studied promoters (Supplementary Figure S11). The data therefore support a specific response to Top1ccs at divergent CGI promoters.

CPT induces a rapid burst of Top1ccs that correlates with a transient RNAPII block at promoters

Our results clearly showed that CPT-induced Top1ccs increase antisense transcript levels selectively at active divergent promoters containing a CGI. Gene Set Enrichment analyses (28–30) of this set of genes provided no significant finding beyond the generic features of housekeeping genes and regulatory activity in nuclear compartments. We however noticed that 62% of the 256 unknown tag clusters increased by Top1ccs (Table 1) overlap with sites of DNase I hypersensitivity in HCT116 cells (elaborated at the UCSC Genome Browser), whereas almost the rest (36%) was within 3000 nt from such regions, documenting that they map at or close to nucleosome-free regions. As Top1 has been mapped to open chromatin sites of promoters (22,51), we then investigated the kinetics of Top1cc formation at the selected CGI promoters following CPT treatment.

The kinetics of Top1cc formation overall in the genome was determined by CsCl gradient purification following an immediate lysis of cells with a strong detergent, to prevent the reversion of Top1cc. Interestingly, Top1ccs are increased at high levels after only 2 min of CPT treatment, and then remain stable with slight variations up to 20 min of CPT treatment (Figure 6A). Interestingly, the bound but not the free Top1 molecules showed high molecular weight forms likely due to ubiquitination or sumoylation after only 2 min of drug treatment (Figure 6A), showing a rapid response of the cells to CPT-stabilized Top1ccs. Then, we used a K$^+$SDS precipitation assay of SDS cell lysates (30) to determine Top1cc levels at the selected CGI promoters. Along with them, we investigated silent promoter regions, centromeric α-satellite DNA [a negative region for Top1 binding, see (28)], and a known Top1 binding site at the origin of replication located at the 3′ end of the human lamin B2 gene (52). As expected, the findings showed that the lamin B2 region was enriched >20-fold in K$^+$SDS precipitates over α-satellite DNA (Figure 6D). We then investigated TSS-proximal regions of 7 active promoters (GAPDH, POLR2A, MYC, VEGFA, SP2, TDG, ATF1), and the DNA recovery in the precipitates showed a bell-shaped curve over time of CPT treatment with an initial increase at short times (0.5–5 min) followed by a marked decrease at later times (10–20 min) (Figure 6D). CPT increased DNA recovery from 2- to 4-fold with respect to untreated HCT116 cells. The increase was instead much lower in HCT116-shRNATop1 cells (<1.5-fold), consistent with the reduced cellular Top1 levels. CPT did not increase DNA recovery at promoters of three silent genes (MYOD1, LOR and FOXA2 in Figure 6D) showing that Top1 is at low levels, if any, at inactive promoters. We then looked at regions in the body of MYC and HIF-1α genes, and the findings showed similar kinetics of Top1cc formation by CPT at these sites as well (Figure 6E).

The unexpected DNA recovery kinetics in K$^+$SDS precipitates and CsCl gradient results demonstrated that CPT could immediately induce the formation of Top1ccs followed by their rapid removal from active promoters. We then determined kinetics of RNAPII density during
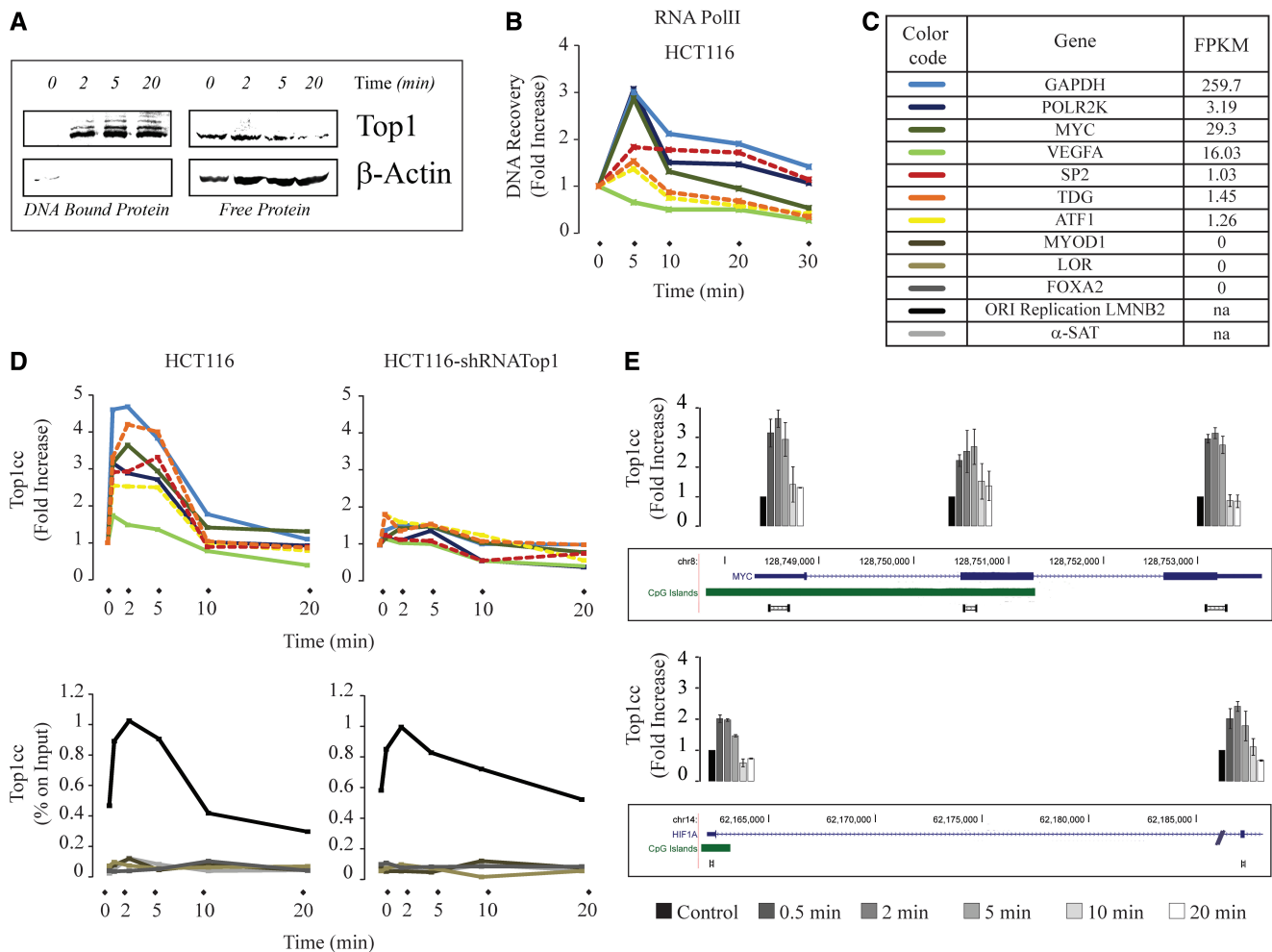
Figure 6. CPT determines a rapid and transient burst of Top1ccs that correlates with transient RNA Pol II block. (**A**) ICE bioassay was performed on HCT116 cells after 2, 5 and 20 min of 10 μM CPT treatment. The gel shows both CsCl gradient fractions containing DNA bound proteins (on the left) and free proteins (on the right). β-Actin was used as loading control. (**B**) ChIP on RNA Pol II. HCT116 cells were treated for 5, 10, 20 and 30 min with CPT 10 μM and the DNA recovery at indicated gene promoters is plotted as fold increase on not treated cells against treatment time. The analyzed gene regions are on TSS or immediately downstream to it and are listed (**C**). (**C**) List of genes observed in ChIP and K⁺SDS Assays [(**B**) and (**D**) panels], with the relative FPKM obtained by RNA-Seq. Used primers are reported in Supplementary Table SVI. (**D**) K⁺SDS precipitation assay on HCT116 and HCT116-shRNATop1 cells. Cells were treated for 2, 5, 10 and 20 min with 10 μM CPT. The recovery of Top1ccs at indicated gene promoters is reported as fold increase or percentage on input on not treated cells against treatment time. Used primers are in Supplementary Table SVI. (**E**) K⁺SDS precipitation assay on HCT116 in different positions along MYC and HIF-1α genes. Below the DNA recovery, a map showing the analyzed region. Standard errors (SEM) are not reported in (**B**) and (**D**) for image clarity. The average SEMs are as follows: 37% of the mean value (panel B); 19 and 37% of mean values for top and lower graphs (panel D HCT116), respectively; 16 and 23% of mean values for top and lower graphs (panel D HCT116-shRNATop1), respectively.

CPT treatment at the same promoters with ChIP (Figure 6B). Interestingly, at active promoters with a high rate of RNAPII recruitment, the results show an accumulation of the polymerase followed by a reduction of it during drug treatment (Figure 6B). The timing of the bell-shaped curve correspond to the timing of Top1cc burst, suggesting that CPT-trapped Top1ccs at promoter sites can block polymerase escape from promoter regions, and their removal can eventually release RNAPII allowing pausing escape (31).

CPT transiently increases R-loops in whole cells paralleling the bursts of Top1ccs

As a rapid increase of CPT-induced Top1ccs may affect DNA supercoiling at transcriptionally active genes (51),

we also investigated the formation of R-loop structures (favored by negatively supercoiled DNA) with an antibody against the DNA/RNA hybrid duplex (53) by immunofluorescence and cell imaging in HCT116 and HCT116-shRNATop1 cells. The kinetics of R-loop formation appear to parallel the burst of Top1ccs reported earlier in the text (Figure 6), as R-loops were significantly increased by 2–10 min of CPT treatment while they were reduced at longer (1 h) times (Figure 7A and B) in HCT116 cells. The fluorescence signal was mainly increased in nucleoli and at cytoplasmic spots localized at the periphery of the nucleus (Figure 7A). These cytoplasmic spots likely correspond to mitochondria where R-loops are present at the DNA origin of replication, as evidenced by mitochondria-specific staining in the same cells (Supplementary Figure S12).
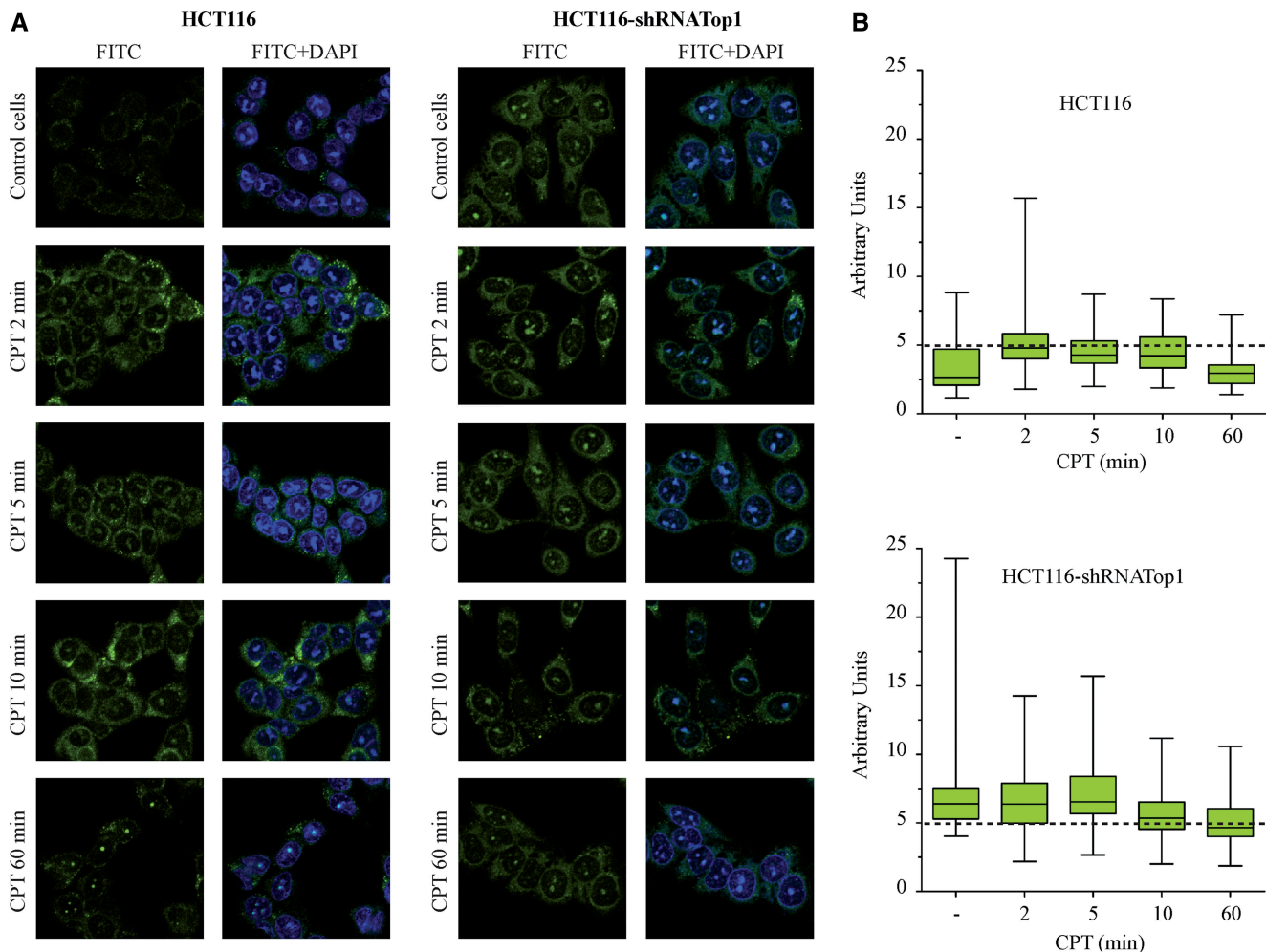
Figure 7. CPT rapidly increases cellular R-loops. (**A**) Immunofluorescence of HCT116 and HCT116-shRNATop1 cells using S9.6 antibody against DNA-RNA hybrid duplex. Cells were treated for 2, 5, 10 and 60 min with 10 μM CPT, 24 h post-seeding. (**B**) Box-plots representing the variation of fluorescence intensity (panel A) after drug treatment. The value is obtained using ImageJ software. Number of observed cells for each samples: 79, 85, 72, 99 and 90, respectively, for HCT116; 91, 185, 159, 126 and 125, respectively, for HCT116-shRNATop1 cells.

After 1 h of CPT, fluorescence distribution was further changed with a single (or few) prominent spot(s) detected in the nucleus (Figure 7), suggesting that R loop structures last longer at some genomic regions. Interestingly, the findings in HCT116-shRNATop1 cells were different. First, in control cells, fluorescence was higher in HCT116-shRNATop1 than HCT116 cells, and in particular, nucleolar signal was marked (Figure 7A and B). Second, CPT did not increase the R-loop signal in HCT116-shRNATop1 cells (Figure 7A and B). However, at longer time fluorescence intensity decreased, and patterns were more similar in both cell lines. Thus, cell images show that CPT can induce a rapid and transient increase of R-loops in a Top1-dependent manner, likely due to negatively superhelical tension of DNA template at actively transcribed regions, such as ribosomal genes.

DISCUSSION

The present findings demonstrate for the first time that Top1ccs have a specific and dynamic impact at divergent CGI promoters in human cells. CPT-stabilized Top1ccs enhance levels of antisense RNAPII transcripts upstream of TSS at divergent CGI promoters of intermediate activity. The CPT effects are dependent on cellular Top1 levels, and independent from DNA replication. Top1ccs immediately form at transcribing regions, but then their levels decrease after only 10 min of drug treatment. The bell-shaped Top1cc kinetic parallels a similar transient increase of RNAPII density at promoters, indicating that Top1ccs transiently block the polymerase. Top1ccs are stabilized not only by CPT but also by oxidative base damage, base mismatches and ribonucleotides incorporated into DNA strands (14,54). Thus, Top1ccs may physiologically occur in living cells at a significant rate and may cause specific dynamical changes leading to alterations of RNAPII recruitment and antisense transcripts.

Divergent promoters have been reported in several eukaryotic cells (46,47,50,55–57) raising the question of how this process contributes to regulation of gene activity and chromatin structure. It is therefore interesting that the majority of TSSs in murine embryonic stem cells

show evidence of bidirectional transcription (47). Previous results showed that divergent RNAPIIs are regulated by P-TEFb at post-initiation pausing sites and that antisense transcript levels are controlled by the exosome (50). Our results are in agreement with a role for Cdk9 (a P-TEFb subunit) in antisense transcription, as Cdk9 inhibition and silencing can reduce the cellular levels of antisense transcripts. We have previously shown that CPT causes a specific transcriptional stress with RNAPII hyperphosphorylation, increased RNAPII escape from promoter pausing sites, chromatin accessibility marks and antisense transcripts at human HIF-1a and c-Myc genes (31). Moreover, inhibition of Cdk9 and Cdk7 activity by DRB could suppress the CPT-induced RNAPII escape and hyperphosphorylation and increased chromatin accessibility. However, we show here that DRB is not able to suppress the CPT-induced increase of antisense transcript levels at the studied divergent CGI promoters. As DRB is an effective transcription inhibitor and it does not suppress Top1ccs induced by CPT (14), we therefore propose that the CPT effect may mainly be due to a reduction of RNA degradation. This hypothesis is also supported by the finding that reducing the RNA degradation pathway by silencing exosome subunits can enhance promoter-associated antisense transcripts in human cells (58).

A recent article (51) showed that human promoters of active genes constitute topological domains of negative supercoiling mainly due to the short range of torsional stress dissipation. As Top2 and Top1 can be functionally redundant, the authors found that Top1 is mainly responsible for supercoil relaxation at promoters of intermediate activity, whereas Top2 becomes more important at highly active promoters (51). Consistently, we have here found that Top1ccs increase antisense transcripts at intermediately active divergent CGI promoters, whereas we did not detect a similar effect by inhibiting TopII with etoposide. Thus, the specific increase of antisense transcripts by CPT likely occurs where Top1 constitutes the main DNA relaxation activity. We cannot however exclude that TopII inhibitors may alter antisense transcript levels in a different set of gene promoters.

In addition, CPT has been shown to increase negative supercoils at promoters of intermediate activity after 5 min of cell treatment (51). In bi-directional promoters, the rate of transcription-generated negative torsional tension is necessarily higher as divergent RNAPIIs elongate in different directions. Thus, divergent CGI promoters may constitute a set of gene control regions highly sensitive to exogenous and endogenous perturbation of Top1 activity. CGI promoters may readily respond to Top1ccs by an increase of negative supercoils leading to strand denaturation and formation of non-B structures. Interestingly, TOP1 gene deletion has been shown to cause an increase of R-loops at actively transcribed rRNA genes in yeast, leading to an accumulation of slowed RNAPs I at the 5' end of the gene (59).

An important finding of the present work is the demonstration that CPT induces a stabilization of R loops at nucleoli and mitochondria in HCT116 cells with a kinetic that closely parallels Top1cc formation (Figures 6 and 7). Interestingly, cells with reduced Top1 levels (HCT116-shRNATop1) show higher levels of R loops at nucleoli and whole cells without CPT treatment. This finding is consistent with a previous investigation showing that overexpression of RNaseH can reduce genomic instability and replication/transcription conflicts due to Top1 silencing (35). R-loops can form physiologically at sequences showing a G-richness in one strand, known as GC skew (53). We noticed that 185 of 246 promoter-associated antisense transcripts overlap with or are within 3000 bp distant from a GC skewed region. In addition, some of our antisense transcripts overlap to DNA segments known to form R-loops in untreated human cells (53). Thus, we propose that Top1ccs may cause an increase of negative supercoils at divergent CGI promoters determining a stabilization of R-loops, and then trigger an accumulation of promoter-associated antisense transcripts. Even though further studies are needed to establish such hypothesis, Top1ccs-dependent genome instability may preferentially occur at divergent CGI promoters of intermediate activity.

Top1cc-triggered antisense transcripts may constitute a specific molecular response to Top1ccs in normal and stress conditions. It remains to be established if they play a role in the maintenance of vital cellular processes under stress conditions such as increased levels of Top1ccs. Antisense transcripts have been shown to contribute to the regulation of gene expression regulation. We recently showed that the 5' aHIF1a antisense RNA of the human HIF-1a gene may have a role in the nucleus-cytoplasm transport machinery (33). In addition, accumulated antisense transcripts may interact with RNAPII or transcription factors to regulate chromatin structure, recruitment and sense transcription. This has been shown for transcripts upstream of the cyclin D1 gene that recruit TLS, a RNA-binding factor sensitive to DNA damage (60), and of Polycomb target genes that bind to Suz12 to repress gene transcription (61). Thus, antisense transcripts may regulate gene expression through different mechanisms at divergent CGI promoters in response to Top1ccs. However, the definition of their functions remains to be established by further investigations.

ACCESSION NUMBERS

The RNA-Seq data have been submitted to the SRA database (http://www.ncbi.nlm.nih.gov/sra) with identifier PRJNA186590.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Baranello,L., Levens,D., Gupta,A. and Kouzine,F. (2012) The importance of being supercoiled: How DNA mechanics regulate dynamic processes. *Biochim. Biophys. Acta*, **1819**, 632–638.
2. 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.
3. Kouzine,F., Liu,J., Sanford,S., Chung,H.J. and Levens,D. (2004) The dynamic response of upstream DNA to transcription-generated torsional stress. *Nat. Struct. Mol. Biol.*, **11**, 1092–1100.
4. Roca,J. (2011) Transcriptional inhibition by DNA torsional stress. *Transcription*, **2**, 82–85.
5. Wang,J.C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.*, **3**, 430–440.
6. Wu,H.Y., Shyy,S.H., Wang,J.C. and Liu,L.F. (1988) Transcription generates positively and negatively supercoiled domains in the template. *Cell*, **53**, 433–440.
7. Naughton,C., Avlonitis,N., Corless,S., Prendergast,J.G., Mati,I.K., Eijk,P.P., Cockroft,S.L., Bradley,M., Ylstra,B. and Gilbert,N. (2013) Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nat. Struct. Mol. Biol.*, **20**, 387–395.
8. Aguilera,A. and Garcia-Muse,T. (2012) R loops: from transcription byproducts to threats to genome stability. *Mol. Cell*, **46**, 115–124.
9. Aguilera,A. and Gomez-Gonzalez,B. (2008) Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.*, **9**, 204–217.
10. Kim,N. and Jinks-Robertson,S. (2012) Transcription as a source of genome instability. *Nat. Rev. Genet.*, **13**, 204–214.
11. Roy,D., Yu,K. and Lieber,M.R. (2008) Mechanism of R-loop formation at immunoglobulin class switch sequences. *Mol. Cell. Biol.*, **28**, 50–60.
12. Brill,S.J., DiNardo,S., Voelkel-Meiman,K. and Sternglanz,R. (1987) Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. *Nature*, **326**, 414–416.
13. Capranico,G., Marinello,J. and Baranello,L. (2010) Dissecting the transcriptional functions of human DNA topoisomerase I by selective inhibitors: implications for physiological and therapeutic modulation of enzyme activity. *Biochim. Biophys. Acta*, **1806**, 240–250.
14. Pommier,Y. (2006) Topoisomerase I inhibitors: camptothecins and beyond. *Nat. Rev. Cancer*, **6**, 789–802.
15. Huang,H.S., Allen,J.A., Mabb,A.M., King,I.F., Miriyala,J., Taylor-Blake,B., Sciaky,N., Dutton,J.W. Jr, Lee,H.M., Chen,X. *et al.* (2011) Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. *Nature*, **481**, 185–189.
16. Champoux,J.J. (2001) DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.*, **70**, 369–413.
17. Pommier,Y., Barcelo,J.M., Rao,V.A., Sordet,O., Jobson,A.G., Thibaut,L., Miao,Z.H., Seiler,J.A., Zhang,H., Marchand,C. *et al.* (2006) Repair of topoisomerase I-mediated DNA damage. *Prog. Nucleic Acid Res. Mol. Biol.*, **81**, 179–229.
18. Staker,B.L., Hjerrild,K., Feese,M.D., Behnke,C.A., Burgin,A.B. Jr and Stewart,L. (2002) The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc. Natl Acad. Sci. USA*, **99**, 15387–15392.
19. Koster,D.A., Palle,K., Bot,E.S., Bjornsti,M.A. and Dekker,N.H. (2007) Antitumour drugs impede DNA uncoiling by topoisomerase I. *Nature*, **448**, 213–217.
20. Duann,P., Sun,M., Lin,C.T., Zhang,H. and Liu,L.F. (1999) Plasmid linking number change induced by topoisomerase I-mediated DNA damage. *Nucleic Acids Res.*, **27**, 2905–2911.
21. Liu,L.F. and Wang,J.C. (1987) Supercoiling of the DNA template during transcription. *Proc. Natl Acad. Sci. USA*, **84**, 7024–7027.
22. Durand-Dubief,M., Persson,J., Norman,U., Hartsuiker,E. and Ekwall,K. (2010) Topoisomerase I regulates open chromatin and controls gene expression in vivo. *EMBO J.*, **29**, 2126–2134.
23. French,S.L., Sikes,M.L., Hontz,R.D., Osheim,Y.N., Lambert,T.E., El Hage,A., Smith,M.M., Tollervey,D., Smith,J.S. and Beyer,A.L. (2011) Distinguishing the roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes. *Mol. Cell. Biol.*, **31**, 482–494.
24. Kretzschmar,M., Meisterernst,M. and Roeder,R.G. (1993) Identification of human DNA topoisomerase I as a cofactor for activator-dependent transcription by RNA polymerase II. *Proc. Natl Acad. Sci. USA*, **90**, 11508–11512.
25. Merino,A., Madden,K.R., Lane,W.S., Champoux,J.J. and Reinberg,D. (1993) DNA topoisomerase I is involved in both repression and activation of transcription. *Nature*, **365**, 227–232.
26. Shykind,B.M., Kim,J., Stewart,L., Champoux,J.J. and Sharp,P.A. (1997) Topoisomerase I enhances TFIID-TFIIA complex assembly during activation of transcription. *Genes Dev.*, **11**, 397–407.
27. Salceda,J., Fernandez,X. and Roca,J. (2006) Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA. *EMBO J.*, **25**, 2575–2583.
28. Khobta,A., Ferri,F., Lotito,L., Montecucco,A., Rossi,R. and Capranico,G. (2006) Early effects of topoisomerase I inhibition on RNA polymerase II along transcribed genes in human cells. *J. Mol. Biol.*, **357**, 127–138.
29. Lavelle,C. (2007) Transcription elongation through a chromatin template. *Biochimie*, **89**, 516–527.
30. Lotito,L., Russo,A., Chillemi,G., Bueno,S., Cavalieri,D. and Capranico,G. (2008) Global transcription regulation by DNA topoisomerase I in exponentially growing Saccharomyces cerevisiae cells: activation of telomere-proximal genes by TOP1 deletion. *J. Mol. Biol.*, **377**, 311–322.
31. Baranello,L., Bertozzi,D., Fogli,M.V., Pommier,Y. and Capranico,G. (2010) DNA topoisomerase I inhibition by camptothecin induces escape of RNA polymerase II from promoter-proximal pause site, antisense transcription and histone acetylation at the human HIF-1alpha gene locus. *Nucleic Acids Res.*, **38**, 159–171.
32. Sordet,O., Larochelle,S., Nicolas,E., Stevens,E.V., Zhang,C., Shokat,K.M., Fisher,R.P. and Pommier,Y. (2008) Hyperphosphorylation of RNA polymerase II in response to topoisomerase I cleavage complexes and its association with transcription- and BRCA1-dependent degradation of topoisomerase I. *J. Mol. Biol.*, **381**, 540–549.
33. Bertozzi,D., Iurlaro,R., Sordet,O., Marinello,J., Zaffaroni,N. and Capranico,G. (2011) Characterization of novel antisense HIF-1alpha transcripts in human cancers. *Cell Cycle*, **10**, 3189–3197.
34. Miao,Z.H., Player,A., Shankavaram,U., Wang,Y.H., Zimonjic,D.B., Lorenzi,P.L., Liao,Z.Y., Liu,H., Shimura,T., Zhang,H.L. *et al.* (2007) Nonclassic functions of human topoisomerase I: genome-wide and pharmacologic analyses. *Cancer Res.*, **67**, 8752–8761.
35. Tuduri,S., Crabbe,L., Conti,C., Tourriere,H., Holtgreve-Grez,H., Jauch,A., Pantesco,V., De Vos,J., Thomas,A., Theillet,C. *et al.* (2009) Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat. Cell Biol.*, **11**, 1315–1324.
36. He,Y., Vogelstein,B., Velculescu,V.E., Papadopoulos,N. and Kinzler,K.W. (2008) The antisense transcriptomes of human cells. *Science*, **322**, 1855–1857.

37. Cox,M.P., Peterson,D.A. and Biggs,P.J. (2010) SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics*, **11**, 485.
38. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
39. Trapnell,C., Pachter,L. and Salzberg,S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105–1111.
40. Trapnell,C., Williams,B.A., Pertea,G., Mortazavi,A., Kwan,G., van Baren,M.J., Salzberg,S.L., Wold,B.J. and Pachter,L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, **28**, 511–515.
41. Hollander,M. and Wolfe,D. (1999) *Nonparametric Statistical Methods*. Wiley, New York, NY.
42. Subramanian,D., Furbee,C.S. and Muller,M.T. (2001) ICE bioassay. Isolating in vivo complexes of enzyme to DNA. *Methods Mol. Biol.*, **95**, 137–147.
43. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
44. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
45. Solier,S., Barb,J., Zeeberg,B.R., Varma,S., Ryan,M.C., Kohn,K.W., Weinstein,J.N., Munson,P.J. and Pommier,Y. (2010) Genome-wide analysis of novel splice variants induced by topoisomerase I poisoning shows preferential occurrence in genes encoding splicing factors. *Cancer Res.*, **70**, 8055–8065.
46. Core,L.J., Waterfall,J.J. and Lis,J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845–1848.
47. Seila,A.C., Calabrese,J.M., Levine,S.S., Yeo,G.W., Rahl,P.B., Flynn,R.A., Young,R.A. and Sharp,P.A. (2008) Divergent transcription from active promoters. *Science*, **322**, 1849–1851.
48. Dunham,I., Kundaje,A., Aldred,S.F., Collins,P.J., Davis,C.A., Doyle,F., Epstein,C.B., Frietze,S., Harrow,J., Kaul,R. *et al.* (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
49. Thurman,R.E., Rynes,E., Humbert,R., Vierstra,J., Maurano,M.T., Haugen,E., Sheffield,N.C., Stergachis,A.B., Wang,H., Vernot,B. *et al.* (2012) The accessible chromatin landscape of the human genome. *Nature*, **489**, 75–82.
50. Flynn,R.A., Almada,A.E., Zamudio,J.R. and Sharp,P.A. (2011) Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proc. Natl Acad. Sci. USA*, **108**, 10460–10465.
51. Kouzine,F., Gupta,A., Baranello,L., Wojtowicz,D., Ben-Aissa,K., Liu,J., Przytycka,T.M. and Levens,D. (2013) Transcription-dependent dynamic supercoiling is a short-range genomic force. *Nat. Struct. Mol. Biol.*, **20**, 396–403.
52. Abdurashidova,G., Radulescu,S., Sandoval,O., Zahariev,S., Danailov,M.B., Demidovich,A., Santamaria,L., Biamonti,G., Riva,S. and Falaschi,A. (2007) Functional interactions of DNA topoisomerases with a human replication origin. *EMBO J.*, **26**, 998–1009.
53. Ginno,P.A., Lott,P.L., Christensen,H.C., Korf,I. and Chedin,F. (2012) R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell*, **45**, 814–825.
54. Kim,N., Huang,S.N., Williams,J.S., Li,Y.C., Clark,A.B., Cho,J.E., Kunkel,T.A., Pommier,Y. and Jinks-Robertson,S. (2011) Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science*, **332**, 1561–1564.
55. Neil,H., Malabat,C., d'Aubenton-Carafa,Y., Xu,Z., Steinmetz,L.M. and Jacquier,A. (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature*, **457**, 1038–1042.
56. Valen,E., Preker,P., Andersen,P.R., Zhao,X., Chen,Y., Ender,C., Dueck,A., Meister,G., Sandelin,A. and Jensen,T.H. (2011) Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. *Nat. Struct. Mol. Biol.*, **18**, 1075–1082.
57. Xu,Z., Wei,W., Gagneur,J., Perocchi,F., Clauder-Munster,S., Camblong,J., Guffanti,E., Stutz,F., Huber,W. and Steinmetz,L.M. (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature*, **457**, 1033–1037.
58. Preker,P., Nielsen,J., Kammler,S., Lykke-Andersen,S., Christensen,M.S., Mapendano,C.K., Schierup,M.H. and Jensen,T.H. (2008) RNA exosome depletion reveals transcription upstream of active human promoters. *Science*, **322**, 1851–1854.
59. El Hage,A., French,S.L., Beyer,A.L. and Tollervey,D. (2010) Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.*, **24**, 1546–1558.
60. Wang,X., Arai,S., Song,X., Reichart,D., Du,K., Pascual,G., Tempst,P., Rosenfeld,M.G., Glass,C.K. and Kurokawa,R. (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*, **454**, 126–130.
61. Kanhere,A., Viiri,K., Araujo,C.C., Rasaiyaah,J., Bouwman,R.D., Whyte,W.A., Pereira,C.F., Brookes,E., Walker,K., Bell,G.W. *et al.* (2010) Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol. Cell*, **38**, 675–688.