Transcription-dependent dynamic supercoiling is a short-range genomic force

Fedor Kouzine1,6, Ashutosh Gupta1,2,6, Laura Baranello1,6, Damian Wojtowicz3, Khadija Ben-Aissa4, Juhong Liu5, Teresa M Przytycka3 & David Levens1

Transcription has the capacity to mechanically modify DNA topology, DNA structure and nucleosome arrangement. Resulting from ongoing transcription, these modifications in turn may provide instant feedback to the transcription machinery. To substantiate the connection between transcription and DNA dynamics, we charted an ENCODE map of transcription-dependent dynamic supercoiling in human Burkitt’s lymphoma cells by using psoralen photobinding to probe DNA topology in vivo. Dynamic supercoils spread ~1.5 kilobases upstream of the start sites of active genes. Low- and high-output promoters handled this torsional stress differently, as shown by using inhibitors of transcription and topoisomerases and by chromatin immunoprecipitation of RNA polymerase and topoisomerases I and II. Whereas lower outputs are managed adequately by topoisomerase I, high-output promoters additionally require topoisomerase II. The genome-wide coupling between transcription and DNA topology emphasizes the importance of dynamic supercoiling for gene regulation.

Chromatin structure, gene-regulatory proteins, histones and DNA modification vary temporally with gene expression. DNA structure and topology also may regulate and be modified by nearby genetic activity1. Transcribing RNA polymerase generates supercoils in the DNA template2, potentially facilitating or impeding DNA-dependent processes3. Thus, besides serving as a passive information repository, DNA could actively participate in the real-time regulation of gene activity. Many studies have focused on the dynamics and functions of proteins regulating transcription; in general, these studies have not considered the role of DNA structure and topology in gene regulation.

In the ‘twin domain’ theory (twin supercoiled domain), threading of DNA through the transcription machinery, along a helical path, dynamically drives positive supercoils ahead and negative supercoils behind the translocating RNA polymerase2. Negative supercoils untwist while positive supercoils overtwist DNA. Proceeding without pause, RNA polymerase would generate approximately seven supercoils per second4; unless dissipated, this torsional stress would rise to enormous levels disruptive to all genetic processes1,5. DNA topoisomerases transiently break and rejoin the DNA backbone to remove positive and negative supercoils6. Depending on the intensity of ongoing transcription and the disposition of topoisomerases, local supercoiling might exceed the relaxation capacity of nearby topoisomerases, leaving the residual DNA torsional stress to propagate through the surrounding chromatin7. This stress might influence the binding of regulatory proteins to DNA, change nucleosome dynamics and reorganize chromatin fibers3. Supercoiling may also drive duplex B-DNA into single-stranded or other non-B-DNA conformations that alter the ability of DNA and chromatin to loop and twist, potentially modifying the function of enhancers and other cis control elements8. Proteins that bind non-B-DNA may require alternative structures, and because non-B-DNA is incompatible with nucleosome binding, such structures may sustain nucleosome-free regions9. As the magnitude and distribution of supercoiling forces throughout the genome are not known, the extent to which any or all of these potential regulatory mechanisms are realized in vivo has been a matter of speculation.

The accumulation and propagation of torsional stress along a DNA fiber depends on the rate of transcription, the length of the transcribed unit and the arrangement of promoters (for example, divergent transcription would generate mutually reinforcing upstream negative supercoils)10,11. How torsional stress is transmitted through DNA will depend on the topological domains formed by protein-DNA interactions or by the anchoring of DNA to immobile nuclear structures12; such domains may concentrate or exclude supercoils. The binding of proteins, nucleosome positioning and histone modifications might all influence the transmission of torsional stress or the activity of topoisomerases. Fundamental to elucidating the role and the control of torsional stress in gene regulation is the understanding of its disposition within chromosomes.

Whether metazoan chromosomes (such as those in bacteria) are organized into supercoiled domains and whether such supercoils are regulated or regulatory remains controversial13. Supercoiling of intracellular DNA has been estimated from the intercalation of

---

1Laboratory of Pathology, National Cancer Institute, Bethesda, Maryland, USA. 2Department of Physics, University of Maryland, College Park, Maryland, USA. 3Computational Biology Branch, National Center for Biotechnology Information, Bethesda, Maryland, USA. 4Experimental Immunology Branch, National Cancer Institute, Bethesda, Maryland, USA. 5Center for Drug Evaluation and Research, Food and Drug Administration, Bethesda, Maryland, USA. 6These authors contributed equally to this work. Correspondence should be addressed to D.L. (levens@helix.nih.gov).

Received 18 April 2012; accepted 17 January 2013; published online 17 February 2013; doi:10.1038/nsmb.2517
psoralen derivatives into DNA; intercalators generally insert between the bases of underwound rather than overwound DNA where the bases are squeezed together\textsuperscript{14}. Recent studies in yeast and flies have provided a coarse-grain view of the distribution of torsional stress along chromosomes, but low resolution has hampered analysis of the factors governing the generation, relaxation and transmission of supercoiling at individual genes \textit{in vivo}\textsuperscript{15,16}. Site-specific experiments using Southern blots at several genes\textsuperscript{14,17\textendash}20 showed that although the genome is generally relaxed, supercoiled DNA exists at a few loci in mammalian and insect cells. This supercoiling remains largely unstudied. Torsional stress has also been measured by monitoring the supercoiling of episomes recovered directly from cells before or after excision from chromosomes, and it has been inferred from supercoil-dependent structural transitions in DNA or from the activity of supercoil-dependent recombinases\textsuperscript{7,12,21}. The low resolution or low throughput of these methods has provided a limited view of the interplay between the factors determining the generation, relaxation and transmission of DNA supercoiling \textit{in vivo}.

To address these issues, genomic oligonucleotide microarrays were probed with DNA photo–cross-linked by psoralen \textit{in vivo}. High-resolution mapping revealed dynamic supercoiling transmitted approximately 1.5 kilobases (kb) upstream from transcription start sites (TSSs) to be a characteristic of virtually every transcribed gene. High-output promoters sustain high levels of supercoiling that are counterbalanced by the recruitment of topoisomerases. Topoisomerase I (Topo I) and topoisomerase II (Topo II) are differentially recruited and distinctly deployed, illustrating the interconnection between DNA supercoiling and gene regulation.

RESULTS

Overview of the approach

As is well established, psoralen permeates cells, intercalates preferentially into underwound DNA and cross-links complementary DNA strands upon exposure to ultraviolet (UV) light\textsuperscript{22} (Fig. 1). Because psoralen intercalation is also favored by high AT content but disfavored by nucleosomes and other DNA-protein interactions\textsuperscript{16}, corrections for sequence and chromatin must be made to estimate supercoiling \textit{in vivo}. To quantify transcription-generated torsional stress—dynamic supercoiling—psoralen intercalation was compared between actively transcribing cells and cells in which transcription was specifically inhibited just before cross-linking. This comparison was intrinsically normalized for the effects of sequence and long-lived DNA-protein interactions (such as nucleosomes) to highlight the effect of dynamic supercoiling on psoralen cross-linking.

To measure cross-linking genome wide, Raji human B cells were treated with psoralen and UV light in G1 phase to minimize the influences of replication and mitosis on DNA topology\textsuperscript{7}. Genomic DNA was recovered, sonicated, denatured and electrophoretically fractionated to resolve slowly migrating cross-linked DNA from the faster-migrating un–cross-linked DNA. The cross-linked fraction was expected to be enriched for DNA negatively supercoiled at the instant of irradiation (Fig. 1a). The separated fractions labeled with Cy5 or Cy3 were hybridized to high-density oligonucleotide microarrays spanning ENCODE regions\textsuperscript{23}. The log ratio (cross-linked/un–cross-linked) of the fluorescent signals, defined as the cross-linking level (CL), provided a continuous picture of psoralen intercalation as a function of genome coordinate. Two computationally smoothed examples are provided (Fig. 1b). The CLs of promoters differed markedly from the CLs of intergenic regions, thus reflecting differences in DNA topology, GC content or specialized chromatin structures. Therefore, the absolute CL from untreated cells was not immediately instructive about the level of dynamic supercoiling.

Because physiologically achievable levels of negative supercoiling increase psoralen intercalation by only about two-fold relative to that for relaxed DNA\textsuperscript{14}, the resulting signal-to-noise ratio demands experimental replication to achieve statistical significance. Three biological replicates were required for significance of the CL profiles and other maps analyzed here (Supplementary Note). The data were averaged across replicates, and high-frequency noise was filtered by Fourier convolution smoothing\textsuperscript{24} (supercoiling levels would be expected to fluctuate on the scale of the torsional (~300 base pair (bp)) and bending (~150 bp) persistence lengths of DNA, not base pair to base pair). To observe the transcription-dependent supercoiling, the CL maps were compared between cells treated with an inhibitor of transcription elongation and untreated cells. Similarly, to reveal the dynamic character of DNA supercoiling and to examine its regulation, untreated versus topoisomerase inhibitor–treated cells were compared. Because different inhibitors act at different points in the topoisomerase reaction cycle, changes in DNA topology subsequent to treatment should reflect their modes of action\textsuperscript{25}.

To relate supercoiling with transcription, nuclear RNA was hybridized with these same microarrays. To correlate with psoralen

---

**Figure 1** Psoralen photobinding is a genome probe to detect DNA supercoiling \textit{in vivo}. (a) Overview of the approach: treatment of cells with psoralen followed by UV irradiation produces DNA interstrand cross-links. Thermal denaturation of genomic DNA fragments results in the formation of two fractions: the cross-linked fraction (XL) migrates slowly in gels, whereas the un–cross-linked (non-XL) population is composed of rapidly migrating single-stranded DNA. After electrophoretic separation, these fractions are purified, fluorochrome labeled and hybridized with densely tiled oligonucleotide arrays. The genomic distribution of the ratio of cross-linked and un–cross-linked DNA (log, scale being 0 at the global mean) represents the efficiency of psoralen intercalation. (b) Representative examples of the psoralen cross-linking map show peculiarities near TSSs. The \( x \) axis shows genomic position. The \( y \) axis shows the CL level, which is the log\(_2\) ratio (cross-linked/un–cross-linked) of the fluorescent signals from the DNA microarray. Negative CL values indicate lower propensity of psoralen intercalation. Curves are computationally smoothed. The breaks in the curve correspond to the sequences that were not profiled on the microarray, owing to uniqueness and conformational issues. Schematics of the genes are represented below each curve, with red arrows showing the TSSs. (c) Composite analysis of psoralen CL levels near the TSSs of ENCODE genes with low (from 0 to 20%) and medium (from 60 to 80%) expression, before and after treatment of cells with DRB.
Figure 2 DRB treatment does not affect nucleosome profiles or binding of non-nucleosomal proteins to DNA. (a) Nucleosome occupancy around the TSSs of genes with low (solid lines) and high (dashed lines) expression in the presence (orange lines) or absence (black lines) of DRB. The y axis shows the level of nucleosome binding expressed as tags per million (a.u., arbitrary units). The x axis shows the genomic coordinate centered on the TSS. For each position, the nucleosome coverage was determined by counting the number of nucleosomes mapping to that position.

(b) Enrichment of histone H1 and HMG14 proteins at promoter regions of selected genes (indicated by alphabetical letters; details in Online Methods) relative to a reference intergenic region before and after DRB treatment. Promoters are ranked in three groups which have, respectively from left to right, low, medium and high expression. Data are shown as percentage of input (n = 3 or 4; error bars, s.d.).

(c) DNA supercoiling around TSSs is determined for genes with low, medium and high expression. The y axis shows the ∆CL, which is the computational difference between CL values derived from DRB-treated and untreated cells. Negative ∆CL values reflect a higher propensity of psoralen to intercalate into the DNA, owing to transcription-dependent negative supercoiling. The zero point represents absence of transcription-dependent supercoils.

Dynamic supercoils upstream of promoters

The dynamic range of gene expression is so large that, at the extremes, mechanistic and structural differences might obscure the elastic response of chromatin to applied torque. Therefore, the smoothed CL profiles of 8-kb windows surrounding TSSs of genes with low (0–20%) and medium (60–80%) expression were compared to determine whether modest differences in torsional stress were detected. Meta-analysis of the data for both sets of promoters revealed troughs of CL (Fig. 1c) at TSSs that largely reside in psoralen-unfriendly CpG islands and are laden with transcription and chromatin complexes. The CL profiles were generated also for cells treated with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). DRB rapidly enters cells, specifically inhibits CDK9 phosphorylation of RNA polymerase II at the C-terminal domain and blocks transcription pause release4. After brief DRB treatment, the CL profile should reflect static properties, such as sequence and chromatin structure, but not dynamic supercoiling that was expected to be drained by topoiso merases and diffusion of torsional stress away from TSSs. (An alternative drug, α-amanitin, permeates cells and inhibits transcription too slowly to allow study of dynamic supercoiling). Indeed, the differences in the CL between DRB-treated and untreated cells were maximal near TSSs and gradually declined up to ~1.5 kb upstream for moderately expressed genes (Fig. 1c), whereas for weakly expressed genes the CL remained small and diminished more quickly upstream of the TSS with or without DRB inhibition of transcription, just as predicted.

To confirm that DRB-elicited CL changes reflected dynamic supercoiling and not large-scale chromatin rearrangements, nucleosome occupancy was monitored across the same gene set analyzed with psoralen cross-linking27. Although the nucleosome profiles of active versus inactive genes were distinct (downstream of TSSs, active genes had fewer but more well-positioned nucleosomes than inactive genes, as previously noted28), neither set was affected by DRB treatment (Fig. 2a). Thus nucleosome rearrangement cannot explain the changes in psoralen cross-linking provoked by DRB. The accessibility of upstream inter nucleosomal linker DNA to non-nucleosomal chromosome proteins H1 and HMG14 was evaluated by chromatin immunoprecipitation (ChIP) assay in the presence and absence of DRB across a panel of 15 genes expressed at different levels (Fig. 2b). As expected, binding of these factors paralleled decreased expression, irrespective of DRB treatment. Dynamic supercoiling seemed likely to explain the difference in psoralen binding upstream regions of active versus inactive genes. We defined a parameter called cross-linking difference (∆CL) to be a metric of dynamic supercoiling: ∆CL = CL_{DRB} – CL_{DRB} (Supplementary Note). ∆CL was defined to give negatively supercoiled regions a negative value; segments devoid of dynamic supercoils have a ∆CL near zero. ∆CL reports only transcription-dependent psoralen cross-linking while subtracting the influence of static DNA-protein interactions and DNA sequence. ∆CL profiles in a 4-kb window surrounding TSSs (Fig. 2c) decayed to baseline at about 1.5 kb from the TSSs of highly expressed genes and were further reduced and absent at weakly and nonexpressed genes, respectively. As the twin-supercoiled-domain theory predicts, ∆CL is diminished within gene bodies, where each RNA polymerase is a node between positive and negative supercoils (Supplementary Fig. 1a), and so between pairs of elongating polymerase, positive and negative supercoils cancel2. In gene bodies, nucleosomes are disrupted and reassembled during transcription; for these reasons, we focused on upstream regions, where interpretation is simpler, dealing mainly with negative supercoils emanating from TSSs.
Dynamic supercoiling is associated with promoters but not enhancers. Although enhancers bind RNA polymerase, the output of enhancer RNAs transcribed from these regions is meager compared to that of promoter-sponsored transcription (Fig. 4c). Concordantly, ΔCL was at background levels for enhancers, irrespective of distance to their associated promoters (Supplementary Fig. 1b). CTCF is a multi–zinc-finger protein mediating looping between remote regions of the genome (30); CTCF sites were not associated with changes in ΔCL, which indicated that CTCF-bounded loops do not comprise a capacitor for transcription-generated supercoils (Supplementary Fig. 1c).

To get a finer view of dynamic supercoiling upstream of TSSs, the ΔCL from −2 kb to +2 kb was graphed across a moving window through the expression spectrum. ΔCL strongly correlated with transcriptional activity, as predicted by the twin-supercoiled-domain theory. Weakly expressed genes showed only a small dip in the ΔCL close to the TSS, but as RNA production increased, the ΔCL signal spread up to 1.5 kb upstream (Fig. 4a). ΔCL became increasingly negative as gene expression increased to up to 80% of maximal expression. Beyond 80%, DNA supercoiling plateaued or even diminished near TSSs, which suggested that special mechanisms are marshaled to contend with the highest levels of torsional stress.

One way to reduce supercoiling near the highest-output promoters would be to recruit Topo I or Topo II more effectively. To test this, upstream regions of promoters with different outputs were analyzed by ChiP and quantitative PCR (qPCR; Fig. 4b) with and without DRB. Topo I and Topo II bound to the DNA were trapped by using very brief treatments with camptothecin (CPT) and β-lapachone (β-LAP), respectively. CPT highly selectively inhibits strand religation during the Topo I catalytic cycle, whereas β-LAP traps Topo II during the formation of the DNA-cleavage complex and inhibits Topo I before the strand cleavage (31–33). For both topoisomerases, the weakly transcribed genes were scarcely enriched relative to a reference intergenic region (Fig. 4c). Whereas Topo II was dramatically recruited to highly active genes, Topo I was most efficiently recruited to the promoters of moderately expressed genes. These results confirmed the differential recruitment of Topo I and Topo II according to promoter output.

Parameters controlling the level of dynamic supercoiling

In the twin-supercoiled-domain theory, three major factors define the DNA topology of regions upstream of promoters: (i) the rate of supercoil generation by RNA polymerase; (ii) how efficiently torsional stress is transported to chromatin locations by twist diffusion or in bloc rotation of chromatin segments; and (iii) the rate of supercoil removal by topoisomerases. The contributions of each of these parameters to upstream supercoiling were examined.

Dynamic supercoiling should increase with transcription; at steady state, the torsional stress generated will be balanced by topoisomerase activity. If torsional stress is freely transmitted through DNA fibers, then increased supercoiling near TSSs will be propagated farther upstream, unless there are barriers to twist and writhe diffusion. To assess the relationship between dynamic supercoiling and promoter output, ENCODE genes were ranked by expression, and the average ΔCLs from the TSS to −800 bp versus −4,800 to −5,600 bp were compared by sliding-window averaging across the expression spectrum (Fig. 3a,b). ΔCL was near zero in far-upstream regions, independent of expression, but at TSSs, ΔCL was clearly more negative with high expression. Thus dynamic supercoiling is a local feature of active promoters and not a characteristic of large chromosomal domains.

Figure 3 Differential patterns of supercoil generation for weakly to moderately versus highly transcribed genes. (a) Schematic representation describing the calculation used to determine the relationship between gene expression and DNA topology. (b) The ΔCL signal of upstream promoter regions was averaged over 800 bp for each gene and plotted against the level of gene expression (black curve). Smoothing of the curve was done. The ΔCL signal between −5,600 bp and −4,800 bp (orange curve) was graphed for comparison.

Figure 4 Perturbing the distribution of supercoils with topoisomerase inhibitors reveals the pattern of Topo I and Topo II recruitment to TSSs. (a) Three-dimensional representation of the ΔCL profiles of genes ranked according to their level of expression in the absence of topoisomerase inhibitors (top) or in the presence of CPT (middle) or β-LAP (bottom). (b) Schematic representation of qPCR design for the ChiP analysis. DNA recovery was determined at promoters and at reference intergenic regions. (c) ChiP in Raji cells treated with CPT or β-LAP in the presence or absence of DRB. The relative enrichment of Topo I (top) or Topo II (bottom) at the promoter areas of genes (indicated by alphabetical letter; details in Online Methods) with different expression levels is shown. Five genes were analyzed in each expression range. Data are normalized to a nonexpressed intergenic region (n = 3 or 4; error bars, s.d.).
To better relate dynamic supercoiling with topoisomerase activity, topoisomerase recruitment was evaluated with and without transcriptional inhibition. Whereas DRB treatment did not perturb Topo II levels, Topo I was reduced at medium-output promoters, those most sensitive to Topo I activity (Fig. 4c). Consistent with Topo I being a supercoil-sensitive enzyme dynamically associated with transcribed loci, its recruitment was dependent on transcription activation and supercoil generation. In contrast, Topo II was recruited by features other than dynamic supercoiling.

Dynamic supercoiling appeared sensitive to the distribution and kinetics of topoisomerases. To confirm this, the ∆CLs of promoter regions were compared with and without topoisomerase-inhibitor treatment. Topo I nicks a single DNA strand, relaxes supercoils by rotating about the intact DNA strand and then closes the nick. CPT at the DNA-protein interface hinders rotation of the nicked DNA. Consequently, in the presence of CPT, negative supercoiling should intensify transiently upstream of promoters. If the relationship between transcription and supercoiling is as hypothesized, then the ∆CL upstream of moderately expressed genes that depend on Topo I should be more sensitive to CPT than are highly expressed genes (Fig. 5). The effect of CPT was stronger for moderately expressed genes (Low β-LAP, which inhibits Topo I before strand nicking and Topo II in the midst of DNA cleavage, was selected to reveal Topo II’s role in resolving topological issues during transcription. Because Topo I and II are functionally partially redundant, and Topo I activity increases along with the dynamic supercoiling (Fig. 4c), it was necessary to inhibit Topo I (in order to blunt a confounding compensatory increase in Topo I activity) and Topo II. Topo II function could be inferred from the difference between inhibition of Topo I and of both Topo I and Topo II. To trap a double-strand break, both Topo II subunits have to interact simultaneously with the drug on each strand. Thus, with low β-LAP concentration and short treatments, nicks rather than double-strand breaks predominate, and diffusion of torsional stress off these nicks should result in the relaxation of regions served by the Topo II. Indeed, 5 min of β-LAP treatment uniformly relaxed upstream DNA, minimizing the ∆CL from the TSS to all upstream points (Figs. 4a and 5). Therefore, Topo II acted close to TSSs, relaxing negatively supercoiled DNA. Topo II inhibitors that evoked a rapid DNA-damage response were not studied (Supplementary Fig. 2).

RNA polymerase II ChIP-sequencing (ChIP-seq) data from Raji cells provided an independent measure of transcriptional activity for the transcriptional activity genes. The ∆CL profiles of these genes were sorted by pausing index that relates RNA polymerase II density at promoters versus gene bodies. Paused and elongating promoters corresponded to moderately and highly expressed genes; genes lacking RNA polymerase were designated as silent. The ∆CLs before and after inhibitor treatment followed the same patterns whether sorted by pausing index or expression (Fig. 5 and Supplementary Fig. 3).

Fine-tuning of DNA supercoiling with topoisomerases

Two scenarios may be hypothesized for the role of topoisomerases in the steady-state regulation of dynamic supercoils. In the first, negative torsional stress generated during transcription spreads into the upstream promoter regions (Fig. 6a), where diffusely recruited Topo I and Topo II relieve the stress. Because the probability that a segment upstream of the TSS remains topoisomerase free decreases exponentially with distance, the level of supercoiling is high near TSSs but decays sharply upstream with this model. Alternatively, if topoisomerases are recruited focally to the most dynamically stressed DNA, that is, TSSs (Fig. 6a), then supercoiling would be reduced right at the TSS, but beyond this zone, any residual supercoiling would decay only gradually. The ∆CL patterns were compared between sets of genes with different expression to discriminate between these possibilities. Whereas the ∆CL upstream of moderately transcribed genes decayed rapidly, suggestive of diffuse topoisomerase recruitment, at high-output promoters, ∆CL diminished at the TSS but thereafter declined.

[Image: Figure 5 ∆CL curves generated in the absence or presence of CPT or β-LAP. Genes with low, medium and high expression are shown in each graph.]

[Image: Figure 6 Differential topoisomerase I and II use in the regulation of transcription-induced torsional stress. (a) Two modes of topoisomerase recruitment at the upstream regions of promoters: the diffuse mode (solid line) suggests that topoisomerases are randomly distributed over the upstream promoter region; the focal mode (dashed line) hypothesizes that topoisomerases work near the TSSs. (b) The two modes can be visualized in the ∆CL curves for genes with medium and high expression. (c) Proposed model of supercoiling regulation by topoisomerases. Dynamic supercoiling near moderately active genes is mainly managed by Topo I, which is distributed over a broad upstream promoter region, whereas highly active promoters recruit Topo II focally near the TSS.]
gradually, consistent with focal recruitment (Fig. 6b). The observed relationship between transcription, supercoiling and the response to topoisomerase inhibition suggested that highly active genes recruit Topo II to their TSSs, whereas weakly expressed genes do not.

Use of topoisomerase-selective inhibitors allowed estimation of the relative relaxation by each enzyme as a function of expression. CPT increased supercoiling more with medium than high expression, but with high expression, supercoils diffused farther upstream (Fig. 5). With β-LAP, upstream supercoiling was lost for highly active genes, whereas with medium expression, ΔCL diminished at TSSs and further declined upstream (Fig. S). So topoisomerases acted redundantly upstream of moderately expressed genes, but when transcription intensified, Topo II was drawn to transcription start sites (Fig. 6c).

DISCUSSION

The role of dynamic supercoiling in the regulation and execution of genetic transactions has been incompletely described. Though torsional stress has been definitively demonstrated in naked DNA in vitro and in episomes in vivo, the pervasiveness and significance of dynamic supercoiling for most chromosomal genes is not known7,40. Recent studies that used psoralen as a probe for supercoiling in yeast revealed that distinct chromosome compartments confine different levels of DNA helical tension, but they lacked sufficient resolution to directly relate DNA topology to gene activity because yeast genes and the yeast genome are too compact to confine torsional stress to single targets, and the DNAs immobilized on the microarrays were insufficiently short to enable finer mapping16. A genome-wide study of psoralen binding to Drosophila polytene chromosomes was limited by the optical resolution of conventional light microscopy15. The present work studied ENCODE genes in their normal chromosomal context in the presence of functional topoisomerases. The resolution of high-density oligonucleotide arrays revealed the fine-grain distribution of dynamic supercoiling near promoters and its control by topoisomerases. These experiments showed that transcription-dependent DNA supercoiling was transmitted locally upstream of promoters but that highly expressed genes relied upon Topo II to dissipate dynamic supercoiling, whereas moderately expressed genes depended on Topo I. The level of supercoiling depends on the introduction of torsional stress into DNA and its removal by topoisomerases or by diffusion to remote regions of the genome3. The dynamics of supercoil diffusion depend on the properties of chromatin fibers: the positions of nucleosomes, the interactions between them, internucleosomal linker–binding proteins and the nucleosome modifications could all influence supercoil propagation. Our analysis revealed that torsional stress is dissipated over a short range (~1.5–2 kb), which suggests that dynamic supercoiling is not usually confined by fixed boundaries in chromatin. After crossing such boundaries, focal high levels of negative supercoiling decrease abruptly12,19. Alternatively, topological domain boundaries for each particular gene may be heteromorphic or transient in nature, thus resulting in variable domain sizes between similar cells. The simplest interpretation of our data is that DNA supercoiling upstream of the active promoter is established mostly by frictional restraint to DNA twist diffusion through chromatin. Context-dependent supercoiling associated with transgene location supports the idea that chromatin features may modify the distribution of supercoils17. Even without fixed boundaries, other architectural features, for example, divergent transcription, could modify the generation and propagation of dynamic supercoils (though the number of divergent promoters expressing both partners was too low to provide more than a hint of mutually reinforcing supercoils; Supplementary Fig. 1d and Supplementary Note). Meta-analysis of the ΔCL between convergent promoters was further complicated by variable gene length, multiple termination sites and engaged RNA polymerases downstream of termination sites.

As suggested by transcription defects in cells with mutant topoisomerases, proper supercoil management is required for efficient transcription41. Topo I and Topo II, which relax positive and negative supercoils, are fully redundant for transcription in yeast, but double mutants are severely impaired for elongation. However, in mammalian cells the topoisomerases only partially compensate for each other, which suggests specific functions for each in transcription42. The topological problems of gene expression may dictate specialized roles for each enzyme because the positive and negative supercoils generated during transcription distort DNA differently and reside in different molecular domains43. So recruitment of each topoisomerase to active genes may be context dependent10,44. The results of this study reveal two characteristics of the relaxation of transcription-induced DNA torsional stress by topoisomerases.

First, both Topo I and Topo II prevent the buildup of negative supercoiling in promoter regions. Topo I is a rapid, processive, torque-sensitive enzyme with low activity on nucleosomal templates45,46 that is well evolved to act ahead of elongating RNA polymerase, where accumulated positive torsional stress and histone modifications mobilize nucleosomes47. Inhibiting Topo I slows elongation in vivo5. In yeast, Topo II binds nucleosome-free regions near active TSSs48, and in mammalian cells the enzyme is enriched near the promoters of Topo II–sensitive genes49. In addition, Topo II activity would be favored by the crossed DNA segments50, as occurs when plectonemes form in negatively supercoiled DNA unbuffered by chromatin rearrangement1,51. Because all elongating transcription complexes impose a 90° bend in the template, as downstream DNA is screwed into the RNA polymerase active site, the upstream DNA is translationally rotated, thus generating writhe52. Therefore, Topo II would efficiently relax negative supercoiling behind the transcribing RNA polymerase. The twin-supercoiled-domain model predicts that dynamic negative supercoiling is highest at the promoter4. Accordingly, the activity of Topo II should be localized near the TSSs of highly active genes as demonstrated here. Some of the changes in ΔCL after β-LAP treatment may arise secondarily from Topo I inhibition. Because Topo I and Topo II are partially compensatory, combinations of inhibitors and rapidly inactivateable mutants will be required to separate the contribution of each topoisomerase to normal transcription.

Second, besides draining supercoils, it may be important to sustain a steady-state level of torsional stress in upstream regions to manage supercoil-driven structural transitions that respond to ongoing transcription2,40,53,54. The processive, fast but difficult to control Topo I is reduced at promoters of highly versus moderately expressed genes, DNA relaxation at these genes is accomplished by using step-by-step DNA transport by Topo II, in which ATP-driven conformational changes accompany very transient DNA breakage55. Thus at highly active genes, topological homeostasis could be better enforced by Topo II. Coordinating the rates of transcription and DNA relaxation would adjust supercoiling at different genes; if so, relaxing activity would be essential not only to remove topological impediments to transcription but also to establish a steady level of negative supercoiling within the regulatory regions of active genes.

DNA supercoiling in regulatory pathways

Much evidence supports the idea that DNA mechanics serves a variety of regulatory functions3,8. In vitro studies functionally couple chromatin structure with DNA topology47. DNA supercoiling may
assist chromatin remodeling, influence chromatin and DNA structure and modify DNA–transcription factor interactions. Torsional stress transported through DNA by changing the energy landscape and topology of the chromatin fiber may signal to remote sites. This signal could restrict DNA topology may be adapted for next-generation sequencing and influence RNA production.

Finally, because the structure and mechanics of cellular RNA polymerase is conserved across eukaryotes and prokaryotes, many of the DNA topology–sensitive regulatory mechanisms in bacteria may also operate in higher organisms. Despite differences in the complexity of the transcription machinery between kingdoms, both are forced to contend with the same polymer physics: the requirement to strongly bend DNA for preinitiation-complex formation and the need to locally melt DNA during initiation. Negative supercoiling facilitates both bending and melting; consequently, this fundamental linkage between DNA topology and transcription is conserved. In gyrase-containing bacteria, genomic DNA is globally undertwisted to optimize the transcription of many genes, but in higher eukaryotes, where genes are often separated by large reaches of inactive DNA, each gene may define its own topology. By coordinating topoisomerase activity with transcription, supercoiling in regulatory regions is dynamically buffered. Supercoiled-assisted melting might adjust the early rate-limiting steps of transcription and influence RNA production.

As costs decline, this psoralen-based procedure for the analysis of DNA topology may be adapted for next-generation sequencing and will help to uncover other DNA topology–related mechanisms in genome functioning.

METHODS
Methods and any associated references are available in the online version of the paper.

Accession codes
Data have been deposited in the GEO database under accession code GSE43752.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS
Our research is supported by the Intramural Research Program of the US National Institutes of Health, National Cancer Institute and Center for Cancer Research, National Library of Medicine and Food and Drug Administration. This study used the high-performance computational capabilities of the Helix Systems at the National Institutes of Health, Bethesda, Maryland (http://helix.nih.gov/). We thank R. Caaselas and A. Yamane for Illumina sequencing, and D. Clark and E. Batchelor for critical comments.

AUTHOR CONTRIBUTIONS
F.K., L.B., A.G., K.B.-A., J.L. and D.L. designed research; F.K. and L.B. performed experiments; A.G. developed analytical and computational tools; A.G., F.K., L.B., D.W., T.M.P. and D.L. analyzed data; F.K., A.G., L.B. and D.L. wrote the paper; A.G. and D.L. wrote the supplementary material.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nmb.2517. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Lavelle, C. Forces and torques in the nucleus: chromatin under mechanical constraints. Biophys. J. Cell Biol. 87, 307–322 (2009).
2. Liu, J.-F. & Wang, J.C. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA 84, 7024–7027 (1987).
3. Baranello, L. et al. The importance of being supercoiled: how DNA mechanisms regulate dynamic processes. Biophys. Biophys. Acta 1819, 632–638 (2012).
4. Darzacq, X. et al. In vivo dynamics of RNA polymerase II transcription. Nat. Struct. Mol. Biol. 14, 796–806 (2007).
5. Rocca, J. Transcriptional inhibition by DNA torsional stress. Transcription 2, 82–85 (2011).
6. Rocca, J. The torsional state of DNA within the chromosome. Chromosoma 120, 323–334 (2011).
7. Kouzine, F. et al. The functional response of upstream DNA to dynamic supercoiling in vivo. Nat. Struct. Mol. Biol. 15, 146–154 (2008).
8. Kouzine, F. & Lewens, D. Supercoil-driven DNA structures regulate genetic transactions. Front. Biosci. 12, 4409–4423 (2007).
9. Wong, B. et al. Characterization of Z-DNA as a nucleosome-boundary element in yeast Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA 104, 2229–2234 (2007).
10. Durand-Dubief, M. et al. Topoisomerase I regulates open chromatin and controls gene expression in vivo. EMBO J. 29, 2126–2134 (2010).
11. Seila, A.C. et al. Divergent transcription: a new feature of active promoters. Nucleic Acids Res. 29, 2578–2590 (2001).
12. Gartenberg, M.R. & Wang, J.C. Identification of barriers to rotation of DNA segments in yeast from the topology of DNA rings excised by an inducible site-specific recombinase. Proc. Natl. Acad. Sci. USA 90, 10514–10518 (1993).
13. Travasso, A. & Muskovitchlisvili, O. A common topology for bacterial and eukaryotic transcription initiation? EMBO Rep. 8, 147–151 (2007).
14. Kramer, P.R., Bat, O. & Sinden, R.R. Measurement of localized DNA supercoiling and topological domain size in eukaryotic cells. Methods Enzymol. 304, 639–650 (1999).
15. Matsumoto, K. & Hirose, S. Visualization of unstrained negative supercoils of DNA on polytene chromosomes of Drosophila. J. Cell Sci. 117, 3797–3805 (2004).
16. Bermúdez, I. et al. A method for genome-wide analysis of DNA helical tension by means of psoralen-DNA photobinding. Nucleic Acids Res. 38, e162 (2010).
17. Kramer, P.R. & Sinden, R.R. Measurement of unstrained negative supercoiling and topological domain size in living human cells. Biochemistry 36, 3151–3158 (1997).
18. Jupe, E.R., Sinden, R.R. & Cartwright, I.L. Stably maintained microdomain of localized unstrained supercoiling at a Drosophila heat shock gene locus. EMBO J. 12, 1067–1075 (1993).
19. Jupe, E.R., Sinden, R.R. & Cartwright, I.L. Specialized chromatin structure domain boundary elements flanking a Drosophila heat shock gene locus are under torsional strain in vivo. Biochemistry 34, 2628–2633 (1995).
20. Ljungman, M. & Hanawalt, P.C. Presence of negative torsional tension in the promoter region of the transcriptionally poised dihydrofolate reductase gene in vivo. Nucleic Acids Res. 23, 1780–1789 (1995).
21. Schwikardi, M. & Droge, P. Site-specific recombination in mammalian cells catalyzed by λ resolvase mutants: implications for the topology of episomal DNA. FEBS Lett. 471, 147–150 (2000).
22. Sinden, R.R., Bat, O. & Kramer, P.R. Psoralen cross-linking as probe of torsional tension and topological domain size in vivo. Methods 17, 112–124 (1999).
23. ENCODE Project Consortium. A user’s guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 9, e1001046 (2011).
24. Raghuraman, M.K. et al. Replication dynamics of the yeast genome. Science 294, 115–121 (2001).
25. Vos, S.M. et al. All tangled up: how cells direct, manage and exploit topoisomerase function. Nat. Rev. Mol. Cell Biol. 12, 827–841 (2011).
26. Nguyen, V.T. et al. In vivo degradation of RNA polymerase II largest subunit triggered by α-amanitin. Nucleic Acids Res. 24, 2924–2929 (1996).
27. Wei, G. et al. Genome-wide mapping of nucleosome occupancy, histone modifications, and gene expression using next-generation sequencing technology. Methods Enzymol. 513, 297–313 (2012).
28. Schones, D.E. et al. Dynamic regulation of nucleosome positioning in the human genome. Cell 132, 887–898 (2008).
29. Kim, T.K. et al. Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–187 (2010).
30. Ohlhson, R., Lobanenkov, V. & Kleinová, E. Does CTCF mediate between nuclear organization and gene expression? Bioessays 32, 37–50 (2010).
31. Friedman, B. et al. Induction of DNA topoisomerase II-mediated DNA cleavage by β-lapachone and related naphthoquinones. Cancer Res. 57, 620–627 (1997).
32. Li, C.J., Averboukh, L. & Pardee, A.B. β-Lapachone, a novel DNA topoisomerase I inhibitor with a mode of action different from camptothecin. J. Biol. Chem. 268, 22463–22468 (1993).
33. Pommer, Y. Topoisomerase I inhibitors: camptothecins and beyond. Nat. Rev. Cancer 6, 789–802 (2006).
34. Zebeck, K.L. et al. Recruitment timing and dynamics of transcription factors at the Hsp70 loci in living cells. Mol. Cell 40, 965–975 (2010).
35. Koster, D.A. et al. Antifourm drugs impede DNA uncoupling by topoisomerase I. Nature 448, 213–217 (2007).
36. Khobta, A. et al. Early effects of topoisomerase I inhibition on RNA polymerase II along transcribed genes in human cells. J. Mol. Biol. 357, 127–138 (2006).
37. Muslimovic, A. et al. Numerical analysis of etoposide induced DNA breaks. PLoS ONE 4, e5859 (2009).
38. Barski, A. et al. High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837 (2007).
39. Core, L.J., Waterfall, J.J. & Lis, J.T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322, 1845–1848 (2008).
40. Kouzine, F. et al. The dynamic response of upstream DNA to transcription-generated torsional stress. Nat. Struct. Mol. Biol. 11, 1092–1100 (2004).
41. Brill, S.J. & Sternglanz, R. Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. Cell 54, 403–411 (1988).
42. Baranello, L. et al. 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-1α gene locus. Nucleic Acids Res. 38, 159–171 (2010).
43. Timsit, Y. & Varnai, P. Helical chirality: a link between local interactions and global topology in DNA. PLoS ONE 5, e9326 (2010).
44. French, S.L. et al. 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 (2011).
45. Liu, J. et al. Role of topoisomerase IIβ in the expression of developmentally regulated genes. Mol. Cell Biol. 26, 7929–7941 (2006).
46. Lyu, Y.L. et al. Role of topoisomerase IIβ in the expression of developmentally regulated genes. Mol. Cell Biol. 26, 7929–7941 (2006).
47. Roca, J. et al. The path of the DNA along the dimer interface of topoisomerase II. J. Biol. Chem. 279, 25783–25788 (2004).
48. Lesne, A., Becavin, C. & Victor, J.M. The condensed chromatin fiber: an allosteric chemo-mechanical machine for signal transduction and genome processing. Phys. Biol. 9, 013001 (2012).
49. Unniraman, S. & Nagaraja, V. Axial distortion as a sensor of supercoil changes: a molecular model for the homeostatic regulation of DNA gyrase. J. Genet. 80, 119–124 (2001).
50. Revyakin, A., Ebright, R.H. & Strick, T.R. Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. Proc. Natl. Acad. Sci. USA 101, 4776–4780 (2004).
51. Parvin, J.D. & Sharp, P.A. DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. Cell 73, 533–540 (1993).
Microarrays used in this study were designed with each probe having 50-bp and a 12-bp overlap. Thus any given region of the genome is represented by a 38-bp non–cross-linked and cross-linked DNA fragments were separated by electrophoresis, the cells is within the linear range of DNA binding. After electrophoresis, the gel was incubated with denaturing solution (0.5 M NaOH, 1.5 M NaCl) at 65 °C for 3 h to reverse psoralen cross-links and stained with SYBR green 15. Cross-linked non–cross-linked and cross-linked DNA fragments were purified by electrophoresis and hybridized in different combinations to NimbleGen ENCODE arrays (tiled with 12-bp-overlapping, 50-mer probes across unique regions of ENCODE as well as 200 kb centered on the MYC TSS). Three biological replicates, each hybridized to a new array, were assayed. DNA labeling, hybridization, detection, data extraction and quality assessment were performed at NimbleGen.

Data analysis. To smooth the data, the Fourier convolution technique was used. The technique uses a window size of 400 data points (equivalent to 561 bp; Fig. 1c and subsequent figures). NimbleGen ENCODE (hg18) microarrays had usable data for a total of 855 transcriptional regions (Supplementary Note). In order to avoid multiple counting of same gene, clusters were identified that were either overlapping or had a TSS within 50 bp of each other. Only the largest 'transcribed region' from each of these groups was used. This brought down the total number of transcribed regions to 445 (with 415 unique genes). The complete list of the transcribed regions is provided in Supplementary Table 1.

Sequence-dependent background correction. To remove the sequence bias of psoralen intercalation, the transcribed regions were sorted on the basis of AT content within ±3,000 bp of TSSs. To reduce systematic errors, these 445 transcribed regions were divided into 10 groups. A correction term, for each of the decades, was calculated by averaging the raw ratios in the flanking regions, –8 kb to –2 kb and +2 kb to +8 kb of each transcribed region. The data for each of the constituent transcribed regions were then baseline shifted by using this base-composition correction term to yield the corrected data, which were used for further analysis.

Supplementary Note gives a detailed description.

ChIP and qPCR. ChIP assays were performed with Raji cells as described, with minor changes. Briefly, 5 × 10⁶ cells were cross-linked with 1% formaldehyde and sonicated in TE to produce chromatin fragments of 800 bp on average. Immunoprecipitations with Dynabeads Protein A (Invitrogen) were carried out, using 4 µg of antibodies. Immunoprecipitates were washed twice with RIPA buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, pH 8.0, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, 200 mM NaCl); twice with RIPA buffer plus 300 mM NaCl; twice with LiCl buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% Na deoxycholate); and twice with TE. The beads were then resuspended in TE plus 0.25% SDS supplemented with proteinase K (500 µg/ml, Roche) and incubated overnight at 65 °C. The DNA was recovered from the eluate by phenol-chloroform extraction followed by ethanol precipitation and was dissolved in TE. For qPCR detection, performed using the LightCycler 480 and FastStart DNA Master SYBR Green I kit (Roche Diagnostics), the percentage of IP enrichment as compared to input was calculated, and data are presented as fold change with respect to a reference intergenic region of drug-treated cells. Quantification and melting-curve analyses were performed by using the Roche LightCycler software by the crossing-point method, as indicated by the supplier. Fifteen genes were analyzed in total; five genes in each group ranked according to the RNA production: 0–20%, 60–80% and 80–100%. All detection primers are listed in Supplementary Table 2. Antibodies used were as follows: anti-Topo I (Epitomics, 3552-1); anti-Topo II-β (Epitomics, 3747-1); anti-histone H1 (Santa Cruz Biotech., sc-8030) and anti-HMG14 (Abcam, ab5212). For all the antibodies used in this study, the DNA-recovered values were around 100-fold higher than for the IgG controls (Santa Cruz Biotechnology: normal mouse IgG, sc-2025; normal rabbit IgG, sc-2027).

Total protein extraction and western blot. Cells were lysed with RIPA buffer, briefly sonicated and then pelleted by centrifugation. The supernatant was saved for western blotting. Proteins (40 µg) were subjected to SDS-PAGE and then blotted onto a nitrocellulose membrane for incubation with specific antibodies: anti-CHK2 (Cell Signaling, 2662); anti–phospho-CHK2 (Cell Signaling, 2661); anti-γ-H2AX (Millipore, 05-636); and anti-histone H3 (Abcam, ab1797), all used at 1:2,000 dilutions.

ChIP-seq and MNase-seq data processing. ChIP-seq and MNase-seq on the for maldehyde cross-linked cells were performed as described. Raw sequencing data from Pol II ChIP-seq (single-end tags) and MNase-seq (pair-end tags) libraries were processed by using Illumina Analysis Pipeline, and the acquired reads were mapped to the human genome (hg18) by using Bowtie (http://bowtie-bio.sourceforge.net/index.shtml). Only uniquely mapped reads located within the ENCODE regions were retained. For MNase-seq, we used only those pairs of reads that were between 100 bp and 250 bp apart. The nucleosome coverage map was calculated for each position by counting the number of nucleosome-sized fragments (centered at the midpoint of each pair of reads) that cover that position.