Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures

Catherine Naughton1,2, Nicolaos Avlonitis3, Samuel Corless1, James G Prendergast1, Ioulia K Mati3, Paul P Eijk4, Scott L Cockroft3, Mark Bradley3, Bauke Ylstra4 & Nick Gilbert1,2

DNA supercoiling is an inherent consequence of twisting DNA and is critical for regulating gene expression and DNA replication. However, DNA supercoiling at a genomic scale in human cells is uncharacterized. To map supercoiling, we used biotinylated trimethylpsoralen as a DNA structure probe to show that the human genome is organized into supercoiling domains. Domains are formed and remodeled by RNA polymerase and topoisomerase activities and are flanked by GC-AT boundaries and CTCF insulator protein–binding sites. Underwind domains are transcriptionally active and enriched in topoisomerase I, ‘open’ chromatin fibers and DNase I sites, but they are depleted of topoisomerase II. Furthermore, DNA supercoiling affects additional levels of chromatin compaction as underwind domains are cytologically decondensed, topologically constrained and decompacted by transcription of short RNAs. We suggest that supercoiling domains create a topological environment that facilitates gene activation, providing an evolutionary purpose for clustering genes along chromosomes.

DNA has a helical structure that is influenced by the localized sequence context and remodeling by cellular machines, but at the level of DNA the net state of the genome is torsionally relaxed1. However, negative DNA supercoils are detected in localized regions2,3 where they can alter the topology of the chromatin fiber4. This altered structure is maintained by transcription, which generates both positive (overwind) and negative (underwind) supercoils ahead of and behind RNA polymerase5,6, respectively, and in the absence of transcription the chromatin returns to a ground state7. In general, genes that are supercoiled are more efficiently transcribed8,9, and as tension may be a prerequisite for transcription initiation10–13, supercoiling domains might influence the gene expression of surrounding genes. However, excessive supercoiling inhibits elongation14,15, topoisomerases are preferentially associated with transcribed regions, where they can relieve tension16,17. It is unclear how negative supercoils are introduced into loci before transcriptional activation, but certain inducible genes are probably organized into a transcriptionally poised chromatin environment17,18. At other genes, the transcription initiation complex transcribes short stretches of RNA, and the polymerase synthesizes both sense and antisense RNAs19–21. Furthermore, transcription of short upstream RNAs might be important for DNA supercoiling19.

The human genome is organized into GC- and AT-rich sequence regions termed isochores22, which were originally defined by cesium chloride ultracentrifugation. Genes are preferentially found in GC-rich isochores, and gene clustering may reflect the impact of a higher-order chromatin organization that influences gene expression23–25. In support of the existence of this genome-wide chromatin fiber, analysis has shown that gene-rich domains of the human genome are enriched in disrupted (open) chromatin fibers and that these regions have a decompacted large-scale chromatin structure26 that is remodeled by transcription27. Within gene clusters, open chromatin can spread to surrounding regions27,28, suggesting that chromatin states can propagate. In mammalian cells nucleosome arrays29 fold to form chromatin fibers that are subsequently organized into large-scale chromatin structures in an unknown manner. In metaphase chromosomes, after gentle lysis, chromatin loops can be seen to emanate from a proteinaceous scaffold-like structure30 that might persist in interphase31–33. Consistent with this, in interphase, chromosome loops tens to hundreds of kilobases long have been seen to form thick “chromonema” fibers34, and nuclease digestion of cells indicates that the chromatin is organized into 50-kilobase (kb) loops, forming higher-order 300-kb structures35. In contrast, the random walk or giant loop model of chromatin organization predicts a flexible backbone to which giant loops, each comprising several megabases of DNA, are attached36. Much of our understanding of genome organization has depended on cytological approaches, but recently molecular techniques such as Hi-C have shown that the interphase genome is organized into self-interacting topological domains31–33 and these may correspond to cytologically visible large-scale chromatin structures. However, the factors that regulate or determine these structural domains and their relationship to function are poorly understood.

To investigate DNA supercoiling on a genomic scale and to study its impact on higher levels of chromatin organization, we developed a new approach using biotinylated 4,5,8-trimethylpsoralen (TMP) as a probe to map DNA supercoiling in cells. We found that transcription and topoisomerase activity alter DNA supercoiling around
transcription start sites (TSSs), and that this alteration is propagated along the chromatin fiber, creating heterogeneously sized supercoiling domains. These domains are delimited by GC-AT boundaries and correspond to ‘underwound,’ ‘overwound’ or topologically ‘stable’ regions. Underwound domains are GC rich, highly transcribed and enriched in open chromatin fibers, TSSs, RNA polymerase and topoisomerase I binding, but they are depleted in topoisomerase II. Underwound domains also have a decompacted large-scale chromatin structure that is maintained by a balance between transcription and topoisomerase activities, suggesting that changes in DNA topology can be propagated through the chromatin fiber, affecting further levels of chromatin organization.

RESULTS
Mapping DNA supercoiling at chromosomal loci
Despite the role of DNA supercoiling in transcriptional regulation and DNA replication, its organization, remodeling and impact across the genome are unknown. As the introduction of negative DNA supercoils untwists DNA strands, TMP intercalation has been used to monitor superhelicil tension on a global scale and at specific genes. We further developed this approach to monitor changes in DNA supercoiling at genomic loci, around genes and at regulatory elements in vivo, by attaching a biotin, via a linker, to TMP (bTMP) for use as a DNA structure probe (Fig. 1a). bTMP permeates the cell and preferentially intercalates into regions of DNA enriched in negative supercoils (Fig. 1a). After intercalation, bTMP is cross-linked by 360-nm light to form adducts between either its 3,4-pyrone or 4,5-furan bond and the 5,6 double bond of pyrimidine bases. We monitored bTMP incorporation in human retinal pigment epithelial (RPE1) cells by using a NeutrAvidin-fluorescein conjugate and found that it is evenly bound across the nucleus, with no pronounced enrichment or depletion at heterochromatin or in the vicinity of nucleoli (Fig. 1b). Torsional stress can be rapidly released by treating cells with bleomycin, a glycopeptide antibiotic that introduces DNA strand breaks (Supplementary Fig. 1). Treatment of cells with bleomycin led to a substantial decrease in bTMP binding, showing that the bTMP probe can discriminate between different DNA structures.

To map DNA supercoiling at high resolution, we briefly treated RPE1 cells with bTMP (20 min) and exposed them to UV light (10 min). bTMP was quantitatively incorporated into DNA (Fig. 1c) and bound in a UV-dependent manner. DNA bound to bTMP was fragmented by sonication, purified, enriched using streptavidin-coated magnetic beads, labeled and hybridized to custom-made tiling microarrays covering large regions of the human genome (Fig. 1d and Supplementary Table 1). We used a concentration of bTMP that introduced only one cross-linked psoralen molecule every 1–5 kb. To correct for topology-independent effects of the bTMP drug, we analyzed bTMP binding on naked genomic DNA. bTMP binding to genomic DNA was subtracted from bTMP binding in cells, that is, log2(bTMP Cells/input) – log2(bTMP genomic DNA/input), giving the normalized log2(bTMP/input binding).

To investigate DNA supercoiling on a genomic scale, we analyzed bTMP binding across human chromosome 11 using a custom-designed Nimblegen tiling array (Fig. 1e). Chromosome 11 is representative of the genome as it covers 135 megabases (Mb; 5% of the genome) and comprises regions of differing gene densities and base compositions. For example, 11p15.5 and 11p15.1 are gene rich whereas 11p14.1 is gene poor. Control cells showed differential binding of bTMP across the chromosome, with gene-rich T-bands being more negatively supercoiled than gene-poor G-bands; we found pronounced peaks of negative supercoiling at highly expressed genes such as NEAT1 and MALAT1 at 11q13.1 (Supplementary Fig. 2). To investigate the factors regulating supercoiling, we treated cells with α-amanitin, an inhibitor of transcription (Supplementary Fig. 3).
Figure 2 Organization and boundaries of supercoiling domains. (a) Microarray data of bTMP binding, indicative of DNA supercoiling, at HSA 11p15.4, spanning two topological domains. Distribution of DNase I–sensitive sites and CTCF-binding sites in RPE1 cells obtained from the ENCODE project. (b) Overlap between topological domain boundaries and supercoiling (SC) boundaries across HSA 11. Overlap was determined by taking a ±20 kb window at each topological boundary and assessing whether this overlapped with a SC boundary (P < 0.01 by random permutation). (c) Base composition around SC boundaries. (d) Number of CTCF sites on HSA 11 near a SC boundary. Overlap was determined by counting number of CTCF-binding sites within a ±20 kb window at each SC boundary. (e) Number of CTCF-binding sites surrounding a SC boundary (Kolmogorov-Smirnov test compared with randomly generated data, P < 2.2 × 10^{-16}).

After 5 h of treatment, α-amanitin was washed out and cells were allowed to recover for 2 h. Transcription inhibition substantially altered bTMP binding at all chromosomal loci (Fig. 1e), indicating considerable DNA remodeling. After drug washout, bTMP binding was rapidly returned to its starting state, showing that DNA supercoiling was remodeled by transcription.

We examined chromosomal loci and found that DNA supercoiling was organized into domains (Fig. 1e). We reasoned that supercoiling domains would be regions that could change their structure independently of surrounding regions. We therefore calculated a difference profile between the supercoiling signal for control and α-amanitin–treated cells and used an edge filter to identify supercoiling boundaries (Fig. 1f; see Online Methods). From this we identified 606 domains across chromosome 11 with a median size of ~100 kb (Fig. 1g) and characterized them as being underwound, overwound or stable on the basis of differences in supercoiling between control cells and α-amanitin–treated cells. Underwound domains corresponded to 52% of the total domains, but by size corresponded to only 30% of the domains, indicating they were generally smaller than overwound domains.

Organization of DNA supercoiling domains

To better understand the arrangement of supercoiling domains, we examined large chromosomal regions. Locus 11p15.4 (Fig. 2a) was both enriched and depleted in bTMP binding in underwound and overwound domains, respectively. Underwound domains were more often enriched in DNase I–binding sites (according to data from the Encode project, http://genome.ucsc.edu/ENCODE/)—for example, around 6.5 Mb. But there were also stable and overwound regions enriched in DNase I–sensitive sites, indicating that there was not a simple relationship between chromatin structure and DNA supercoiling. Recently Hi–C^{31} studies have shown that the human genome is organized into large topological domains. These domains are large (~900 kb) and conserved among cell types and species and might correspond to a higher level of genome organization. Our analysis showed that ~30% of topological boundaries lay within ±20 kb of supercoiling domain boundaries (P < 0.01 by simulation; Fig. 2b). We noticed that many of these supercoiling boundaries were located at a transition between GC-rich and GC-poor regions. To confirm this, we aligned the boundaries and analyzed the base composition across the regions to show that supercoiling domains were flanked by GC-AT boundaries (Fig. 2c). Topological domain boundaries are associated with the CTCF insulator protein^{31}. Similarly 25% of CTCF-binding sites were found within ±20 kb of a supercoiling boundary (Fig. 2d), compared with 15% expected by chance, with a pronounced peak of CTCF-binding sites at the boundaries (Fig. 2e; simulated data gives flat line at 0.32 below data shown; Kolmogorov-Smirnov test compared with randomly generated data, P < 2.2 × 10^{-16}). This analysis indicates that large topological domains are subdivided into dynamic supercoiling domains flanked by CTCF proteins, showing an additional level of transcription-dependent genome organization.

Remodeling of DNA supercoiling domains

To investigate the formation and remodeling of supercoiling domains, we examined 17.5 Mb of the human genome across seven loci at high resolution using Agilent tiling path arrays (Supplementary Table 1). The Xq13.1 locus (Fig. 3a) has a GC-rich cluster of genes around 69.5 Mb that was negatively supercoiled. In contrast, two large gene bodies (EDA (expressed) and TEX11 (not expressed)) located near 69 and 70 Mb were AT rich and positively supercoiled, indicating there might be a relationship between GC content and supercoiling. However, this result was not due to a localized (<100 bp) GC effect as we found no relationship between localized GC content and DNA supercoiling (data not shown), and mapping DNA supercoiling using GC- or AT-rich probes gave similar results. Similarly, there were very pronounced changes at the 11p15.1 locus (Fig. 3b). This region is gene rich and predominantly consists of underwound domains; transcription inhibition here promotes a rapid change in supercoiling—for example, around 17.5 Mb—that is reversed upon drug washout. Generally genes were located within single domains; however, some genes, such as SERGEF (at 18 Mb), were broken into both underwound and overwound regions. Supercoiling generated by transcription of one gene may be canceled out by transcription of a neighboring gene in the same direction or be enhanced by convergent or divergent
transcription of neighboring genes. The active SCYL1 and LTBP3 genes at 11q13.1 are convergent (Supplementary Fig. 2), but there were no pronounced changes in DNA supercoiling, suggesting that some mechanism removed extraneous supercoils.

Bleomycin treatment of cells introduced DNA nicks (Supplementary Fig. 1) and diminished differences in DNA supercoiling across chromosomal loci, showing that the different supercoiling structures we observed depended on DNA topology. Furthermore, treatment of cells with the topoisomerase inhibitors camptothecin (topoisomerase I) and ICRF193 (topoisomerase II) gave similar results, indicating that topoisomerases were also required to maintain the supercoiling state. However, treatment of cells with both α-amanitin and topoisomerase inhibitors had a limited effect, showing that changes in supercoiling after transcription inhibition required topoisomerase activity.

To quantify changes in supercoiling at both gene-rich and gene-poor chromosomal regions, we summed the amounts of bTMP binding across the loci (Fig. 3c). Gene-rich regions (11p15.5, Xq13.1, 11p15.1, Enr312 and Enr332) were significantly more negatively supercoiled (underwound) than gene-poor regions (11p14.1 and Xq25; P < 2.2 × 10^{-16}, t-test). At gene-rich loci, bTMP binding decreased after transcription inhibition and then increased, consistent with an increase in negative supercoiling, after transcription recovery. Notably, gene-poor regions behaved in the opposite manner, with an increase in positive supercoiling after transcription initiation. As total bTMP binding in cells does not change with transcription inhibition (data not shown), this suggests that supercoiling may be propagated between chromosomal regions.

### Properties of supercoiling domains

Supercoiling domain boundaries could be physical structures enriched in CTCF insulator proteins (Fig. 2) and topoisomerase II that bind to a proteinaceous scaffold. Alternatively, the domains themselves might be enriched in enzymatic activities that facilitate domain formation. To analyze these possibilities, we mapped transcription, RNA polymerase and topoisomerase binding across the chromosomal loci (Fig. 4a). Underwound domains were highly transcribed, and there was a strong relationship between transcription and supercoiling (Fig. 4b). Underwound domains were significantly enriched in DNase I-sensitive sites, disrupted (open) chromatin fibers and TSSs, and were GC rich (Fig. 4c). Underwound domains were also enriched in RNA polymerase and topoisomerase I, which were correlated with each other (Fig. 4d). Furthermore, underwound domains were significantly depleted in topoisomerases IIα and IIβ (Fig. 4a,c). Underwound domains generally had an organization opposite to that of underwound domains, whereas stable domains had properties intermediate to both. These data suggest that polymerases and topoisomerases located in domains determine their topology; this
would allow supercoiling to propagate between domains but be controlled by the underlying organization of genes, base composition and flanking by GC-AT boundaries and CTCF insulator proteins (Fig. 2). We therefore suggest that domain structure varies depending on the transcriptional environment; this indicates that supercoiling domains are dynamic properties of the chromatin fiber rather than static, structural features.

**DNA supercoiling at promoters and regulatory elements**

As some gene promoters are negatively supercoiled\(^2\),\(^3\)\(^\text{a}\), we analyzed DNA supercoiling upstream and downstream of TSSs of coding and noncoding genes (Fig. 5a). TSSs were negatively supercoiled in a region that extends ~20 kb into the body of the gene and 10 kb upstream. DNA nicking with bleomycin reduced psoralen binding, a region that extends ~20 kb into the body of the gene and 10 kb upstream. DNA nicking with bleomycin reduced psoralen binding, allowing supercoiling to propagate between domains but being constrained by transcription. To investigate other enzymatic activities required for remodeling DNA supercoiling at TSSs, we treated cells with inhibitors of topoisomerases I and II (Fig. 5b). We observed changes in TSS supercoiling with topoisomerase treatment, but these were abrogated when cells were treated with both transcription and topoisomerase inhibitors; thus topoisomerase activity is required for remodeling DNA supercoiling at promoters and chromosomal domains (Fig. 3). To further investigate DNA supercoiling at other features, we analyzed bTMP binding at DNase I sites and found that these were negatively supercoiled in a transcription-dependent manner (Fig. 5c). In contrast, randomly generated DNA shows no peaks around genomic features (data not shown). Furthermore, CTCF sites that can act as insulators and facilitate chromatin–chromatin interactions\(^4\)\(^\text{a}\) are negatively supercoiled, as are p300-CBP enhancer binding sites (data from A549 cells; Fig. 5d); this indicates that factor binding might be influenced by underlying DNA topology, facilitated by topoisomerase I activity (Fig. 5e).

**Large-scale chromatin compaction of supercoiling domains**

The mechanisms responsible for folding large-scale chromatin structures are poorly characterized. As packaging of large-scale chromatin structures can be transcription dependent\(^2\),\(^4\),\(^1\),\(^4\),\(^\text{a}\) we hypothesized that an alteration in DNA supercoiling may affect higher levels of chromatin organization. To investigate a relationship between supercoiling and compaction of chromatin structures, we analyzed underwound (11p15.5 and 11p15.1) and overwound regions (11p14.1; Fig. 6a, b) and then used three-dimensional (3D) DNA fluorescence in situ hybridization (FISH) to study large-scale chromatin compaction of these genomic loci. Labeled fosmid probes (Supplementary Table 2) separated by ~1.5 Mb were differently labeled and hybridized to RPE1 cells (Fig. 6c). Compaction of different regions was analyzed by measuring the distance between probes. The gene-rich, underwound 11p15.5 and 11p15.1 loci (Fig. 6d) were more cytologically decompacted than the overwound 11p14.1 locus, indicating a relationship between DNA supercoiling and compaction. To test whether large-scale chromatin...
structures are under tension, we treated cells with bleomycin. Loci 11p15.1 (1.5 Mb) and Xq13.1 (2 Mb) were compacted by bleomycin treatment (Fig. 6e,f), showing that the cytological structure of chromosomal regions was under topological strain.

Remodeling of large-scale chromatin structures
To assess the effect of transcription inhibition on large-scale chromatin structures, we analyzed the compaction of the 11p15.5 and 11p15.1 loci. After α-amanitin treatment, the loci became positively supercoiled (overwound, Fig. 3c) and the chromosomal structures concomitantly became compacted (Fig. 7a), suggesting that transcription can affect large-scale chromatin structures by altering DNA topology.

To investigate the kinetics of transcription responsible for altering DNA supercoiling, we inhibited transcription with α-amanitin, washed the drug out and monitored changes in RNA polymerase and RNA transcription during recovery (Fig. 7b). Western blots showed a substantial, proteasome-dependent (data not shown) loss in elongating RNA polymerase that persisted through recovery, whereas the initiating form of the polymerase was more resistant to degradation (Fig. 7c). After transcriptional activation, we reasoned that elongating RNA polymerase would produce large RNA fragments consistent with gene transcription, whereas a polymerase engaged in nonproductive transcription would produce shorter transcripts. Pulse labeling of the cells with [3H]uridine showed that production of long RNAs (>200 nucleotides (nt)) was substantially reduced after 5 h of α-amanitin treatment and continued to decrease after α-amanitin washout (Fig. 7d). In contrast, short RNAs (<200 nt) were...

Figure 5 DNA supercoiling around TSSs and regulatory elements. (a) bTMP binding ±20 kb around TSSs with or without bleomycin for active and inactive genes. (b) Changes in DNA supercoiling around TSSs after inhibiting RNA polymerases and topoisomerases. (c) Distribution of DNA supercoiling ±5 kb of DNase I–sensitive sites before and after transcription inhibition. (d) DNA supercoiling ±10 kb around CTCF- and p300-CBP-binding sites. CTCF-binding sites for RPE1 cells were obtained from the ENCODE project. p300-CBP-binding sites are for A549 cells from ENCODE project. (e) Topoisomerase binding ±20 kb around CTCF-binding sites determined by ChIP microarray. Peak signal versus randomly generated background data, P < 2.2 × 10^{-16} (f-test).

Figure 6 Underwound domains are cytologically decondensed and torsionally constrained. (a) Ideogram of human chromosome (HSA) 11 with T-bands, R-bands, G-bands and probe positions at underwound and overwound 1.5 Mb chromosomal loci studied. (b) DNA supercoiling at underwound and overwound loci, as measured by bTMP binding. (c) Representative images of 3D DNA FISH of pairs of labeled fosmid probes (red and green spots) positioned 1.5 Mb apart at underwound and overwound loci, as measured by bTMP binding. (d) Distance between pairs of fosmid probes at underwound and overwound loci. (e) Representative images of 3D DNA FISH of pairs of labeled fosmid probes (red and green spots) positioned 1.5 Mb apart at 11p15.1 with or without bleomycin. Nuclei were counterstained with DAPI. Scale bar, 5 µm. (f) Change in large-scale chromatin compaction at the 11p15.1 and Xq13.1 loci after treatment with bleomycin. Box plots and P values (Wilcoxon test) are described in Figure 4. n is the number of separate probe pairs examined.
produced in abundance by 1 h after α-amanitin washout and continued to increase at 2 h. These results show that changes in DNA supercoiling precede transcriptional elongation; at this time, the cell is synthesizing short RNAs and only the initiating form of RNA polymerase is present. On the basis of this, we hypothesized that the initiating form of RNA polymerase synthesizes short RNAs. To further investigate this, we inhibited transcription elongation with flavopiridol, a kinase inhibitor that blocks phosphorylation of the C-terminal domain of RNA polymerase II, leaving it in a nonprocessive state. Elongating polymerase decreased substantially (Supplementary Fig. 4a). We monitored RNA synthesis (Supplementary Fig. 4b) and found a rapid reduction in long RNA production, but a lag before a drop in the synthesis of short RNAs, consistent with initiating RNA polymerase producing short RNAs. To assess how large-scale decompacted chromatin structures were formed, we treated cells with α-amanitin followed by drug washout, and monitored the compaction of 11p15.1 by FISH. Upon transcription inhibition the locus was rapidly compacted, but it was decompacted after α-amanitin washout with the same kinetics (Fig. 7e) as for the remodeling of DNA supercoiling (Fig. 3) and the transcription of short RNAs (Fig. 7d).

Our previous experiments showed that topoisomerase inhibition caused a change in DNA supercoiling (Fig. 3). Therefore, to further

Figure 7 Transcription and topoisomerase dependence of large-scale chromatin structures. (a) Change in large-scale chromatin compaction using 3D DNA FISH at 11p15.5 and 11p15.1 loci after 5 h α-amanitin treatment to inhibit transcription. (b) Experimental approach for investigating changes in chromatin structure after transcription inhibition and recovery. (c) Western blot of global RNA polymerase using antibodies to differently activated forms of polymerase after transcription inhibition. GAPDH, loading control. (d) Incorporation of [3H]uridine into short (<200 nt) and long (>200 nt) RNA after 30 min pulse labeling to measure RNA synthesis after inhibition of transcription by α-amanitin followed by recovery. Data are mean ± s.d. (n = 3). (e) Compaction of 11p15.1 locus after transcription inhibition and drug washout. (f) Western blot of topoisomerase (topo) I and II proteins after topoisomerase RNAi. GAPDH, loading control. (g) Distance between pairs of fosmid probes at 11p15.1 and 11p15.5 loci after topoisomerase RNAi (neg., negative; topo, topoisomerase). (h) Distance between pairs of fosmid probes at the 11p15.5 locus after transcription inhibition by α-amanitin with or without topoisomerase inhibitors ICRF193 or camptothecin. Box plots and P values (Wilcoxon test) are described in Figure 4. n is number of separate probe pairs examined. In e and h, P values were calculated compared with control.

Figure 8 Relationship among transcription, DNA supercoiling and large-scale chromatin structures. Transcriptionally inactive chromatin is topologically overwound and has a cytologically compact large-scale chromatin structure. In contrast, a transcriptionally active region or transcriptional activation alters DNA supercoiling, remodeling supercoiling domains; this is accompanied by decompaction of large-scale chromatin structures. Therefore, large structural domains, for example as described by Hi-C1, are subdivided into smaller transcription-dependent supercoiling domains, providing an additional level of functional organization within the human genome.
test the relationship between changes in DNA supercoiling and large-scale chromatin structures, we ablated topoisomerases with RNA interference (RNAi; Fig. 7f). Cells transfected with short interfering RNA (siRNA) and with a fluorescent oligonucleotide were sorted using fluorescence-activated cell sorting onto slides (Supplementary Fig. 5a,b), and chromatin compaction was measured at the 11p15.1 and 11p15.5 loci. Depletion of topoisomerases for 48 h did not affect transcription (Supplementary Fig. 5c) but, as with α-amanitin treatment (Fig. 7a), the loci were compacted, showing that large-scale chromatin structures are also topoisomerase dependent (Fig. 7g).

To confirm this result, we inhibited topoisomerases using drugs that also promoted the compaction of large-scale chromatin structures (Fig. 7h). Treatment of cells with both α-amanitin and topoisomerase inhibitors abrogated any change in supercoiling (Fig. 3) or in large-scale chromatin structure (Fig. 7f), showing that compaction of these structures by transcription inhibition requires topoisomerase activity. Taken together, these data support a model that links transcription to the alteration of supercoiling domains and the subsequent decompaction of large-scale chromatin structures (Fig. 8).

**DISCUSSION**

**Formation and functional role of supercoiling domains**

This study shows that RNA polymerase and topoisomerase activities remodel DNA supercoiling, creating supercoiling domains (Figs. 1–4) and influencing the folding of large-scale chromatin structures (Figs. 6 and 7). As transcription is more efficient on a supercoiled template, genes located within underwound domains are activated more efficiently than those located in a less transcriptionally permissive chromatin environment. Supercoiling introduced at transcriptionally active loci could be transmitted to adjacent gene regions, creating an environment that promotes transcriptional activation of surrounding genes. This could provide an additional level of gene regulation and create an evolutionary constraint for maintaining genes with similar properties together.

It is unclear how genes initially adopt this underwound configuration. Some genes, such as HSP70, are in a transcriptionally poised state with the polymerase already engaged at the locus, whereas other genes are probably in a torsionally relaxed state that is less able to form an initiation complex. Our data suggest that DNA supercoiling is initially introduced at these loci when the initiating form of RNA polymerase generates short RNA transcripts, creating a favorable environment for the generation of a stable initiation complex. After transcription initiation, polymerases in collaboration with topoisomerases and helicases can then maintain the supercoiling state of the locus in a regulated manner. Genes such as pS2 (TFF1) cyclically recruit factors for transcription, and these bursts of transcriptional activity generate supercoiling that needs to be dissipated before the next round of transcription. This suggests that there must be mechanisms for monitoring the amount of DNA supercoiling, so recruitment and regulation of topoisomerases is essential to maintain a suitable level of torsion and ensure optimal transcription.

Our results also suggest that topoisomerases are actively required for remodeling DNA supercoiling, supercoiling domains (Figs. 3 and 5) and large-scale chromatin structures (Fig. 7). We were surprised to observe pronounced changes in supercoiling that occur after transcription inhibition and to a lesser extent in topoisomerase inhibition (Fig. 3a,b). As the DNA microarrays show relative bTMD binding, it is necessary to consider these changes relative to each other. It is therefore possible that most changes in supercoiling occur in the gene-rich underwound domains whereas the structure of the overwound domains remains relatively constant.

**Supercoiling domains and large-scale chromatin structures**

Polymerases introduce supercoils into DNA, increasing the free energy available in the chromatin fiber. Therefore, release of supercoils is thermodynamically favored: it releases energy that can be used by transcription, DNA replication and chromatin remodeling. Unconstrained DNA, which is found between nucleosomes, partially unwinds in response to negative supercoiling, whereas nucleosomal DNA behaves differently. As this DNA is tightly associated with nucleosomes, untwisting of the DNA causes unconstrained chromatin fibers to rotate, enabling superhelical turns to be transmitted through the domain. However, it is unknown how DNA supercoiling affects chromatin fiber structures; the fiber can probably absorb a certain amount of supercoiling but eventually becomes distorted, influencing its conformation.

We have shown earlier that gene-rich regions of the human genome are enriched in disrupted (open) chromatin fibers; however, these do not correspond directly to gene transcription, as inactive genes can be located in open domains and active genes can be located in compact domains. We have also shown that open chromatin can be propagated to adjacent regions and that disrupted chromatin from neighboring active genes has an additive effect. As there is a net change in supercoiling across chromosomal loci (Fig. 3c), we suggest that superhelical tension can be propagated from one region to another. This is supported by data, as supercoiling domains cover larger regions than the promoters of genes, suggesting that supercoils must be able to ‘flow’ along the chromatin fiber. This would explain why transcriptionally inactive genes can be found in open chromatin, if superwound genes are also active, and provides an explanation for the propagation of open chromatin states.

Earlier studies have shown that large-scale chromatin domains are ~150 kb in size; these might correspond to structures constrained by topoisomerases or by an attachment to an underlying matrix. Our study indicates that supercoiling domains are heterogeneous, are ~100 kb in size and are generated by polymerase and topoisomerase activities. We suggest that supercoiling domains also have a structural role: underwound domains have a decompacted large-scale chromatin structure and overwound domains have a compact structure, suggesting they affect higher levels of chromatin organization. Although supercoiling domains are formed by enzymatic activities, other proteins, such as HMG1A1, could be required to stabilize these structures. Alternatively, as an underlying nuclear matrix has not been well defined, domain boundaries may interact with each other, generating a ‘chromatin matrix.’ These dynamic interactions could stabilize the domains but simultaneously provide sufficient flexibility to allow domain remodeling as transcription profiles change.

The genome is organized into isochores with a mean size of 0.9 Mb (ref. 47). GC-rich isochores are gene rich and are also more negatively supercoiled than GC-poor regions, but what is the important determinant of the supercoiling state, genes or GC content? These two properties have evolved hand in glove, but as GC-rich sequences position nucleosomes more tightly, these strongly positioned nucleosomes create a fiber that is relatively irregular. This may in turn facilitate negative supercoiling, providing a permissive environment for genes and gene activation. In contrast, regions that bind nucleosomes less tightly enable nucleosomes to slide or move to evenly spaced positions, forming a regular chromatin fiber that is less readily supercoiled. Therefore, different isochores would be expected to have different properties, affecting the behavior of genes within them. As gene-rich (and GC-rich) regions of the genome are also more likely to contain chromosomal aberrations, and genes misregulated in cancer are more often found in GC-rich, open chromatin (ref. 49, data not shown), alterations in
DNA supercoiling and changes in supercoiling domains could influence genome stability. DNA supercoiling influences gene transcription and DNA replication in normal cells, but it may also be important in disease as many tumors overexpress topoisomerases. Therefore, topoisomerase misregulation might affect supercoiling, giving rise to chromosomal aberrations or a greater mutation rate, which could increase cellular heterogeneity and provide a driving force for tumor evolution.

**METHODS**

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

**Accession codes.** Gene Expression Omnibus database: GSE43451 (bTIMP binding, RNA polymerase chromatin immunoprecipitation (ChIP), topoisomerase ChIP and RNA expression).

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

**ACKNOWLEDGMENTS**

We thank J. Reek for help with bioinformatic analysis, F. Rustenburg and D. Israel for microarray hybridizations and H. Rasmahove, J. Allan, W. Bickmore, J. Caceres and N. Hastie for helpful discussions. This work was funded by the Wellcome Trust 078219/Z/05/Z (N.G.) and Breast Cancer (N.G.). N.G. is a recipient of a UK Medical Research Council senior fellowship (MR/J00913X/1).

**AUTHOR CONTRIBUTIONS**

C.N., S.C., M.B.R. and N.G. conceived, designed and interpreted experiments. C.N., S.C. and N.G. did biological experiments. N.A., I.K.M. and N.G. carried out chemical synthesis, and P.P. did microarray experiments. J.G.P., S.C. and N.G. did all bioinformatic analysis. N.G. supervised the project and all authors wrote the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nsmb.2509. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.
ONLINE METHODS

Cell culture. RPE1 cells (ATCC CRL-4000) were cultured in DMEM-F12, 3 mM glutamine and 15 mM HEPES supplemented with 0.34% sodium bicarbonate, 10% FBS, penicillin (100 U ml−1), streptomycin (100 µg ml−1) and phenol red (8.1 mg l−1). Transcription was blocked by adding flavopiridol (100 µM) or et-aminatin (50 µg ml−1) to cells for the times indicated. et-aminatin treatment was reversed by three PBS washes and addition of fresh medium. Topoisomerase I and II were inhibited using camptothecin (5 µM) or ICRF-193 (35 µM; Biomol), respectively.

Three-dimensional DNA fluorescence in situ hybridization. Cells were grown overnight on glass slides. Slides were rinsed with PBS and fixed in 4% paraformaldehyde for 10 min. Slides were rinsed with PBS and cells were permeabilized for 10 min on ice with PBS supplemented with 0.2% triton. After rinsing, slides were stored in 70% ethanol at 4 °C.

For processing, slides were dehydrated through an ethanol series and incubated with 2× SSC supplemented with 100 µg ml−1 RNase A (Invitrogen) at 37 °C for 60 min. Slides were then rinsed briefly with 2× SSC, dehydrated through an ethanol series and air dried. Slides were warmed by incubation in a 70 °C oven for 5 min before denaturation for 1 min in 70% formamide in 2× SSC, pH 7.5, at 70 °C. Slides were then transferred to 70% ethanol on ice, dehydrated through an ethanol series and air dried before overnight hybridization at 37 °C with pairs of probes (listed in Supplementary Table 2) labeled in digoxigenin-11-dUTP or biotin-16-dUTP. Some 150 ng of each labeled probe was hybridized with 5 µg salmon sperm and 10 µg human CotI DNA. Slides were washed four times for 3 min in 2× SSC at 45 °C and four times for 3 min in 0.1× SSC at 60 °C before being transferred to 4× SSC with 0.1% Tween 20 at room temperature. Digoxigenin-labeled probes were detected by using one layer of rhodamine-conjugated sheep anti-digoxigenin and a second layer of Texas red–conjugated anti-sheep (Vector Laboratories). Biotin-labeled probes were detected by using one layer of FITC-conjugated streptavidin followed by a layer of biotin-conjugated anti-avidin and a second layer of FITC-conjugated streptavidin (Vector Laboratories). Slides were counterstained with 0.5 µg ml−1 DAPI.

Image capture and analysis. Four-color 3D RNA-DNA FISH slides were imaged using a Hamamatsu Orca AG CCD camera (Hamamatsu Photonics) Zeiss Axiosplan II fluorescence microscope with Plan-Neofluor objectives, a 100-W Hg source (Carl Zeiss) and Chromo 86000v2 quadruple band-pass filter set (Chroma Technology) with the single excitation and emission filters installed in motorized filter wheels (Prior Scientific Instruments). Image capture and analysis were done using in-house scripts written for Iola Spectrum (Scanalytics). For three-color imaging, the microscope was the same but a Chroma 83000 triple-band-pass filter set was used with single excitation filters installed. For FISH, images were collected from at least 50 randomly selected nuclei for each experiment and then analyzed by using custom Iola scripts that calculate the distance between two probe signals. The significance of compartment between pairs of probes was tested using the nonparametric Wilcoxon test for paired samples, whereas P values between chromosomal loci were calculated by using the nonparametric Mann-Whitney U test (using R programming). P < 0.05 was considered significant.

RNA extraction and analysis. Total RNA (large RNAs (>200 nt) and small RNAs (<200 nt)) was extracted from cells using Tri Reagent (Sigma) and purified on silica matrices (miRNeasy kit). Residual DNA was removed by on-column DNase I (Roche) treatment. RNA size fractionation was monitored by analyzing RNA species on 10 or 15% TBE-urea gels (Invitrogen).

[3H]Uridine incorporation. Global transcription was determined by measuring [5–3H]uridine incorporation for a 30 min time period. Cold dA, dG, dC, dT and C (37 nM final) were added to cells to suppress label incorporation into DNA with 185 nM [5–3H]uridine. After 30 min incubation, cells were rinsed with PBS and RNA was extracted by selective binding to a silica matrix (miRNeasy Kit, Qiagen). Residual DNA was removed by on-column DNase I treatment. RNA was quantified using a Nanodrop and [3H] incorporation was measured by scintillation counting.

Western blotting. Cells were suspended in 2× SDS lysis buffer, incubated at 100 °C for 5 min and sonicated briefly. For RNA polymerase blots, protein samples were resolved on 8% bis-tris gels and transferred to PVDF membrane by wet transfer. For histone blots, protein samples were resolved on 12% bis-tris gels and transferred to PVDF membrane by semi-dry transfer. Membranes were probed with antibodies using standard techniques and detected by enhanced chemiluminescence. RNA polymerase II antibodies were H-224, 1:200 (Santa Cruz, sc-9001); 8WG16, 1:200 (Covance, MMS-126R); H5, 1:500 (Covance, MMS-129R); H14, 1:500 (Covance, MMS-134R); CTD418, 1:200,000 (Upstate 15-623) as described52. We also used antibody to GAPDH, 1:1,000 (Cell Signaling, 2118).

Chromatin immunoprecipitation. ChiP was done as described53 except that magnetic protein A beads (Invitrogen) were used for rabbit antibodies and protein G beads were used for mouse IgM antibodies. Antibodies were to topoisomerase I (Abcam, ab3825), topoisomerase IIR (Santa Cruz, sc-5347), topoisomerase IIβ (BD Transduction Laboratories, 611493) and total RNA polymerase II (Santa Cruz, sc-9001). All antibodies were characterized with western blots, and ChiP was optimized using quantitative PCR assays. For microarray hybridization, immunoprecipitated DNA was amplified using whole-genome amplification (Sigma).

Short interfering RNA transfections. Cells in six-well plates (for protein extraction) or on slides (for FISH analysis) were transfected with topoisomerasesII siRNA (ON-TARGETplus SMARTpool siRNA, Dharmacon) using Lipofectamine 2000 reagent (Invitrogen) for 4 h with a fluorescent marker oligo (BLOCK-iT fluorescent oligo, Invitrogen). siRNA sequences were TOP1, GAAAAGGGCUUCUAGUACGU, GAUUUCGGAUAUGAUGAU, GCAG AUCAUAUCACCCCA and CGAAGAGGUGAUGAGGCUP; TOP2A, CGAA AAGGAUGGUAGUACUA, GAUAGAACUGAGGCGUCA, GAGAAGA AUAUCAGUAUG and GGUGAUCUCUGAAAGUAUA; TOP2B, GAAGUUG GUUCCUGUGAGAGA, CGAAGACUCCACUAAUCAGCA, GAUCAGAUUG AACUGCA and GGUGUAGUAUGAGAUGAU; NEG, ON-TARGETplus nontargeting Pool D-001810-10-05. All siRNAs were used at 100 nM.

Biotinylated psoralen. BTMP was synthesized as described54 and purified by HPLC giving >95% by thin-layer chromatography and ELSD (evaporative light scattering). The structure was confirmed by 1H NMR.

Analyzing changes in DNA supercoiling. Cells or control genomic DNA were treated with various concentrations (50–500 µg ml−1) of bTMP for 20 min at room temperature in the dark. bTMP was UV cross-linked to DNA at 360 nm for 10 min. DNA was purified from cells using SDS and protease K digestion followed by phenol-chloroform-isoamyl alcohol extraction. DNA was fragmented by sonication (ten times for 20 s at 2 µm). Biotin incorporation into DNA was detected by dot blotting using alkaline phosphatase–conjugated avidin as a probe. The bTMP–DNA complex in TE was immunoprecipitated using avidin conjugated to magnetic beads for 2 h at room temperature and then overnight at 4 °C. Beads were washed sequentially for 5 min each at room temperature with TSEI (20 mM Tris, pH 8.1, 2 mM EDTA, 500 µM NaCl, 1% Triton X-100 and 0.1% SDS), TSEII (20 mM Tris, pH 8.1, 2 mM EDTA, 500 µM NaCl, 1% Triton X-100 and 0.1% SDS) and buffer III (10 mM Tris, pH 8.1, 0.25 M LiCl, 1 mM EDTA, 1% NP40 and 1% deoxycholate). Beads were then washed twice with TE buffer for 5 min. To extract DNA and to release psoralen adducts, the samples were boiled for 10 min at 90 °C in 50 µl of 95% formamide with 10 mM EDTA. Samples were then made up to 200 µl with water, and the DNA was purified using a Qiagen PCR purification kit. To release torsional stress in the chromatin fiber, the DNA was
cleaved by treatment of cells with 100 µM bleomycin (Sigma B2434) for 10 min at 37 °C. DNA was extracted and DNA damage was analyzed by nondenaturing gel electrophoresis.

**Biotinylated trimethyl psoralen fluorescence detection.** Cells were grown overnight on glass slides, rinsed with 3x PBS and then treated with bTMP (1 mg ml⁻¹) for 20 min at room temperature in the dark. bTMP was UV cross-linked to the DNA at 360 nm for 10 min. Slides were rinsed three times with PBS and fixed in 4% paraformaldehyde. Slides were rinsed three times with PBS, and cells were permeabilized for 10 min with PBS supplemented with 0.2% Triton. After rinsing with PBS, slides were blocked 5% horse serum in PBS for 15 min at room temperature in a humidified chamber. bTMP was detected by incubation overnight with NeutrAvidin FITC (25 µg ml⁻¹; Invitrogen).

**Microarray hybridization, data processing and analysis.** See Supplementary Note.

52. Stock, J.K. *et al.* Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat. Cell Biol.* 9, 1428–1435 (2007).
53. Naughton, C. *et al.* Progressive loss of estrogen receptor alpha cofactor recruitment in endocrine resistance. *Mol. Endocrinol.* 21, 2615–2626 (2007).