DNA conformation regulates gene expression: the MYC promoter and beyond

Olga Zaytseva^{1,2} and Leonie M. Quinn^{1,2}

¹ACRF Department of Cancer Biology and Therapeutics, The John Curtin School of Medical Research, The Australian National University, Canberra City ACT 2600, AUSTRALIA.

²School of Biomedical Sciences, University of Melbourne, Parkville 3010, AUSTRALIA.

Corresponding Author:

Olga Zaytseva

E-mail: olga.zaytseva@anu.edu.au

Keywords: *MYC* transcription, non-B-DNA, ssDNA/RNA-binding proteins

Abbreviations: Top I/II, topoisomerase; RNA Pol II, RNA Polymerase II; FUSE, Far Upstream Sequence Element; ChIP, Chromatin Immunoprecipitation; NHE, nuclease hypersensitive element

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/bies.201700235

This article is protected by copyright. All rights reserved.

Abstract

Emerging evidence suggests that DNA topology plays an instructive role in cell fate control through regulation of gene expression. Transcription produces torsional stress, and the resultant supercoiling of the DNA molecule generates an array of secondary structures. In turn, local DNA architecture is harnessed by the cell, acting within sensory feedback mechanisms to mediate transcriptional output. *MYC* is a potent oncogene, which is upregulated in the majority of cancers; thus numerous studies have focused on detailed understanding of its regulation. Dissection of regulatory regions within the *MYC* promoter provided the first hint that intimate feedback between DNA topology and associated DNA remodeling proteins is critical for moderating transcription. As evidence of such regulation is also found in the context of many other genes, here we expand on the prototypical example of the *MYC* promoter, and also explore DNA architecture in a genome-wide context as a global mechanism of transcriptional control.

1. Introduction

Accurate expression of genetic material carried by a cell's DNA is essential for correct function of individual cells and development of all multicellular organisms. Despite gene expression requiring a reciprocal interplay between DNA and proteins, the weight of studies have focused on proteins binding DNA as the masters of transcriptional control. However, characterization of protein binding sites and their intra-regulation, although significant, does not comprehensively portray the nuances of gene control. Despite this, the role of DNA structure has been largely overlooked, with DNA viewed as playing primarily a passive role as a mere scaffold for protein complex assembly. While it has long been known that DNA can adopt complex three-dimensional structures,^[1] the role of a given DNA conformation in cell function has remained elusive. Pioneering studies, that somewhat went against the grain, revealed that conformational changes in DNA structure within the *MYC* promoter were critical for modulating transcription of this significant oncogene.^[2-4] With more recent advances in genome-wide profiling of DNA architecture,^[5-8] the instructive role of DNA structure in gene regulation can no longer be ignored.

In the normal cellular context, the majority of DNA is found in B-form, a right-handed double helix, where each turn comprises 10-10.5 bases per rotation. However, the presence of high salt, ethanol, or other solvents can induce an altered orientation of bases, resulting in the A-

DNA or Z-DNA forms. These alternative DNA forms have also been described *in vivo*, indicating potential roles in living cells. For example, A-DNA exists at the sites of interaction with transcription factors,^[9] including the TATA-binding protein (TBP), which enables bending of the DNA molecule.^[10] Z-DNA, a left-handed helix, tends to be formed by sequences with alternating purine and pyrimidine bases, such as GCGCGC motifs found at promoter regions ^[11,12]. Z-DNA has also been detected at nucleosomes as a result of SWI/SNF remodeling activity.^[13] Additional non-B-DNA secondary structures can form as a result of base pairing of opposite DNA strands not occurring sequentially; these include melted single-stranded DNA, hairpin, triplex and quadruplex structures.^[14] Importantly, we are beginning to understand the roles of such non-canonical structures in the normal cellular context, and association between aberrant DNA conformation, genome instability and development of disease.^[15,16] Here, we discuss the involvement of non-B-DNA in regulation of gene expression, with the paradigm of the *MYC* promoter providing a basis for integrating the emerging genome wide studies.

2. MYC is a potent oncogene requiring precise transcriptional regulation

Since its discovery, MYC has featured prominently in cancer biology, and has been detected at elevated levels in nearly all cancer types.^[17-19] Unlike most oncogenic alterations, which occur through an activating mutation, MYC merely requires increased abundance for tumorigenic activity. For example, one of the earliest oncogenic mechanisms, found to drive elevated MYC in Burkitt's lymphoma, involves a translocation of Chromosome 8 that results in fusion of the immunoglobulin locus upstream of the wild type *MYC* coding region.^[20] Subsequent studies demonstrated that elevated MYC promotes tumourigenesis by activating genes required for cell growth, including protein synthesis^[21-24] and cell proliferation programs.^[25,26] During normal development, most signaling pathways converge on the *MYC* promoter to regulate *MYC* expression and pattern tissue growth, thus dissection of *MYC* regulation has provided a valuable model for studying the mechanisms that accomplish complex integration of signals to precisely fine-tune gene expression.^[17,27,28]

The sheer number and variety of MYC targets has sparked intense interest in both regulatory mechanisms and function. Detailed analysis of MYC function have been reviewed comprehensively elsewhere,^[29-31] and although a complete discussion is beyond the scope of this review, we provide some brief background. MYC is a member of the basic helix-loop-helix (bHLH) protein family, which heterodimerizes with its binding partner MAX to stimulate transcription.^[32,33] Significantly, recent genome-wide Chromatin

Immunoprecipitation (ChIP)-sequencing studies have revealed that when elevated, MYC loads onto promoters, enhancers and regions of open chromatin structure i.e. transcriptionally active sites.^[34-37] Thus, in the context of increased MYC abundance, as in cancer, MYC acts as a broad transcriptional amplifier rather than a sequence specific transcription factor.

3. Transcription applies force and generates supercoiling of DNA

Early crystallographic structures indicated that nucleosomes interact with B-DNA, generating stable complexes for secondary DNA structure and compacting the lengthy DNA molecule into the nucleus.^[38] Regulatory DNA regions in human cells enriched for high GC content tend to display high nucleosome occupancy, suggesting intrinsic binding affinity encoded into the DNA sequence, possibly the first step in subduing gene expression.^[39] However, nucleosome structure has to be remodeled at promoters, in order to initiate transcription and permit entry of sequence-specific binding proteins and the RNA Polymerase II (RNA Pol II) machinery. Pioneer factors initiate transcription by recognizing DNA bound to nucleosomes, driving nucleosome displacement to enable interaction between the DNA strand and core RNA Pol II machinery.^[40,41] Assembly of the RNA Pol II transcriptional complex modifies the DNA structure further, enabling binding of the general transcription factors including TFIID, which contains the helicase subunits required to melt DNA. Thus, the single-stranded transcription bubble necessary for binding of the nascent mRNA nucleotides emerges.^[42]

As mRNA synthesis proceeds, RNA Pol II requires access to distal base pairs and further unwinding of the DNA helix. Compared with the width of the DNA molecule the protein component of the transcription pre-initiation complex (PIC) is massive; in addition to RNA Pol II it comprises more than 100 proteins, with six general transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, and TFIIF) and regulatory complexes (such as the 30 subunit Mediator coactivator complex and chromatin remodeling machinery).^[43-52] Thus, rather than the protein complex moving along the length of the DNA molecule, a more tenable scenario would be the transcription complex maintaining a relatively constant position and the DNA molecule actively feeding through the protein interface.^[49-53] Naturally, as a result of PIC formation and early transcription, torsional strain is applied to the DNA molecule to result in DNA supercoiling and formation of various non-B DNA structures.

Strain on the DNA molecule in the same direction as the helix is referred to as positive supercoiling, winding in the opposite direction produces negative supercoils.^[54-56] Furthermore, studies from the late 70s demonstrated that negative supercoiling, or twisting

DNA against the helical direction, can melt the DNA duplex giving rise to single-stranded DNA.^[57] The outcome is not only relief of torsion, but the energy released further drives helix opening, which *in vivo* could be envisaged to grant the transcriptional machinery access to the coding strand. Despite these significant early advances, the focus on DNA topology was somewhat averted with the advent of DNA sequencing technology. Indeed, with the focus turning more to enzymatic processes, the dogma shifted and DNA architectural studies were relegated to the realm of interesting *in vitro* exercises with little physiological significance. Moreover, given the observation that Topoisomerase (Top I and Top II) enzymes were found to relax both positive and negative supercoils,^[58-61] it was thought that DNA stress should not be a feature of normal cells.

However, the changes in DNA structure in a living cell occur far too rapidly for topoisomerases to detect and completely remove all supercoils. With the introduction of new technologies it has become apparent that supercoiling exists in cells and plays key regulatory roles, thus shifting the focus back to structural roles played by DNA in gene regulation.^[62,63] Improved *in vitro* single-molecule techniques such as optical and magnetic tweezers have allowed quantification of the forces and torque generated by transcription.^[64,65] Furthermore, introduction of a defined amount of positive or negative torsion into DNA minicircles and imaging by cryo-EM has revealed the diversity of structures formed due to the ability of DNA to conduct force throughout its length.^[66] Given that these processes occur *in vivo*, a rational conclusion is that the energy generated by DNA supercoiling has the potential to instruct gene regulation.^[67] In fact, mapping of supercoiling across the human genome has revealed that supercoiled domains are organized into discrete units bounded by insulators, providing a structural basis for co-expression of genes located in close proximity.^[68]

The rigidity of B-DNA confers capacity to transfer torsional stress along the length of the molecule. Naturally, the act of transcription significantly changes DNA architecture, and can cause alterations extending up to 2kb.^[62] The twin-supercoiled-domain models developed in the late 1980s detailed the effect of transcription on DNA: positive supercoiling was generated in front of the transcription bubble, and negative supercoiling behind^[69,70] (Figure 1). The modification of DNA conformation, in response to the torsional stress of transcription, results in local melting of B-DNA, non-canonical interactions between base pairs and generation of non-B-DNA secondary structures. Some of the first indications that transcriptional activation could remodel DNA structure to modulate rates of transcription in a physiological setting came from studies of the human *MYC* promoter.^[2,3,71]

4. DNA structure and supercoiling regulate *MYC* transcription

MYC provides the best-characterized promoter in regard to mechanism of gene expression control dependent on modulation of DNA structure via supercoiling. Early studies mapping single-stranded elements by permanganate sensitivity revealed a region of the *MYC* gene hyperreactive with this unpaired-base selective reagent. This region 1.7kb upstream of the active *MYC* promoter, called the Far Upstream Sequence Element (FUSE), melts into single-stranded DNA as a result of transcriptionally generated negative supercoiling.^[3] At the FUSE, local enrichment of A-T bases facilitates conversion to single-stranded DNA, and the capacity to create sequence-dependent sites of DNA opening, providing conformational flexibility compared with double-stranded DNA.

Mapping of the *MYC* promoter, using constructs containing regions of sequence upstream of the *MYC* coding region, identified two regulatory sites necessary for downregulation of *MYC* upon differentiation.^[72] The identified distal region contains the FUSE, while the proximal site corresponds to the nuclease hypersensitive element (NHE) III₁, also called the CT element as it contains five imperfect repeats of a CCCTCCCC sequence. The ability to form multiple non-B-DNA structures enables this region to play significant roles in regulating *MYC* transcription.^[73]

The NHE III₁ can dissociate into single-stranded DNA, permitting digestion by the single-strand specific S1 nuclease and the DNase I nuclease.^[71] The nuclease sensitivity is only evident in cell lines expressing *MYC*, indicating that non-B-DNA structure plays a role in *MYC* activation.^[71] Furthermore, owing to the underlying homopurine-homopyridine mirror repeat in the underlying sequence of the NHE III₁, negative supercoiling at the promoter can induce a triplex or H-DNA structure.^[74,75] Formation of the triplex results in a sharp bend of the DNA molecule (illustrated in Figure 1), which can act as a steric block to prevent rotation of DNA relative to RNA Pol II. The observation that transcription is unhindered when the NHE III₁ sequence is mutated to prevent triplex formation, confirms this bending is important for *MYC* regulation.^[75] In addition, H-DNA is mutagenic, promoting formation of chromosomal breakpoints such as those observed in Burkitt's lymphoma.^[76,77]

The underlying GC-rich sequence of the NHE III₁ enables formation of additional DNA secondary structures at this site. Four adjacent guanine bases can form a square planar structure called a G-tetrad^[78], while the C-rich i-motif on the complementary strand can form

an i-tetraplex^[79,80] (Figure 1). The G-quadruplex is a stable 3-dimensional DNA structure formed by multiple G-tetrads arranged in a stack. The mammalian *MYC* promoter harbors one of the most well-studied G-quadruplex structures, located in the Guanine-rich nuclease hypersensitive element (NHE) III₁ region.^[79,81,82] Although supercoiling present in circular plasmids is not sufficient for G-quadruplex formation,^[83] *in vitro* studies have demonstrated that negative supercoiling generated by upstream transcription can promote formation of G-quadruplexes thousands of base pairs away, illustrating the ability of this DNA structure to convey considerable mechanical force.^[84] G-quadruplex structures can also influence promoter-proximal Pol II pausing. For example, stabilization of the G-quadruplex, by treatment of lymphoma cell lines with oligonucleotides targeting the C-rich strand, results in reduced proliferation and decreased *MYC* expression.^[85] The i-motif structure was originally characterized *in vitro*, due to the highly acidic conditions required for formation.^[80] However, in the presence of negative supercoiling, stabilization of the i-motif is possible under physiological pH, further highlighting the importance of torsional stress to DNA conformation.^[86] Thus the NHE III₁ region exists in a dynamic equilibrium between double-stranded DNA and supercoiling-induced secondary structures that include the G-quadruplex, i-motif, H-DNA and single stranded DNA, together providing a regulatory feedback on transcriptional output of the *MYC* promoter.

5. Supercoiling plays genome-wide roles in transcriptional regulation

Recent genome-wide studies indicate that supercoiling is expected to modulate transcription of many, if not all, genes. As outlined above, coiling of DNA and the energy generated as a result of promoter unwinding can be harnessed as a productive force in transcription.^[67] When unresolved by the action of topoisomerases, positive DNA supercoiling that accumulates ahead of the transcriptional complex increases the rigidity of DNA, reducing the spacing between turns and acting as a block to transcription. Indeed, studies in human colorectal cancer cell lines deficient for Top I indicate that topoisomerase is not active during RNA Pol II initiation.^[87] This allows transcription to proceed, but become increasingly stalled proximal to the promoter. Pause release requires activity of numerous factors, including TEFb and BRD4, which phosphorylate transcriptional machinery and as a result promote Top I activation. The resolution of accumulated tension allows RNA Pol II progression and productive transcriptional elongation. In *Drosophila* cells, topoisomerase inhibition has similarly been found to stall RNA Pol II downstream of the transcription start site; however the ensuing elevation of torsional strain facilitates nucleosome destabilization to increase elongation.^[88] Nucleosomal remodeling by the SWI/SNF complex also changes local DNA

structure and further twists the DNA; the resultant force is harnessed to shift nucleosome positions.^[89]

Negative supercoiling is also significant for regulation of transcription. DNA under negative supercoiling tension is more likely to melt into single-stranded DNA. It is therefore more flexible and has a higher capacity to bend and form non-B-DNA structures. Genes that are highly transcribed, such as the ribosomal RNA genes, display negative supercoiling.^[90] On the other hand, excessive negative supercoiling, unresolved by topoisomerase, also acts to dampen RNA Pol II transcriptional activity.^[91] Thus, the amount of torsional force that accumulates as a result of transcription, acting as a negative feedback regulatory mechanism, maintains transcriptional output within a narrow range. Recent mathematical and physical models, which account for transcriptionally generated supercoiling and neighbourhood gene activity, similarly predict a dynamic regulatory feedback.^[92,93]

6. Non-B-DNA is found throughout the genome

Recently, permanganate mapping has revealed single-stranded DNA and other non-B-DNA structures across the genome.^[5] This method effectively detected the predicted single-stranded DNA generated at the transcription bubble. Additionally, non-B-DNA was detected upstream in enhancer regions of transcriptionally active genes. Comparison with RNA Pol II loading by ChIP-sequencing (ChIP-seq) revealed that the quantity of non-B-DNA correlated with gene expression, indicating that negative supercoiling likely forms as a result of transcriptional activity. Similarly, ChIP-seq using a subunit of human double-stranded RNA adenosine deaminase (hADAR1), which recognizes Z-DNA, identified enrichment at active genes occupied by RNA Pol II.^[7] Analysis of DNA sequence motifs failed to detect enrichment for consensus sequences predicted to form non-B-DNA structures.^[5] Gene Ontology analysis indicated that non-B-DNA is frequently detected upstream of developmentally regulated genes, suggesting that expression of these genes requires additional layers of fine-tuning by noncanonical DNA structures, perhaps as a mechanism to integrate multiple signaling inputs.

Thousands of sequences with predicted capacity to form G-quadruplexes have been identified throughout the genome. Using oligonucleotide probes, sequences similar to the *MYC* G-quadruplex were observed at multiple promoters, indicating broader regulatory roles in transcription.^[85] Moreover, ChIP-seq analysis using an antibody that recognizes G-quadruplex structure revealed enrichment of peaks at nucleosome-depleted regions.^[6] Interestingly, the sequences with predicted capacity to form G-quadruplexes are particularly enriched at

promoters of oncogenes, while tumour suppressor promoters are underrepresented, indicating that gene function is linked with local genomic structure.^[94] On the *MYC* promoter, the G-quadruplex plays inhibitory roles when stabilized. Several factors, including the transcription factor and master tumour suppressor TP53^[95] and Nucleolin (the nucleolar factor implicated in chromatin decondensation required for rRNA transcription^[96]) can bind to the G-quadruplex to stabilize its structure and negatively regulate transcription.

7. Single-stranded nucleic acid binding proteins regulate conformation of non-B-DNA at the *MYC* promoter

Both the FUSE and the NHE III₁ element form a complex feedback system to modulate *MYC* transcription through interactions with single-stranded DNA binding proteins. Although the nuclease-sensitive FUSE element is essential for proper *MYC* transcription, as deletion impairs *MYC*-reporter activity, insertion of FUSE region alone without neighbouring regulatory elements is insufficient to stimulate *MYC*.^[72] This was the first indication that the FUSE was not a traditional enhancer, but likely provided a non-canonical mechanism for controlling transcription. The most abundant protein associated with FUSE (revealed by affinity purification) was Fuse Binding Protein (initially FBP, now FUBP1), which preferentially binds the single-stranded noncoding strand of the FUSE.^[97] By simultaneously binding FUSE and the XPB helicase subunit of the general transcription factor TFIID complex, FUBP1 is able to generate a specific DNA conformation and form a loop between FUSE and RNA Pol II at the transcription start site within the *MYC* promoter.^[98]

Subsequently another protein, FBP-Interacting Repressor (FIR), was isolated on the basis of its capacity to interact with FUBP1.^[99] When bound to FUSE, FUBP1 facilitates recruitment of FIR^[100], and upon binding FIR similarly interacts with the XPB helicase subunit of TFIID, strengthening the DNA loop conformation.^[101] However, binding of FIR promotes the eviction of FUBP1, and repression of *MYC* expression.^[100,101] Thus, FUBP1 and FIR are predicted to act antagonistically with non-B-DNA to provide a “cruise control” on the *MYC* promoter (Figure 2).

More recent studies revealed that this mechanism of fine-tuning *MYC* transcription is conserved in invertebrates. Indeed, studies in the *Drosophila* model have not only confirmed the existence of a FUBP1-FIR-XPB/TFIID network, but have extended our understanding of single-stranded DNA binding proteins during *MYC* promoter regulation. The *Drosophila* ortholog of FUBP1, called Psi, is required for transcriptional initiation and elongation of RNA

Pol II across the *MYC* gene. Additionally, Psi interacts with the Mediator transcriptional complex, likely playing a role to integrate multiple signaling inputs on the *MYC* promoter.^[102] The ortholog of FIR, encoded by the *Half pint (Hfp)* gene in *Drosophila*, is required for repression of *Drosophila MYC*.^[103] Furthermore, interaction between Hfp and Haywire (the ortholog of XPB subunit of TFIIH) is essential for repression of *MYC* and patterning of cell and tissue growth. In contrast, mutants with a deleted C-terminal domain of Haywire are unable to restrict *MYC*-dependent overgrowth.^[104] These *in vivo* studies have illuminated connections between DNA topology and developmental growth signaling pathways; however, further studies in *Drosophila* are required to map the DNA topology across the *MYC* promoter and genome-wide.

Melting of DNA in the NHE III₁ region of the *MYC* promoter provides a scaffold for numerous proteins that promote *MYC* transcriptional modulation. While in the double stranded conformation, the NHE III₁ contains multiple binding sites for the Sp1 and Sp3 transcription factors, although only Sp1 acts as a transcriptional activator.^[105] Following introduction of negative supercoiling by transcription, hnRNPK is recruited to newly formed open regions of pyrimidine-rich single-stranded DNA, thus activating *MYC* expression.^[106,107] Furthermore, hnRNPK facilitates binding of Cellular nucleic acid protein (CNBP) to the opposite DNA strand, further stabilizing open DNA at this site and promoting *MYC* transcription.^[71,108]

8. Single-stranded DNA binding proteins also have roles in RNA processing

Interestingly, RNA processing functions have also been reported for both FUBP1 and FIR. FUBP1 has been implicated in splicing and translational control; by binding the 3'UTR of mRNAs for *GAP-43*,^[109] *nucleophosmin*,^[110] *p27*,^[111] *DMD*^[112] and the RNA of hepatitis C virus.^[113] PUF60, which is differentially spliced to form FIR, plays more general roles in splicing by associating with the spliceosome and cooperating with U2AF to promote removal of introns with weak 3' splice sites.^[114,115] The FUBP1 protein contains K-homology domains, possessing sequence similarity with the RNA-binding domains of the hnRNPK protein, which also has the ability to interact with single-stranded DNA.^[116] Like FUBP1, hnRNPK plays roles in both RNA biology and transcription.^[2] *In vitro* studies have also demonstrated that hnRNPK and hnRNPA1 proteins interact with the NHE III₁ region 100-150 base pairs upstream of the *MYC* transcription start site, and binding via the CT-sequence element is required for promoter activation.^[117] Further support for transcriptional roles have been provided by overexpression studies, which revealed physical interaction between

hnRNPK and TFIID.^[107] hnRNPK can also activate reporter constructs downstream of the CT-element *in vitro*, but can only do so for circular not linear plasmids, indicating that the negative supercoiling generated in circular plasmids is required to expose hnRNPK's single-stranded DNA binding target.^[2]

Thus, these single-stranded nucleic acid binding proteins elicit great flexibility, depending on their interacting partners: they can behave as either RNA-processing proteins or bind single-stranded DNA to modulate transcription. This raises the possibility that competing interactions between RNA and single-stranded DNA could provide these proteins with the capacity to act as sensors of the cellular state. For example, decreased total RNA could release these single-stranded nucleic acid binding proteins for interaction with single-stranded DNA at promoters, thus establishing a feedback loop through activation with genes such as *MYC*.

The versatility of single-stranded nucleic acid binding domains raises the question, is there evidence that proteins characterized for RNA binding are involved more broadly in transcription? Non-B-DNA structures are detected throughout the genome; hence it is logical that single-stranded nucleic acid binding proteins might play similar roles in regulating genes other than *MYC*. Indeed, purification of proteins bound to single-stranded DNA sequences upstream of the *BRCA1* gene have identified a number of interactors, including FUBP1 and hnRNP, and additional single-stranded nucleic acid binding proteins PTBP, DAZ AP, TIA-1 and PCBP.^[118] Thus, regulation of *BRCA1* transcription likely also harnesses DNA supercoiling by binding factors that recognize, and possibly modify, promoter DNA topology.

There are further examples of proteins with a variety of functions in both single-stranded DNA and RNA biology. Recently, additional members of the hnRNP RNA-binding protein family have also been implicated in promoter binding. Namely, hnRNP LL was observed to interact with the single-stranded DNA i-motif of the *BLC2* promoter sequence to stimulate transcription in a DNA conformation-dependent manner.^[119] In HeLa cells, hnRNP RALY can bind chromatin, via both RNA-dependent and independent mechanisms, to regulate cell cycle signatures and proliferation.^[120] Moreover, the purine-rich element binding protein (Pur) family is a highly conserved group of proteins with a variety of functions, including regulation of RNA transport and translation, and transcriptional regulation of genes including *MYC* through binding and remodeling of single stranded DNA.^[121-124] In *Drosophila*, regulation of *MYC* transcription by the RNA-binding protein B52 has been reported,^[125]

although it is unclear whether this occurs as a result of direct interaction with single-stranded DNA in the promoter. The SSBP family of single-stranded DNA binding proteins also modulates transcription, for example, Sspb3 activates transcription of trophoblast cell fate genes during mouse development.^[126,127]

9. Does non-B-DNA regulate transcription by impacting nuclear organization?

Many components of the transcriptional machinery consist of intrinsically disordered proteins, suggesting inherently flexible interactions and the possibility of promiscuous binding.^[128] For example, the MYC protein possesses a wide variety of binding partners^[129] containing disordered regions, thus enabling the MYC-complex to adopt context-specific conformations,^[130] likely a feature key for MYC's global amplifier capacity. Although these interactions have been primarily studied in the context of protein-protein interactions, transition from rigid B-DNA to flexible non-B-DNA conformation would also increase versatility of binding, providing further nuances to transcriptional regulation.

Interestingly, computational simulations of transcriptionally generated supercoiling in fission yeast induces organization of self-interacting domains; that these are also observed in living cells, highlights the significance of dynamic DNA architecture.^[131] Moreover, regions of highly active transcription have recently been found to segregate within the nucleus, forming phase-separation domains.^[132] It is possible that intrinsically disordered proteins may also contribute to phase separation by interacting with unwound DNA at sites of highly active transcription. Nevertheless, remodeling of DNA requires highly ordered nucleic acid binding domains, such as those found in many single-stranded DNA/RNA binding proteins.^[133] Thus altering the structure of DNA, via single-stranded DNA binding proteins, could regulate gene expression by facilitating interaction between highly transcribed regions through DNA bending. On the other hand, stabilization of B-DNA, through torsion and supercoiling, might lead to increased rigidity and prevent generation of the conformation required for phase separation. Further studies to address these questions are awaited with great interest.

10. Conclusions and Outlook

With improved computational analysis, advanced sequencing and imaging technology, we will undoubtedly augment our understanding of the interplay between DNA and protein components during transcriptional control. However, generation of models to investigate mechanisms of gene control will also require careful consideration of three-dimensional orientation and the forces generated by all cellular activities.

Evolutionary perspective of non-B-DNA as a regulatory mechanism can be derived from the FUBP family, an ancient group of proteins with conserved roles in DNA and RNA biology. Given the prevailing RNA world hypothesis we envisage a likely scenario, where organisms repurposed many single-stranded nucleic acid binding proteins during the switch from an RNA to the DNA world. This adaptation makes evolutionary sense, i.e. rather than make an entirely new set of proteins, the move to DNA-based life would be expedited by repurposing existing RNA binding factors with the capacity to remodel single-stranded DNA. Therefore, we predict that one key adaption of proteins, including KH-domain proteins, was the capacity to modulate single-stranded DNA structure at key promoters to control transcription. Given the prevalence of single-stranded and non-B-DNA upstream of active promoters, we envisage a burgeoning field, as we further unveil the capacity of the cell to harness alternate DNA structures, and forces transmitted by the DNA molecule, to fine-tune gene expression and control animal development.

Acknowledgements – We thank David Levens for valuable feedback on the manuscript.

The authors declare no conflict of interest

Figure Legends

Figure 1. Supercoiling generated by transcription impacts DNA topology and provides feedback mechanisms. As a result of transcription, positive supercoiling (+) is generated in front of RNA polymerase, inhibiting its progression and impacting DNA conformation to destabilize nucleosomes. In order to stimulate transcription, positive supercoiling is resolved by topoisomerase. Negative supercoiling (-) is generated behind the transcription bubble, promoting formation of non-B-DNA structures, which play regulatory roles to modulate transcription.

Figure 2. Non-B-DNA-interacting proteins modulate MYC transcription. The FUSE and the NHEIII₁ elements are regulatory structures that change their conformation based on context. Basal level of transcription and supercoiling promote single-stranded DNA formation and binding by FUBP1, hnRNP and CNBP proteins, which stimulate transcription of *MYC*. Stabilization of the G-quadruplex at the NHEIII₁ element by proteins such as Nucleolin, and

recruitment of FIR by further remodeling of the promoter act to downregulate expression of *MYC*.

References

- [1] R. D. Wells, R. W. Blakesley, S. C. Hardies, G. T. Horn, J. E. Larson, E. Selsing, J. F. Burd, H. W. Chan, J. B. Dodgson, K. F. Jensen, I. F. Nes, R. M. Wartell, *CRC Crit. Rev. Biochem.* **1977**, 4, 305.
- [2] T. Tomonaga, D. Levens, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 5830.
- [3] M. I. Avigan, B. Strober, D. Levens, *J. Biol. Chem.* **1990**, 265, 18538.
- [4] L. Bazar, D. Meighen, V. Harris, R. Duncan, D. Levens, M. Avigan, *J. Biol. Chem.* **1995**, 270, 8241.
- [5] F. Kouzine, D. Wojtowicz, L. Baranello, A. Yamane, S. Nelson, W. Resch, K.-R. Kieffer-Kwon, C. J. Benham, R. Casellas, T. M. Przytycka, D. Levens, *Cell Syst* **2017**, 4, 344.
- [6] R. Hänsel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill, S. Balasubramanian, *Nat. Genet.* **2016**, 48, 1267.
- [7] S.-I. Shin, S. Ham, J. Park, S. H. Seo, C. H. Lim, H. Jeon, J. Huh, T.-Y. Roh, *DNA Research* **2016**, 23, 477.
- [8] F. Kouzine, D. Wojtowicz, A. Yamane, W. Resch, K.-R. Kieffer-Kwon, R. Bandle, S. Nelson, H. Nakahashi, P. Awasthi, L. Feigenbaum, H. Menoni, J. Hoeijmakers, W. Vermeulen, H. Ge, T. M. Przytycka, D. Levens, R. Casellas, *Cell* **2013**, 153, 988.
- [9] X. J. Lu, Z. Shakked, W. K. Olson, *Journal of Molecular Biology* **2000**, 300, 819.
- [10] Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *Nature* **1993**, 365, 512.
- [11] A. Herbert, A. Rich, *J. Biol. Chem.* **1996**, 271, 11595.
- [12] A. R. Rahmouni, R. D. Wells, *Science* **1989**, 246, 358.
- [13] N. Mulholland, Y. Xu, H. Sugiyama, K. Zhao, *Cell Biosci* **2012**, 2, 3.
- [14] M. Kaushik, S. Kaushik, K. Roy, A. Singh, S. Mahendru, M. Kumar, S. Chaudhary, S. Ahmed, S. Kukreti, *Biochem Biophys Rep* **2016**, 5, 388.
- [15] G. Wang, K. M. Vasquez, *DNA Repair* **2014**, 19, 143.
- [16] J. Zhao, A. Bacolla, G. Wang, K. M. Vasquez, *Cell. Mol. Life Sci.* **2010**, 67, 43.
- [17] D. Levens, *Genes & Cancer* **2010**, 1, 547.
- [18] M. Gabay, Y. Li, D. W. Felsher, *Cold Spring Harbor Perspectives in Medicine* **2014**, 4, DOI 10.1101/cshperspect.a014241.
- [19] C. V. Dang, *Cell* **2012**, 149, 22.
- [20] C. A. Spencer, M. Groudine, *Ann. N. Y. Acad. Sci.* **1990**, 599, 12.
- [21] C. Grandori, N. Gomez-Roman, Z. A. Felton-Edkins, C. Ngouenet, D. A. Galloway, R. N. Eisenman, R. J. White, *Nat Cell Biol* **2005**, 7, 311.
- [22] G. Poortinga, M. Wall, E. Sanij, K. Siwicki, J. Ellul, D. Brown, T. P. Holloway, R. D. Hannan, G. A. McArthur, *Nucleic Acids Res.* **2011**, 39, 3267.
- [23] S. S. Grewal, L. Li, A. Orian, R. N. Eisenman, B. A. Edgar, *Nat Cell Biol* **2005**, 7, 295.
- [24] G. Poortinga, K. M. Hannan, H. Snelling, C. R. Walkley, A. Jenkins, K. Sharkey, M. Wall, Y. Brandenburger, M. Palatsides, R. B. Pearson, G. A. McArthur, R. D. Hannan, *EMBO J.* **2004**, 23, 3325.
- [25] G. Bretones, M. D. Delgado, J. León, *Biochim. Biophys. Acta* **2015**, 1849, 506.
- [26] M. Shichiri, K. D. Hanson, J. M. Sedivy, *Cell Growth Differ.* **1993**, 4, 93.

- [27] J. Er Amanda Lee, L. May Parsons, L. M Quinn, *AIMS Genetics* **2014**, *1*, 81.
- [28] I. Wierstra, J. Alves, *Adv. Cancer Res.* **2008**, *99*, 113.
- [29] E. Wolf, C. Y. Lin, M. Eilers, D. L. Levens, *Trends Cell Biol.* **2015**, *25*, 241.
- [30] C. V. Dang, *Cold Spring Harbor Perspectives in Medicine* **2013**, *3*, DOI 10.1101/cshperspect.a014217.
- [31] C. J. Poole, J. van Riggelen, *Genes (Basel)* **2017**, *8*, 142.
- [32] E. M. Blackwood, R. N. Eisenman, *Science* **1991**, *251*, 1211.
- [33] R. N. Eisenman, *Genes Dev.* **2001**, *15*, 2023.
- [34] Z. Nie, G. Hu, G. Wei, K. Cui, A. Yamane, W. Resch, R. Wang, D. R. Green, L. Tessarollo, R. Casellas, K. Zhao, D. Levens, *Cell* **2012**, *151*, 68.
- [35] C. Y. Lin, J. Lovén, P. B. Rahl, R. M. Paranal, C. B. Burge, J. E. Bradner, T. I. Lee, R. A. Young, *Cell* **2012**, *151*, 56.
- [36] A. Sabo, B. Amati, *Cold Spring Harbor Perspectives in Medicine* **2014**, *4*, a014191.
- [37] S. Walz, F. Lorenzin, J. Morton, K. E. Wiese, B. von Eyss, S. Herold, L. Rycak, H. Dumay-Odelot, S. Karim, M. Bartkuhn, F. Roels, T. Wüstefeld, M. Fischer, M. Teichmann, L. Zender, C.-L. Wei, O. Sansom, E. Wolf, M. Eilers, *Nature* **2014**, *1*.
- [38] K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, T. J. Richmond, *Nature* **1997**, *389*, 251.
- [39] D. Tillo, N. Kaplan, I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, Y. Field, J. D. Lieb, J. Widom, E. Segal, T. R. Hughes, *PLoS ONE* **2010**, *5*, DOI 10.1371/journal.pone.0009129.
- [40] C. Chronis, P. Fiziev, B. Papp, S. Butz, G. Bonora, S. Sabri, J. Ernst, K. Plath, *Cell* **2017**, *168*, 442.
- [41] A. Soufi, M. F. Garcia, A. Jaroszewicz, N. Osman, M. Pellegrini, K. S. Zaret, *Cell* **2015**, *161*, 555.
- [42] F. C. Holstege, U. Fiedler, H. T. Timmers, *EMBO J.* **1997**, *16*, 7468.
- [43] N. A. Woychik, R. A. Young, *Trends in Biochemical Sciences* **1990**, *15*, 347.
- [44] T. I. Lee, R. A. Young, *Annu. Rev. Genet.* **2000**, *34*, 77.
- [45] R. D. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12955.
- [46] T. K. Kim, T. Lagrange, Y. H. Wang, J. D. Griffith, D. Reinberg, R. H. Ebright, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12268.
- [47] T. K. Kim, R. H. Ebright, D. Reinberg, *Science* **2000**, *288*, 1418.
- [48] B. L. Allen, D. J. Taatjes, *Nature Publishing Group* **2015**, *16*, 155.
- [49] C. Bernecky, J. M. Plitzko, P. Cramer, *Nat Struct Mol Biol* **2017**, *24*, 809.
- [50] S. Schilbach, M. Hantsche, D. Tegunov, C. Dienemann, C. Wigge, H. Urlaub, P. Cramer, *Nature* **2017**, *551*, 204.
- [51] K. Nozawa, T. R. Schneider, P. Cramer, *Nature* **2017**, *545*, 248.
- [52] E. Nogales, R. K. Louder, Y. He, *Annu Rev Biophys* **2017**, *46*, 59.
- [53] H. Boeger, D. A. Bushnell, R. Davis, J. Griesenbeck, Y. Lorch, J. S. Strattan, K. D. Westover, R. D. Kornberg, *FEBS Letters* **2005**, *579*, 899.
- [54] J. Bednar, P. Furrer, A. Stasiak, J. Dubochet, E. H. Egelman, A. D. Bates, *Journal of Molecular Biology* **1994**, *235*, 825.
- [55] A. V. Vologodskii, S. D. Levene, K. V. Klenin, M. Frank-Kamenetskii, N. R. Cozzarelli, *Journal of Molecular Biology* **1992**, *227*, 1224.
- [56] C. J. Benham, *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2397.
- [57] P. P. Lau, H. B. Gray, *Nucleic Acids Res.* **1979**, *6*, 331.
- [58] Y. Seol, K. C. Neuman, *Biophys Rev* **2016**, *8*, 221.
- [59] J. J. Champoux, *Annu. Rev. Biochem.* **2001**, *70*, 369.
- [60] L. Stewart, M. R. Redinbo, X. Qiu, W. G. Hol, J. J. Champoux, *Science* **1998**, *279*, 1534.
- [61] L. F. Liu, C. C. Liu, B. M. Alberts, *Cell* **1980**, *19*, 697.
- [62] F. Kouzine, A. Gupta, L. Baranello, D. Wojtowicz, K. Ben-Aissa, J. Liu, T. M.

- Przytycka, D. Levens, *Nat Struct Mol Biol* **2013**, *20*, 396.
- [63] L. Baranello, D. Levens, A. Gupta, F. Kouzine, *Biochim. Biophys. Acta* **2012**, *1819*, 632.
- [64] S. Forth, M. Y. Sheinin, J. Inman, M. D. Wang, *Annu Rev Biophys* **2013**, *42*, 583.
- [65] J. Ma, M. D. Wang, *Biophys Rev* **2016**, *8*, 75.
- [66] R. N. Irobalieva, J. M. Fogg, D. J. Catanese, T. Sutthibutpong, M. Chen, A. K. Barker, S. J. Ludtke, S. A. Harris, M. F. Schmid, W. Chiu, L. Zechiedrich, *Nature Communications* **1AD**, *6*, 1.
- [67] F. Kouzine, S. Sanford, Z. Elisha-Feil, D. Levens, *Nature Publishing Group* **2008**, *15*, 146.
- [68] C. Naughton, N. Avlonitis, S. Corless, J. G. Prendergast, I. K. Mati, P. P. Eijk, S. L. Cockroft, M. Bradley, B. Ylstra, N. Gilbert, *Nature Publishing Group* **2013**, *20*, 387.
- [69] L. F. Liu, J. C. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7024.
- [70] G. N. Giaever, J. C. Wang, *Cell* **1988**, *55*, 849.
- [71] G. A. Michelotti, E. F. Michelotti, A. Pullner, R. C. Duncan, D. Eick, D. Levens, *Molecular and Cellular Biology* **1996**, *16*, 2656.
- [72] C. Skerka, P. F. Zipfel, U. Siebenlist, *Oncogene* **1993**, *8*, 2135.
- [73] V. González, L. H. Hurley, *Annual Review of Pharmacology and Toxicology* **2010**, *50*, 111.
- [74] A. J. Kinniburgh, *Nucleic Acids Res.* **1989**, *17*, 7771.
- [75] B. P. Belotserkovskii, E. De Silva, S. Tornaletti, G. Wang, K. M. Vasquez, P. C. Hanawalt, *J. Biol. Chem.* **2007**, *282*, 32433.
- [76] G. Wang, K. M. Vasquez, *Mutat. Res.* **2006**, *598*, 103.
- [77] G. Wang, K. M. Vasquez, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13448.
- [78] D. Sen, W. Gilbert, *Nature* **1988**, *334*, 364.
- [79] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593.
- [80] T. Simonsson, M. Pribylova, M. Vorlickova, *Biochemical and Biophysical Research Communications* **2000**, *278*, 158.
- [81] T. A. Brooks, L. H. Hurley, *Genes & Cancer* **2010**, *1*, 641.
- [82] T. Simonsson, P. Pecinka, M. Kubista, *Nucleic Acids Res.* **1998**, *26*, 1167.
- [83] D. A. T. Sekibo, K. R. Fox, *Nucleic Acids Res.* **2017**, DOI 10.1093/nar/gkx856.
- [84] C. Zhang, H.-H. Liu, K.-W. Zheng, Y.-H. Hao, Z. Tan, *Nucleic Acids Res.* **2013**, *41*, 7144.
- [85] F. Rezzoug, S. D. Thomas, E. C. Rouchka, D. M. Miller, *PLoS ONE* **2016**, *11*, e0161588.
- [86] D. Sun, L. H. Hurley, *J. Med. Chem.* **2009**, *52*, 2863.
- [87] L. Baranello, D. Wojtowicz, K. Cui, B. N. Devaiah, H.-J. Chung, K. Y. Chan-Salis, R. Guha, K. Wilson, X. Zhang, H. Zhang, J. Piotrowski, C. J. Thomas, D. S. Singer, B. F. Pugh, Y. Pommier, T. M. Przytycka, F. Kouzine, B. A. Lewis, K. Zhao, D. Levens, *Cell* **2016**, *165*, 357.
- [88] S. S. Teves, S. Henikoff, *Nat Struct Mol Biol* **2014**, *21*, 88.
- [89] I. Gavin, P. J. Horn, C. L. Peterson, *Molecular Cell* **2001**, *7*, 97.
- [90] S. L. French, M. L. Sikes, R. D. Hontz, Y. N. Osheim, T. E. Lambert, A. El Hage, M. M. Smith, D. Tollervey, J. S. Smith, A. L. Beyer, *Molecular and Cellular Biology* **2011**, *31*, 482.
- [91] R. S. Joshi, B. Piña, J. Roca, *Nucleic Acids Res.* **2012**, *40*, 7907.
- [92] C. A. Brackley, J. Johnson, A. Bentivoglio, S. Corless, N. Gilbert, G. Gonnella, D. Marenduzzo, *Phys. Rev. Lett.* **2016**, *117*, 018101.
- [93] S. A. Sevier, H. Levine, *Phys. Rev. Lett.* **2017**, *118*, 268101.
- [94] J. Eddy, N. Maizels, *Nucleic Acids Res.* **2006**, *34*, 3887.

- [95] M. Petr, R. Helma, A. Polášková, A. Krejčí, Z. Dvořáková, I. Kejnovská, L. Navrátilová, M. Adámik, M. Vorlíčková, M. Brazdova, *Biosci. Rep.* **2016**, DOI 10.1042/BSR20160232.
- [96] V. González, L. H. Hurley, *Biochemistry* **2010**, *49*, 9706.
- [97] R. Duncan, L. Bazar, G. Michelotti, T. Tomonaga, H. Krutzsch, M. Avigan, D. Levens, *Genes Dev.* **1994**, *8*, 465.
- [98] H.-J. Chung, J. Liu, M. Dundr, Z. Nie, S. Sanford, D. Levens, *Molecular and Cellular Biology* **2006**, *26*, 6584.
- [99] J. Liu, L. He, I. Collins, H. Ge, D. Libutti, J. Li, J. M. Egly, D. Levens, *Molecular Cell* **2000**, *5*, 331.
- [100] C. D. Cukier, D. Hollingworth, S. R. Martin, G. Kelly, I. Díaz-Moreno, A. Ramos, *Nature Publishing Group* **2010**, *17*, 1058.
- [101] J. Liu, F. Kouzine, Z. Nie, H.-J. Chung, Z. Elisha-Feil, A. Weber, K. Zhao, D. Levens, *EMBO J.* **2006**, *25*, 2119.
- [102] L. Guo, O. Zaysteva, Z. Nie, N. C. Mitchell, J. E. Amanda Lee, T. Ware, L. Parsons, R. Luwor, G. Poortinga, R. D. Hannan, D. L. Levens, L. M. Quinn, *Nucleic Acids Res.* **2016**, *44*, 7646.
- [103] N. C. Mitchell, T. M. Johanson, N. J. Cranna, A. L. J. Er, H. E. Richardson, R. D. Hannan, L. M. Quinn, *Development* **2010**, *137*, 2875.
- [104] J. E. A. Lee, N. C. Mitchell, O. Zaytseva, A. Chahal, P. Mendis, A. Cartier-Michaud, L. M. Parsons, G. Poortinga, D. L. Levens, R. D. Hannan, L. M. Quinn, *Nature Communications* **2015**, *6*, 7404.
- [105] B. Majello, P. De Luca, G. Suske, L. Lania, *Oncogene* **1995**, *10*, 1841.
- [106] T. Tomonaga, D. Levens, *J. Biol. Chem.* **1995**, *270*, 4875.
- [107] E. F. Michelotti, G. A. Michelotti, A. I. Aronsohn, D. Levens, *Molecular and Cellular Biology* **1996**, *16*, 2350.
- [108] E. F. Michelotti, T. Tomonaga, H. Krutzsch, D. Levens, *J. Biol. Chem.* **1995**, *270*, 9494.
- [109] N. Irwin, V. Baekelandt, L. Goritchenko, L. I. Benowitz, *Nucleic Acids Res.* **1997**, *25*, 1281.
- [110] M. E. Olanich, B. L. Moss, D. Piwnica-Worms, R. R. Townsend, J. D. Weber, *Oncogene* **2011**, *30*, 77.
- [111] Y. Zheng, W. K. Miskimins, *The International Journal of Biochemistry & Cell Biology* **2011**, *43*, 1641.
- [112] J. Miro, A. M. Laaref, V. Rofidal, R. Lagrèfeuille, S. Hem, D. Thorel, D. Méchin, K. Mamchaoui, V. Mouly, M. Claustres, S. Tuffery-Giraud, *Nucleic Acids Res.* **2015**, *43*, 2378.
- [113] Z. Zhang, D. Harris, V. N. Pandey, *J. Virol.* **2008**, *82*, 5761.
- [114] M. L. Hastings, E. Allemand, D. M. Duelli, M. P. Myers, A. R. Krainer, *PLoS ONE* **2007**, *2*, e538.
- [115] P. S. Page-McCaw, K. Amonlirdviman, P. A. Sharp, *RNA* **1999**, *5*, 1548.
- [116] D. T. Braddock, J. M. Louis, J. L. Baber, D. Levens, G. M. Clore, *Nature* **2002**, *415*, 1051.
- [117] M. Takimoto, T. Tomonaga, M. Matunis, M. Avigan, H. Krutzsch, G. Dreyfuss, D. Levens, *J. Biol. Chem.* **1993**, *268*, 18249.
- [118] S. Thakur, T. Nakamura, G. Calin, A. Russo, J. F. Tamburrino, M. Shimizu, G. Baldassarre, S. Battista, A. Fusco, R. P. Wassell, G. Dubois, H. Alder, C. M. Croce, *Molecular and Cellular Biology* **2003**, *23*, 3774.
- [119] L. Lannes, P. Young, C. Richter, N. Morgner, H. Schwalbe, *Chembiochem* **2017**, *18*, 2033.
- [120] N. Cornella, T. Tebaldi, L. Gasperini, J. Singh, R. A. Padgett, A. Rossi, P. Macchi, *Journal of Biological Chemistry* **2017**, jbc.M117.795591.

- [121] H. Liu, S. M. Barr, C. Chu, D. S. Kohtz, Y. Kinoshita, E. M. Johnson, *Biochemical and Biophysical Research Communications* **2005**, 328, 851.
- [122] G. L. Gallia, E. M. Johnson, K. Khalili, *Nucleic Acids Res.* **2000**, 28, 3197.
- [123] M. K. White, E. M. Johnson, K. Khalili, *cc* **2009**, 8, 1.
- [124] N. Darbinian, G. L. Gallia, K. Khalili, *J. Cell. Biochem.* **2001**, 80, 589.
- [125] C. Fernando, A. Audibert, F. Simon, J. Tazi, F. Juge, *Genetics* **2015**, 199, 1201.
- [126] P. Castro, H. Liang, J. C. Liang, L. Nagarajan, *Genomics* **2002**, 80, 78.
- [127] J. Liu, X. Luo, Y. Xu, J. Gu, F. Tang, Y. Jin, H. Li, *Stem Cell Res Ther* **2016**, 7, 79.
- [128] M. Fuxreiter, P. Tompa, I. Simon, V. N. Uversky, J. C. Hansen, F. J. Asturias, *Nat. Chem. Biol.* **2008**, 4, 728.
- [129] W. B. Tu, S. Helander, R. Pilstål, K. A. Hickman, C. Lourenco, I. Jurisica, B. Raught, B. Wallner, M. Sunnerhagen, L. Z. Penn, *Biochim. Biophys. Acta* **2015**, 1849, 469.
- [130] R. Bayliss, S. G. Burgess, E. Leen, M. W. Richards, *Biochem. Soc. Trans.* **2017**, 45, 709.
- [131] F. Benedetti, D. Racko, J. Dorier, Y. Burnier, A. Stasiak, *Nucleic Acids Res.* **2017**, 45, 9850.
- [132] D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, P. A. Sharp, *Cell* **2017**, 169, 13.
- [133] T. H. Dickey, S. E. Altschuler, D. S. Wuttke, *Structure* **2013**, 21, 1074.

Figure 1

Figure 2

Graphical Abstract

Torsional stress, and supercoiling generated by transcription, shape the topology of DNA. Furthermore, the resultant secondary DNA structures and their interacting partners provide feedback to regulate gene expression. Emerging evidence suggests these mechanisms, identified through analysis of the *MYC* promoter, are applicable genome-wide.

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Zaytseva, O;Quinn, LM

Title:

DNA Conformation Regulates Gene Expression: The MYC Promoter and Beyond

Date:

2018-04-01

Citation:

Zaytseva, O. & Quinn, L. M. (2018). DNA Conformation Regulates Gene Expression: The MYC Promoter and Beyond. BIOESSAYS, 40 (4), <https://doi.org/10.1002/bies.201700235>.

Persistent Link:

<http://hdl.handle.net/11343/283699>