Annual Review of Biochemistry

The Semantics and Mechanisms of Enhancers and Promoters: "What Is True for *E. coli* Is True for the Elephant, Only More So"

Kevin Struhl

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA; email: kevin@hms.harvard.edu

Annu. Rev. Biochem. 2026. 95:199–219

First published as a Review in Advance on March 20, 2026

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

https://doi.org/10.1146/annurev-biochem-080125-014150

Keywords

gene regulation, transcription, RNA polymerase, enhancers, transcriptional activation, activator proteins

Abstract

Three fundamental classes of gene regulatory elements were identified in *Escherichia coli* in the 1960s: operators bound by repressor proteins, promoters bound by the basic transcription machinery, and enhancers bound by activator proteins. Promoters mediate constitutive gene expression, whereas operators and enhancers regulate expression in response to physiological conditions. As discovered in the 1980s, interactions between proteins associated with spatially separated elements can loop out the intervening DNA and regulate transcription. Eukaryotic gene regulation is mediated by the same three classes of genetic elements, but due to semantic confusion, it is often believed the eukaryotic enhancers activate transcription from long distances via enhancer–promoter loops. However, eukaryotic enhancers mediate only local changes in chromatin, and they stimulate transcription via directional and short-range interactions with the basic RNA polymerase II machinery at promoters. Enhancer action at a distance is mediated by loops between proximal and distal enhancers that bring activator proteins associated with distal enhancers in proximity to promoters. Thus, Jacques Monod's 1954 conjecture that "what is true for *E. coli* is true for the elephant, only more so" has proven correct.

Contents

JACQUES MONOD'S COMMENT

"What is true for *E. coli* is true for the elephant, only more so". Jacques Monod's prescient comment about gene regulation was apparently never published but rather said after a lecture in 1954. It was a radical conjecture because the groundbreaking work of Monod and François Jacob on *Escherichia coli* gene regulation was in its early stages, and virtually nothing was known about gene regulation in eukaryotes. Notably, many references to Monod's comment omit the essential qualifier "only more so." This has caused misunderstanding of the key concept, namely that there are universal principles of gene regulation with elaborations for increased biological complexity. In addition, semantic confusion between promoters and enhancers has resulted in a common misunderstanding of the mechanism of transcriptional activation in eukaryotes. Here, I provide a conceptual and mechanistic perspective that links the original ideas of Jacob and Monod to the current understanding of eukaryotic gene regulation.

FUNDAMENTAL CONCEPTS OF TRANSCRIPTIONAL REGULATION AS REVEALED BY GENETIC ELEMENTS

Fundamental concepts of gene regulation were established in the early 1960s by the brilliant work of Jacob & Monod (1) performed at the Institut Pasteur in Paris. Their studies on the *E. coli lac* operon were designed to explain coregulated expression of linked protein-coding genes that perform a common biological function (lactose utilization). Nevertheless, as prophesized by Monod's comment that "what is true for *E. coli* is true for the elephant, only more so," the concepts derived from this single experimental model are relevant to prokaryotic and eukaryotic genes. It was the first case in which detailed study of a single experimental system yielded fundamental principles, something that is often forgotten in the genomic era. Their essential ideas can be divided into three parts.

First, genetic elements, and hence DNA sequences, involved in gene expression are functionally distinct and physically separate from the protein-coding (previously called structural) genes (**Figure 1*a***). While it seems obvious now, the distinction between sequences that encode proteins from those involved in expression and regulation of those proteins was a major conceptual breakthrough. Even now in the era of chromatin, three-dimensional genome structure, and biological condensates, it is often unappreciated that regulated gene expression ultimately depends

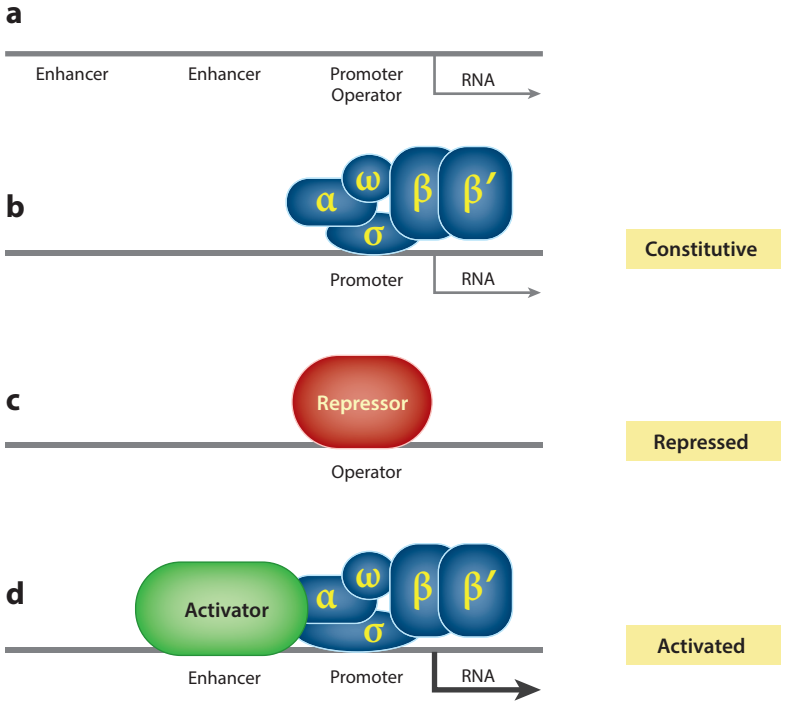


Figure 1

Transcriptional regulatory elements in *E. coli*. (*a*) A DNA region with enhancer, promoter, and operator elements shown upstream of a transcription start site (*arrow*) and mRNA coding region. (*b*) RNAP (5 subunits indicated by *blue ovals*) interacts with promoter DNA sequences via the σ subunit and mediates constitutive levels of transcription depending on the quality of the promoter. (*c*) A repressor protein (*red oval*) interacts with operator sequences to inhibit transcription, typically by blocking association of RNAP with the promoter. (*d*) An activator protein (*green oval*) stimulates transcription by directly interacting with subunits (typically α or σ) of RNAP. Abbreviation: RNAP, RNA polymerase.

on the DNA sequences that determine the identity and biochemical properties of the proteins that recognize these sequences.

Second, genetic elements necessary for constitutive expression of structural genes, which occurs at the same level under all conditions, are distinct from genetic elements that regulate the expression level of the gene under different conditions (2, 3). Genetic elements that mediate constitutive expression, but not the regulation of expression, are termed promoters (**Figure 1*b***). In molecular terms, promoters are recognized by basic transcription machineries (e.g., the 5-subunit RNA polymerase in the case of *E. coli*) that initiate RNA synthesis from nearby sites. Depending on the DNA sequence, promoters can be strong or weak, supporting high or low levels, respectively, of constitutive expression. Bacterial RNA polymerases typically exist in multiple forms depending on the σ subunit that confers DNA sequence specificity and hence promoter choice (4).

Third, regulatory elements, termed operators (5), inhibit expression of the gene via *trans*-acting repressors that inhibit expression via the operators (**Figure 1*c***) (1). Repressor function is regulated by environmental conditions (e.g., lactose inhibits the function of the *lac* repressor), thereby resulting in regulation of the gene. Jacob and Monod initially suggested that the repressor was RNA, imagining that it would act through base pairing to the operator DNA. However, it soon became clear that repressors are DNA-binding proteins (6, 7), which indicated that biological specificity occurs between very different biochemical entities.

Although Jacob and Monod were rather dogmatic in insisting that gene regulation was exclusively mediated by repressors, Ellis Englesberg and colleagues (8, 9) discovered a regulatory element in the arabinose operon that mediates positive control (**Figure 1*d***). This positive control element, unfortunately termed the initiator upon its discovery, was the functional (and subsequently shown to be the physical) target of the AraC activator protein. As the term initiator now describes sequences around the mRNA start site that contribute to initiation (10–12), the positive control sequence that activates the arabinose operon is analogous to, and hence best termed, an enhancer (discussed below). Activation by AraC requires a direct interaction with arabinose, thereby limiting expression of the protein-coding genes to the appropriate physiological conditions. AraC also functions as a repressor in the absence of arabinose via operator sequences that are distinct from enhancer sequences (13). Englesberg's work on the arabinose operon (14) and details of AraC function (15) have been described elsewhere.

THE EXISTENCE OF GENETIC ELEMENTS IS INDEPENDENT OF THEIR FUNCTIONAL ACTIVITIES

Although regulatory elements differentially affect gene expression depending on physiological or developmental conditions, genetic elements per se are physically and conceptually present under all conditions (**Figure 1*a***). As such, it is important to distinguish between the existence of a regulatory element and its activity under a specific condition. For example, the *lac* operator is a genetic element, whether or not lactose is in the growth medium. In this regard, identification of gene regulatory elements via mutation, molecular genetic approaches, or epigenomic profiling is limited by their functional properties under the specific conditions examined. Conversely, while computational approaches to identify regulatory elements via their DNA sequences are independent of their function, they are only predictive and do not account for the remarkable complexity and diversity of gene regulatory patterns.

THE QUALITATIVE NATURE OF GENETIC ELEMENTS IS LINKED TO BIOLOGICAL FUNCTION

A genetic element is a qualitative concept, ultimately defined by DNA sequences linked to a biological function; it is distinguished from the vast majority of sequences that do not encode that

function. A semantic issue arises because genetic elements are typically recognized by proteins whose DNA binding is quantitative in nature. DNA-binding proteins do not simply have target and nontarget sites; rather, they interact with all DNA sequences with a continuum of affinities (16). So-called nonspecific binding to any DNA sequence typically occurs with 10^4–10^5 times lower affinity than binding to an optimal target site represented by a sequence motif. Moreover, DNA sequences have a wide range of intermediate binding affinities that are linked to their level of similarity to the optimal motif (16). This continuum of affinities means that the distinctions between a genetic element, a weak binding site, and a nonspecific sequence are somewhat arbitrary. Furthermore, the number of nonspecific or weak binding sites vastly outnumber the strong binding sites that are more likely to mediate biological functions.

In a similar vein, ~90% of elongating RNA polymerase II (Pol II) molecules in yeast are not involved in making conventional mRNA and hence represent transcriptional noise (17). Indeed, Pol II transcription initiates at numerous sites in both random-sequence and genomic DNA (18, 19), and it is presumed that all such transcription is mediated by the same basic transcription machinery used at classically defined promoters. The specificity difference between transcription from classical promoter elements and that from random-sequence DNA is ~10^4 fold, similar to the difference for individual DNA-binding proteins (17). For these reasons, the original Jacob–Monod concept that genetic elements are linked to biological function (e.g., protein- or RNA-coding sequences) remains most useful, although it is becoming increasingly clear that biological function is not easily defined (17).

DIFFERENT LOGIC BETWEEN PROKARYOTES AND EUKARYOTES

Before addressing the semantics and mechanisms of promoters and enhancers, it is important to note that the logic of gene regulatory elements is fundamentally different in prokaryotic and eukaryotic organisms (20). The key difference lies in the transcriptional ground state, which is defined by the activity of a promoter in vivo (20). In *E. coli*, and presumably other prokaryotes, a promoter is sufficient for transcription because the genome is not coated with histones and hence is generally accessible to proteins (21). Consequently, to block promoter function under appropriate conditions, it is necessary to have repressor proteins. Conversely, activator proteins are not necessary for transcription, but they can stimulate transcription from weak promoters under appropriate conditions. In eukaryotes, a promoter is essentially inactive because chromatin blocks accessibility of the basic Pol II machinery. As a consequence, activator proteins are essential for transcription from promoters, whereas repressor proteins are not required to keep expression low. However, eukaryotic repressors can inhibit the function of activators and/or silence gene expression below the transcriptional ground state.

THE CONFUSING SEMANTICS OF EUKARYOTIC PROMOTERS AND ENHANCERS

The fundamental concepts of gene regulation established by Jacob, Monod, and Englesberg have been muddled by terminology that, as is often the case, leads to mechanistic misunderstanding. In contrast to the Jacob–Monod definition, metazoan promoters were, and often still are, defined by short DNA segments that permit accurate initiation in vitro from sites used in vivo. Unfortunately, some so-called promoter DNA segments also include DNA sequences recognized by specific transcription factors that vary among genes. The adjacent sequences recognized by specific DNA-binding transcription factors lie upstream of what is termed the core promoter.

The term core promoter is unnecessary, and hence confusing, because it identical to the Jacob–Monod-defined promoter that is recognized by the basic Pol II machinery and hence required

for constitutive expression but not regulation of expression. Unlike bacterial promoters, which comprise sequences directly recognized by RNA polymerase, Pol II promoters contain sequences that are recognized by general transcription factors [e.g., TATA-binding protein (TBP) and the associated factors TFIIA and TFIIB], not Pol II itself. Most importantly, the common terminology for metazoan promoters blurs the fundamental distinction between genetic elements mediating constitutive expression versus regulation of expression.

The term enhancer is also a source of semantic confusion. Enhancers were originally defined in mammalian cells as genetic elements that stimulate transcription when located at long and variable distances upstream or downstream from a promoter (22–24). They activate transcription in response to environmental and developmental signals (25–27) and hence determine gene regulatory patterns (28, 29). Eukaryotic enhancers are bound by sequence-specific activator proteins, and hence they are conceptually analogous to the AraC-dependent, positive control element described by Englesberg. The semantic confusion arises from the fact that the activator proteins that bind enhancers also bind to promoter-proximal sequences. Furthermore, enhancers typically express noncoding enhancer RNAs that are generated by the basic Pol II machinery and emanate bidirectionally from the enhancer (30–32).

These observations have led to the confusing idea that enhancers and promoters are very similar (30). Conceptually, as recognized decades ago, genetic elements that mediate constitutive expression should be distinguished from those that mediate regulation. Mechanistically, enhancers are bound by activator proteins, whereas promoters are bound by the basic transcription machinery. Enhancers and promoters are conceptually and functionally distinct entities, and the semantics of these terms have mechanistic consequences.

DISTAL AND PROXIMAL ENHANCERS: RESOLUTION OF THE SEMANTIC CONFUSION WITH PROMOTERS

The semantic confusion described above is resolved by the terms distal enhancers and proximal enhancers. Distal enhancers are located far from the promoters they affect (i.e., the original definition of enhancers), whereas proximal enhancers are located near the promoter and are often present on DNA fragments that contain promoters. Both types of enhancers are bound by sequence-specific activator proteins and hence are regulatory in nature. An individual activator protein can bind to either proximal or distal enhancers or both types, depending on where the cognate sequences are located. Thus, distal and proximal enhancers are similar functional entities that differ only by their location with respect to the promoter.

Metazoan, yeast, and prokaryotic enhancers are similar in that they bind activator proteins that stimulate transcription from a promoter. However, unlike prokaryotic and yeast (33) enhancers that can and often do contain a single activator-binding site, metazoan enhancers require multiple activator-binding sites. The requirement for multiple binding sites mechanistically imposes combinatorial activation, the fundamental principle for how multicellular organisms generate the extraordinary diversity of gene regulatory patterns (28, 29). Multiple distal enhancers with different regulatory specificities can activate an individual promoter (28, 29).

ENHANCERS GENERATE LOCALIZED REGIONS OF ACTIVE CHROMATIN AND BIDIRECTIONAL ENHANCER RNAs IN THE APPARENT ABSENCE OF A PROMOTER

In accordance with the fundamentally different logic of gene regulation in eukaryotes and prokaryotes, eukaryotic transcriptional activator proteins have a distinct biochemical activity that is lacking in prokaryotic activators. Specifically, eukaryotic activators bound to

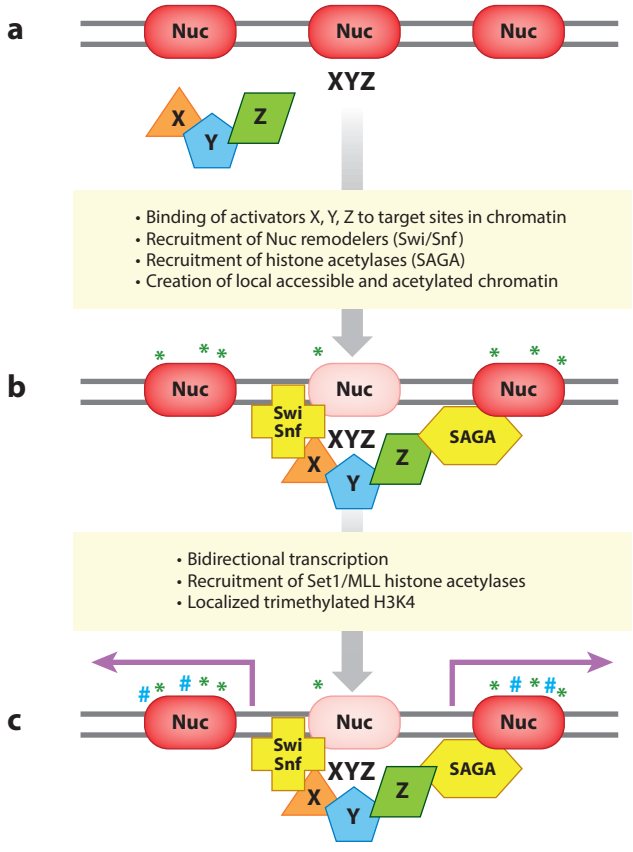


Figure 2

Local effects of activator proteins bound to enhancers. (*a*) Activator proteins X, Y, and Z (*colored polygons*) bind to their cognate sequences in chromatin (Nucs are depicted as *red ovals*). (*b*) The bound activators recruit Nuc remodelers and histone acetylases via protein–protein interactions, thereby creating localized regions of altered chromatin that are accessible (*light pink ovals*) and acetylated (*green asterisks*). (*c*) Bidirectional transcription of enhancer RNAs emanates from the altered chromatin region and leads to recruitment of Set1/MLL histone methylases and localized trimethylated H3K4 (*blue hash symbols*). Abbreviations: H3K4, histone 3 lysine 4; Nuc, nucleosome.

enhancers recruit nucleosome-remodeling (e.g., Swi/Snf or BRG1) and histone acetylase (e.g., SAGA/Gcn5 or p300) complexes, thereby altering chromatin structure at the enhancer (20, 34–36) (**Figures 2*a*,*b*** and **3**). As recruitment is mediated by direct interactions between activator proteins and chromatin-modifying complexes, promoters and transcription are not required for recruitment. Thus, both distal and proximal enhancer regions are depleted of nucleosomes and hence accessible to transcription factors, and they have hyperacetylated histones that contribute to transcriptional activation (**Figures 2*b*** and **3**).

Enhancer-bound activator proteins invariably generate enhancer RNAs that initiate bidirectionally from the edges of the nucleosome-depleted region that they created (**Figures 2*c*** and **3**) (30–32). Enhancer RNAs are generated by standard transcriptional initiation by the basic Pol II machinery, and they come from proximal enhancers, distal enhancers, and fortuitous enhancers occurring at evolutionarily irrelevant sequences in yeast (37, 38). Distal enhancers, by definition, are not linked to nearby promoters. As bidirectional transcription from essentially all fortuitous

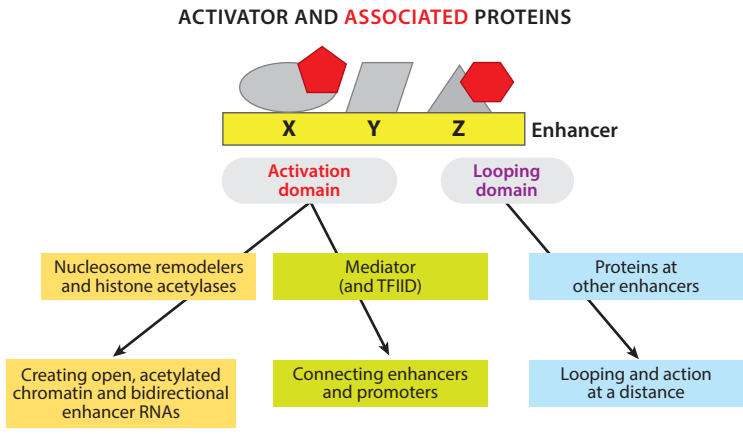


Figure 3

Functions of activator and associated coactivator proteins bound at enhancers. Activator proteins X, Y, and Z (*dark gray polygons*) and associated coactivator proteins (*red polygons*) bound to enhancer elements X, Y, and Z mediate three classes of protein–protein interactions via activation and looping domains (*light gray ovals*). Activation domains interact with and recruit nucleosome remodelers and histone acetylases to create localized regions of altered chromatin and generate bidirectional enhancer RNAs (see **Figure 2**); they also interact with Mediator (and to a lesser extent TFIID) to connect proximal enhancers and promoters. Looping domains connect proteins bound at distal and proximal enhancers to permit enhancer-mediated activation at long distances from the promoter.

and newly evolved enhancers occurs at equal (and typically low) levels in each direction (37, 38), the Pol II machinery presumably initiates transcription in a nonsequence-specific manner from accessible DNA. Despite the apparent lack of sequence specificity typical of promoters, enhancer RNAs initiate transcription at a small number of sites. This initiation-site specificity is likely due to the location of the +1 nucleosome flanking the accessible region and/or distance-constrained interactions between activator proteins and the basic Pol II machinery on DNA (discussed below). Thus, the sequence specificity mediating enhancer RNAs is determined primarily by the enhancer, not the promoter.

The behavior of fortuitous enhancers and newly evolved enhancers indicates that bidirectional enhancer RNAs are an inevitable mechanistic consequence of activator functions (37, 38). Hence, the mere existence of enhancer RNAs does not necessarily imply any biological function. Many enhancer RNAs are unstable and likely functionally insignificant, although some play functional roles. In addition, the effects of Pol II transcription on chromatin structure might be functionally significant even when the enhancer RNAs are not. As bidirectional transcription from enhancers represents the mechanistic ground state, preferential transcription in the coding direction requires evolutionary selection (38). Proximal enhancers, which are located near a promoter on one side of the enhancer region, provide a simple mechanism for biased unidirectional transcription of coding regions.

At well-transcribed promoters, Pol II and the associated Paf1 complex efficiently recruit Set1/MLL histone methylase complexes that generate trimethylated lysine 4 of histone H3 (H3K4me3) nucleosomes near the initiation site (**Figure 2*c***) (39, 40). In contrast, recruitment of these histone methylases at distal enhancers is typically inefficient because bidirectional transcription, and hence Pol II occupancy, is relatively low. Consequently, unlike proximal enhancers, distal enhancers typically have low levels of H3K4me3 nucleosomes. This epigenomic (41) distinction between distal and proximal enhancers does not reflect functional differences but rather

their location near a promoter. Thus, the epigenomic signature around promoters is due to separate contributions of proximal enhancers (low histone occupancy and histone hyperacetylation) and promoters (H3K4 trimethylation).

ENHANCERS, VIA CHROMATIN, REGULATE DNA-BASED PROCESSES OTHER THAN POL II TRANSCRIPTION

The localized chromatin-modifying function of enhancers is linked to, but mechanistically distinct from, the process of transcriptional activation. It also plays a role in DNA-based processes such as Pol III transcription, DNA replication, and V(D)J recombination that lack dedicated regulatory proteins but nevertheless are subject to cell-type-specific regulation (42). Mammalian cell types differ in their constellation of expressed tRNA genes even though the Pol III transcription machinery is present in all cells (43, 44). V(D)J recombination of antibody and T cell receptor genes in lymphoid cells expressing the V(D)J recombinase proteins RAG1 and RAG2 is highly regulated by developmental state (45). In metazoans, the origin replication complex (ORC) that initiates DNA replication is present in all cells and lacks inherent DNA-binding specificity, yet it binds to discrete genomic locations that differ among cell types (46). In all these cases, cell-type- and gene-specific regulation requires sequence-specific DNA-binding proteins.

Enhancers, via sequence-specific activator proteins, are the only known entities for targeted and localized nucleosome depletion. Such nucleosome-depleted regions underlie how enhancers facilitate TFIIIC binding and Pol III transcription, ORC binding and initiation of DNA replication, and RAG1/RAG2 binding and V(D)J recombination (42). It is unclear whether enhancer RNAs per se influence Pol III transcription or ORC binding and function, although blocking Pol II transcription with α-amanitin does not affect Pol III transcription (47). As generation of enhancer RNAs requires nucleosome depletion, any role of enhancer RNAs would be linked to the enhancer-mediated effects on chromatin structure. In this regard, enhancers also affect V(D)J recombination via transcription of noninformational RNAs that is coupled to generation of H3K4me3 nucleosomes near RAG target sites. The PHD domain of RAG2 interacts with H3K4me3 nucleosomes, thereby increasing RAG binding and V(D)J recombination (48) in an enhancer-dependent manner.

Thus, regulation of DNA-mediated processes other than Pol II transcription is mediated by the regulatory properties of the enhancers, indirectly via their effects on chromatin. In contrast, as discussed in the next section, activator proteins directly interact with the basic Pol II machinery, and such interactions are crucial for transcriptional activation. Thus, enhancers represent the first step in decoding the genome through their unique ability to alter local chromatin structure via specific DNA sequences and their cognate activator proteins.

ENHANCERS STIMULATE POL II TRANSCRIPTION BY DIRECTIONAL AND SHORT-RANGE INTERACTIONS WITH THE BASIC POL II MACHINERY AT PROMOTERS

Prokaryotic activator proteins such as AraC (49), λcI (50), CAP (51), and NtrC (52, 53) function via direct interactions with various RNA polymerase subunits. Depending on the activator and the target subunit within RNA polymerase, the spacing between the enhancer and promoter can be highly constrained or relatively flexible. In general, the enhancer–promoter distance is short (<100 bp), but NtrC can activate transcription when bound >1 kb upstream or downstream of the promoter (52, 53). Two activator proteins bound to distinct enhancers can synergistically activate transcription, particularly when they interact with different RNA polymerase subunits (54).

By analogy, eukaryotic activator proteins interact directly with components of the Pol II transcription machinery, although not with Pol II itself (55). The major target of activator proteins is Mediator, a multiprotein complex that also interacts directly with Pol II (56–61) and hence acts as a physical and dynamic bridge between enhancers and promoters (**Figure 3**). Interestingly, the nature of the bridge differs between yeast and metazoan cells (61). Mediator is preferentially associated with enhancers in yeast, whereas it is preferentially associated with the promoter in metazoans. A few activator proteins interact directly with TFIID (**Figure 3**) (62–64), a component of the basal Pol II machinery consisting of TBP and TBP-associated factors.

The communication between activators bound to enhancers and Mediator (or TFIID) bound to promoters occurs only over short distances (~50–300 bp) (**Figure 4*a***). In other words, promoters can be stimulated directly by proximal enhancers but not by distal enhancers (**Figure 4*b***; see the next section for how distal enhancers act at a distance). For example, in transient transfection experiments, the β-interferon enhancer strongly stimulates transcription from the β-interferon promoter when the two elements are 75 bp apart, but it weakly stimulates transcription when they are 220 bp apart, and it fails to stimulate transcription when they are >500 bp apart (65). The β-interferon enhancer also fails to stimulate transcription of the heterologous thymidine kinase promoter when located 2.3 kb upstream (65). Similarly, at the endogenous locus in GATA1-deficient cells, the β-globin enhancer within the locus control region (LCR) stimulates transcription only when located close to the β-globin promoter (66). In addition, systematic distance experiments between the yeast *GAL* enhancer and the *bis3* promoter (67) yield results that are remarkably similar to those described for the β-interferon enhancer. Yeast enhancers activate transcription poorly when located downstream of the promoter (67–69).

DISTAL ENHANCERS ACTIVATE TRANSCRIPTION VIA LOOPS WITH PROXIMAL ENHANCERS

The short-range nature of direct enhancer–promoter communication appears to go against the original definition and common view that enhancers act at long distances from promoters. It also appears to go against the numerous examples of long-distance loops between enhancers and promoters. Indeed, DNA loops between spatially separated genomic regions have been identified on a genomic scale by chromosome conformation capture methodologies (70, 71).

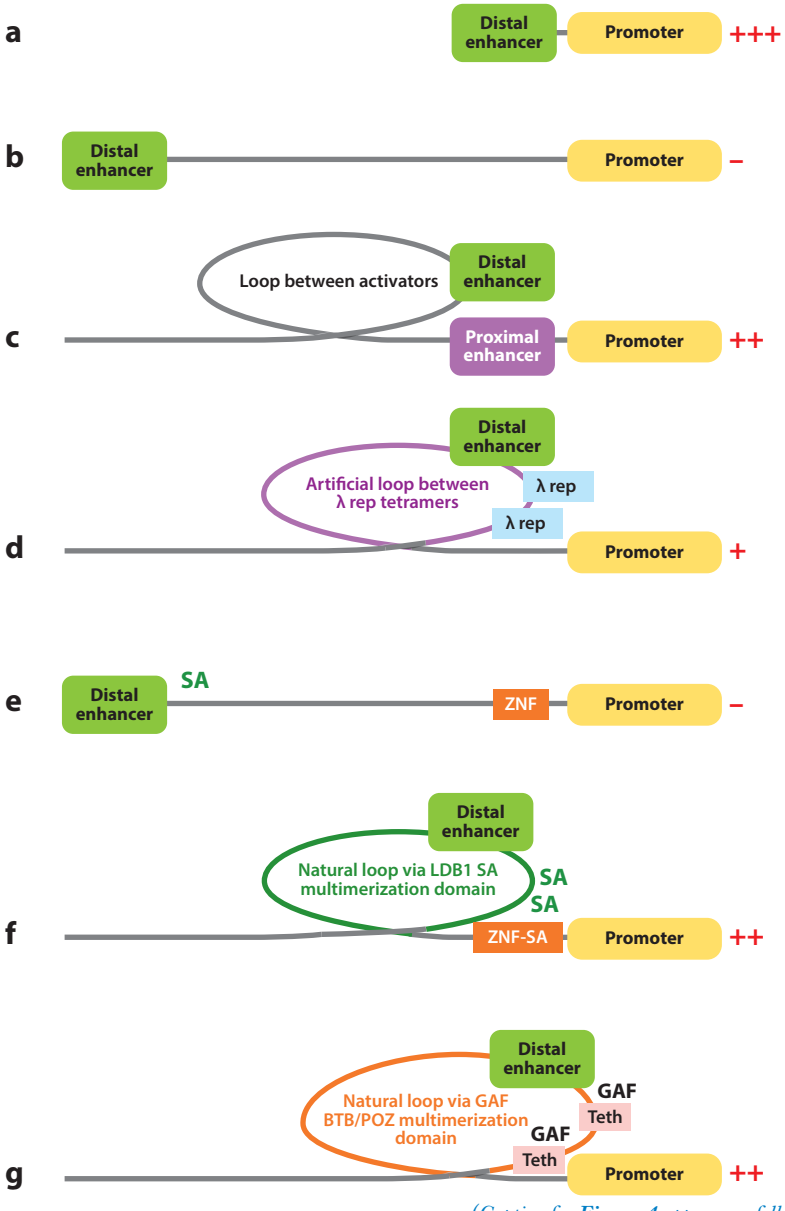
DNA looping and its mechanistic basis was demonstrated nearly 40 years ago in *E. coli*. In four initial examples—the Gal repressor (72, 73), Lac repressor (74), λ repressor (75), and AraC activator (76, 77)—looping involves a tetramerization domain that is required for protein interactions between dimers bound to separated target sites. AraC forms alternative loops that use distinct enhancers or operators and that involve distinct AraC conformations depending on the presence or absence of arabinose (77). In general, these DNA loops repress transcription, although some loops can indirectly result in activation dependent on a proximal enhancer (78). The loop between the NtrC activator and the σ54 subunit of RNA polymerase has been visualized (79), but the structural details of the protein–protein interaction have not been elucidated.

The common misconception about enhancer action at a distance in eukaryotes arises from semantic confusion between enhancers (activator-binding sites) and promoters (bound by the basic Pol II machinery). Instead, as discussed elsewhere (80), two key studies demonstrate that long-range activation by a distal enhancer occurs only when transcription factor–binding sites are located shortly upstream of the promoter (**Figures 3** and **4**).

First, in transient transfection experiments, the β-interferon enhancer does not stimulate transcription when located 2.3 kb upstream from the thymidine kinase promoter (65). However, this enhancer strongly activates transcription when located upstream of a slightly larger promoter

fragment that includes binding sites for the transcriptional activator Sp1 (**Figure 4*c***). Thus, long-range activation requires a functional connection between proteins bound at the distal enhancer and Sp1 bound to promoter-proximal sites (i.e., a proximal enhancer). Strikingly, the requirement for Sp1 can be bypassed by artificially connecting the distal and proximal regions via interaction between λ repressor dimers at target sites through their tetramerization domains (**Figure 4*d***) (65).

Second, in related experiments performed at the endogenous β-globin locus control region [LCR; more recently termed a superenhancer (81)] cannot stimulate β-globin transcription when the proximal enhancer is genetically inactivated via GATA-1 knockout (the activator GATA-1



(*Caption for **Figure 4** appears on following page*)

Figure 4 (*Figure appears on preceding page*)

Long distance activation by distal enhancers requires loops to proximal enhancers, not promoters. (*a*) A distal enhancer (*green box*) located near a promoter (*yellow box*) leads to high levels of activation (*red plus symbols*). (*b*) A distal enhancer located far from a promoter fails to activate transcription (*red minus symbol*). (*c*) A distal enhancer forming a loop to a proximal enhancer (*purple box*) supports high levels of activation. (*d*) The requirement for a loop between distal and proximal enhancers to mediate transcriptional activation at a distance can be met when λ rep dimers bound to distal and proximal target sites interact to form a tetramer and thus generate an artificial loop. (*e*) A distal enhancer associated with the SA domain of LDB1 cannot activate transcription of a promoter with an adjacent site bound by an artificial ZNF (*orange box*). (*f*) A distal enhancer associated with the SA domain of LDB1 can activate transcription from a distance if the ZNF contains an SA domain and generates a loop via SA multimerization. (*g*) In *Drosophila*, activation by a distal enhancer requires a loop mediated by the BTB/POZ multimerization domain of GAF bound to tethering elements located near a distal enhancer and promoter. Loss of the BTB/POZ domain abolishes looping and therefore activation. Abbreviations: λ rep, λ repressor; GAF, GAGA-associated protein; SA, self-association; Teth, tethering element; ZNF, zinc finger protein.

is required for proximal enhancer function) (**Figure 4*b***) (66). However, an artificial zinc finger protein designed to bind sequences near the inactivated proximal enhancer can activate β-globin transcription but only if it also contains the LDB1 self-association (SA) multimerization domain. At the endogenous β-globin locus, LDB1 interacts with LMO1, a protein recruited to both the proximal enhancer and the distal superenhancer in a manner that depends on binding of the GATA-1 activator to these enhancers (**Figure 4*e*,*f***) (66). This observation indicates that enhancer action at a distance involves looping that occurs between proximal and distal enhancers, and the LDB1 SA domain is critical for loop formation (66).

In a conceptually similar, but mechanistically distinct, manner, *Drosophila* tethering elements play an analogous role in looping between proximal and distal enhancers. However, unlike enhancers, tethering elements do not activate transcription on their own (82) (**Figure 4*g***). These tethering elements are bound by the *Drosophila* GAGA-associated protein (GAF) via GAGA sequence elements, and they are located near distal enhancers and promoters (83). Like activator proteins, GAF generates local nucleosome-depleted regions via recruitment of nucleosome remodelers (84), but it appears to lack an activation domain. GAF contains a BTB/POZ oligomerization domain that mediates long-range looping interactions between tethering elements (85). As such, the BTB/POZ oligomerization domain of GAF performs the same looping function as the SA oligomerization domain of LDB1. Lastly, GAF-mediated loops permit long-distance activation in yeast (86), something not normally observed in this unicellular organism. Thus, GAF effectively creates loops that bring distal enhancers closer to promoter-proximal regions, thereby allowing transcriptional activation at a distance. Mutations of the relevant tethering elements block looping and transcriptional activation (83, 85). To date, tethering elements have been identified only in *Drosophila*, but it seems likely that they exist in other metazoans.

Systematic analysis of artificially targeted (via dCas9 and guide RNAs) activator proteins indicates that long-range activation scales with genomic distance (87). This observation, which resembles that of long-range activation by NtrC in *E. coli* (78), suggests that the efficiency of loop formation decreases with distance, presumably due to the looping proteins being present at lower local concentrations. The relationship between genomic distance and transcriptional activity depends on the gene (87), but it is unknown whether this is due to differences in proximal enhancers or the response of different promoters to the loop. Interestingly, cooperativity between different enhancers can compensate for the loss of activity over large genomic distances, possibly reflecting the formation of loops between distal enhancers (88). In this regard, LDB1 can establish multienhancer networks (89), although the detailed nature of the loops within these networks remains to be established. Importantly, LDB1- and GAF-mediated loops do not involve CTCF

insulator protein, cohesin, or loop extrusion (83, 89) and hence are distinct from the loops involved in topologically associating domains (90–92).

Activators bound to distal enhancers do not form transcription-competent loops by interacting with Mediator or other general transcription factors that associate with promoters. In accord with this conclusion, looping is subtly affected on a genomic scale only when Pol II transcription is eliminated via inactivation of Mediator, TFIID, or Pol II (93–96). In contrast, loss of LDB1 drastically reduces a subset of loops, indicating that LDB1 is critical for some enhancer-mediated loops (89). Thus, looping to proximal enhancers brings activator proteins bound at distal enhancers close to the promoter, whereupon their activation domains can mediate short-range interactions with Mediator (or TFIID). Importantly, protein–protein interactions that mediate looping are distinct from those mediating transcriptional activation, although both types of interactions are critical for enhancer action at a distance (**Figure 3**).

POTENTIAL MECHANISMS FOR ACTIVATION AT A DISTANCE

Long-distance activation of a promoter requires that activators bound at distal enhancers stimulate transcriptional activity beyond the level achieved by activators bound at proximal enhancers alone. At least three non–mutually exclusive mechanisms could be imagined. First, distally and proximally bound activators might synergistically activate transcription via the presence of multiple activation domains near the promoter (**Figure 5*a***). Second, proteins like GAF, when bound near proximal promoters, recruit chromatin-modifying enzymes but lack activation domains to recruit Mediator, thereby requiring looped activators bound to (or, in the case of GAF, near) the distal enhancer to mediate transcriptional activation (**Figure 5*b***). Third, looping interactions could permit cooperative binding of activators to their separated target sites (**Figure 5*c***). As enhancers are typically bound by multiple proteins, looping interactions may or may not involve proteins that interact with Mediator (or TFIID) and mediate transcriptional activation per se.

THE DIFFERENCE BETWEEN YEAST AND METAZOAN ENHANCERS REFLECTS EVOLUTIONARY CHOICE, NOT A DIFFERENCE IN ACTIVATION MECHANISM

Initial descriptions indicated that metazoan enhancers could activate transcription from long and variable distances upstream or downstream of promoters (22, 23, 25), whereas yeast enhancers could not (67, 68). This dichotomy led to a long-standing view that yeast and metazoan enhancers are fundamentally different. However, creating an artificial loop in yeast via GAF or other methods permits activation at longer distances than usual and at locations downstream of the promoter (86), suggesting that yeast and metazoan enhancers are mechanistically similar. In this view, the failure of natural yeast enhancers to activate transcription from long and variable distances is due to the absence of looping domains in natural yeast activators. This evolutionary choice for yeast activators to lack looping domains may reflect the shorter distances between functionally unrelated genes in yeast as well as the lower regulatory and biological complexity of a unicellular organism.

THE MOLECULAR NATURE OF INTERACTIONS BETWEEN PROXIMAL AND DISTAL ENHANCERS

Long-distance activation by enhancers is of fundamental importance, and DNA loops involving enhancers have been identified on a genomic scale in numerous studies. It is therefore remarkable that there is very limited information on the molecular interactions that mediate enhancer-dependent DNA loops. The LDB1 coactivator protein interacts with LMO1, a protein associated with, but not directly bound to, both the proximal enhancer and the distal LCR/superenhancer

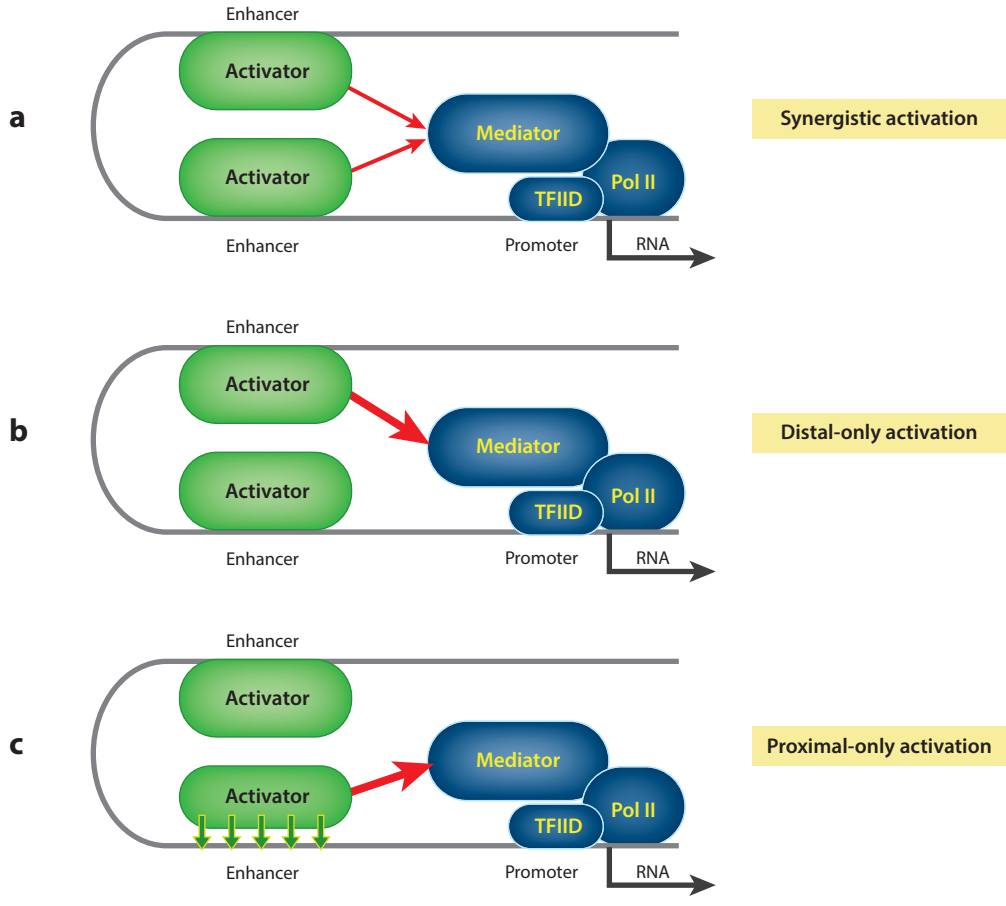


Figure 5

Models for transcriptional activation at a distance. Activator proteins bound at proximal and distal enhancers form a loop bringing the distal enhancer close to the promoter and the Mediator and TFIID components of the RNA polymerase II (Pol II) machinery. (*a*) Synergistic activation occurs when proximally and distally bound activators contact the Pol II machinery via their activation domains. (*b*) Distal-only activation is mediated primarily or solely by the activation domain of the distally bound activator. The proximally bound activator presumably recruits chromatin-modifying activities but contributes little via direct contact to the Pol II machinery. (*c*) Proximal-only activation is mediated primarily or solely by the proximally bound activator. The distal enhancer helps recruit the activator proteins to the proximal enhancer.

(66). Loss of GATA-1 results in selective loss of LDB1 at the proximal enhancer and transcriptional activation of the β-globin promoter. The looping interaction is mediated by the LDB1 SA domain, which generates higher-order complexes. Thus, even though LDB1 and LMO1 do not directly bind DNA, the SA domain is analogous to the tetramerization domains of bacterial proteins that mediate DNA loops. LDB1, through its interactions with the pioneer factor Zelda, also has a role in early embryogenesis and zygotic gene activation in *Drosophila* (97).

As mentioned earlier, *Drosophila* GAF contains an N-terminal BTB/POZ oligomerization domain that is important for long-range interactions (82, 83, 85). However, unlike LDB1, GAF does not appear to bind enhancers but rather tethering elements that are unable to directly activate transcription. GAF and Vostok, which binds GCAACA sequences, affect a largely nonoverlapping subset of looping interactions in the *Drosophila* brain (98). DNA looping via a multimerization

domain can also be mediated by FOXP3, a transcription factor that binds DNA through a forkhead domain (99). The atomic structure of FOXP3-mediated loops has been described (99), but the role of these loops in transcriptional regulation has yet to be clarified.

As mentioned above, protein–protein interactions that mediate looping are separate from those mediating transcriptional activation. This separation is obvious for *Drosophila* tethering elements in which GAF mediates the looping required for transcriptional activation but not transcriptional activation per se. It is unclear whether LDB1 contains an activation domain or whether activation is mediated by a different protein (or proteins) associated with the proximal and distal enhancers. A key distinction between looping and activation domains is that looping interactions occur over large distances whereas enhancer–promoter interactions are only short range. This observation suggests that interactions between activation domains and Mediator (or TFIID) are weaker than those mediated by the oligomerization domains of LDB1 or GAF. In this regard, the eukaryotic oligomerization domains may permit longer-distance loops than bacterial looping domains that do not appear to oligomerize.

It is virtually certain that long-distance activation involves DNA loops mediated by interactions involving proteins associated with proximal and distal enhancers. Nevertheless, there are several classes of looping mechanisms. First, by analogy to loops in bacteria, activator proteins directly bound to proximal and distal enhancers could interact via domains that mediate tetramerization or higher-order oligomerization (**Figure 6*a***). Second, looping could involve heterotypic interactions between two distinct proteins (**Figure 6*b***). Heterotypic looping interactions could contribute to specificity between proximal and distal enhancers and serve as a new mechanism that contributes to the diversity of gene regulatory patterns. Third, as for LDB1, loops could involve oligomerization domains of proteins associated with enhancer sequences but not directly bound to them. Such oligomerization domains could link multiple genomic regions into a specific hub (**Figure 6*c***). In a related mechanism, proteins like GAF can bind to sequences near promoters and distal enhancers, thereby creating a conceptually similar loop that permits activation by the distal enhancer. Fourth, looping interactions could involve a combined surface or separate interactions of multiple proteins associated with the distal and/or proximal enhancer. Fifth, the critical interactions may not be simple adhesive surfaces but rather may involve condensate-forming domains (**Figure 6*d***). Such condensates would have the potential to link many separate genomic regions to a hub in which key transcriptional regulatory proteins are in high concentration, thereby increasing transcriptional activation. At present, it is generally unknown (except for LDB1 and GAF) which of these mechanisms are used at individual genes and which predominate.

While looping between proteins associated with enhancers is the only plausible mechanism to explain the specificity of gene regulation at a distance, the dynamics of loop formation and the relationship to transcriptional activation are unclear. Although typically portrayed as stable loops (e.g., **Figure 5**), there is accumulating evidence that these loops are highly dynamic and relatively infrequent. In addition, the temporal relationship between loop formation and transcriptional activation is controversial, and it may differ among the enhancers and promoters involved.

"WHAT IS TRUE FOR *E. COLI* IS TRUE FOR THE ELEPHANT, ONLY MORE SO"

What is true for *E. coli* and the elephant? Gene regulatory information is ultimately provided by genetic elements composed of short DNA sequences that are specifically recognized by transcriptional regulatory proteins. There are three types of regulatory elements: promoters recognized by the basic transcription machinery, operators recognized by repressor proteins, and enhancers recognized by activator proteins. Repressors can inhibit transcription by directly blocking the binding of the basic transcription machinery or activators. A crucial aspect of activator function

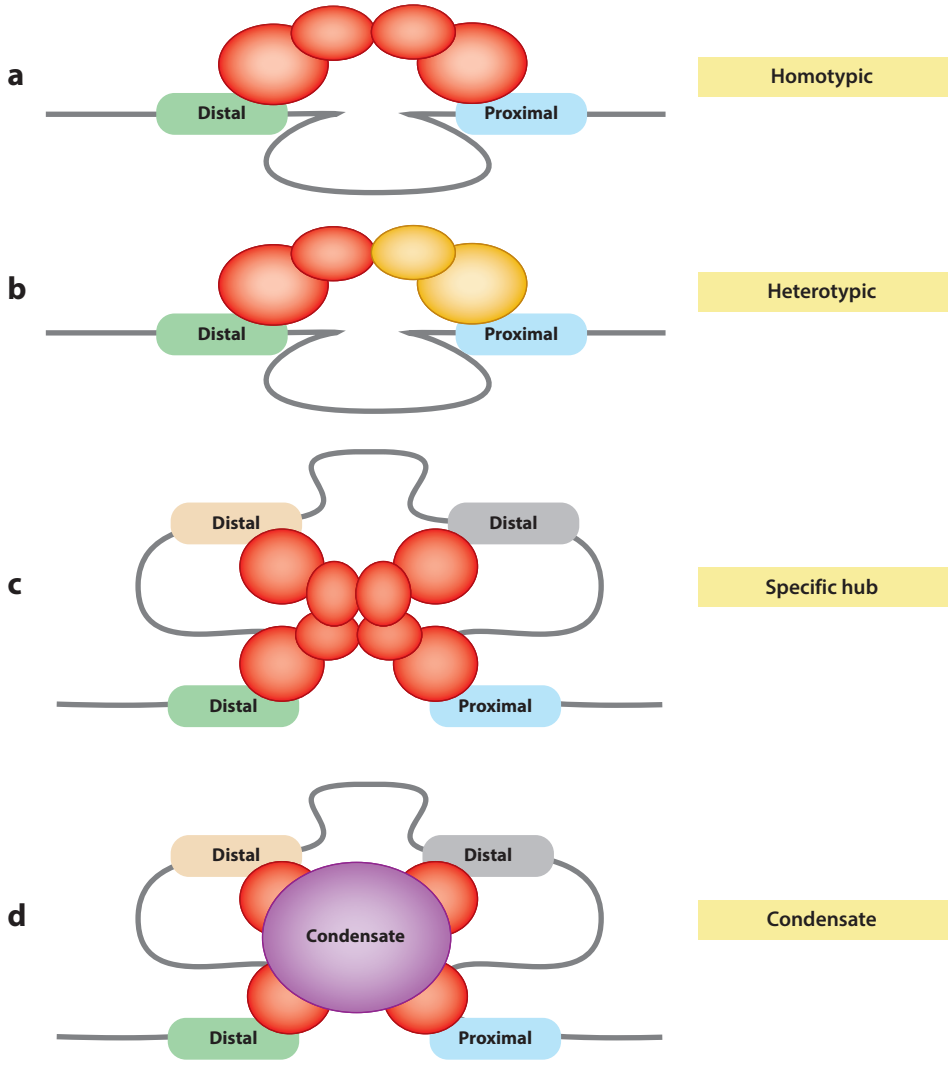


Figure 6

Models for DNA looping. (*a*) Homotypic interactions involve adhesive surfaces between different molecules of the same protein, such as those mediated by the λ repressor tetramerization domain. (*b*) Heterotypic interactions involve adhesive surfaces between different proteins. (*c*) Specific hubs involve oligomeric interactions via adhesive surfaces of proteins (e.g., LDB1) bound to multiple enhancers. (*d*) Condensates involve multivalent, weak interactions between proteins bound to multiple enhancers.

is a direct and relatively short-range interaction with the basic transcription machinery, which is mediated by a protein surface distinct from that mediating DNA binding. DNA looping is required for action at a distance, and it occurs via interactions between proteins (either activators or repressors) associated with spatially separate sites. Transcriptional regulation is not mediated through promoters and the basic transcription machinery but rather through functional changes in activators or repressors in response to different physiological conditions. Such changes can involve interactions with small molecules or protein modifications that alter (either positively or negatively) functional activity.

How are elephants (and by extension all eukaryotes including unicellular yeasts) "more so" than *E. coli*? An incomplete list of differences follows below.

1. Eukaryotic genomes are packaged into chromatin, which affects the fundamental logic of gene regulation (20) (see the previous section). Activators, repressors, and the Pol II elongation machinery, via recruitment of chromatin-modifying activities, are ultimately responsible for histone modifications and DNA methylation. This leads to ever-changing epigenomes that have profound effects on gene regulation.
2. The eukaryotic transcription machinery is not merely RNA polymerase but rather a remarkably large entity (>50 proteins) that includes general transcription factors. Different general factors, not Pol II itself, mediate specific binding to promoter sequences, promoter melting, and strand separation, unlike *E. coli* RNA polymerase, which mediates all these activities.
3. Unlike bacterial RNA polymerases, Pol II contains a C-terminal tail that is differentially phosphorylated during initiation, elongation, and termination. This phosphorylation is required for cotranscriptional recruitment of factors that mediate mRNA capping, splicing, polyadenylation, and histone methylation at specific lysine residues.
4. Transcriptional activators interact directly with coactivators, not Pol II itself. Some coactivators (Mediator and TFIID) are components of the basic Pol II machinery, while others have chromatin-modifying activities.
5. Enhancers in metazoans (but not yeast) require multiple activators that bind to nearby, but distinct, sites that function combinatorially and synergistically.
6. DNA looping in eukaryotes occurs over much longer distances and, to the extent known, does not involve bimolecular interactions between DNA-binding activators or repressors via tetramerization domains. Instead, looping involves interactions between multimerization domains associated with, but not directly bound to, distal and proximal enhancers. Compared to bimolecular interactions, multimerization domains likely mediate stronger interactions that are perhaps important for looping over long distances. Though not yet demonstrated, heterotypic interactions between different proteins or condensate-forming domains are plausible mechanisms for looping.

Thus, the fundamental concepts of gene regulation appear to be universal, but the "more so" differences between elephants (and more generally eukaryotes) and *E. coli* underlie the greatly increased complexity of gene regulatory patterns that is essential for the extraordinary biological diversity on earth.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank Steve Buratowski for many useful comments on the manuscript. This work was supported by grant GM131801 to K.S. from the National Institutes of Health up until the time it was terminated by the Trump administration.

LITERATURE CITED

1. Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3:318–56
2. Jacob F, Ullman A, Monod J. 1964. The promoter, a genetic element necessary for expression of an operon. *C.R. Hebd. Seances Acad. Sci.* 258:3125–28

3. Scaife J, Beckwith JR. 1966. Mutational alteration of the maximal level of *Lac* operon expression. *Cold Spring Harb. Symp. Quant. Biol.* 31:403–8
4. Feklistov A, Sharon BD, Darst SA, Gross CA. 2014. Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* 68:357–76
5. Jacob F, Perrin D, Sanchez C, Monod J. 1960. Operon: a group of genes with the expression coordinated by an operator. *C.R. Hebd. Seances Acad. Sci.* 250:1727–29
6. Ptashne M. 1967. Specific binding of the λ phage repressor to λ DNA. *Nature* 214:232–34
7. Gilbert W, Muller-Hill B. 1967. The lac operator is DNA. *PNAS* 58:2415–21
8. Englesberg E, Irr J, Power J, Lee N. 1965. Positive control of enzyme synthesis by gene C in the L-arabinose system. *J. Bacteriol.* 90:946–57
9. Sheppard DE, Englesberg E. 1967. Further evidence for positive control of the L-arabinose system by gene *araC*. *J. Mol. Biol.* 25:443–54
10. Chen W, Struhl K. 1985. Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. *EMBO J.* 4:3273–80
11. Smale ST, Baltimore D. 1989. The "initiator" as a transcription control element. *Cell* 57:103–13
12. Struhl K. 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* 49:295–97
13. Englesberg E, Squires C, Meronk F Jr. 1969. The L-arabinose operon in *Escherichia coli* B-r: a genetic demonstration of two functional states of the product of a regulator gene. *PNAS* 62:1100–7
14. Hahn S. 2014. Ellis Englesberg and the discovery of positive control in gene regulation. *Genetics* 198:455–60
15. Schleif R. 2010. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiol. Rev.* 34:779–96
16. Tanay A. 2006. Extensive low-affinity transcriptional interactions in the yeast genome. *Genome Res.* 16:962–72
17. Struhl K. 2007. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* 14:103–5
18. Gvozdenov Z, Barcutean Z, Struhl K. 2023. Functional analysis of a random-sequence chromosome reveals a high level and the molecular nature of transcriptional noise in yeast. *Mol. Cell* 83:1786–97
19. Camellato BR, Brosh R, Ashe HJ, Maurano MT, Boeke JD. 2024. Synthetic reversed sequences reveal default genomic states. *Nature* 628:373–80
20. Struhl K. 1999. Fundamentally different logic of gene expression in eukaryotes and prokaryotes. *Cell* 98:1–4
21. Wade JT, Reppas NB, Church GM, Struhl K. 2005. Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites. *Genes Dev.* 19:2619–30
22. Banerji J, Rusconi S, Schaffner W. 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27:299–308
23. Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP, Chambon P. 1981. The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* 9:6047–68
24. Fromm M, Berg P. 1982. Deletion mapping of DNA regions required for SV40 early promoter function in vivo. *J. Mol. Appl. Genet.* 1:457–81
25. Banerji J, Olson L, Schaffner W. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33:729–40
26. Chandler VL, Maler BA, Yamamoto KR. 1983. DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. *Cell* 33:489–99
27. Gillies SD, Morrison SL, Oi V, Tonegawa S. 1983. A tissue-specific enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33:717–28
28. Struhl K. 1991. Mechanisms for diversity in gene expression patterns. *Neuron* 7:177–81
29. Levine M. 2010. Transcriptional enhancers in animal development and evolution. *Curr. Biol.* 20:R754–63
30. Core LJ, Martins AL, Danko CG, Waters CT, Siepel A, Lis JT. 2014. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* 46:1311–20

31. Henriques T, Scruggs BS, Inouye MO, Muse GW, Williams LH, et al. 2018. Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev*. 32:26–41
32. Mikhaylichenko O, Bondarenko V, Harnett D, Schor IE, Males M, et al. 2018. The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev*. 32:42–57
33. Iyer V, Struhl K. 1995. Mechanism of differential utilization of the *his3* T_R and T_C TATA elements. *Mol. Cell. Biol.* 15:7059–66
34. Utley RT, Ikeda K, Grant PA, Cote J, Steger DJ, et al. 1998. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* 394:498–502
35. Kingston RE. 1999. A shared but complex bridge. *Nature* 399:199–200
36. Cosma MP, Tanaka T, Nasmyth K. 1999. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle– and developmentally regulated promoter. *Cell* 97:299–311
37. Hughes A, Jin Y, Rando OJ, Struhl K. 2012. A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. *Mol. Cell* 48:5–15
38. Jin Y, Eser U, Struhl K, Churchman LS. 2017. The ground state and evolution of promoter regions directionality. *Cell* 170:889–98
39. Ng HH, Robert F, Young RA, Struhl K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11:709–19
40. Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, et al. 2003. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1. Linking transcriptional elongation to histone methylation. *Mol. Cell* 11:721–29
41. Struhl K. 2024. The distinction between epigenetics and epigenomics. *Trends Genet*. 40:995–97
42. Struhl K. 2024. Non-canonical function of enhancers: regulation of RNA polymerase III transcription, DNA replication, and V(D)J recombination. *Trends Genet*. 40:471–79
43. Moqtaderi Z, Wang J, Raha D, White RJ, Snyder M, et al. 2010. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat. Struct. Mol. Biol.* 17:635–40
44. Oler AJ, Cairns BR. 2010. Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat. Struct. Mol. Biol.* 17:620–28
45. Gellert M. 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu. Rev. Biochem.* 71:101–32
46. Miotto B, Ji Z, Struhl K. 2016. Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancer. *PNAS* 113:E4810–19
47. Raha D, Wang Z, Moqtaderi Z, Wu L, Zhong G, et al. 2010. Close association of RNA polymerase II and many transcription factors with Pol III genes. *PNAS* 107:3639–44
48. Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, et al. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450:1106–10
49. Reeder T, Schleif R. 1993. AraC protein can activate transcription from only one position and when pointed in only one direction. *J. Mol. Biol.* 231:205–18
50. Hochschild A. 1994. Transcriptional activation. How λ repressor talks to RNA polymerase. *Curr. Biol.* 4:440–42
51. Busby S, Ebright RH. 1999. Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* 293:199–213
52. Reitzer LJ, Magasanik B. 1986. Transcription of *glnA* in E. coli is stimulated by activator bound to sites far from the promoter. *Cell* 45:785–92
53. Ninfa AJ, Reitzer LJ, Magasanik B. 1987. Initiation of transcription at the bacterial *glnAp2* promoter by purified E. coli components is facilitated by enhancers. *Cell* 50:1039–46
54. Joung JK, Koepp DM, Hochschild A. 1994. Synergistic activation of transcription by bacteriophage λ cI protein and *E. coli* cAMP receptor protein. *Science* 265:1863–66
55. Malik S, Roeder RG. 2023. Regulation of the RNA polymerase II pre-initiation complex by its associated coactivators. *Nat. Rev. Genet.* 24:767–82
56. Allen BL, Taatjes DJ. 2015. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell. Biol.* 16:155–66

57. Plaschka C, Lariviere L, Wenzeck L, Seizl M, Hemann M, et al. 2015. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* 518:376–80
58. Robinson PJ, Trnka MJ, Pellarin R, Greenberg CH, Bushnell DA, et al. 2015. Molecular architecture of the yeast Mediator complex. *eLife* 4:e08719
59. Jeronimo C, Langelier MF, Bataille AR, Pascal JM, Pugh BF, Robert F. 2016. Tail and kinase modules differently regulate core Mediator recruitment and function in vivo. *Mol. Cell* 64:455–66
60. Petrenko N, Jin Y, Wong K-H, Struhl K. 2016. Mediator undergoes a compositional change during transcriptional activation. *Mol. Cell* 64:443–54
61. Petrenko N, Struhl K. 2021. Comparison of transcriptional initiation by RNA polymerase II across eukaryotic species. *eLife* 10:e67964
62. Mencia M, Moqtaderi Z, Geisberg JV, Kuras L, Struhl K. 2002. Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast. *Mol. Cell* 9:823–33
63. Li XY, Bhaumik SR, Zhu X, Li L, Shen WC, et al. 2002. Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure. *Curr. Biol.* 12:1240–44
64. Chen WY, Zhang J, Geng H, Du Z, Nakadai T, Roeder RG. 2013. A TAF4 coactivator function for E proteins that involves enhanced TFIID binding. *Genes Dev*. 27:1596–609
65. Nolis IK, McKay DJ, Mantouvalou E, Lomvardas S, Merika M, Thanos D. 2009. Transcription factors mediate long-range enhancer–promoter interactions. *PNAS* 106:20222–27
66. Deng W, Lee J, Wang H, Miller J, Reik A, et al. 2012. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149:1233–44
67. Struhl K. 1984. Genetic properties and chromatin structure of the yeast *gal* regulatory element; an enhancer-like sequence. *PNAS* 81:7865–69
68. Guarente L, Hoar E. 1984. Upstream activation sites of the *cyc1* gene of *Saccharomyces cerevisiae* are active when inverted but not when placed downstream of the "TATA box." *PNAS* 81:7860–64
69. Brent R, Ptashne M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43:729–36
70. Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, et al. 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res*. 16:1299–309
71. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326:289–93
72. Irani MH, Orosz L, Adhya S. 1983. A control element within a structural gene: the *gal* operon of Escherichia coli. *Cell* 32:783–88
73. Semsey S, Geanacopoulos M, Lewis DE, Adhya S. 2002. Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *EMBO J.* 21:4349–56
74. Choy HE, Adhya S. 1992. Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *PNAS* 89:11264–68
75. Hochschild A, Ptashne M. 1986. Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. *Cell* 44:681–87
76. Dunn T, Hahn S, Ogden S, Schleif R. 1984. An operator at −280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *PNAS* 81:5017–20
77. Huo L, Martin KJ, Schleif R. 1988. Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. *PNAS* 85:5444–48
78. Hao N, Shearwin KE, Dodd IB. 2019. Positive and negative control of enhancer-promoter interactions by other DNA loops generates specificity and tunability. *Cell Rep*. 26:2419–33.e3
79. Su W, Porter S, Kustu S, Echols H. 1990. DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. *PNAS* 87:5504–8
80. Struhl K. 2025. Distal enhancers loop to proximal enhancers, not to promoters. *Nat. Rev. Mol. Cell. Biol.* 26:730–31
81. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, et al. 2013. Super-enhancers in the control of cell identity and disease. *Cell* 155:934–47

82. Li X, Levine M. 2024. What are tethering elements? *Curr. Opin. Genet. Dev.* 84:102151
83. Batut PJ, Bing XY, Sisco Z, Raimundo J, Levo M, Levine MS. 2022. Genome organization controls transcriptional dynamics during development. *Science* 375:566–70
84. Judd J, Duarte FM, Lis JT. 2021. Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. *Genes Dev.* 35:147–56
85. Li X, Tang X, Bing X, Catalano C, Li T, et al. 2023. GAGA-associated factor fosters loop formation in the *Drosophila* genome. *Mol. Cell* 83:1519–26.e4
86. Petrascheck M, Escher D, Mahmoudi T, Verrijzer CP, Schaffner W, Barberis A. 2005. DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res.* 33:3743–50
87. Jensen CL, Chen LF, Swigut T, Crocker OJ, Yao D, et al. 2025. Long-range regulation of transcription scales with genomic distance in a gene-specific manner. *Mol. Cell* 85:347–61
88. Thomas HF, Feng S, Haslhofer F, Huber M, Garcia Gallardo M, et al. 2025. Enhancer cooperativity can compensate for loss of activity over large genomic distances. *Mol. Cell* 85:362–75
89. Aboreden NG, Lam JC, Goel VY, Wang S, Wang X, et al. 2025. LDB1 establishes multi-enhancer networks to regulate gene expression. *Mol. Cell* 85:376–93
90. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, et al. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376–80
91. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2016. Formation of chromosomal domains by loop extrusion. *Cell Rep.* 15:2038–49
92. Hansen AS. 2020. CTCF as a boundary factor for cohesin-mediated loop extrusion: evidence for a multi-step mechanism. *Nucleus* 11:132–48
93. Sun F, Sun T, Kronenberg M, Tan X, Huang C, Carey MF. 2021. The Pol II preinitiation complex (PIC) influences Mediator binding but not promoter–enhancer looping. *Genes Dev.* 35:1175–89
94. Ramasamy S, Aljahani A, Karpinska MA, Cao TBN, Velychko T, et al. 2023. The Mediator complex regulates enhancer-promoter interactions. *Nat. Struct. Mol. Biol.* 30:991–1000
95. Barshad G, Lewis JJ, Chivu AG, Abuhashem A, Krietenstein N, et al. 2023. RNA polymerase II dynamics shape enhancer–promoter interactions. *Nat. Genet.* 55:1370–80
96. Zhang S, Ubelmesser N, Barbieri M, Papantonis A. 2023. Enhancer–promoter contact formation requires RNAPII and antagonizes loop extrusion. *Nat. Genet.* 55:832–40
97. Galouzis CC, Kherdjemil Y, Forneris M, Viales RR, Marco-Ferreres R, Furlong EEM. 2025. Chip (Ldb1) is a putative cofactor of Zelda forming a functional bridge to CBP during zygotic genome activation. *Mol. Cell* 85:2425–41.e9
98. Hu J, Li X, Lomaev D, Vorobyeva NE, Levine M, et al. 2025. Vostok: a looping factor for the organization of the regulatory genome in the *Drosophila* brain. *Mol. Cell* 85:2442–51.e5
99. Zhang W, Leng F, Wang X, Ramirez RN, Park J, et al. 2023. FOXP3 recognizes microsatellites and bridges DNA through multimerization. *Nature* 624:433–41