Introduction to Theme
“Chromatin, Epigenetics, and Transcription”

Joan W. Conaway
Stowers Institute for Medical Research, Kansas City, Missouri 64110;
email: JLC@stowers.org

Abstract
Transcriptional regulation in eukaryotes depends on a complex network of interactions between RNA polymerases and a host of transcription factors and coregulators that control their activity during transcription initiation and elongation. Among these are an enormous variety of enzymes and proteins that modulate chromatin structure via changes in DNA methylation, histone modifications, and nucleosome location. This volume of the Annual Review of Biochemistry contains a set of four reviews addressing the interplay between mechanisms that regulate DNA methylation, chromatin structure, and transcription.

Keywords
DNA methylation, noncoding RNA, histone modification, RNA polymerase II, elongation
The development of a multicellular organism from a fertilized egg requires proper execution of a transcription program that results in synthesis of many thousands of messenger RNAs (mRNAs), at exactly the right time, in the right set of cells, and in the right amount. Misregulation of the transcriptional program is associated not only with developmental defects but also with many disease states. Accordingly, major goals of research over the past several decades have been to define the components of the transcriptional regulatory apparatus and to work out the details of how they function together to attain the exquisitely precise control of the transcription output required for normal development and function. This work has led over time to the revelation that regulation of eukaryotic mRNA synthesis is a multi-tiered process governed by an intricate interplay between chromatin structure and the RNA polymerase II (Pol II) transcription machinery. Chromatin structure is controlled locally, at the gene level, and globally through the action of a withering collection of enzymes and proteins that regulate both the methylation status of chromosomal DNA and the distribution and modification of nucleosomes along it. Like the cell’s machinery for regulating chromatin structure, the RNA Pol II transcription machinery is elaborate and composed of a large collection of transcription factors that control not only transcript initiation, but also transcript elongation through chromatin. This volume of the Annual Review of Biochemistry includes four articles addressing the interplay between chromatin structure and RNA Pol II transcription.

At its most fundamental level, the transcription of a gene by any RNA polymerase can be described in terms of a transcription cycle that can be divided into multiple, mechanistically distinct stages. In the first stage, the enzyme must be able to identify, bind to, and initiate transcription correctly at promoters in the background of the large amount of nonpromoter DNA in the cell. Having initiated transcription, an RNA polymerase must elongate the nascent transcript until it reaches the gene’s 3’ end, a process that in higher eukaryotes can require it to traverse hundreds of kilobases before it terminates transcription and is released from the gene.

For many years, it was assumed that the majority of Pol II transcriptional regulation was accomplished at the earliest stages of the transcription cycle by processes that affect the efficiency with which the enzyme is recruited to promoters and initiates transcription; however, hints that transcript elongation could also be a key site for regulation came from evidence that most Pol II elongation in vitro or in cells is sensitive to inhibition by the protein kinase inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole or DRB (1, 2). Subsequently, seminal studies of heat shock gene expression in Drosophila and of c-myc, c-fos, and other oncogenes in human cells revealed that transcription of some genes can be regulated by promoter-proximal pausing in which Pol II pauses shortly after initiating transcription and is released into productive elongation only upon gene activation (3–5). Over the next decade or so, biochemical studies identified a collection of transcription elongation factors, including negatively acting factors such as DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF), which induce transcriptional pausing, and positively acting factors, such as the protein kinase P-TEFb and eleven-nineteen lysine rich in leukemia (ELL), which promote efficient elongation. With the advent of genome-wide methods for measuring nascent transcription synthesis as well as the location of Pol II and elongation factors on chromatin, it has become clear that control at the level of transcript elongation is a general feature of Pol II transcription and plays a critical role in development and disease. In their review, entitled “RNA Polymerase II Elongation Control,” Zhou and colleagues (6) discuss our current understanding of the mechanisms and transcription factors that control transcript elongation and associated processes, such as RNA capping, splicing, and polyadenylation.

In eukaryotic cells, chromosomal DNA is packaged into arrays of nucleosomes, each of
which contains ~146 bp of DNA wrapped around a histone octamer containing two copies each of histones H2A, H2B, H3, and H4 (7). These arrays are, in turn, folded into chromatin fibers made up of higher-order nucleosomal structures, allowing the several meters of DNA that make up the human genome to fit into nuclei with diameters of just a few microns. Although this compaction presents a serious impediment to the RNA Pol II transcription machinery, as well as to enzymes involved in other nuclear processes such as DNA replication and repair, it also provides remarkable opportunities for gene regulation. Indeed, it is now clear that establishment and maintenance of distinct patterns of gene expression in different cell lineages are due in significant part to epigenetic mechanisms that involve alterations in chromatin structure driven in part by post-translational modifications of histones and by DNA methylation.

Regions of transcriptionally silenced chromosomal DNA are enriched in CpG dinucleotides containing cytosine residues methylated at position 5 (8). In addition, the chromatin of silenced regions contains high levels of histone H3 dimethylated on lysine 9 (H3K9me2) and histone H3 trimethylated on lysine 27 (H3K27me3) (9). In contrast, transcriptionally active regions are associated with hypomethylated DNA and with chromatin with a different constellation of histone marks, including histone H4 trimethylated on lysine 4 (H3K4me3) (9). Two reviews in this series, “Programming of DNA Methylation Patterns” by Cedar & Bergman (10), and “The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis,” by Shilatifard (11), review our current understanding of how these epigenetic marks are established, maintained, and function to control chromatin structure and gene regulation.

In the final review in the series, “Genome Regulation by Long Noncoding RNAs,” Rinn & Chang (12) describe the identification and functional analysis of a large set of long noncoding RNAs (lncRNAs), at least some of which are proving to have roles in establishing gene expression patterns important for stem cell maintenance, cell fate determination during development, and diseases such as cancer. Until recently, it was thought that the vast majority of higher eukaryotic genomes are made up of nontranscribed, so-called junk DNA that lacks known protein coding sequences and has no clear function. With the advent of high-throughput methods for analyzing the entire population of RNAs present in a given cell type, however, it has become clear that a much larger fraction of the genome is transcribed than previously appreciated. Indeed, mammalian cells express thousands of lncRNAs that are encoded by previously unannotated regions of the genome. That these lncRNAs often exhibit sequence conservation across species suggested that they have some function linked to their structures (13, 14), and the challenges are now to determine which of them contribute to regulatory processes and how. In their review, Rinn & Chang discuss emerging evidence that lncRNAs can assemble into ribonucleoprotein complexes and contribute to gene regulation by mechanisms almost as diverse as those employed by more conventional protein regulators. Some lncRNAs serve as decoys that mimic the DNA-binding sites of transcriptional regulators and prevent them from binding to their normal targets in promoter DNA. Among these are the lncRNA Gas5, which binds the the glucocorticoid receptor, and another called PANDA, which binds the transcription factor NF-Y. Others can serve as scaffolds or adapters to promote interactions between multiple proteins or protein complexes that regulate chromatin structure. For example, an lncRNA called HOTAIR can bind to both the histone H3 K27 methyltransferase polycomb repressive complex 2 (PRC2) and the histone H3 K4 demethylase LSD1-CoREST to coordinate a repressive chromatin methylation state. In addition, lncRNAs can serve as guides that target their protein partners to specific chromosomal locations, much like the DNA-binding domain of a sequence-specific DNA-binding protein. Thus, HOTAIR as well
as a host of other lncRNAs have been shown to bind PRC2, and in several cases, these have been shown to contribute to PRC2 recruitment and histone H3K27 methylation at specific loci.

A few decades ago, we knew virtually nothing about the enzymes and proteins that coordinate epigenetic modifications of chromatin and DNA with transcription or about how their functions, and malfunctions, contribute to normal development and disease. Taken together, these four reviews highlight just how much has been learned. But as they also make clear, each new answer raises new and unexpected questions, and research in this area should be fruitful for many years to come.

**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.

**LITERATURE CITED**