Transcriptional regulation in the context of chromatin structure

Alan P. Wolfe

Laboratory of Molecular Embryology, National Institute of Child Health and Human Development, NIH, Bethesda, MD 20892-5431, U.S.A.

Introduction

Biophysicists and biochemists have often focused on chromatin structure simply as an engineering problem. The highly conserved histones are thought to wrap DNA around themselves to assemble monotonous arrays of nucleosomes, which in turn self-associate to form semi-crystalline higher-order structures. This is very convenient for scientists who require structurally uniform samples to interpret their experiments on the physical properties of biological material. However, this homogeneous packaging of DNA within the chromosome is not consistent with the experience of those molecular geneticists and cell biologists who investigate chromatin function. These investigators find the organization of DNA within the chromosome to be highly varied, reflecting the functional differentiation of the chromosome into distinct domains.

Recent observations provide a molecular explanation that reconciles the dual requirements of packaging of DNA into chromatin via nucleosomal structures and the functional differentiation of the chromosome into distinct

1Address for correspondence: Sangamo BioSciences Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, CA 94804, U.S.A. (e-mail: awolffe@sangamo.com).
domains. The assembly of variants of the histones themselves into chromatin, and the inclusion of specific trans-acting factors with striking similarities to the histones, can retain some of the architectural features of canonical nucleosomes, yet provide highly differentiated structures necessary to facilitate transcription. Moreover, chromatin structure is dynamic, accommodating the various functions that require DNA as a template. In the past, changes in chromatin conformation were usually attributed to DNA being used for some purpose, but it is now clear that alterations in chromatin conformation can have a causal role in gene control. Recent progress in the identification and characterization of chromatin-modifying proteins, particularly acetyltransferases, deacetylases and members of the SWI/SNF (mating type switching/sucrose non-fermenting in *Saccharomyces cerevisiae*) superfamily of proteins, provides mechanistic insight into exactly how chromatin remodelling is targeted to regulate transcription. Evolution has been remarkably successful in shaping chromatin such that it can become alternately transparent or opaque, facilitating or restricting the access of transcription factors and RNA polymerase to DNA. The molecular mechanisms controlling this access are central to gene regulation.

**Principles of gene regulation in chromatin**

Metazoan genes are regulated by a series of cis-acting elements and trans-acting factors. The cis-acting elements include locus control regions, enhancers and proximal promoter elements (Figure 1a). These are each assembled into precise nucleoprotein architectures that communicate with each other and with the local chromatin environment (see later) and which can promote transcription at several levels (Figure 1b). The formation of chromatin loops including 10–100 kb of DNA can facilitate interactions between locus control regions and enhancers or promoters. Higher-order chromatin structure can promote interactions over the 1–10 kb range, and positioned nucleosomes can allow enhancers or promote proximal regulatory elements to interact with the basal transcriptional machinery. For example, an oestrogen-responsive enhancer in the *Xenopus* vitellogenin promoter lies 300 bp upstream of the TATA box and 180 bp upstream of liver-specific proximal promoter element. Nucleosome assembly on the DNA between the enhancer and the proximal promoter element stimulates transcription more than 100-fold when compared with a non-specific chromatin structure (Figure 2). Histones can also repress transcription by preventing general transcription factors and the TATA box from recruiting RNA polymerase (see later). Sequence-specific transcription factors bound at locus control regions, enhancers and promoters mediate the recruitment of a complex chromatin remodelling machinery that controls the interplay of histones, transcription factors and RNA polymerase at the TATA box through both positive and negative effects.
Core histones and related regulatory proteins

The four core histones (H2A, H2B, H3 and H4) have been highly conserved throughout eukaryotic evolution. Each has a very similar C-terminal domain structure, known as the histone-fold, that directs the formation of specific heterodimers between the histones and also determines the path of DNA in the nucleosome [2,3]. Each histone-fold contains a long central helix bordered on each side by a loop segment and a shorter helix. The long central helix acts as a heterodimerization interface. Dimerization creates three DNA-binding surfaces through the interaction of loop segments at each end of the long central helix and through the juxtaposition of the two short α-helices flanking the N-termini of the long central helix.

Figure 1. Regulation of eukaryotic gene expression

(a) Three types of cis-acting elements exist, arbitrarily defined by their distance from the site of TATA-box-binding protein (TBP) binding at the TATA box. These are locus control regions, enhancers and promoter proximal elements, at which sequence-specific transcription factors (SSTFs) bind. Communication between transcription factors bound at these elements may be facilitated by packaging of the intervening DNA by histones. SSTFs influence the binding of general transcription factors (GTFs), TBP and TBP-associated factors (TAFs). The TAFs, GTFs, and TBP recruit RNA polymerase to the promoter and facilitate transcription. They may also direct the displacement of histones from the TATA box. (b) Several levels of nucleoprotein higher-order structure also allow communication between the trans-acting factors bound to cis-acting sequences.
Figure 2. Role of a static loop created by a nucleosome including part of the *Xenopus* vitellogenin B1 promoter in potentiating transcription

From a specific chromatin template that allows positioned assembly of nucleosomes, transcription increases with increasing numbers of nucleosomes whereas transcription from a non-specific chromatin template in which nucleosomes are not positioned decreases with increasing numbers of nucleosomes. ORU, oestrogen response unit; NRE, negative regulatory element; NF-1, nuclear factor-1; Vit, vitellogenin mRNA; tk, thymidine kinase mRNA; xER, *Xenopus* oestrogen receptor. Reproduced from Schild, C., Claret, F.-X., Wahli, W. & Wolffe, A.P. A nucleosome-dependent static loop potentiates estrogen-regulated transcription from the *Xenopus* vitellogenin B1 promoter *in vitro*. *EMBO J.* 12, 423–433 © (1993), with permission from the European Molecular Biology Organization.
The N-terminal tails of the core histones and the C-terminal tail of histone H2A protrude on the outside of the nucleosome, where they can potentially contact other nucleoprotein complexes, leading to activation or repression of transcription in a promoter-specific manner. The N-terminal tail domains are also the targets of specific enzymes that covalently modify conserved lysine and serine residues. Both the phosphorylation of serine residues and the acetylation of lysine residues are associated with the modulation of transcriptional activity (see later). Ubiquitinylation of the C-terminal tail of H2A is also correlated with transcriptional activation. These modifications can be targeted through unknown mechanisms to particular chromosomal domains.

Several core histone variants exist that have specific alterations from normal histones in both the N-terminal tails and the DNA-binding surfaces of the histone-fold domains. For example, differences in amino acid sequence from the normal somatic H2A are conserved in a particular H2A variant (H2A.Z) from ciliate protozoa (*Tetrahymena*) to humans, and the histone H2A.Z variant in *Drosophila* has been shown to be essential for early development.

The core histones appear to have evolved from a DNA-binding protein that contained only the three α-helices of the histone-fold domain, lacking any tail domains. The archaeabacterial protein, HMf, consists of only the histone-fold domain and has the capacity to wrap DNA around itself within nucleosome-like structures. The eukaryotic core histones have retained this property, but added the capacity of the assembled nucleoprotein complex to interact with other proteins outside the nucleosome through their additional tail domains.

The histone-fold is also found in some transcriptional regulatory proteins including TAFII40 and TAFII60 (TATA-box-binding protein associated factors) and the related CCAAT-box-binding proteins, NF-Y/CFB and HAP2/3/5. These use the histone-fold both to direct protein–protein interactions and to bind to DNA. TAFII40 and TAFII60 exist as a heterodimer in the general transcription factor TFIIID. TFII40 resembles H3 and TAFII60 resembles H4. Both proteins have extended C-terminal tails that interact with other components of TFIIID and transcriptional activators. It has been proposed that TAFII40 and TAFII60 participate in the assembly of nucleosome-like structures, excluding normal histones from the TATA box, yet maintaining DNA in a semi-compacted state competent for transcription [4]. Metazoan NF-Y/CFB and *Saccharomyces cerevisiae* HAP2/3/5 are highly related trimeric transcriptional activators. The evolutionarily conserved peptide sequences of two of the subunits (CBF-A and CBF-C or HAP3 and HAP5) resemble the histone-fold domains of histones H2B and H2A, respectively. These domains are essential for DNA binding in the presence of the third protein (CBF-B or HAP2) that confers sequence specificity.
Deviant nucleosomes — the ‘winged helix’ connection

Linker histones such as histone H1 or H5 contain a structured nucleic acid-binding domain known as the ‘winged helix’. This domain is found in a variety of sequence-specific transcriptional regulators, including the prokaryotic catabolite gene activator protein (CAP) and the eukaryotic hepatocyte nuclear factor 3 (HNF3) protein. The winged helix consists of a bundle of three α-helices attached to a three-stranded anti-parallel β-sheet [5]. HNF3 binds within the major groove of DNA via one of the α-helices, suggesting that the structured domain of the linker histone will contact nucleosomal DNA in the same way (Figure 3). The linker histone contains additional basic N- and C-terminal domains that influence the path of linker DNA between adjacent nucleosomes.

Linker histones, such as histone H1, have been shown to direct the exact positioning of nucleosomes with respect to DNA sequence [6]. This positioning relies on the sequence- and structure-selective recognition of DNA by the linker histone, and protein–protein contacts made between the winged helix domain of the linker histone and the histone-fold domains of the core histones. During *Xenopus laevis* development, early embryonic variants of linker histones are replaced progressively with normal somatic histone H1. The inclusion of histone H1 in nucleosomes directs the positioning of the histone–DNA contacts over key promoter elements and represses transcription [7]. These transitions in chromatin composition and organization contribute directly to establishing the body plan of the embryo.

The mouse serum albumin enhancer exists in the active state within an array of precisely positioned nucleosome-like particles. Specific enhancer-binding factors, including HNF3, are part of the nucleosome-like particles, and HNF3 can actively direct their positioning with respect to DNA sequence [8]. These observations led to the hypothesis that HNF3 replaces linker histones within chromatin containing the serum albumin enhancer, thereby establishing a precise regulatory nucleoprotein architecture (Figure 4).

Transcription acetyltransferases

The synthesis of mRNA requires co-ordination of the activity of many transcription factors and enzymes. Specificity in transcriptional control relies upon the combinatorial binding of sequence-selective transcription factors to regulatory elements flanking the transcription start site (Figure 1). Activation domains within these factors recruit transcriptional co-activators that in turn enhance the activity of the basal transcriptional machinery. The role of co-activators is to integrate the signals from the various sequence-selective factors so that a final level of transcription can be determined. The simplest mechanism by which this might be achieved is for the co-activators to act as a scaffold between sequence-selective factors and the basal machinery [9].
However, many co-activators also function as histone acetyltransferases. This introduces an alternative and potentially more dynamic model.

The Gcn5p–Ada2p–Ada3p complex in yeast is the archetypical transcription acetyltransferase and is a transcriptional co-activator targeted by transcription factors with acidic activation domains. The Gcn5p component of the complex has the capacity to acetylate specific lysines in histones H3 and H4 which are known to be associated with transcriptional activity [10]. Similar enzymic activities have now been found for P/CAF (a human homologue of Gcn5p), which associates with the p300/CBP [cAMP-response-element-binding protein (CREB)-binding protein] co-activator [11], and for p300/CBP

Figure 3. Model for the interaction of the histones with DNA in the nucleosome
This view is of one turn of DNA. For clarity only one molecule each of H2A, H2B and H4 is shown. Two molecules of H3 are shown to indicate the interface between the two (H3, H4) heterodimers. The numbers round the periphery of the diagram show turns of DNA away from the dyad axis. The structured domain of the linker histone (H5) is shown associating with nucleosomal DNA inside the super-helical turns of DNA around the histone octamer. The C- and N-termini of specific histones are shown by C$^{\text{H2B}}$, N$^{\text{H2B}}$, etc.
itself [12]. p300/CBP interacts with a variety of sequence-selective DNA-binding transcription factors, including nuclear hormone receptors, c-Jun/v-Jun, c-Myb/v-Myb, c-Fos and MyoD. A core component of TFIID TAF\textsubscript{II}250 also has histone acetyltransferase activity [13].

Since core histone acetylation greatly facilitates the access of transcription factors to the DNA in a nucleosome [14,15], and transcriptional co-activators are histone acetyltransferases, a model can be proposed for transcriptional regulation in which histone acetylation directs the local destabilization of repressive histone–DNA interactions (Figure 1). Targeted acetylation allows the basal transcriptional machinery to displace nucleosomes, assemble a functional transcription complex and never have to deal with chromatin again. However, a more interesting possibility follows from the discovery that transcriptional

---

**Figure 4. A specialized nucleosome on the mouse serum albumin enhancer**

Two nucleosomes are shown positioned on the enhancer (numbers are relative to the 5' end). The boundaries of micrococcal nuclease digestion are indicated by the brackets. The positions of transcription factor-binding sites are shown, as is the potential site of HNF3 or linker histone H1 interaction with the nucleosomal structures. The helix that interacts with DNA is shaded. C/EBP, (CAAT/enhancer-binding protein; HNF3, hepatocyte nuclear factor3; NF-1, nuclear factor-1. Reproduced from Wolfe, A.P. & Pruss, D. Deviant nucleosomes: the functional specialization of chromatin. Trends Genet. 12, 58–62 © (1996), with permission from Elsevier Science.
regulators that deacetylate the histones exist [16]. This provides a molecular mechanism whereby transcription might be controlled continually. Core histones remain associated with DNA in the vicinity of a promoter in spite of the recruitment of the basal transcriptional machinery [17]. Thus the targeted or general activity of histone deacetylases will tend to return nucleosomes to their repressive configuration. The maintenance of gene activity would therefore require the continued activity of the co-activators as acetyltransferases. In this way transcriptional activity could be modulated continually through variation in chromatin conformation (Figure 5). These observations further emphasize that the eukaryotic transcriptional machinery is not only adapted to function in a chromatin environment, but also has the potential to make use of the packaging of DNA to regulate genes.

The SWI/SNF2 superfamily

Within S. cerevisiae the outcome of the interplay of transcription factors and histones at specific sites within gene promoters is influenced by the products of the SWI1/ADR6, SWI/SNF2, SWI3, SNF5, and SNF6 genes [18]. All five of these proteins are found within a single ‘general activator’ complex required for the transcriptional induction of many yeast genes [18]. Genetic and biochemical studies of these yeast proteins, and their larger eukaryotic homologues, suggest that the general activator complex serves as a molecular machine to help transcription factors overcome the specific repressive effects of nucleosome assembly on transcription.

Figure 5. Transcriptional regulation in chromatin
A major clue into the molecular mechanism by which the general activator complex exerts its function came from a genetic screen for mutations of genes that would allow transcription of the \( HO \) endonuclease gene (involved in yeast mating type switching) in the absence of SWI1 [19]. Two genes, \( SIN1 \) and \( SIN2 \), were identified which, when mutated, led to SWI-independent transcription. Both of the \( SIN \) genes isolated in this way encode components of chromatin. \( SIN1 \) is a highly charged nuclear protein similar to mammalian high-mobility group proteins 1 and 2. A more direct association with nucleosomal structure is found for \( SIN2 \), which encodes histone H3 or H4 [19]. The \( SIN \) mutants in histones H3 and H4 cluster in one \( \beta \)-bridge motif within the heterodimer of H3 and H4. Due to the juxtaposition of two (H3, H4) heterodimers at the dyad axis of the nucleosome, the \( SIN \) mutations have the potential to disrupt histone–DNA interactions involving the central turn of DNA at the dyad axis (Figure 3). This could have a major impact on the integrity of both the nucleosome and higher-order chromatin structures.

The SWI2/SNF2 subunit of the general activator complex has a potent DNA-dependent ATPase activity [20]. Similar properties are associated with some DNA helicases; however, the general activator complex does not possess helicase activity, but may retain the capacity to track along DNA without unwinding the double helix. Such a processive movement may disrupt histone–DNA interactions. It is also possible that the general activator complex might mediate the removal of \( SIN1 \) or facilitate the displacement of histones (H2A, H2B) or (H3, H4), through direct protein–protein or protein–DNA interactions. These will be progressively more difficult, as the stability of their interaction with DNA increases with their position towards the centre of the nucleosome. Biochemical analysis of chromatin following the activation of transcription by a mammalian steroid receptor, in a general-activator-dependent pathway, reveals only depletion of histone H1 [21]. Moreover any disruption of chromatin is rapidly reversible [22]. These results argue strongly against complete displacement of the core histones from DNA. The potential weakening of interaction between histones (H3, H4) and DNA in the \( SIN2 \) mutants may facilitate local sliding of the entire histone octamer relative to regulatory elements, thereby facilitating access of the basal transcriptional machinery without octamer displacement [23]. \( SIN1 \) may interact with linker DNA in yeast, so removal of \( SIN1 \) might facilitate a comparable increase in histone octamer mobility and \( trans \)-factor access.

A major problem for the SWI/SNF family of ATPases has been the lack of proven mechanisms that might target their activities to known genes. However, it has now been shown that the association of the \( Brahma \)-related gene 1/BRG1-associated factor (BRG1–BAF) complex (a mammalian SWI/SNF homologue) with chromatin relies on the ligand-responsive glucocorticoid receptor [24]. The BRG1–BAF complex was known to facilitate transcriptional activation by ligand-bound glucocorticoid receptor, yet the molecular mechanisms by which this would be targeted were unclear. The
BRG1–BAF complex will disrupt nucleosomes \textit{in vitro} and facilitate the binding of transcription factors to their recognition elements within chromatin, independent of any targeted transactivation domains. The recruitment of the BRG1–BAF complex to the glucocorticoid receptor \textit{in vivo} is dependent upon a functional ligand-binding transactivation domain in the receptor. This result provides a direct connection between the \textit{in vitro} chromatin-remodelling activities of the BRG1–BAF complex and the well established restructuring of nucleosomes on the glucocorticoid-responsive mouse mammary tumour virus chromatin used in these studies. It is presently unclear whether specific components of the BRG1–BAF complex make direct contact with the glucocorticoid receptor or if the association is indirect.

Additional support for the targeted association of SWI/SNF family ATPases comes from experiments in \textit{Drosophila} and \textit{Xenopus}. A distinct member of the SWI/SNF family, dMi-2, contributes to the determination of segmental identity during \textit{Drosophila} development \cite{25} by interacting with the Hunchback protein, which in turn binds directly to the regulatory elements of homoeotic genes to repress transcription. Biochemical experiments in \textit{Xenopus} demonstrate that Mi-2 is part of a protein complex containing histone deacetylase \cite{26}. Thus repression of homoeotic genes may involve histone deacetylation.

\textbf{Future prospects}

Histone acetyltransferases and deacetylases are now implicated in the fundamental mechanisms of transcription control. In many instances the proteins with these activities had already been characterized as having important regulatory functions. A focus for current research is to determine the exact consequence of histone acetylation for these specific regulatory functions. It is also important to recognize that these functions are likely to reveal a close relationship with the role in histone acetylation. Our knowledge of these issues is still far from complete; nevertheless the study of regulated histone acetylation has opened a window for visualizing chromatin in action. With respect to the SWI/SNF complex, unsolved questions include: determining exactly how the targeting of the complex to a specific promoter is directed, the nature of the molecular mechanisms directing chromatin disruption and how the process of gene activation is reversed. Resolution of these issues will require considerable progress in determining both the structural and functional role of SIN1 within chromatin, and the structural consequences of the SIN2 mutations in histones H3 and H4.

The discovery of novel structural subunits and regulatable physical interactions between chromatin components and SWI/SNF ATPase family members promises to integrate this interesting family of enzymes into diverse signal transduction pathways. The genomes of yeast and \textit{Caenorhabditis elegans} suggest that these organisms contain 17 and 21 family members, respec-
tively [27]. Many of these proteins will be found in distinct regulatory complexes that may be involved in gene activation, repression or both. It is also undoubtedly true that SWI/SNF ATPases will contribute to chromatin and chromosomal dynamics associated with the many other nuclear events that use DNA as a template.

**Summary**

- A wide variety of histone-like proteins can be assembled into nucleosomal structures.
- Core and linker histone variants, proteins of the histone-fold and winged-helix families can all contribute to the local differentiation of functional chromosomal domains.
- It is very difficult to disrupt core histone interactions within a nucleosome in vivo. Histones H3 and H4 do not exchange out of chromatin outside S-phase. Histones H2A and H2B do exchange out of chromatin, but do so predominantly during transcription. This confers stability on the nucleosome during the cell cycle. Linker histones have a much less stable association with nucleosomal DNA, allowing for reversible activation of transcription.
- A distinct feature of histone interactions with nucleosomal DNA is the exposure of DNA on the surface of the nucleosome. One side of DNA is occluded on the histone surface, but the other is exposed and potentially accessible to other regulatory proteins.
- A major contributory factor to the functional specialization of chromatin is the capacity to target nucleosome modification and disruption.

**References**