CellPress

Review

The Hierarchy of Transcriptional Activation: From Enhancer to Promoter

Douglas Vernimmen[1,*] and Wendy A. Bickmore[2]

Regulatory elements (enhancers) that are remote from promoters play a critical role in the spatial, temporal, and physiological control of gene expression. Studies on specific loci, together with genome-wide approaches, suggest that there may be many common mechanisms involved in enhancer–promoter communication. Here, we discuss the multiprotein complexes that are recruited to enhancers and the hierarchy of events taking place between regulatory elements and promoters.

Trends

Enhancers are first primed by pioneer transcription factors (TFs).

Other TFs are likely required for subsequent events.

There is a hierarchy between enhancers and the promoters that they regulate.

Enhancers and promoters share similar properties, but differ in the characteristics and the abundance of the RNAs that they produce.

By recruiting the preinitiation complex and other proteins, enhancers have a role of increasing the concentration of the transcription machinery at target promoters.

Core Promoter

Genes transcribed by RNA polymerase (Pol)II usually have two distinct families of *cis*-acting elements: the promoter [≤1 kb from the **transcription start site (TSS)**] – composed of a core promoter [1,2] and nearby (proximal) regulatory elements [3,4], and more remote (distal) *cis*-regulatory elements (≥1 kb from TSS), which can be **enhancers**, silencers, insulators, or **locus control regions (LCRs)** [3]. The exact composition of core promoter elements may be a key determinant of enhancer–promoter specificity [5,6]. In mammalian genomes, enhancers are enriched in core promoter elements but are CpG poor, whereas promoters are generally CpG rich [7,8]. Beside the CpG content, enhancers and promoters have broad similarities and overlapping functional properties, and have been considered to form a single class of regulatory element [9].

The core promoter represents the docking site for the **general transcription factors (GTFs)**, including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH, which, together with PolII, form the **preinitiation complex (PIC)** [10]. The PIC is thought to assemble on the core promoter in a specific and sequential order that directs PolII to the nearby TSS [10]. However, this is only sufficient to direct low levels of accurately initiated transcription from DNA templates *in vitro*; a process generally referred to as basal transcription.

The first step in PIC assembly is binding of TFIID, a multisubunit complex consisting of TATA-box-binding protein (TBP) and a set of 14 TBP-associated factors (TAFs) [10]. Transcription then proceeds through a series of steps, including promoter melting, clearance, and escape, before fully functional PolII elongation is achieved. Alternative core promoter complexes may help to maintain specific transcriptional programmes in terminally differentiated cell types [11–14].

Models of transcription regulation view this as a cycle, in which complete PIC assembly is stimulated only once. After PolII escapes from the promoter, TFIID, TFIIE, TFIIH and the **mediator complex** remain on the core promoter; subsequent reinitiation then only requires *de novo* recruitment of subcomplexes comprising PolII–TFIIF and TFIIB (reviewed in [15]). The various steps of PIC assembly on a core promoter can occur with different timings during differentiation. For example, TBP is already bound to the promoters of *α1-AT*, *HNF-4α*, *VpreB1*,

[1]The Roslin Institute, Developmental Biology Division, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK
[2]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK

*Correspondence: douglas.vernimmen@roslin.ed.ac.uk (D. Vernimmen).

 http://dx.doi.org/10.1016/j.tig.2015.10.004

and λ5, long before differentiation and the transcriptional activation of these genes [16,17]. Additional transcription factors and PolII are recruited later when the genes are transcribed. The one-step recruitment of a (pre-)formed **holocomplex** at promoters has been also described [18–21]. However, it is worth noting that the right temporal window to appreciate the dynamics of PIC recruitment is often missing from most studies.

In metazoans, the transition from initiation to productive elongation is another important step that involves several levels of regulation. In a region between 30 and 60 nucleotides downstream of the TSS, PolII is often found stalled and thus paused at this site, awaiting additional signals for full elongation [22]. The release of **paused PolII** is controlled by several TFs such as the negative elongation factor (NELF), the DRB sensitivity-inducing factor (DSIF) and the transcription elongation factor P-TEFb complex (CDK9 and cyclin T). P-TEFb is part of a larger multisubunit complex, called super elongation complex (SEC) [23]. The **C-terminal domain (CTD)** of PolII plays an important role in elongation by its phosphorylation at several residues. Recently, a new multiprotein complex, termed **Integrator**, has been shown to regulate elongation by recruiting the SEC [24].

Large Protein Complexes are Bound to Promoters and Enhancers

Transcription is greatly stimulated by a second class of TFs, termed **activators**. In general, activators are sequence-specific DNA-binding proteins whose recognition sites are usually present near the core promoter and/or at enhancers. Binding of TFs at these elements usually corresponds to nucleosome-free regions (NFRs) characterised by hypersensitivity to digestion by nucleases (**DNase hypersensitive sites, DHSs**) [2,25,26]. This open-chromatin structure can be facilitated by chromatin remodelling factors, which are recruited by TFs and modify histones of the nearby nucleosomes.

Binding of activators does not stimulate transcription from chromatinised templates *in vitro*. The search for factors that stimulate activator-dependent transcription led to the identification of **co-activators** including; Mediator complexes [27,28], CREB-binding protein (CBP) [29], p300 [30], and Brahma-associated factor (BAF) [31]. TFs recruit coactivators that can then modify chromatin and/or interact with the core transcription machinery.

The large multiprotein Mediator complex can act as a bridge between transcription activators and components of the PIC [32] (see below). It appears to play important roles in many steps of transcription, including PIC formation and the transition to elongation [32]. Mediator is >1 MDa in size and >30 nm in length, with distinct structural modules and a flexible structure that changes in response to the binding of different TFs [33]. TF binding seems to induce a conformational change in Mediator that facilitates PolII binding. Different TFs bind different Mediator subunits, and Mediator complexes that lack a specific subunit can still activate transcription in response to TFs that bind to other subunits. Therefore, among other proteins (e.g., CTCF and cohesin complex) not described in this review, Mediator provides an important bridge for integrating information coming from different signalling pathways. Mediator might also provide an important binding surface for noncoding RNAs, including enhancer RNAs (eRNAs) (see below).

Other coactivators are ATP-dependent chromatin remodelling factors (such as BAF), or histone acetyltransferases (HATs) – p300/CBP. These can be part of the same complexes. ATP-dependent chromatin-remodelling families form different complexes by a combinatorial assembly of many subunits, to produce biological specificity [34]. BAF complexes, which belong to the SWI/SNF family of ATPase-dependent chromatin remodelling complexes, are involved in the relaxation of higher-order chromatin structures and in nucleosome movement and exchange [35]. The p400 SWI/SNF is associated with a HAT (TIP60) in the Tip60/p400 complex that is involved in histone (H2A/H2A.Z) exchange. CBP and its paralogue p300 are coactivator HATs

Glossary

Activator: *trans*-acting factor binding a DNA sequence to activate the transcriptional activity of a target gene.
Coactivator: non-DNA binding protein that associates with an activator and enhances transcription.
C-terminal domain (CTD): The C-terminal domain (CTD) of the largest subunit of RNA Polymerase II (PolII) consists of an array of repeats of a heptapeptide sequence (52 repeats in mammals). Amino acids in these repeats are targets for post-translational modification such as phosphorylation of serine 5 (Ser5P) - associated with early elongation (or the paused state of PolII) and Ser2P associated with full elongation.
DNase hypersensitive site (DHS): open region in the genome with increased chromatin accessibility to DNaseI that may reflect the occupation by a TF or the disruption of nucleosome structure. DHSs form NFRs.
Enhancer: regulatory sequence that increases the rate, or the probability, of transcription of a target gene. An enhancer may lie far away, upstream or downstream from the gene it regulates or may be located in an intron of its target gene or indeed in an intron of another gene.
General transcription factor (GTF): also referred to as basal transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH) that bind to core promoters.
Integrator: large coactivator complex containing at least 14 subunits with a total MW over 1 MDa. Integrator is restricted to metazoans.
Locus control region (LCR): genomic region that has the ability to confer physiological levels of tissue-specific expression on a gene linked in *cis*, independent of the integration site of the gene. An LCR can open silent chromatin.
Mediator: large coactivator complex containing 30 subunits in metazoans distributed in three modules: the head, middle, and tail. Mediator is conserved throughout all eukaryotes.
Paused PolII: after promoter escape, the engaged PolII is stalled at a pause site, waiting for further signals to progress during elongation.
Pioneer transcription factors: TFs that can bind their target sites at nucleosomal DNA. This facilitates chromatin remodelling and the