Perspectives on the RNA Polymerase II Core Promoter

James T. Kadonaga
Section of Molecular Biology, 0347, University of California, San Diego, La Jolla, CA 92093-0347, USA

Abstract

The RNA polymerase II core promoter is sometimes referred to as the gateway to transcription. The core promoter is generally defined to be the stretch of DNA that directs the initiation of transcription. This simple description belies a complex multidimensional regulatory element, as there is considerable diversity in core promoter structure and function. Core promoters can be viewed at the levels of DNA sequences, transcription factors, and biological networks. Key DNA sequences are known as core promoter elements, which include the TATA box, initiator (Inr), TCT, BRE, MTE, and DPE motifs. There are no universal core promoter elements that are present in all promoters. Different types of core promoters are transcribed by different sets of transcription factors and exhibit distinct properties, such as specific interactions with transcriptional enhancers, that are determined by the presence or absence of particular core promoter motifs. Moreover, some core promoter elements have been found to be associated with specific biological networks. For instance, the TCT motif is dedicated to the transcription of ribosomal protein genes in Drosophila and humans. In addition, nearly all of the Drosophila Hox genes have a DPE motif in their core promoters. The complexity of the core promoter is further seen in the relation among transcription initiation patterns, the stability or lability of transcriptional states, and the organization of the chromatin structure in the promoter region. Hence, the current data indicate that the core promoter is a critical component in the regulation of gene activity.

Introduction

The RNA polymerase II core promoter is generally defined to be the minimal stretch of DNA that is sufficient to direct the accurate initiation of transcription by RNA polymerase II.[1–4] Traditionally, the core promoter has been commonly thought to be a transcriptional element that functions universally and indiscriminately with different enhancers and in multiple contexts. It is now apparent, however, that the core promoter is a structurally and functionally diverse transcriptional regulatory element. Hence, in the analysis of gene activity, it is necessary to understand and to incorporate the specific components of the core promoter. In this review, I will discuss the core promoter from perspectives that have been acquired from research on transcription in Drosophila and mammals, wherein many aspects of the basic mechanisms of transcription are nearly identical.

Transcription at the core promoter is mediated by the basal transcription machinery. RNA polymerase II is a multisubunit enzyme that can synthesize RNA from the DNA template, but the purified polymerase is unable to recognize the core promoter. This process requires additional basal transcription factors that are sometimes referred to as the "general transcription factors" or "GTFs", although they do not function universally at all core promoters.[5] These basal transcription factors include TFIIA (transcription factor, RNA polymerase II, A), TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. The recognition of core

*Correspondence to: jkadonaga@ucsd.edu.

promoter sequence motifs is often performed by TFIID, which comprises the TATA-box-binding protein (TBP) and TBP-associated factors (TAFs). During the transcription process, the basal factors and RNA polymerase II assemble into a large complex at the core promoter that is termed the transcription preinitiation complex (PIC). Then, upon addition of nucleotides, transcription initiates and the polymerase elongates down the DNA template. It is important to note, however, that there are different mechanisms of basal transcription from different types of core promoters. Our current knowledge of basal transcription factors is based primarily on the analysis of TATA-dependent core promoters. It is thus likely that many other basal transcription factors remain to be identified.

CORE PROMOTER ELEMENTS

The analysis of the core promoter can be performed at the levels of the DNA sequences, protein factors, and biological functions. From the perspective of the DNA sequences, core promoters can be viewed as comprising specific DNA sequence motifs, which are known as core promoter elements or motifs, that confer distinct properties to the promoter. There are no universal core promoter elements that are found in all promoters. Some of the known core promoter motifs are shown in Figure 1. The best known core promoter element is the TATA box[6]; however, the TATA box is found in only about 10 to 20% of metazoan core promoters. It is likely that many more core promoter elements remain to be discovered, as a substantial fraction of core promoters (perhaps about one-third) lack any of the known core promoter motifs. It might also be noted that the organization of core promoter motifs in the yeast Saccharomyces cerevisiae is quite different from that in metazoans (see, for example, Ref. 7). I will summarize some of the key features of several core promoter elements that are present in metazoans. More detailed descriptions of their properties can be found elsewhere.[1–4]

The Initiator (Inr)

The Initiator (Inr) encompasses the +1 transcription start site and is probably the most commonly occurring core promoter element. The existence of a distinct sequence at the start site was described by Corden et al.,[8] and the function of the Inr as a transcriptional element was articulated by Smale and Baltimore.[9] Although many factors have been found to interact with Inr sequences, it appears that basal transcription involves the binding of TFIID to the Inr.[1] Based on functional studies that measure TFIID binding or transcriptional activity, the consensus sequences (IUPAC single nucleotide code) for the Inr are TCAKTY (A designated as +1) in Drosophila[10,11] and YYANWYY (A+1) in humans.[12] Inr consensus sequences based on computational analyses are TCAGTY (A+1) for focused promoters in Drosophila[13,14], TCA (A+1) for dispersed promoters in Drosophila[15], and YR (R+1) in humans.[16,17] [Focused and dispersed promoters are discussed below.] As a point of reference, the A in the Inr consensus is commonly designated as the +1 position in the promoter whether or not the predominant site of initiation is at this nucleotide. This convention is useful because other core promoter motifs, particularly the DPE and MTE, function cooperatively with the Inr in a manner that is sensitive to their distance to the A+1 position.

The TATA Box and BRE Motifs

The TATA box sequence is the first core promoter motif to be identified.[6] In metazoans, the TATA box consensus is TATAWAAR, with the 5′ T generally being located at −30 or −31 relative to the A+1 in the Inr. The TATA-box is bound by the TBP subunit of TFIID. It is interesting to note that the TATA box and TBP are both conserved from archaebacteria to humans.[18]

There are two BRE (TFIIB recognition element) motifs, which are located either upstream (BREu) or downstream (BREd) of a subset of TATA box elements. It is uncommon for a promoter to contain both BREu and BREd motifs. The BRE motifs act in conjunction with the TATA box. Like TBP and the TATA box, the TFIIB and the BRE are also conserved from archaebacteria to humans. The BREu (which was originally termed "BRE") as well as the BREd have been found to have both positive and negative effects on basal transcription activity.[19–21] In studies of the Drosophila Caudal protein, which is a sequence-specific enhancer-binding factor, it was found that the presence of a BREu upstream of a TATA box suppresses transcriptional activation by Caudal.[22] Thus, it would likely be interesting and important to explore the function of the BRE motifs in the context of transcriptional activation and enhancer function.

The DPE and MTE Motifs

The DPE (downstream core promoter element) was discovered in the analysis of TATA-less promoters in Drosophila.[23] The MTE (motif ten element), which is located immediately upstream of the DPE, was identified as an overrepresented core promoter sequence termed "motif 10"[13] and then found to be a functional core promoter element.[24] The MTE and the DPE are conserved from Drosophila to humans, and both motifs appear to be in close proximity to the TFIID subunits TAF6 and TAF9, which resemble histones H4 and H3.[25–27] Because the MTE and the DPE were adjacent or potentially overlapping, the relationship between these two motifs was investigated in greater detail. This study revealed that there are three key downstream functional subregions that are of particular importance for MTE and DPE function.[27] The first and second subregions (18-22 and 27–29, which are named for their downstream positions relative to the A+1 in the cognate Inr) are essential for the MTE, whereas the second and third subregions (27–29 and 30–33) constitute the DPE. Favored nucleotides for each subregion are: 18–22 (CGANC), 27–29 (CGG), and 30–33 (WYGT). Disfavored nucleotides are: 18–22 (RT - - W), 27–29 (GYA), and 30–33 (SRW -). It is interesting to note that the combination of the first and third subregions will also yield an active core promoter, although this particular arrangement, termed "Bridge", appears to occur rarely in natural promoters. From one perspective, the downstream core promoter region from +18 to +33 (relative to A+1) could be envisioned as being one continuous downstream core promoter region. However, it is likely that more MTE-like promoters will have transcriptional properties that are distinct from those of more DPE-like promoters.[24] Therefore, it is probably best to maintain the current designations of the downstream core promoter motifs.

DIFFERENT MODES OF TRANSCRIPTION INITIATION

Focused versus Dispersed Transcription Initiation

In the analysis of eukaryotic transcription, it became apparent at a relatively early stage that transcription initiation at different genes occurs in two distinct patterns, which I will refer to as "Focused" and "Dispersed" (Figure 2). In focused transcription initiation (also known as "Peaked", "Single Dominant Peak", "Narrow Peak", "Sharp Peak"), there is either a single transcription start site or a cluster of start sites in a small region of several nucleotides. Focused transcription was observed, for instance, in most viral promoters, ß-globin promoters, and histone gene promoters.[6] In dispersed transcription initiation (also known as "Broad", "Broad with Peak", "Weak Peak"), there are multiple weak start sites that are distributed over a region of about 50–100 nt. Dispersed transcription was originally seen in the promoters of housekeeping genes, which are involved in the general maintenance of cellular functions. One early example of a dispersed promoter is the mammalian HMG-CoA reductase promoter.[28] In vertebrates, another feature of housekeeping gene promoters is that they are typically located in a G/C-rich stretch of DNA known as a CpG island.[29,30] These

observations have led to the hypothesis that dispersed promoters are associated with CpG islands and housekeeping genes. As discussed below, this theory will likely be revised, particularly with respect to CpG islands.

It is important to note that the terms "Focused" and "Dispersed" refer to opposite ends of a continuum of transcriptional patterns. Hence, some promoters exhibit a dispersed pattern with one major transcription start site. Some data suggest that such promoters are more closely related to dispersed than to focused promoters.[31] For simplicity, however, I will restrict the discussion to differences between more purely focused versus dispersed promoters.

Regulated vs. Constitutive Transcription from Focused vs. Dispersed Promoters

In the initial analysis of eukaryotic transcription start sites, there appeared to be a general correlation of focused transcription with regulated genes and dispersed transcription with housekeeping genes, which typically have steady levels of transcription (Figure 2). This model was further reinforced with genome-wide analyses of promoter shape and gene expression (for example, see Ref 32). Hence, it is probably generally true that regulated genes use focused promoters, whereas housekeeping genes use dispersed promoters. From a teleological standpoint, this arrangement makes sense because a regulated gene could be turned on or off most effectively through the control of a single start site, whereas constitutive genes could maintain a steady level of transcription through the use of multiple weak start sites, any of which could be varied with only a minimal effect on the overall expression.

Postulated Mechanistic Distinctions between Focused vs. Dispersed Transcription

Ultimately, the distinctions between the different transcriptional patterns are due to differences in the mechanisms by which transcription occurs at the different promoter types. Based on biochemical studies of the transcription process (see, for example, Refs 22–24,33,34), I envision focused transcription to occur when RNA polymerase II is positioned (*i.e.*, "focused") at a specific location of the core promoter as it is assembled into the PIC. One of the key factors in the positioning of the polymerase would be TFIID, which interacts with several core promoter motifs. Then, for each round of transcription, the positioned RNA polymerase initiates transcription from either one highly preferred nucleotide or one of multiple preferred nucleotides in the immediate vicinity (Figure 3(a)). It might be noted that the preference of RNA polymerase II for initiation at specific sites can be altered by the availability of nucleotides[33] or by the introduction of specific dinucleotide primers,[35,36] which is consistent with the theory of multiple potential initiation sites from a positioned polymerase in the PIC. Hence, in this model, the RNA polymerase is positioned and start site selection is dictated by the preference of the polymerase for initiation from particular nearby nucleotides. Transcription from focused promoters typically does not involve a continuous rise in the transcription signal from each position that peaks at a particular nucleotide. Hence, the term "focused" may be a more accurate representation of the phenomenon than "peaked".

Dispersed transcription probably occurs by a substantially different mechanism than focused transcription. Dispersed transcription involves a scattered set of transcription start sites rather than a broad continuous distribution of start sites. For this reason, the term "dispersed" may be a more accurate description of the pattern of start sites than "broad". Dispersed transcription is unlikely to be due to a tandem arrangement of multiple weak focused promoters, because the characteristic sequences in focused core promoters, such as the TATA box, are typically not found in dispersed promoters. Thus, instead of positioning of the polymerase by TFIID as in focused promoters, dispersed transcription may be driven

by promoter-proximal sequence-specific DNA-binding proteins such as Sp1 and NF-Y (see, for example, Ref 37). The combination of an Sp1 binding site and an Initiator element (Inr, which encompasses the transcription start site; discussed below) can drive a low level of transcription in vitro and in cells (see, for example, Refs 38,39). Hence, dispersed promoters, which are known to contain many binding sites for factors like Sp1, may be multiple tandem and interdigitated combinations of binding sites for Sp1 (or related sequence-specific factors) and Inr-like sequences (Figure 3(b)). With regard to this model, it should be further noted that a broad range of sequences can function as Inr elements.[12,16,17] Moreover, in Drosophila, the Inr consensus for dispersed promoters is more relaxed than the Inr consensus for focused promoters.[15]. Therefore, it is likely that loose Inr-like sequences can be used in conjunction with Sp1 and other sequence-specific DNA-binding proteins in dispersed promoters. This model for dispersed transcription is not, to my knowledge, formally published but is based on experiments and ideas from scientists such as Stephen T. Smale and Robert G. Roeder.

CORE PROMOTER ELEMENTS IN BIOLOGICAL NETWORKS

The TCT Motif and the Synthesis of Ribosomal Proteins

The TCT motif (polypyrimidine initiator) (Figure 3) is a core promoter element that is dedicated to the synthesis of ribosomal proteins as well as other proteins involved in translation.[40] The TCT motif is conserved from Drosophila to humans and has a consensus of YYC+1TTTYY in Drosophila and YC+1TYTYY in humans, where C+1 is the transcription start site. The name "TCT motif" derives from the TCT nucleotides that are commonly found from positions −1 to +2. This sequence has also been referred to as the "polypyrimidine initiator" in earlier studies of the mouse *rpS16* gene promoter.[41,42] The TCT motif overlaps with but is distinct from the sequences that encode the 5′ terminal oligopyrimidine tract (5′TOP), which is an oligopyrimidine stretch at the 5′ end of ribosomal protein mRNAs that is involved in the regulation of translation.[43]

The TCT motif appears to be present in nearly all ribosomal protein gene promoters in Drosophila and humans.[40] Out of 52 Drosophila ribosomal gene promoters with high confidence transcription start site data, 44 promoters have a perfect match and alignment to the YYC+1TTTYY consensus, seven promoters have a single mismatch with perfect alignment, and one promoter has two mismatches with perfect alignment.[40] Given that the YYC+1TTTYY consensus would occur once in every 4096 nt in a random sequence, it is remarkable that this motif is found in nearly all ribosomal gene core promoters.

The TCT motif consensus sequence (for example, in Drosophila: YYCTTTYY) is similar to the Inr consensus (in Drosophila: TCAKTY) (Figure 4(a)). The Inr consensus fits within the TCT consensus except that the TCT motif has a TCT (or CCT) trinucleotide where the Inr has a TCA trinucleotide. Given that the TCT motif and Inr both encompass the transcription start site and are nearly identical in sequence, it seemed possible that these two elements are functionally related. Contrary to this hypothesis, the TCT motif cannot substitute for an Inr for basal transcription activity (Figure 4(b)). If, however, the TCT trinucleotide in the TCT motif is mutated to TCA to make it match the Inr consensus (Figure 4(c)), then the mutant TCA version of the TCT motif can substitute for an Inr (Figure 4(b)). In addition, TFIID, which is known to bind to the Inr, is not able to bind to promoters containing a TCT motif, but is able to bind to promoters containing the TCA mutant version of the TCT motif (Figure 4(b)). These results indicate that the TCT-based transcription system functions by a fundamentally different mechanism than the well-established Inr-based transcription system that involve TFIID binding to the Inr. It is also notable that a single T versus A difference in the sequences of the TCT motif and Inr is responsible for their distinct transcriptional properties.

From a broader perspective, these studies show that the TCT-based transcription system is dedicated to the synthesis of ribosomal proteins (Figure 5(a)). Thus, the TCT transcription system complements the RNA polymerase I and RNA polymerase III transcription systems, which are primarily dedicated to the synthesis of ribosomal RNAs and transfer RNAs. Together, these three transcription systems provide the RNA species that are needed for synthesis of the ribosome and transfer RNAs.

From the more limited perspective of the core promoter, the TCT system provides an example in which a core promoter element functions as a key component of a biological regulatory network – in this case, transcription of the ribosomal protein genes. Because the synthesis of the ribosomal protein genes must be coordinated, it makes sense, from a teleological standpoint, for these genes to contain a shared core promoter element.

The TCT motif is also an excellent example of a rare but biologically important core promoter element. It is estimated that there are about 120 TCT motif sequences in Drosophila. There may be many other rare and yet-to-be-identified core promoter elements with important functions in other gene networks.

The Drosophila Hox Genes Are a DPE-based Network

The Drosophila homeotic (Hox) genes, which are involved in the specification of segment identity, are a DPE-based network of genes.[22] Nearly all of the core promoters of the Hox genes contain DPE motifs (Figure 5(b)). The promoters lacking a DPE motif are those associated with the evolutionarily most recent genes, *Ubx* and *Abd-A*. These findings suggested that transcription factors that regulate the Hox genes may act in a DPE-specific manner. The investigation of this hypothesis led to the finding that Caudal protein, which is a sequence-specific enhancer-binding factor and a key regulator of the Hox genes, is a DPE-specific activator. Caudal mediates strong activation of DPE-dependent promoters, weak activation of TATA-dependent promoters, and little or no activation of promoters containing both a BREu and a TATA box.[22] These findings suggest that early versions of the Hox gene network contained only DPE-dependent core promoters and may have been regulated by a set of transcription factors, such as Caudal, that function specifically with the DPE motif.

Caudal provides an example of a sequence-specific transcriptional activator that functions specifically with the DPE. The mechanism by which Caudal functions as a DPE-specific activator is an interesting question that remains to be solved. This process may involve the recruitment of transcription factors such as NC2 (also known as Dr1-Drap1) and Mot1 (also known as BTAF1 and Hel89B), which have been found to promote DPE-dependent transcription and to inhibit TATA-dependent transcription.[44,45]

From a more general perspective, enhancer-core promoter specificity provides a mechanism by which sequence-specific transcription factors can be linked to their desired target genes in regulatory networks. Aside from the studies on Caudal, enhancer elements have been found to activate transcription in a DPE- or TATA-specific manner,[46,47] but the DPE- and TATA-specific activators were not identified. In the future, it will be particularly informative to find additional core-promoter-specific factors like Caudal.

ADDITIONAL PERSPECTIVES ON THE CORE PROMOTER

The Core Promoter and the Response of Target Genes to p53

The tumor suppressor protein p53, which is a sequence-specific transcription factor, activates its various target genes with different kinetics. Target genes that are involved in the control of the cell cycle, such as *CDKN1A* (formerly known as *p21*), are induced rapidly, whereas target genes that promote apoptosis, such as *FAS*, are induced slowly. The rapidly

transcribed *CDKN1A* promoter has a TATA box that facilitates efficient PIC assembly, whereas the slow responding *FAS* promoter lacks a TATA box and displays inefficient PIC assembly.[48] These findings suggest that core promoter structure is an important component of the p53 transcriptional response. It will be interesting and important to see whether other p53 target genes exhibit analogous properties.

RNA Polymerase II Pausing/Stalling and the Pause Button

Studies by Lis and coworkers revealed the existence of paused, transcriptionally-engaged RNA polymerase II approximately 25 nt downstream of the transcription start site of the Drosophila hsp70 and other genes.[49,50] More recently, genome-wide analyses have shown that the pausing of polymerase at about 20–50 nt downstream of the start site occurs commonly in many genes (for reviews, see Refs 51–53). A sequence, termed the pause button (PB), has been found to be enriched in core promoters of genes that exhibit a stalled polymerase.[54] The location and consensus of the PB are similar to those of the DPE.[54,55] It is thus possible that the DPE participates in pausing by RNA polymerase II in downstream promoter region. In this regard, it should be noted that the TAF6 and TAF9 subunits of TFIID are in close proximity to the DPE.[36] TAF6 and TAF9 resemble histones H4 and H3, and may interact with the DNA in the downstream core promoter region like an H3-H4 tetramer and promote pausing by the polymerase.

It is also useful to note that promoters with a stalled/paused polymerase exhibit some enrichment for focused promoters.[31,55] In Drosophila, promoters with a paused/stalled polymerase have an approximately two-fold higher percentage of focused promoters (55%) than that seen in all promoters.[31] Nevertheless, a moderate proportion (44%) of paused/stalled promoters have a dispersed transcription pattern.

CpG Islands and Transcriptional Patterns

In vertebrates, the cytosine bases in CG ("CpG" in the classic nucleotide nomenclature) dinucleotides are generally methylated to give 5-methylcytosine. When such methylated CpG dinucleotides are deaminated, the 5-methylcytosine is converted into thymine, which is not efficiently repaired. Hence, in vertebrates, CpG dinucleotides are generally underrepresented with a notable exception being "CpG islands" (also known as "CGI"), which are stretches of DNA with an average length of about 1 kbp that are rich in G+C bases with CpG dinucleotides that are commonly unmethylated.[30]

In vertebrates, it appears that over half of known promoters are in a CpG island. CpG island promoters often have a dispersed transcription initiation pattern and are also found with housekeeping genes. Hence, the hypothesis emerged that dispersed promoters are associated with housekeeping genes that are present in CpG islands. This model was generally accepted until the analysis of focused versus dispersed transcription was examined on a genome-wide scale in Drosophila, an organism that is mostly deficient in CpG methylation. Because of the low levels of CpG methylation in Drosophila, CpG dinucleotides occur at a normal frequency throughout the genome and CpG islands (*i.e.*, regions of DNA that are enriched for CpG dinucleotides) do not exist. In Drosophila, studies of transcription patterns revealed that both focused and dispersed promoters exist in the absence of CpG islands.[31,32] Hence, there is not an obligatory link between CpG islands and dispersed transcription.

In the absence of CpG methylation, there are distinguishing features between focused and dispersed promoters in Drosophila.[15,32] Relative to dispersed promoters, focused promoters are enriched in regulated genes as well as in core promoter motifs such as the TATA, Inr, and DPE. On the other hand, relative to focused promoters, dispersed promoters are

enriched in housekeeping genes and exhibit a higher degree of chromatin organization. This latter point is discussed in the following section.

Chromatin Structure and the Lability of Gene Expression States

Genome-wide analyses of chromatin structure have shown that many promoters exhibit a distinct pattern of positioned nucleosomes relative to the transcription start site (see, for example, Refs 56,57) (Figure 6). Different studies have examined the localization of bulk nucleosomes, histone variants such as H2A.Z, and covalent histone modifications such as trimethylated lysine residue 4 of histone H3 (H3K4me3), but for the present discussion, I will not make a distinction between these different forms of histones and chromatin.

The combined analysis of nucleosome positioning and transcription patterns revealed that dispersed promoters generally have strong nucleosome positioning, whereas focused promoters generally exhibit little or no nucleosome positioning[31] (Figure 6). Further comparison of chromatin structure and specific core promoter elements showed that promoters with TATA box, Inr, MTE, or DPE motifs have little or no nucleosome positioning. Given that these motifs are generally associated with focused promoters, these results are consistent with the observation that focused promoters generally do not have an ordered chromatin structure. On the other hand, core promoters with a TCT motif, which are mostly focused promoters, are associated with a highly ordered chromatin structure.[31] These results further reinforce the differences between TCT-based versus Inr-based core promoters. The apparently anomalous property of the focused TCT-containing promoters may be due to the function of the TCT motif in the transcription of ribosomal protein genes, which are housekeeping genes. These findings therefore suggest that a highly ordered nucleosome structure is associated with genes in which there is a stable program of expression. In such cases, the promoters might exist as steady state systems. Conversely, regulated genes that have changeable expression programs have little or no order in their chromatin structure.

A summary of chromatin structure, transcription patterns, core promoter elements, and gene expression properties is shown in Figure 6. The term "Stable" reflects a kinetic resistance to change in the level of gene expression, whereas "Labile" indicates a facile change in the state of gene expression. One might describe a "labile" promoter as being "poised" for rapid activation or repression. The mechanisms by which the different chromatin states are established remain to be determined. Nevertheless, it is reasonable to hypothesize that transcription factors recognize the promoter sequences and establish the basic transcriptional state and that the presence or absence of curved DNA sequences that promote nucleosome formation contributes to whether or not there is an ordered array of nucleosomes in the promoter region. In addition, interactions between transcription factors and chromatin, such as the binding of the TAF3 subunit of TFIID to trimethylated histone H3 lysine 4[58], are likely to be important factors that influence the chromatin structure in the vicinity of the promoter.

It is also relevant to note that an extensive analysis of chromatin-associated proteins in Drosophila led to the classification of five types of chromatin, which were termed BLACK, BLUE, GREEN, RED, and YELLOW.[59] BLACK, BLUE, and GREEN chromatin are different forms of heterochromatin, whereas RED and YELLOW chromatin are different types of euchromatin. Notably, RED chromatin is enriched in a Drosophila sequence-specific promoter-binding protein termed the GAGA factor (or GAF), whereas YELLOW chromatin is deficient in GAGA factor.[59] In addition, it was found that focused promoters in Drosophila are enriched for GAGA factor and GAGA binding sites, whereas dispersed promoters are depleted in GAGA factor and GAGA binding sites.[31] Altogether, these data suggest that labile promoters are found in RED chromatin, while inert promoters are found

in YELLOW chromatin. Consistent with this hypothesis, the ribosomal protein genes, which contain the TCT motif in their core promoters, are found in YELLOW chromatin.

TFIID and Cellular Differentiation

There is diversity not only in the constituents of the core promoter, such as the TATA box, Inr, and DPE, but also in the factors that interact with the core promoter (for reviews, see Refs 60,61). The variation of the basal transcriptional machinery adds another dimension to the complexity of regulation of the transcription process. For example, the TBP subunit of TFIID is the canonical TATA box-binding protein, but there are also TBP-related factors (TRFs). TRF3 (also known as TBP2 and TBPL2) is present in vertebrates, and can bind to TATA box sequences. In the analysis of the differentiation of myoblasts to myotubes, it was found that the canonical TBP-containing TFIID complex is replaced by a complex containing TRF3 and TAF3[62] (but also see Refs 61,63). In addition, studies of liver development revealed a substantial depletion of TFIID (TBP and TAFs) but not TFIIB, TFIIE, TFIIH, and RNA polymerase II during differentiation of hepatoblasts to hepatocytes.[64] These findings indicate that the factors that mediate basal transcription can change upon cellular differentiation. In the particular case of TFIID, it is possible that TFIID may function in a more "pluripotent" manner in different cellular contexts, and that more specialized factors may substitute for TFIID in differentiated cells with more restricted and strictly defined transcriptional programs.

Practical Applications of the Core Promoter

There are practical applications of the core promoter. First, because of the potential for enhancer-core promoter specificity, the study of any transcriptional enhancer would be best carried out in conjunction with its cognate core promoter. Second, high levels of transcription can be achieved with an optimized core promoter that has been termed the super core promoter.[65] Third, a reporter gene with multiple core promoter elements can be used in studies involving the characterization of a large number of transcriptional enhancers.[66] For example, the presence of both TATA and DPE motifs in the core promoter would ensure that both TATA-specific and DPE-specific enhancers would function in conjunction with the reporter gene. In the future, the range of applications of the core promoter in research, biotechnology, and medicine will undoubtedly increase.

Conclusion

The core promoter is sometimes referred to as the "gateway to transcription", as it is the site at which RNA polymerase II initiates its journey downstream the DNA template. Some of the many facets of the core promoter are as follows.

- The core promoter itself is a structurally and functionally diverse transcriptional element (Figure 3). Different types of core promoters function with different sets of transcription factors and hence have distinct transcriptional properties (Figure 4). There are probably many core promoter motifs that remain to be discovered. As shown in the example of the TCT motif, rare core promoter elements can be of considerable biological importance. In addition, the identities and functions of many of the basal transcription factors that are used at different types of core promoters remain to be determined.

- Specific core promoter motifs can have key roles in biological networks (Figure 5). The TCT-based transcription system for the synthesis of ribosomal proteins complements the RNA polymerase I and RNA polymerase III transcription systems for the synthesis of ribosomal and transfer RNAs. The Drosophila Hox genes are a DPE-based network, and Caudal, a regulator of the Hox genes, is a DPE-specific

transcription factor. It is likely that other core promoter motifs, many of which remain to be discovered, will exhibit specific functions in other biological networks.

- The core promoter is probably involved in nearly every transcriptional process that occurs in the promoter region. One example is transcriptional pausing and the potential relation between the pause button and the DPE core promoter motif. The core promoter is also likely to have a key role in divergent transcription, wherein transcription has been observed to initiate in both directions from the core promoter region (for review, see Ref 67).

- The focused and dispersed transcription initiation patterns (Figure 1) probably reflect two distinct mechanisms of transcription that occur from two different types of core promoters (Figure 2). There is a general correlation of the transcription of housekeeping genes (a stable transcriptional state) with an ordered chromatin structure as well as a correlation of the transcription of regulated genes (a labile transcriptional state) with a lack of an ordered chromatin structure (Figure 6). Current data suggest that CpG islands are not functionally linked to dispersed transcription or to transcription of housekeeping genes. The TATA, Inr, MTE, DPE, and TCT core promoter motifs are found in focused promoters. These core promoter elements are associated with regulated transcription, with the exception of the TCT motif, which is associated with housekeeping genes (Figure 6).

The core promoter is a diverse and complex regulatory element that is at the heart of the transcription process. At the present time, we have identified many sequences, factors, and biological systems that are associated with core promoter function, yet many fundamental questions remain unanswered. For instance, what is the basic mechanism of transcription from dispersed promoters? How many other core promoter elements remain to be discovered? What are the factors that mediate transcription from TCT-dependent core promoters? How many sequence-specific enhancer-binding factors function specifically with a particular core promoter motif? What is the mechanism by which enhancer-core promoter specificity is achieved? The pursuit of these questions will lead to many exciting and important discoveries and will ultimately provide a framework upon which our knowledge of transcriptional regulation can be assembled.

Acknowledgments

References

1. Smale ST, Kadonaga JT. The RNA polymerase II core promoter. Annu Rev Biochem. 2003; 72:449–479. [PubMed: 12651739]

2. Juven-Gershon T, Kadonaga JT. Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol. 2010; 339:225–229. [PubMed: 19682982]

3. Ohler U, Wassarman DA. Promoting developmental transcription. Development. 2010; 137:15–26. [PubMed: 20023156]

4. Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA. Mammalian RNA polymerase II core promoters: insights from genome-wide studies. Nat Rev Genet. 2007; 8:424–436. [PubMed: 17486122]

5. Thomas MC, Chiang CM. The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol. 2006; 41:105–178. [PubMed: 16858867]

6. Goldberg, ML. PhD Dissertation. Stanford University; 1979. Sequence analysis of Drosophila histone genes.

7. Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene. 2007; 389:52–65. [PubMed: 17123746]

8. Corden J, Wasylyk B, Buchwalder A, Sassone-Corsi P, Kedinger C, Chambon P. Promoter sequences of eukaryotic protein-coding genes. Science. 1980; 209:1406–1414. [PubMed: 6251548]

9. Smale ST, Baltimore D. The "initiator" as a transcription control element. Cell. 1989; 57:103–113. [PubMed: 2467742]

10. Purnell BA, Emanuel PA, Gilmour DS. TFIID sequence recognition of the initiator and sequences farther downstream in Drosophila class II genes. Genes Dev. 1994; 8:830–842. [PubMed: 7926771]

11. Chalkley GE, Verrijzer CP. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J. 1999; 18:4835–4845. [PubMed: 10469661]

12. Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol. 1994; 14:116–127. [PubMed: 8264580]

13. Ohler U, Liao GC, Niemann H, Rubin GM. Computational analysis of core promoters in the Drosophila genome. Genome Biol. 2002; 3:RESEARCH0087. [PubMed: 12537576]

14. FitzGerald PC, Sturgill D, Shyakhtenko A, Oliver B, Vinson C. Comparative genomics of Drosophila and human core promoters. Genome Biol. 2006; 7:R53. [PubMed: 16827941]

15. Ni T, Corcoran DL, Rach EA, Song S, Spana EP, Gao Y, Ohler U, Zhu J. A paired-end sequencing strategy to map the complex landscape of transcription initiation. Nat Methods. 2010; 7:521–527. [PubMed: 20495556]

16. Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engström PG, Frith MC, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet. 2006; 38:626–635. [PubMed: 16645617]

17. Frith MC, Valen E, Krogh A, Hayashizaki Y, Carninci P, Sandelin A. A code for transcription initiation in mammalian genomes. Genome Res. 2008; 18:1–12. [PubMed: 18032727]

18. Reeve JN. Archaeal chromatin and transcription. Mol Microbiol. 2003; 48:587–598. [PubMed: 12694606]

19. Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev. 1998; 12:34–44. [PubMed: 9420329]

20. Evans R, Fairley JA, Roberts SG. Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. Genes Dev. 2001; 15:2945–2949. [PubMed: 11711430]

21. Deng W, Roberts SG. A core promoter element downstream of the TATA box that is recognized by TFIIB. Genes Dev. 2005; 19:2418–2423. [PubMed: 16230532]

22. Juven-Gershon T, Hsu JY, Kadonaga JT. Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. Genes Dev. 2008; 22:2823–2830. [PubMed: 18923080]

23. Burke TW, Kadonaga JT. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev. 1996; 10:711–724. [PubMed: 8598298]

24. Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT. The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev. 2004; 18:1606–1617. [PubMed: 15231738]

25. Burke TW, Kadonaga JT. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes Dev. 1997; 11:3020–3031. [PubMed: 9367984]

26. Shao H, Revach M, Moshonov S, Tzuman Y, Gazit K, Albeck S, Unger T, Dikstein R. Core promoter binding by histone-like TAF complexes. Mol Cell Biol. 2005; 25:206–219. [PubMed: 15601843]

27. Theisen JW, Lim CY, Kadonaga JT. Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. Mol Cell Biol. 2010; 30:3471–3479. [PubMed: 20457814]

28. Reynolds GA, Basu SK, Osborne TF, Chin DJ, Gil G, Brown MS, Goldstein JL, Luskey KL. HMG CoA reductase: a negatively regulated gene with unusual promoter and 5' untranslated regions. Cell. 1984; 38:275–285. [PubMed: 6088070]

29. Zhu J, He F, Hu S, Yu J. On the nature of human housekeeping genes. Trends Genet. 2008; 24:481–484. [PubMed: 18786740]

30. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011; 25:1010–1022. [PubMed: 21576262]

31. Rach EA, Winter DR, Benjamin AM, Corcoran DL, Ni T, Zhu J, Ohler U. Transcription initiation patterns indicate divergent strategies for gene regulation at the chromatin level. PLoS Genet. 2011; 7:e1001274. [PubMed: 21249180]

32. Hoskins RA, Landolin JM, Brown JB, Sandler JE, Takahashi H, Lassmann T, Yu C, Booth BW, Zhang D, Wan KH, et al. Genome-wide analysis of promoter architecture in Drosophila melanogaster. Genome Res. 2011; 21:182–192. [PubMed: 21177961]

33. Kadonaga JT. Assembly and disassembly of the Drosophila RNA polymerase II complex during transcription. J Biol Chem. 1990; 265:2624–2631. [PubMed: 1689291]

34. Kamakaka RT, Tyree CM, Kadonaga JT. Accurate and efficient RNA polymerase II transcription with a soluble nuclear fraction derived from Drosophila embryos. Proc Natl Acad Sci USA. 1991; 88:1024–1028. [PubMed: 1992453]

35. Samuels M, Fire A, Sharp PA. Dinucleotide priming of transcription mediated by RNA polymerase II. J Biol Chem. 1984; 259:2517–2525. [PubMed: 6583201]

36. Zenzie-Gregory B, O'Shea-Greenfield A, Smale ST. Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. J Biol Chem. 1992; 267:2823–2830. [PubMed: 1733976]

37. Kadonaga JT. Regulation of RNA polymerase II transcription by sequence-specific DNA-binding factors. Cell. 2004; 116:247–257. [PubMed: 14744435]

38. Smale ST, Schmidt MC, Berk AJ, Baltimore D. Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proc Natl Acad Sci USA. 1990; 87:4509–4513. [PubMed: 2141169]

39. Emami KH, Burke TW, Smale ST. Sp1 activation of a TATA-less promoter requires a species-specific interaction involving transcription factor IID. Nucleic Acids Res. 1998; 26:839–846. [PubMed: 9443978]

40. Parry TJ, Theisen JW, Hsu JY, Wang YL, Corcoran DL, Eustice M, Ohler U, Kadonaga JT. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. Genes Dev. 2010; 24:2013–2018. [PubMed: 20801935]

41. Hariharan N, Perry RP. Functional dissection of a mouse ribosomal protein promoter: significance of the polypyrimidine initiator and an element in the TATA-box region. Proc Natl Acad Sci USA. 1990; 87:1526–1530. [PubMed: 2304915]

42. Chung S, Perry RP. Cell-free transcription of a mouse ribosomal-protein-encoding gene: the effects of promoter mutations. Gene. 1991; 100:173–180. [PubMed: 2055468]

43. Hamilton TL, Stoneley M, Spriggs KA, Bushell M. TOPs and their regulation. Biochem Soc Trans. 2006; 34:12–16. [PubMed: 16246169]

44. Willy PJ, Kobayashi R, Kadonaga JT. A basal transcription factor that activates or represses transcription. Science. 2000; 290:982–985. [PubMed: 11062130]

45. Hsu JY, Juven-Gershon T, Marr MT 2nd, Wright KJ, Tjian R, Kadonaga JT. TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription. Genes Dev. 2008; 22:2353–2358. [PubMed: 18703680]

46. Ohtsuki S, Levine M, Cai HN. Different core promoters possess distinct regulatory activities in the Drosophila embryo. Genes Dev. 1998; 12:547–556. [PubMed: 9472023]

47. Butler JE, Kadonaga JT. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. Genes Dev. 2001; 15:2515–2519. [PubMed: 11581157]

48. Morachis JM, Murawsky CM, Emerson BM. Regulation of the p53 transcriptional response by structurally diverse core promoters. Genes Dev. 2010; 24:135–147. [PubMed: 20040571]

49. Rougvie AE, Lis JT. The RNA polymerase II molecule at the 5' end of the uninduced *hsp70* gene of D. melanogaster is transcriptionally engaged. Cell. 1988; 54:795–804. [PubMed: 3136931]

50. Rougvie AE, Lis JT. Postinitiation transcriptional control in Drosophila melanogaster. Mol Cell Biol. 1990; 10:6041–6045. [PubMed: 2172790]

51. Fuda NJ, Ardehali MB, Lis JT. Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature. 2009; 461:186–192. [PubMed: 19741698]

52. Levine M. Paused RNA polymerase II as a developmental checkpoint. Cell. 2011; 145:502–511. [PubMed: 21565610]

53. Li J, Gilmour DS. Promoter proximal pausing and the control of gene expression. Curr Opin Genet Dev. 2011; 21:231–235. [PubMed: 21324670]

54. Hendrix DA, Hong JW, Zeitlinger J, Rokhsar DS, Levine MS. Promoter elements associated with RNA pol II stalling in the Drosophila embryo. Proc Natl Acad Sci USA. 2008; 105:7762–7767. [PubMed: 18505835]

55. Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell. 2010; 143:540–551. [PubMed: 21074046]

56. Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, et al. Nucleosome organization in the Drosophila genome. Nature. 2008; 453:358–362. [PubMed: 18408708]

57. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. Dynamic regulation of nucleosome positioning in the human genome. Cell. 2008; 132:887–898. [PubMed: 18329373]

58. Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HTM. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell. 2007; 131:58–69. [PubMed: 17884155]

59. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell. 2010; 143:212–224. [PubMed: 20888037]

60. D'Alessio JA, Wright KJ, Tjian R. Shifting players and paradigms in cell-specific transcription. Mol Cell. 2009; 36:924–931. [PubMed: 20064459]

61. Goodrich JA, Tjian R. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. Nat Rev Genet. 2010; 11:549–558. [PubMed: 20628347]

62. Deato MD, Tjian R. Switching of the core transcription machinery during myogenesis. Genes Dev. 2007; 21:2137–2149. [PubMed: 17704303]

63. Gazdag E, Santenard A, Ziegler-Birling C, Altobelli G, Poch O, Tora L, Torres-Padilla ME. TBP2 is essential for germ cell development by regulating transcription and chromatin condensation in the oocyte. Genes Dev. 2009; 23:2210–2223. [PubMed: 19759265]

64. D'Alessio JA, Ng R, Willenbring H, Tjian R. Core promoter recognition complex changes accompany liver development. Proc Natl Acad Sci USA. 2011; 108:3906–3911. [PubMed: 21368148]

65. Juven-Gershon T, Cheng S, Kadonaga JT. Rational design of a super core promoter that enhances gene expression. Nat Methods. 2006; 3:917–922. [PubMed: 17124735]

66. Pfeiffer BD, Jenett A, Hammonds AS, Ngo TT, Misra S, Murphy C, Scully A, Carlson JW, Wan KH, Laverty TR, et al. Tools for neuroanatomy and neurogenetics in Drosophila. Proc Natl Acad Sci USA. 2008; 105:9715–9720. [PubMed: 18621688]

67. Seila AC, Core LJ, Lis JT, Sharp PA. Divergent transcription: a new feature of active promoters. Cell Cycle. 2009; 8:2557–2564. [PubMed: 19597342]
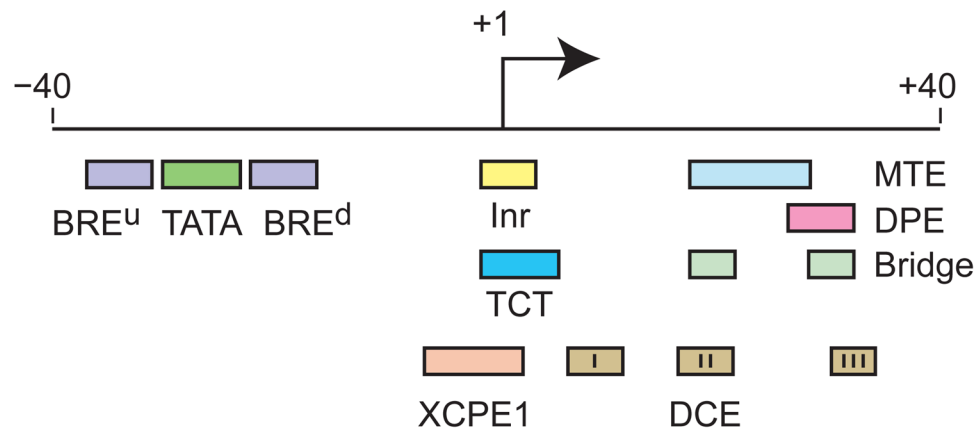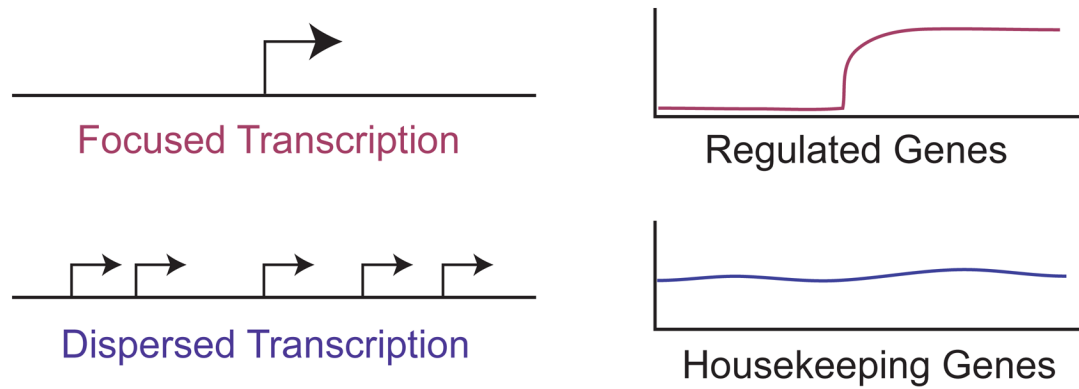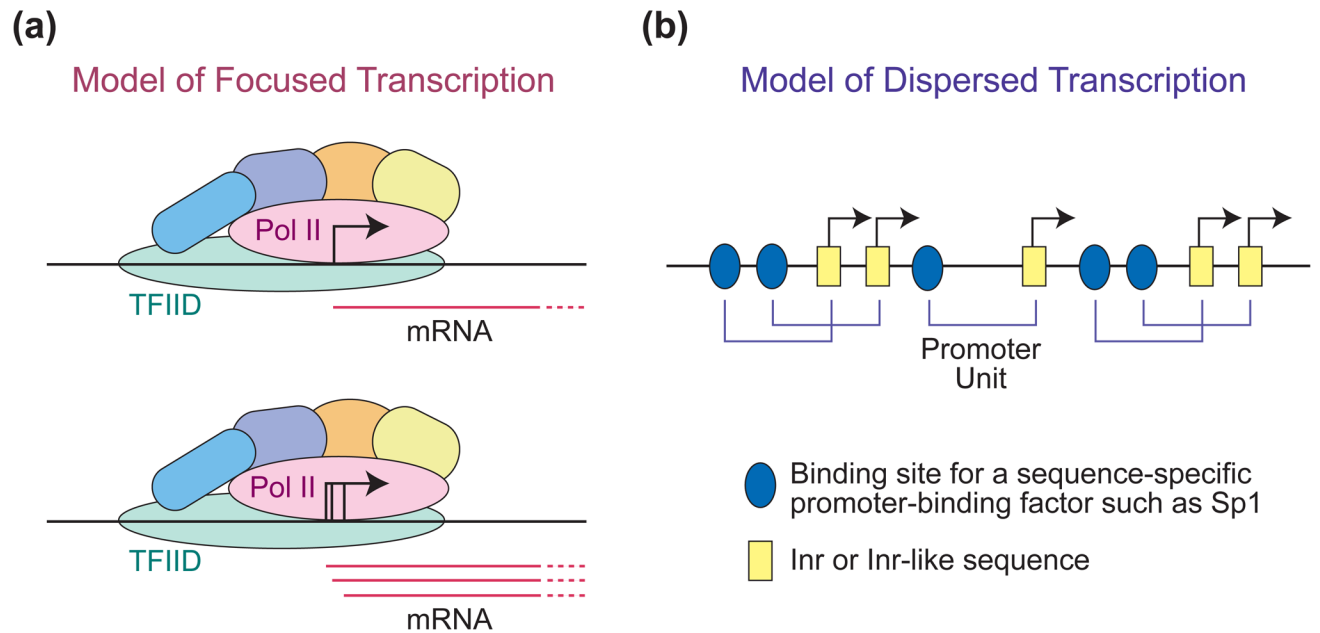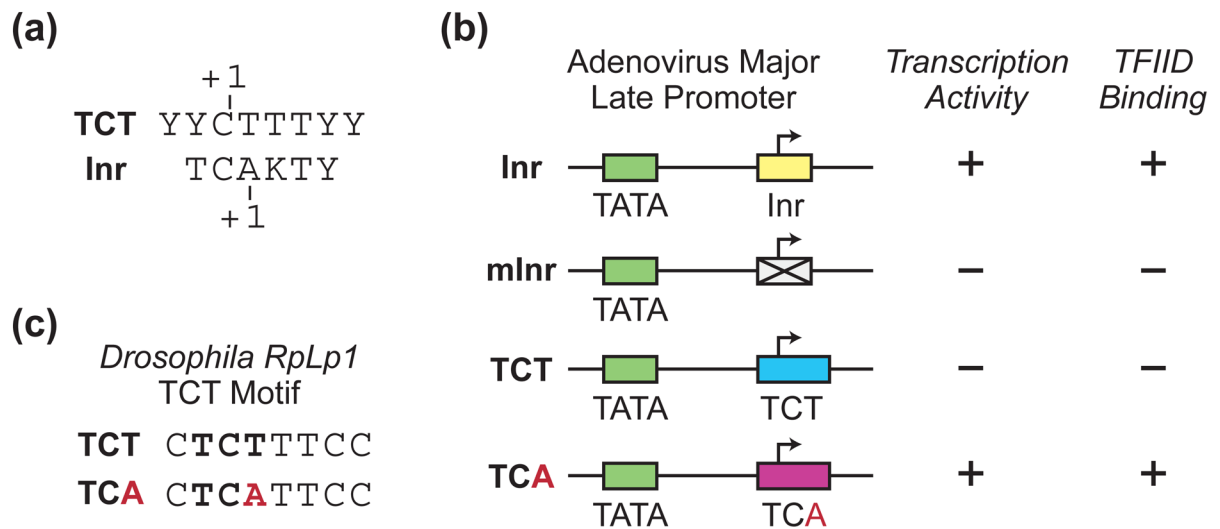
Core Promoter Elements

FIGURE 1.

Some core promoter elements for transcription by RNA polymerase II. The locations of the motifs are drawn roughly to scale. The BREu, TATA, Inr, MTE, DPE, and TCT motifs have been found in both Drosophila and humans. These motifs are typically found in focused core promoters, although there are probably Inr-like elements in dispersed promoters. There are no universal core promoter elements that are found in all promoters. Moreover, it is likely that many other core promoter motifs remain to be discovered. The functional properties of a core promoter are determined by the presence or absence of specific core promoter motifs. For example, some enhancers will activate transcription from DPE-dependent core promoters but not from TATA-dependent core promoters.
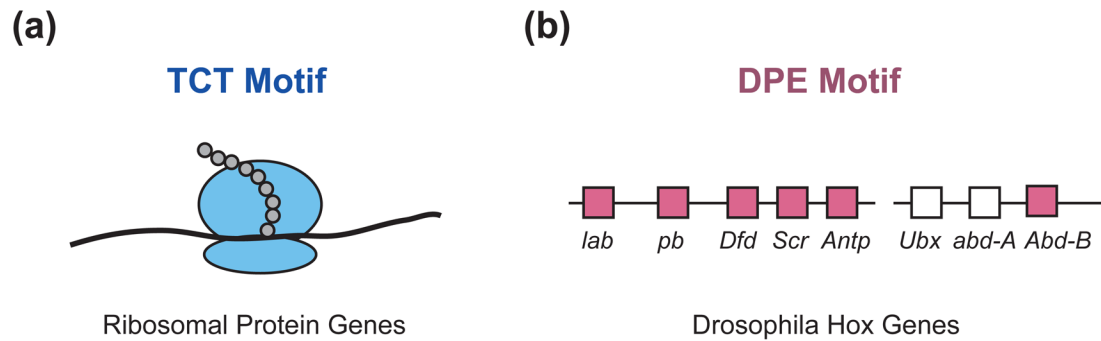
FIGURE 2.
Focused and dispersed modes of transcription initiation. In focused transcription, there is either a single predominant transcription start site or a cluster of start sites in a small region of several nucleotides. In dispersed transcription, there are multiple weak start sites that are distributed over a larger region of about 50 to 100 nucleotides. Focused promoters are generally associated with regulated genes, whereas dispersed promoters are typically found in housekeeping genes, which maintain steady levels of transcription.

(a)

Model of Focused Transcription

(b)

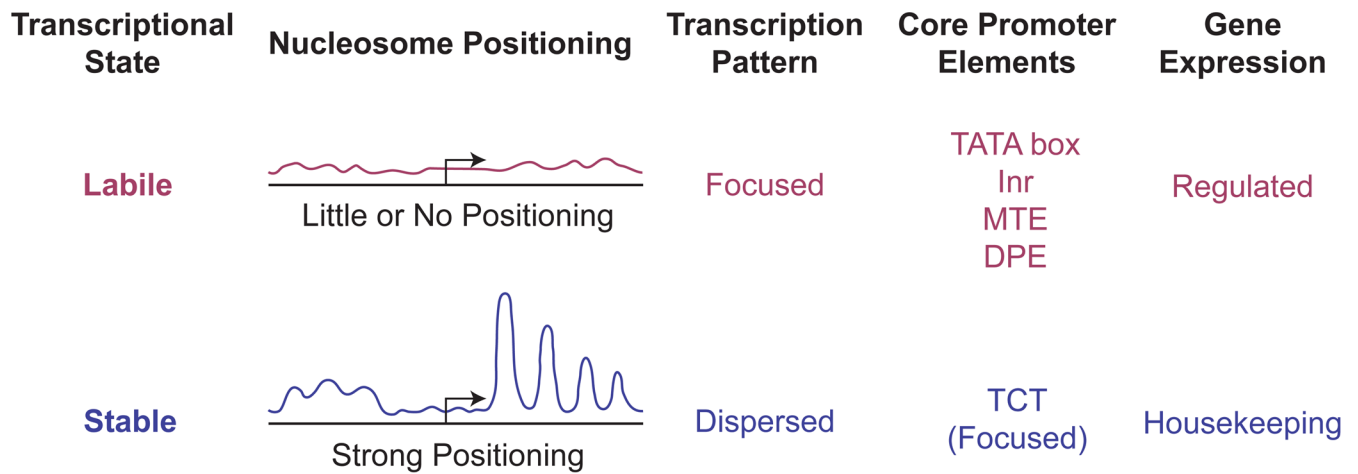Model of Dispersed Transcription

FIGURE 3.
Hypothetical mechanisms for transcription from focused and dispersed promoters. **(a)** In the model for focused promoters, the transcription preinitiation complex (PIC) is assembled at a single location in the core promoter, and start site selection is determined by the preference of RNA polymerase II to initiate transcription at nearby nucleotides. In some cases, there may be a strongly favored nucleotide and a single predominant start site is observed. In other cases, transcription might initiate at multiple nucleotides that are not necessarily adjacent to one another. **(b)** In the model for dispersed transcription, there are multiple weak promoters that are created by the combination of a binding site of a sequence-specific DNA-binding factor, such as Sp1 or NF-Y, and an Inr-like sequence. Thus, a single promoter unit would be a sequence-specific factor binding site and an Inr-like sequence. These promoter units can be arranged in tandem as well as interdigitated.

(a)

```
       +1
        |
TCT  YYCTTTYY
Inr   TCAKTY
        |
       +1
```

(c)

Drosophila RpLp1
TCT Motif

```
TCT  CTCTTTCC
TCA  CTCATTCC
```

(b)

FIGURE 4.
A single T versus A nucleotide distinguishes the TCT motif from the Inr. **(a)** The TCT consensus sequence closely resembles the Inr consensus. The +1 start sites that are typically used with TCT-dependent and Inr-dependent core promoters are indicated. **(b)** The TCT motif cannot substitute for an Inr. The wild-type adenovirus major late core promoter has an Inr motif. Mutation of the Inr results in the loss of transcriptional activity as well as binding by purified TFIID. Substitution of the Inr with the TCT motif also leads to the loss of transcriptional activity and TFIID binding. However, mutation of the TCT motif to fit the Inr consensus (TCA) results in the restoration of transcriptional activity and TFIID binding. These experiments reveal that a single T versus A nucleotide results in different activities of the TCT and Inr motifs.[40] The inability of the TCT motif to substitute for an Inr was also observed in two different DPE-containing promoters.[40] **(c)** Sequence of the mutant TCA version of the TCT motif.

(a)

TCT Motif

Ribosomal Protein Genes

(b)

DPE Motif

lab   *pb*   *Dfd* *Scr* *Antp*   *Ubx* *abd-A* *Abd-B*

Drosophila Hox Genes

FIGURE 5.
Core promoter motifs in biological networks. **(a)** The TCT motif is dedicated to the synthesis of ribosomal proteins.[40] The TCT-based transcription system for the synthesis of ribosomal proteins complements the RNA polymerase I and RNA polymerase III transcription systems for the synthesis of ribosomal RNAs and transfer RNAs. **(b)** The Drosophila Hox genes, except for *Ubx* and *abd-A*, have TATA-less, DPE-dependent core promoters.[22] The DPE motifs in the indicated genes are conserved from Drosophila melanogaster to Drosophila virilis. *Ubx* and *abd-A*, the two most evolutionarily recent Hox genes, lack both TATA and DPE motifs in their core promoters. *lab*, *labial; pb*, *proboscipedia; Dfd*, *Deformed; Scr*, *Sex combs reduced; Antp*, *Antennapedia; Ubx*, *Ultrabithorax; abd-A*, *abdominal-A;* Abd-*B*, *Abdominal-B*.

Transcriptional State	Nucleosome Positioning	Transcription Pattern	Core Promoter Elements	Gene Expression
Labile	Little or No Positioning	Focused	TATA box Inr MTE DPE	Regulated
Stable	Strong Positioning	Dispersed	TCT (Focused)	Housekeeping

FIGURE 6.

Chromatin structure and transcriptional state. Promoters can be viewed from the perspective of the transcriptional state of the gene. In the case of housekeeping genes, the transcriptional state is stable and the flow of factors through the promoter region can be viewed as that of a steady state system. This stable transcriptional state correlates with an ordered chromatin structure in the promoter region, wherein a positioned array of nucleosomes is typically observed. Regulated genes are in a labile transcriptional state in which they are poised for rapid activation or repression. This labile state correlates with little or no apparent order in the chromatin structure in the promoter region. Focused promoters, particularly those with TATA, Inr, MTE, or DPE motifs, are generally associated with regulated genes and a lack of positioned nucleosomes in the promoter region. Dispersed promoters, on the other hand, are commonly found in housekeeping genes with positioned nucleosomes in the promoter region. One notable exception to this general trend is the TCT motif-based transcription system, in which focused promoters mediate the transcription of ribosomal protein genes, which are housekeeping genes. Unlike the TATA-, Inr-, MTE-, and DPE-containing promoters, TCT-dependent promoters exhibit strong nucleosome positioning.[31] These findings suggest that the key feature that dictates the chromatin structure in the promoter region is the stability or lability of the transcriptional state.