understand how different permutations of the same regulatory elements alter gene expression. An understanding of how the combinatorial organization of a promoter encodes regulatory information first requires an overview of the proteins that constitute the transcriptional machinery.

THE EUKARYOTIC TRANSCRIPTIONAL MACHINERY

Factors involved in the accurate transcription of eukaryotic protein-coding genes by RNA polymerase II can be classified into three groups: general (or basic) transcription factors (GTFs), promoter-specific activator proteins (activators), and coactivators (**Figure 2**). GTFs are necessary and can be sufficient for accurate transcription initiation in vitro (reviewed in 141). Such factors include RNA polymerase II itself and a variety of auxiliary components, including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. In addition to these "classic" GTFs, it is apparent that in vivo transcription also requires Mediator, a highly conserved, large multisubunit complex that was originally identified in yeast (reviewed in 38, 119).

GTFs assemble on the core promoter in an ordered fashion to form a transcription preinitiation complex (PIC), which directs RNA polymerase II to the transcription start site (TSS). The first step in PIC assembly is binding of TFIID, a multisubunit complex consisting of TATA-box-binding protein (TBP) and a set of tightly bound TBP-associated factors (TAFs). Transcription then proceeds through a series of steps, including promoter melting, clearance, and escape, before a fully functional RNA polymerase II elongation complex is formed. The current model of transcription regulation views this as a cycle, in which complete PIC assembly is stimulated only once. After RNA polymerase II escapes from the promoter, a scaffold structure, composed of TFIID, TFIIE, TFIIH, and Mediator, remains on the core promoter

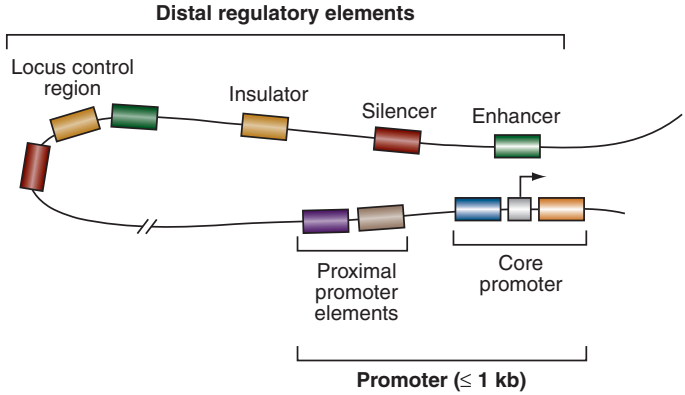

Figure 1

Schematic of a typical gene regulatory region. The promoter, which is composed of a core promoter and proximal promoter elements, typically spans less than 1 kb pairs. Distal (upstream) regulatory elements, which can include enhancers, silencers, insulators, and locus control regions, can be located up to 1 Mb pairs from the promoter. These distal elements may contact the core promoter or proximal promoter through a mechanism that involves looping out the intervening DNA.

(73); subsequent reinitiation of transcription then only requires rerecruitment of RNA polymerase II-TFIIF and TFIIB.

The assembly of a PIC on the core promoter is sufficient to direct only low levels of accurately initiated transcription from DNA templates in vitro, a process generally referred to as basal transcription. Transcriptional activity is greatly stimulated by a second class of factors, termed activators. In general, activators are sequence-specific DNA-binding proteins whose recognition sites are usually present in sequences upstream of the core promoter (reviewed in 149). Many classes of activators, discriminated by different DNA-binding domains, have been described, each associating with their own class of specific DNA sequences. Examples of activator families include those containing a cysteine-rich zinc finger, homeobox, helix-loop-helix (HLH), basic leucine zipper (bZIP), forkhead, ETS, or Pit-Oct-Unc (POU) DNA-binding domain (reviewed in 142). In addition to a sequence-specific DNA-binding domain, a typical activator also contains a separable activation domain that is required for the activator to stimulate transcription (149). An

General transcription factor (GTF): a factor that assembles on the core promoter to form a preinitiation complex and is required for transcription of all (or almost all) genes

Coactivators: adaptor proteins that typically lack intrinsic sequence-specific DNA binding but provide a link between activators and the general transcriptional machinery

PIC: preinitiation complex

TSS: transcription start site