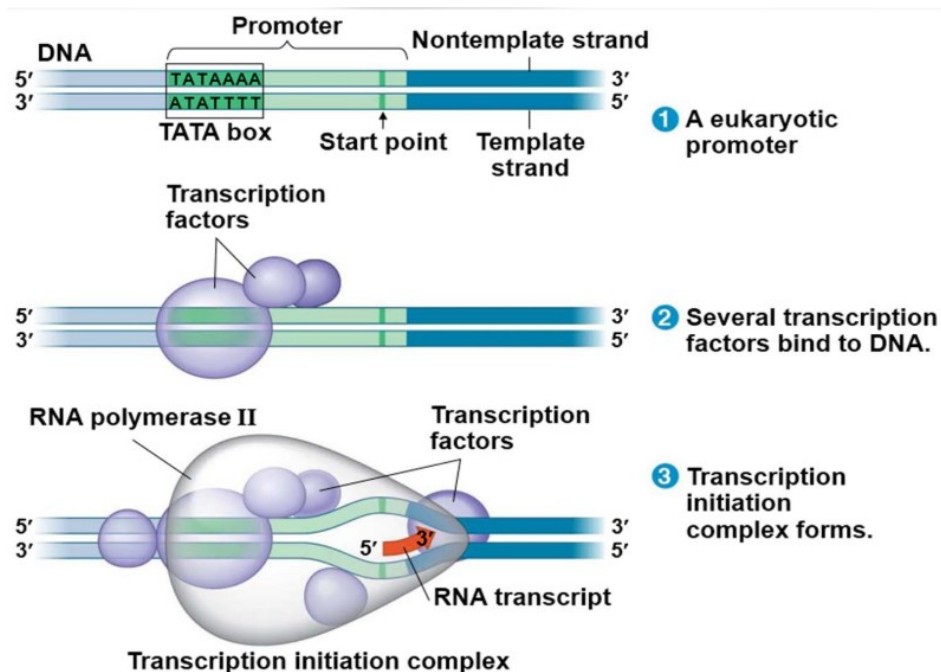


PROMOTERS

3.1 Definition and general organisation

A promoter is a DNA region that supports the initiation of transcription. In protein-coding genes, the promoter generally lies at or near the transcription start site and contains binding sites for proteins involved in assembling and regulating the transcription machinery.

A promoter should not be regarded as a single short DNA motif. In many eukaryotic genes, it comprises a core promoter and surrounding regulatory sequences, including proximal promoter elements. These components work together to determine where transcription begins and how strongly a gene is expressed.



3.2 Core promoter

The core promoter is the minimal DNA region capable of supporting accurate transcription initiation in an appropriate biochemical context. It is usually located around the transcription start site and provides the binding or recognition platform for general transcription factors and RNA polymerase II.

The core promoter is not identical in all genes. Different genes use different combinations of sequence elements, and some lack the commonly described TATA box.

The major core promoter elements are described below.

A. TATA box

The TATA box is an AT-rich DNA sequence, often represented by the consensus sequence TATAAA. In many TATA-containing promoters, it lies approximately 25–30 base pairs upstream of the transcription start site, although the exact position varies.

The TATA box is recognised by the TATA-binding protein (TBP), a component of the general transcription factor TFIID. TBP binding bends the DNA and helps organise the assembly of the transcription pre-initiation complex.

Its functions include:

- Helping position the transcription machinery.
- Contributing to the selection of the transcription start site.
- Facilitating the assembly of general transcription factors.
- Supporting efficient transcription initiation in promoters that contain a functional TATA box.

The TATA box is not universal. Many genes are transcribed from TATA-less promoters.

B. Initiator element (Inr)

The initiator is a core promoter element that overlaps the transcription start site. It helps specify the initiation region and can function in conjunction with other promoter elements.

The Inr is recognised through interactions involving components of TFIID, including TBP-associated factors. It is important in certain promoters that lack a TATA box.

C. TFIIB recognition element (BRE)

The BRE is a promoter sequence recognised by transcription factor IIB (TFIIB). Depending on its position and sequence context, it can contribute to the recruitment and organisation of the transcription initiation complex. Different BRE variants may have different regulatory effects.

D. Downstream promoter element (DPE)

The DPE is located downstream of the transcription start site. It can cooperate with the initiator and other core promoter components to support transcription initiation, particularly in certain TATA-less promoters.

These elements do not necessarily occur together. Their combinations differ among genes and contribute to promoter diversity.

3.3 Proximal promoter elements

The proximal promoter is the region surrounding the core promoter that contains additional binding sites for regulatory transcription factors. It can influence the efficiency of transcription and the response of a gene to cellular signals.

Important examples include:

- CAAT box: A sequence motif found in many promoters. It can bind transcription factors such as NF-Y and contribute to promoter activity.
- GC box: A GC-rich sequence that can bind transcription factors such as Sp1. GC boxes are common in many promoters, including some promoters that lack a TATA box.
- Other transcription factor-binding sites: These may recognise proteins that respond to hormones, growth factors, developmental signals or tissue-specific regulatory programmes.

The term proximal promoter does not refer to a universally fixed length of DNA. The boundaries vary among genes and experimental definitions.

3.4 Mechanism of transcription initiation at a promoter

Transcription initiation involves several coordinated molecular events.

1. Recognition of promoter DNA: General transcription factors and regulatory proteins recognise promoter elements.
2. Pre-initiation complex assembly: TFIID and other general transcription factors help recruit and position RNA polymerase II.
3. Pre-initiation complex assembly: TFIID and other general transcription factors help recruit and position RNA polymerase II.
4. DNA opening: TFIIH and associated factors contribute to local DNA unwinding.
5. RNA synthesis begins: RNA polymerase II initiates RNA synthesis at the transcription start site.
6. Promoter escape and elongation

The formation of the pre-initiation complex is influenced by the local chromatin structure and by activators or repressors associated with the promoter. Transcription is therefore not determined by promoter sequence alone; it also depends on the regulatory proteins available in the cell.

3.5 Types of promoters

Constitutive promoters: These support relatively continuous gene expression in many cell types. However, constitutive does not mean that expression is completely constant or unaffected by cellular conditions.

Inducible promoters: These respond to a particular stimulus. For example, heat-shock promoters may become more active when cells experience elevated temperatures.

Tissue-specific promoters: These are preferentially active in particular tissues because of the transcription factors and regulatory environment present there.

Developmentally regulated promoters: Their activity changes during embryonic development, growth or cellular differentiation.

TATA-containing and TATA-less promoters: This classification is based on the presence or absence of a conventional TATA box. It reflects differences in promoter architecture rather than a simple distinction between strong and weak promoters.

3.6 Biological significance of promoters

Promoters are essential for defining the initiation of transcription. They help determine which RNA transcript is produced, where transcription begins and how the transcriptional machinery responds to regulatory signals.

Alternative promoters can also allow a single gene to produce different transcript variants. Depending on the promoter used, transcription may begin at different sites, resulting in transcripts with different first exons or 5' untranslated regions. This can influence transcript stability, translation and, in some cases, the protein product.

Promoter mutations may reduce or increase transcription, alter transcription start sites or change the responsiveness of a gene to regulatory factors. Such mutations can cause disease even when the coding sequence remains unchanged.

4. ENHANCERS

4.1 Definition and discovery

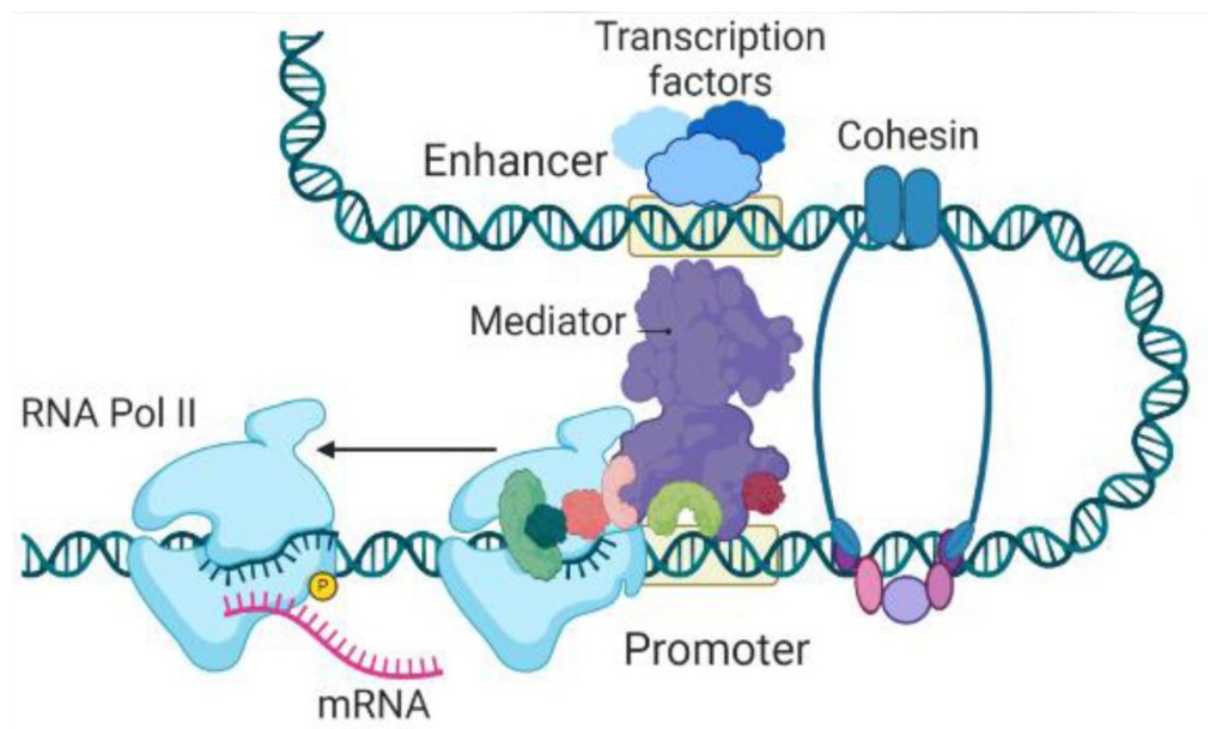
Enhancers are cis-regulatory DNA sequences that increase transcription from compatible target promoters through interactions with sequence-specific transcription factors and associated regulatory proteins.

Enhancers were first identified in studies of viral gene regulation, including work on simian virus 40 (SV40) in 1981. Researchers found that certain DNA sequences stimulated transcription even when placed at a distance from a promoter or in a reversed orientation. Subsequent studies demonstrated that enhancer-mediated regulation is widespread in eukaryotic genomes.

4.2 Structural organisation of enhancers

An enhancer commonly consists of multiple short DNA motifs that bind different transcription factors. These motifs are arranged in a modular fashion, so the enhancer's function depends on the combination, arrangement and activity of the bound proteins.

An enhancer is not defined by a single universal consensus sequence. Rather, its regulatory activity results from the interaction of several DNA-binding proteins and cofactors.



4.3 Important characteristics of enhancers

1. Distance independence

An enhancer can influence a promoter located far away in the linear DNA sequence. Enhancers may lie several kilobases from their target genes, and some act across much greater distances.

2. Orientation independence

Many enhancers retain their activity when their sequence is reversed. This property distinguishes them from the typical directional organisation of promoters, although enhancer function is not identical in every genomic context.

3. Flexible genomic location

Enhancers can occur upstream or downstream of genes, within introns, and sometimes within other transcribed regions. Their genomic position alone does not establish which gene they regulate.

4. Tissue specificity

Some enhancers are active only in particular cell types because they require specific transcription factors. For example, an enhancer that depends on muscle-specific regulatory proteins may be active in skeletal muscle but inactive in many other tissues.

5. Combinatorial control

Enhancers integrate information from several transcription factors. Some factors activate transcription, while others restrict activity in inappropriate cells or developmental stages.

6. Long-range communication

Enhancers frequently interact with promoters through three-dimensional chromatin organisation. These interactions can be dynamic and may be influenced by transcription factors, chromatin-associated proteins and genomic boundaries.

4.4 Mechanism of enhancer-mediated gene activation

The activity of an enhancer can be explained in the following sequence.

Step 1: Recognition of enhancer DNA. Activator proteins bind to specific DNA motifs within the enhancer. Pioneer transcription factors can, in some cases, bind relatively inaccessible chromatin and help establish a regulatory region.

Step 2: Recruitment of coactivators. Bound transcription factors recruit proteins such as p300/CBP and other coactivator complexes. These proteins may facilitate histone acetylation, chromatin remodelling and interactions with the transcriptional machinery.

Step 3: Formation of an active regulatory region. Nucleosome positioning and histone modifications change in ways that can support transcription factor binding and enhancer activity. The exact changes vary with the cell type and enhancer.

Step 4: Communication with the promoter. The enhancer interacts functionally with its target promoter. DNA looping is a major model for explaining how distant DNA regions can communicate. Mediator and cohesin can contribute to this organisation in particular contexts.

Step 5: Stimulation of transcription. Enhancer-associated factors influence the assembly or activity of the transcriptional machinery at the promoter. The outcome may involve increased

initiation, changes in promoter-proximal pausing or improved productive elongation, depending on the gene.

4.5 Enhancer–promoter interaction

The relationship between enhancers and promoters is central to eukaryotic gene regulation. A gene may have several enhancers, each contributing to a different aspect of its expression.

For example, one enhancer may regulate gene expression in the developing limb, while another controls expression in the nervous system. Both can act on the same promoter, but they respond to different transcription factors and developmental signals.

The specificity of enhancer–promoter interactions depends on several factors:

- The transcription factors present in a cell.
- Compatibility between enhancer-bound factors and the target promoter.
- The three-dimensional arrangement of chromatin.
- Regulatory boundaries and neighbouring genomic regions.
- The activity of chromatin-remodelling and transcriptional cofactor complexes.

An enhancer located near a gene does not necessarily regulate that gene. Some enhancers skip the nearest promoter and regulate a more distant gene, making functional testing important.

4.6 Types of enhancers based on biological function

Tissue-specific enhancers: These activate genes preferentially in a particular tissue. They are important for establishing the distinct gene-expression programmes of muscle, liver, brain and other tissues.

Developmental enhancers: These regulate genes at particular stages of development. They contribute to the formation of body structures and the differentiation of embryonic cells.

Hormone-responsive enhancers: These contain binding sites for hormone-regulated transcription factors. For example, glucocorticoid receptor-binding regulatory regions can contribute to transcriptional responses to glucocorticoids.

Signal-responsive enhancers: These respond to signalling pathways activated by growth factors, cytokines or environmental changes.

Super-enhancers: These are large clusters of active regulatory elements that are strongly associated with particular cell identities or cell states. The term describes a functional and analytical category, not a completely separate basic class of DNA sequence. Super-enhancers and LCRs can overlap conceptually or physically, but the terms are not interchangeable.

4.7 Examples of enhancer function

Example 1: SV40 enhancer

SV40 contains an enhancer that stimulates transcription of viral genes. Its activity demonstrated that certain DNA elements can increase transcription even when separated from the promoter or reversed in orientation.

Example 2: Drosophila even-skipped gene

The even-skipped gene encodes a developmental transcription factor required for normal embryonic pattern formation. Different enhancers contribute to expression in particular stripes of the early embryo. This modular regulation is an important example of how enhancers establish spatial patterns of gene expression.

Example 3: Immunoglobulin gene regulation

Enhancers associated with immunoglobulin genes help regulate their expression in B lymphocytes. Their activity depends on combinations of transcription factors associated with B-cell identity and developmental state.

4.8 Biological significance of enhancers

Enhancers provide flexibility and precision to gene regulation. They allow genes to respond to different developmental cues in different tissues without requiring a separate promoter for every expression pattern. They also help integrate multiple signalling pathways and coordinate the expression of genes necessary for cellular differentiation.

Changes in enhancer sequences can cause abnormal gene expression, sometimes resulting in developmental disorders or disease. Enhancer dysfunction may occur through mutations in transcription factor-binding motifs, deletions of regulatory regions or alterations in chromatin organisation.

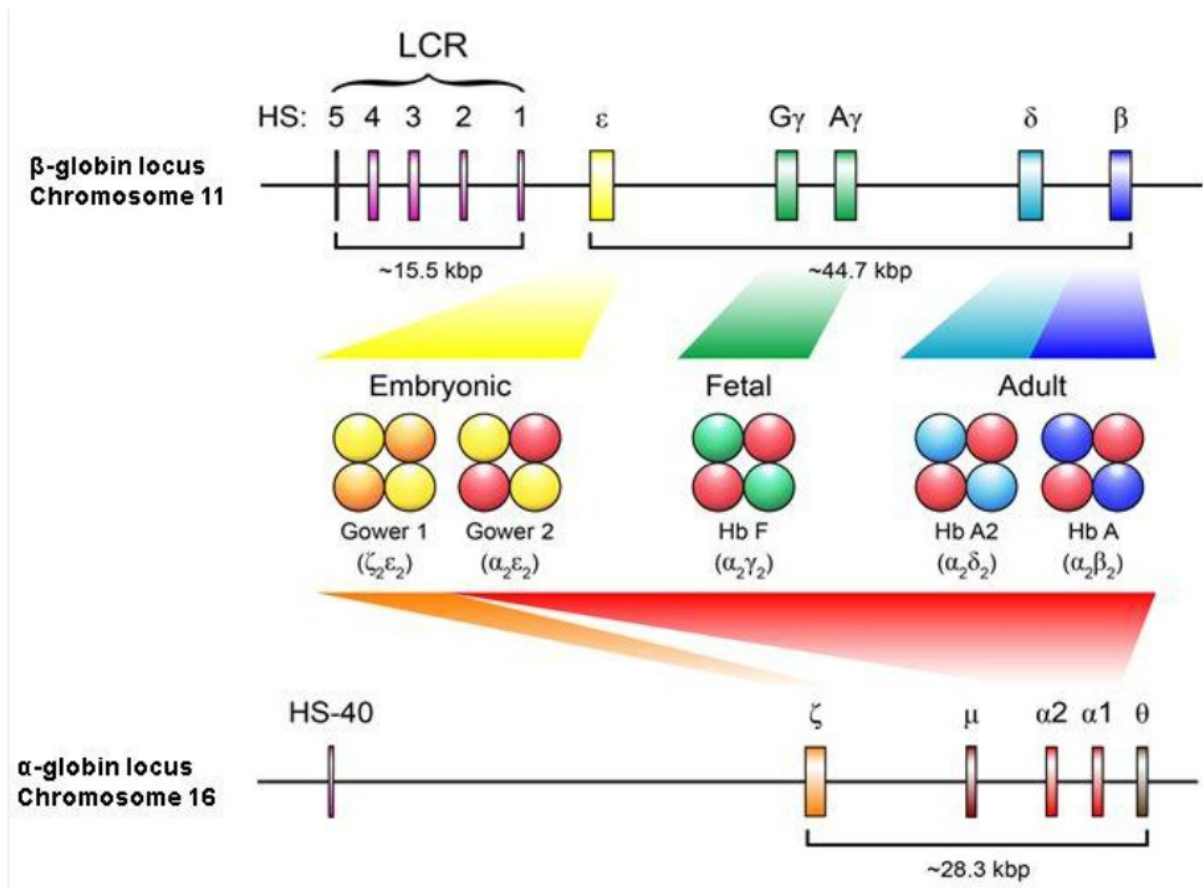
5. LOCUS CONTROL REGIONS (LCRS):

5.1 Definition and discovery

A locus control region (LCR) is a complex cis-regulatory region that helps regulate the expression of one or more linked genes within a genomic locus. LCRs are especially associated with high-level, tissue-specific expression and with reducing certain chromosomal position effects on linked genes.

The concept of LCRs emerged from studies of the mammalian β -globin gene cluster. Researchers observed that the appropriate expression of globin genes depended not only on the individual promoters of these genes but also on a distant regulatory region upstream of the cluster.

LCRs are generally composed of multiple regulatory elements rather than a single short DNA motif. These elements interact with transcription factors and chromatin-associated proteins to establish an appropriate regulatory environment for the linked genes.



5.3 Major characteristics of LCRs

1. Long-range regulatory activity

LCRs can influence genes separated from them by substantial stretches of DNA. Their action involves interactions within the three-dimensional structure of chromatin.

2. High-level gene expression

LCRs can support strong transcription of linked genes in appropriate cell types. However, they do not guarantee maximum expression under all circumstances; the activity of the linked promoters and the availability of regulatory proteins remain important.

3. Tissue-specific activity

Many LCRs function preferentially in particular tissues. For example, the β -globin LCR is associated with strong globin gene expression in erythroid cells, where the relevant transcription factors are present.

4. Chromatin accessibility

LCRs are associated with accessible regulatory chromatin and the recruitment of proteins that support transcription. They can help create a permissive chromatin environment across a region.

5. Position-effect resistance

When a gene is introduced into a chromosome, its expression may be influenced by the surrounding genomic environment. This is called a position effect. In suitable experimental contexts, an LCR can reduce such effects and help support more consistent expression of an associated transgene.

6. Coordination of linked genes

LCRs can participate in regulating a cluster of related genes. However, this does not mean that all genes in the cluster are expressed simultaneously. Developmental switches, promoter compatibility and competition for regulatory interactions help determine which genes are active.

5.4 Mechanism of LCR-mediated regulation

The function of an LCR involves coordinated interactions among DNA sequences, transcription factors, chromatin regulators and gene promoters.

Step 1: Binding of transcription factors. Tissue-specific and general regulatory proteins bind to motifs within the LCR. In erythroid cells, factors such as GATA1 and KLF1 contribute to the regulation of the globin locus.

Step 2: Recruitment of regulatory proteins. The bound transcription factors recruit coactivators and chromatin-remodelling complexes. These proteins contribute to an accessible and transcriptionally permissive regulatory environment.

Step 3: Establishment of regulatory chromatin. Histone modifications, nucleosome organisation and other chromatin features help determine the accessibility and activity of the locus.

Step 4: Long-range communication. The LCR interacts with the promoter of an appropriate target gene through three-dimensional chromatin organisation. These interactions can vary with developmental stage and cell type.

Step 5: Selective gene activation. Developmental regulators and promoter-associated factors determine which linked gene receives productive regulatory input. The LCR contributes to strong expression of the selected gene.

Step 6: Maintenance of appropriate expression. Interactions among the LCR, target promoters and chromatin-associated proteins help sustain the appropriate gene-expression programme for the cell's developmental or physiological state.

This mechanism is a general model rather than a universal sequence of obligatory steps. Different LCRs can operate through distinct combinations of regulatory processes.

5.5 β -globin LCR: an important example

The β -globin locus is one of the most extensively studied examples of long-range gene regulation. The locus contains globin genes that encode haemoglobin chains expressed at different developmental stages.

The major developmental expression pattern is:

- Embryonic stage: The ϵ -globin gene contributes to embryonic haemoglobin production.

- Fetal stage: The γ -globin genes produce the γ chains required for fetal haemoglobin (HbF), which contains two α and two γ chains.
- Adult stage: The β -globin gene contributes to adult haemoglobin (HbA), while the δ -globin gene contributes to a smaller fraction of adult haemoglobin (HbA₂).

The β -globin LCR contributes to high-level expression of the appropriate globin genes in erythroid cells. It interacts with globin gene promoters as part of the regulatory network controlling developmental gene switching.

It is important to note that the LCR alone does not determine the entire switching process. Promoter sequences, transcription factors, chromatin organisation and additional regulatory elements all contribute.

5.6 Biological significance of LCRs

LCRs are important because they allow coordinated regulation of genes that occupy the same genomic region. They help establish appropriate chromatin conditions and support strong, tissue-specific expression. In experimental genetics, LCRs have been used to improve the reliability of transgene expression in suitable systems.

Mutations or deletions affecting an LCR may reduce expression of linked genes, disrupt normal developmental expression or contribute to inherited disease. In the β -globin locus, regulatory abnormalities can affect haemoglobin production even when the protein-coding sequence of an individual globin gene remains intact.

6. Relationship between promoters, enhancers and LCRs

Promoters, enhancers and LCRs should be understood as interacting components of a regulatory system rather than as completely independent elements.

A promoter establishes the site and basic machinery for transcription initiation. Enhancers provide additional regulatory input, often from a distance, by recruiting activators and coactivators. An LCR comprises multiple regulatory components that can influence chromatin organisation and the expression of genes across a genomic locus.

For example, at the β -globin locus, the promoter of a globin gene provides the transcription initiation platform, while the LCR contributes to high-level erythroid expression through long-range regulatory interactions. Developmental transcription factors determine which globin genes are active at different stages.

The activity of a gene therefore depends on the combined effects of its promoter, distal regulatory sequences, transcription factors and chromatin environment